

# Expanded newborn screening

A review of the evidence



Hilary Burton  
Sowmiya Moorthie

May 2010

[www.phgfoundation.org](http://www.phgfoundation.org)

## PHG Foundation Team

Dr Hilary Burton	<i>Programme Director</i>
Dr Victoria Fearn	<i>Specialist Registrar in Public Health</i>
Ms Alison Hall	<i>Project Manager</i>
Ms Jane Lane	<i>Administrator</i>
Dr Sowmiya Moorthie	<i>Project Co-ordinator</i>
Dr Gurdeep Sagoo	<i>Epidemiologist</i>
Dr Simon Sanderson	<i>Principal Associate</i>

## Authorship

Most of the report was written by Hilary Burton and Sowmiya Moorthie. Simon Sanderson led the design and conduct of the systematic review which was carried out by Gurdeep Sagoo and Sowmiya Moorthie. Andrew Morris and Mark Sharrard contributed case histories for the target conditions, and Anita MacDonald and Hazel Rogozinski provided examples of dietary management. The section on economic evaluation was written by Philip Shackley. Alison Hall contributed the chapter on ethical, legal and social aspects. The chapter on the patient viewpoint was written by Steve Hannigan and colleagues at Climb and the Chapter on EU policy on rare disorders was provided by Alastair Kent, Melissa Hillier and David Brown at the Genetic Interest Group.

## Acknowledgements

The Project Team would like to express their thanks for the expert guidance provided by Dr Jim Bonham, Professor Rodney Pollitt, Ms Melanie Downing, Dr Jacqui Calvin, Dr Andrew Morris and Dr Mark Sharrard on the technological, laboratory and clinical perspectives and by the management committee of the CLAHRC-SY research project on newborn screening. Thanks also to Victoria Fearn for her assistance with the systematic review and Jane Lane for administrative support.

The review was funded by the National Institute of Health Research - Collaboration for Leadership in Applied Health Research and Care - South Yorkshire (CLAHRC-SY)

*This report can be downloaded from our website:*  
**[www.phgfoundation.org](http://www.phgfoundation.org)**

Published by PHG Foundation  
2 Worts Causeway  
Cambridge  
CB1 8RN  
UK  
Tel: +44 (0) 1223 740200  
Fax: +44 (0) 1223 740892

*Second edition, May 2010*

© 2010 PHG Foundation  
ISBN 978-1-907198-03-8

*Cover photo: <http://www.flickr.com/photos/tamakisono>*

The PHG Foundation is the working name of the Foundation for Genomics and Population Health, an independent charitable organisation (registered in England and Wales, charity No. 1118664 company No. 5823194), which works with partners to achieve better health through the responsible and evidence-based application of biomedical science.

# Contents

<b>Executive summary</b>	<b>2</b>
<b>Glossary of acronyms</b>	<b>7</b>
<b>Glossary of terms</b>	<b>8</b>
<b>Chapter 1      Introduction</b>	<b>9</b>
<b>Chapter 2      Tandem mass spectrometry</b>	<b>12</b>
<b>Chapter 3      Clinical and epidemiological overview of the five selected inherited metabolic diseases</b>	<b>17</b>
<b>Chapter 4      Evaluating screening and newborn screening: an overview</b>	<b>42</b>
<b>Chapter 5      Review methods</b>	<b>49</b>
<b>Chapter 6      Summary of findings from Health Technology Assessment reports identified by the systematic review</b>	<b>53</b>
<b>Chapter 7      Evaluation of expanded screening programmes worldwide</b>	<b>63</b>
<b>Chapter 8      Patient perspective</b>	<b>111</b>
<b>Chapter 9      Ethical, legal, and social implications of expanding newborn screening</b>	<b>119</b>
<b>Chapter 10      Expanded newborn screening programme and EU policy towards rare diseases</b>	<b>128</b>
<b>Chapter 11      Expanded newborn screening: discussion and recommendations</b>	<b>131</b>
<b>References</b>	<b>154</b>
<b>Appendices</b>	<b>162</b>
1 UK National Screening Committee criteria	
2 Search strategies	

## Executive summary

An understanding of rare inherited metabolic diseases requires an extraordinary level of detailed knowledge of their underlying molecular and biochemical pathology in individuals. At the same time their rarity and variability make them very hard to study in population terms to provide the sort of robust epidemiological evidence usually demanded for national screening programmes. This report attempts to bridge the gap between detailed evidence and broad conclusions for policy making. However, it is almost impossible to simplify and summarise the content in a satisfactory way. We recommend that readers with limited time will find it helpful to read the Executive Summary and Chapter 11 in which the five conditions studied are reviewed against the National Screening Committee Criteria.

Inherited metabolic disorders (IMDs) are a group of disorders each caused by deficient activity in a single enzyme in a pathway of intermediary metabolism. Interruption of the pathway can lead to accumulation of unwanted or toxic products or, sometimes, a deficiency of products essential for health. These changes typically result in metabolic crisis and death, or damage to many organ systems including the brain, leading to severe learning or physical disability often at an early age in those affected.

The abnormal biochemical pathway produces characteristic molecular markers that can be detected by tandem mass spectrometry (MS/MS). MS/MS is an analytical technique that can directly identify and quantify compounds in a sample based on their molecular mass-to-charge ratio. Thus a large number of IMDs can be detected very rapidly and, with automation, as many, or as few different compounds as are of interest can be analysed.

MS/MS technology was recently introduced in England to implement newborn screening for the inherited metabolic condition medium chain acyl-CoA dehydrogenase deficiency (MCADD). Once in place, this raised the possibility that screening for other conditions which, because of their rarity might not each individually merit a screening programme, would now be cost effective in relation to the marginal cost of their addition to the existing programme. Whilst there is patient/parent, clinician and laboratory support for such expansion, it is critical that the technology should not drive screening policy but that this should follow a comprehensive assessment of the likely benefits, harms and costs of expanding screening, especially as each metabolic disease has unique characteristics.

Funding was, therefore, obtained from the National Institute of Health Research - Collaboration for Leadership in Applied Health Research and Care South Yorkshire (CLAHRC-SY) to support a research project to evaluate screening for five additional diseases using MS/MS in 500,000 newborns over the next two years. As a prelude to the clinical phase of this research, the Foundation for Genomics and Population Health (PHG Foundation) was asked to conduct a systematic review of currently available evidence for expanded newborn screening, as the last Health Technology Assessment (HTA) report, published in 2004, only included data up to 2002. There was clearly a need to examine research and practice over the last seven years, especially as MS/MS screening has now been more widely implemented overseas.

## Aims

The aim of the review was to undertake an evidence-based synthesis examining the effectiveness and appropriateness of expanding newborn screening using MS/MS, based on current guidelines for evaluating screening programmes for the following five conditions, which were chosen by UK laboratory and clinical experts.

- Maple Syrup Urine Disease (MSUD)
- Homocystinuria (pyridoxine unresponsive)
- Glutaric Aciduria Type I (GA1)
- Isovaleric Acidaemia (IVA)
- Long-chain 3-hydroxyacyl CoA dehydrogenase deficiency (LCHADD; includes trifunctional protein deficiency)

A central feature of the work was to systematically review the evidence for newborn screening programmes for this panel of five conditions.

## Findings of the review and main recommendations

There is much evidence to support expanding the existing provision of newborn screening to include screening for a wider range of inherited metabolic conditions within the UK, as a means of preventing death and severe disability. The five conditions that are the specific target of this review are all diagnosable through screening and treatable; there is evidence that outcomes are better if infants are diagnosed early and treatment is commenced before any symptoms occur. The strength of international evidence supports the view that an expanded 'bundle' of conditions is more cost-effective than restricting testing to only one or two tests, as at present.

However, whilst there will be significant benefit to a small number of infants and their families, as well as, in total, to the wider health, social and educational services, the possible harms applied to the entire population of newborns arising from this expanded screening programme must be considered. The most significant harms arise from the diagnostic work-up for those eventually found to be disease free and include those directly impacting on the family (such as anxiety) and 'knock-on' effects of the extra work on laboratory, specialist and general paediatric and primary care services. The evidence suggests that such impacts will be manageable but nevertheless there is a need for them to be investigated and quantified.

The paragraphs below provide some of the main findings of the review together with an outline of gaps in knowledge that should be filled by a pilot research programme. In particular they include a number of areas where the translation of general evidence available internationally appears to raise specific questions for the UK and NHS environment.

**Epidemiology:** International evidence is that each condition is individually rare (birth prevalence from 1 in 100,000 to 1 in 400,000). However some, such as homocystinuria, may be more common in the UK. Because of the rarity of the disorders, it is not expected that a pilot study would provide robust epidemiological information but it could provide a basis for information collection including age at diagnosis, sex, ethnic background of parents, history of consanguinity, and family history for screen and clinically detected cases.



**Clinical validity:** MS/MS as a screening test has high specificity and sensitivity in most settings. However, international studies cannot give definitive advice on how to operate a screening programme in UK laboratories and what test performance will be achieved. Participating laboratories must collectively devise and evaluate their tests with respect to analytical and clinical validity using standardised procedures, in particular to maximise test sensitivity whilst minimising false positives. For each condition this should result in flow-charts that show, initial cut-offs; cut-offs for any repeat testing/ or for urgent assessment (depending on condition); cut-offs for further sampling request and undertaking further testing; diagnostic cut-offs; and expected 'flow' of infants through the various branches of the pathway.

**Clinical utility:** Although the systematic review has provided underlying evidence for clinical utility in terms of reduced mortality and morbidity in newborns and later in life, we should ensure that this is achievable within the UK. Cases detected by newborn screening should be included on a register (subject to appropriate consent) and followed up (eventually long-term). Diagnosis should be recorded with detail of underlying genetic and biochemical abnormalities and presenting clinical features. Patient progress should be monitored with details of treatment provided, centre of treatment and description of clinical progress including acute crises and outcomes in terms of morbidity and disability. Parallel active surveillance through the UK laboratories for cases diagnosed clinically should be put in place and these cases also followed up.

**Possible harm from false positives:** The systematic review showed some evidence of harm from 'false positives' due to parental stress and adverse effects on the parent child relationship. A pilot study needs to investigate false positives and their pathway from flagged test to final negative diagnosis. It needs to devise means of minimising stress by improving education in the antenatal period and around the time of screening and also by providing support in the event of a positive result. Educational support will be needed for parents, health professionals and the general public.

**Possible harm from over diagnosis (i.e. diagnosing the condition in those who would otherwise not have developed clinical symptoms):** Detailed understanding of the pathology of the five target conditions leads us to believe that this will not be a major issue. However, a pilot study should include documentation of biochemistry and clinical assessment of every case diagnosed. It should consider some form of external assessment to confirm whether treatment was strictly necessary for each case and record the reasons behind this.

**Economic analysis:** Although international evidence is favourable, the pilot study needs to quantify the extra costs needed to expand the screening programme to include these conditions. Costs falling on laboratories, specialist clinical, paediatric, community services and primary care need to be estimated.

**Availability of specialist care:** There will be an effect on specialist services required to undertake the necessary diagnostic, clinical assessment and follow-up of patients identified through the screening system. The pilot study needs to assess whether there is sufficient capacity in the specialist system in the UK. A comparison of clinical input for screen detected versus clinically detected cases would be useful. Similarly it will be necessary to look at *organisational aspects* to assess how a pilot programme would be embedded into UK NHS services.

**Guidelines and treatment protocols:** The pilot programme should result in a set of guidelines for laboratory and clinical assessment of screen positive patients and for clinical management and follow-up of patients, where possible depending on initial genetic, biochemical and clinical profile.

**Clinical outcome:** The pilot study should begin the process of tracking health and other outcomes for patients and families which should lead to the development of a system and agreed outcome measures.

**Wider benefits and harms:** The pilot study should provide evidence on wider benefits and harms including those to parents and extended family of cases, false positives and their families, health services, researchers and society in general. Where possible the pilot programme should develop and publish operational protocols and resources that will maximise benefit and minimise harm.

**Clinical, social and ethical acceptability:** The pilot programme should work through a group of stakeholders to collect evidence on the clinical, social and ethical acceptability of expanded screening.

**Opportunity cost of screening:** The pilot programme should study qualitative aspects of the opportunity cost by working with relevant health professionals involved in the whole range of the programme.

**Managing, monitoring and setting quality assurance standards:** A detailed plan, resources for running and monitoring the programme and set of quality assurance standards should be developed as part of a pilot programme.

**Resources for parents and public:** Evidence based information about the conditions, advantages and disadvantages of testing, process of testing, consequences of the test result and expected follow-up should be developed as part of a pilot project. The MCADD resources could be used as a template.

## Recommendations

### *The review*

This Report is presented to the National Screening Committee (NSC) as an independent review of the evidence on newborn screening by MS/MS including a systematic review of evidence from newborn screening programmes worldwide that have included one or more of the five target conditions in their repertoire. The following recommendations are made to the NSC:

#### Recommendation 1

The NSC should study this report in detail and determine whether or not it agrees with the conclusions. If not, it should set out:

- Any factual points of disagreement, where possible indicating how such disagreements can be reconciled
- The areas where it considers the evidence is not strong enough to support a favourable

case for expanded screening in general. To the extent that this evidence is unavailable, it should describe what it considers would constitute suitable evidence, and whether it is reasonably foreseeable that this could be obtained in the UK in the short to medium term

### ***An expanded national newborn screening programme***

Whether or not the NSC newborn screening programme should be expanded to include each of these five conditions requires the NSC to weigh up the conditions against their screening criteria. This review has provided interpretation and discussion against each of the criteria. It has concluded that none of the criteria are unfulfilled but that the criteria in very rare genetic conditions may need to be judged differently; there will be trade-offs between criteria; and judgements about ‘fulfilment’ or otherwise are subjective.

### **Recommendation 2**

The NSC should consider the conditions against each of its screening criteria and decide:

- Whether each criterion is met or not
- Where there is insufficient evidence for a given criterion/condition, what it would consider to be sufficient evidence that could be collected within the UK in the short to medium term

### ***Developing a pilot programme***

Following its conclusions that expanded newborn screening would improve health outcomes without causing undue harm and that this could best be undertaken through a national newborn screening programme, the review group of this study has concluded that the next step should be a large scale pilot study. The aim of such a study would be to place expanded newborn screening into experimental practice on a sufficiently large scale to allow some of the unanswered questions relevant to the programme to be answered. Central questions include whether or not laboratories can develop tests with optimal performance for screening, the actual cost to laboratory and clinical services, the impact on these services, and public and professional acceptability. A more complete set of questions that relate also to the effectiveness of screening programmes is outlined above. The NIHR CLAHRC in the Sheffield region has provided funding for a pilot programme and it is believed that many of these questions could be addressed within the funding available.

### **Recommendation 3**

The NSC should:

- Recommend that a pilot programme should be undertaken to address gaps in our knowledge relevant to the expansion of newborn screening in the UK
- Ask the NIHR funded CLAHRC project to conduct the pilot programme
- Set out a mechanism for agreeing further evidence requirement and a process for obtaining and judging this evidence
- Agree to receive a report from the pilot programme on completion



## Glossary of acronyms

ACCE	Analytical validity, Clinical validity, Clinical utility, ELSI
Ala	Alanine
<i>Allo-Ile</i>	<i>Allo-isoleucine</i>
BIA	Bacterial Inhibition Assay
BCAA	Branched Chain Amino Acids
BCKA	Branched Chain Keto Acids
BCKAD	Branched Chain alpha Keto Acid Dehydrogenase
C5	Isovaleryl carnitine
C5DC	Glutaryl carnitine
C16-OH	3-Hydroxyhexadecanoylcarnitine
CBS	Cystathionine $\beta$ -Synthase
CAH	Congenital Adrenal Hyperplasia
CH	Congenital Hypothyroidism
CNS	Central Nervous System
FPR	False Positive Rate
GA1	Glutaric Aciduria Type I
GCDH	Glutaryl-CoA Dehydrogenase
HCY	Homocystinuria
HPLC	High Performance Liquid Chromatography
HTA	Health Technology Assessment
ICU	Intensive Care Unit
Ile	Isoleucine
IMD	Inherited Metabolic Disease
IVA	Isovaleric Acidaemia
IVD	Isovaleryl CoA Dehydrogenase
LCHAD	Long Chain 3-Hydroxy Acyl-CoA Dehydrogenase
MAT	Methionine Adenosyl Transferase
MCAD	Medium Chain Acyl-CoA Dehydrogenase
Met	Methionine
MRM	Multiple Reaction Monitoring
MS/MS	Tandem Mass Spectrometry
m/z	Mass to Charge Ratio
MCADD	Medium Chain Acyl-CoA Dehydrogenase Deficiency
MSUD	Maple Syrup Urine Disease
MTP	Mitochondrial Trifunctional Protein
NPV	Negative Predictive Value
NSC	National Screening Committee
OH-Pro	Hydroxyproline
Phe	Phenylalanine
PKU	Phenylketonuria
PPV	Positive Predictive Value
SCADD	Short Chain Acyl CoA Dehydrogenase Deficiency
TP	True Positive
TN	True Negative
Val	Valine
VLCADD	Very Long Chain Acyl CoA Dehydrogenase Deficiency

## Glossary of terms

<b>Acylcarnitines</b>	Condensation product of a carboxylic acid and carnitine. The transport form for a fatty acid crossing the mitochondrial membrane
<b>Compound heterozygous</b>	The presence of two different mutant alleles at the same loci
<b>Daughter ion</b>	An electrically charged product of reaction of a particular parent (precursor) ion
<b>Galactosaemia</b>	Metabolic disorder inherited as an autosomal recessive trait in which galactose accumulates in the blood due to deficiency of an enzyme catalyzing its conversion to glucose
<b>Genotype</b>	Specific genetic constitution of an individual
<b>Heterozygous</b>	Two different alleles, one in each of a pair of chromosomes, at a particular position in the genome of an individual
<b>Homozygous</b>	Two identical alleles, one on each of a pair of chromosome, at a particular position in the genome of an individual
<b>Intermediary metabolism</b>	The intermediate steps with cells in which nutrient molecules are metabolized and converted into cellular components catalysed by enzymes
<b>Isobaric compounds</b>	Different compounds that have the same molecular weight
<b>Mass-to-charge ratio</b>	A number defining how a particle will respond to an electric or magnetic field that can be calculated by dividing the mass of a particle by its charge
<b>Parent ion</b>	Synonymous with precursor ion
<b>Phenotype</b>	The observable traits of an organism
<b>Point mutation</b>	A DNA sequence variation that involves substitution, insertion or deletion of a single base A, C, G or T
<b>Precursor ion</b>	Ion that reacts to form particular product ions
<b>Product ion</b>	Synonymous with daughter ion
<b>Sensitivity</b>	Proportion of those with a condition who have a positive test result (inversely related to the proportion of false negative results)
<b>Specificity</b>	Proportion of those without a condition who have a negative test result (inversely related to the proportion of false positive results)
<b>Tay-Sachs disease</b>	A hereditary disorder of lipid metabolism, typically affecting individuals of Ashkenazi Jewish ancestry that is characterized by the accumulation of lipids especially in nervous tissue

# 1 Introduction

Tandem mass spectrometry (MS/MS) is an analytical technique used for identifying and quantifying compounds in samples by separating and quantifying ions based on their molecular mass-to-charge ratio. Although it is not a new technology, only comparatively recently has its potential for newborn screening been realised. A large number of inherited metabolic diseases (IMDs) can be detected very rapidly and with automation, as many, or as few, metabolites as are of interest can be requested. It also requires only very small quantities of blood or urine, making it suitable for blood spot screening.

After a pilot evaluation study, newborn screening in England was expanded in 2007 to include medium chain acyl-CoA dehydrogenase deficiency (MCADD). Successful implementation of this policy has meant that all newborn babies in England are offered screening for this disease. The implementation of MCADD screening necessitated the capital investment of MS/MS technology into screening laboratories, together with software for processing and analysing specimens using this modality as well as supporting metabolic clinical services. Once in place, this raises the possibilities that other conditions which, because of their rarity might not each individually merit a screening programme, become cost effective in relation to the marginal cost of their addition to the existing programme. The United States, Canada, Australia and a number of European countries have already expanded their newborn screening programmes to include an extended range of metabolic conditions and, for the most part, these programmes are thought to be unproblematic and effective. There are currently no examples where these programmes have been withdrawn following evaluation and introduction. This has led to professional and public pressure for the UK to follow suit. However, although the international experience is helpful, it is not directly transferable to the UK for a number of reasons, including: differences in age at blood sampling, differences in populations and the prevalence of specific diseases, and different analytical approaches.

It is also critical that the technology should not drive screening policy but that this should follow a comprehensive assessment of the likely benefits, harms and costs, especially as each metabolic disease is unique. Funding has therefore been obtained from the National Institute of Health Research - Collaboration for Leadership in Applied Health Research and Care - South Yorkshire (CLAHRC-SY) to support a research project to evaluate screening for five additional diseases using MS/MS in 500,000 newborns over the next two years. As a prelude to the clinical phase of this research, the Foundation for Genomics and Population Health (PHG Foundation) was asked to conduct a systematic review of currently available evidence for expanded newborn screening, as the last Health Technology Assessment (HTA) report<sup>1</sup>, published in 2004, only included data up to 2002. There was clearly a need to examine research and practice over the last seven years, especially as MS/MS screening has now been more widely implemented overseas.

## 1.1 Aim, scope, and objectives of the review

The aim, scope and objectives of the review were set out as follows:

### *Aim*

To undertake an evidence-based synthesis examining the effectiveness and appropriateness of expanding newborn screening using MS/MS, based on current guidelines for evaluating screening programmes.

## ***Scope***

1. To systematically review the evidence for the panel of five diseases chosen for the research project:
  - Maple Syrup Urine Disease (MSUD)
  - Homocystinuria (pyridoxine unresponsive)
  - Glutaric Aciduria Type I (GA1)
  - Isovaleric Acidaemia (IVA)
  - Long-chain 3-Hydroxyacyl CoA Dehydrogenase Deficiency (LCHADD; includes trifunctional protein deficiency)
2. To systematically review screening programmes already using MS/MS

## ***Objectives***

1. To provide an introduction to the technology of MS/MS including an outline of how analytical validity is determined, how this technology is used to detect target analytes important in the diagnosis of the five inherited metabolic conditions in question, and, in practical terms, how laboratories establish appropriate cut-offs, second-tier and diagnostic testing, how sensitivity and specificity are balanced and important aspects of quality control are maintained and quality assurance is monitored
2. To provide an overview of the five candidate diseases and their natural history; this will include an assessment of clinical heterogeneity, prognostic subgroups and other factors which may be used to stratify risk, clinical features, diagnostic methods, treatments and their effectiveness
3. To provide an assessment of the epidemiology of each disease in the UK, including estimates of birth prevalence and the likely number of cases detected by MS/MS, using suitable population denominators
4. To assess the clinical validity of MS/MS screening, which is the ability of MS/MS to correctly classify those with, and without, each of the five diseases in the population at risk
5. To assess the clinical utility of expanding newborn screening using MS/MS, evaluating whether screening improves outcomes at an affordable cost (including potential harms arising from the screening process as well as the resources required to develop a screening programme)
6. To detail current gaps in the available evidence in order to direct the research questions examined in the pilot project and to identify emerging gaps in the light of service trends and changing needs
7. To provide a timely report to the pilot project's management committee and to the National Screening Committee in order to shape the clinical phase of the project

In the light of initial findings and the advice from the pilot group it was later agreed to include in the report:

1. A chapter written by the voluntary organisations setting out the viewpoint of families affected by these rare metabolic disorders
2. A brief discussion of European and national policy on the needs of patients and their families with rare disorders and the associated issues relevant to screening, prevention, management and research programmes

## 1.2 Method of operation

The project team at the PHG Foundation was led by Dr Hilary Burton, Consultant in Public Health Medicine and Dr Simon Sanderson, Honorary Consultant in Public Health Medicine and supported by a team from the Foundation (Dr Gurdeep Sagoo, Dr Sowmiya Moorthie, Dr Victoria Fearne, Ms Alison Hall and Ms Jane Lane). Expert guidance was provided by Dr Jim Bonham, Professor Rodney Pollitt, Ms Melanie Downing, Dr Jacqui Calvin, Dr Andrew Morris and Dr Mark Sharrard on the technological, laboratory and clinical perspectives and by the management committee of the Pilot Project. The clinical vignettes submitted to the National Screening Committee (NSC) were also incorporated into this work by PHG.

PHG was responsible for undertaking systematic reviews of the literature, conducting epidemiological analyses, assessing ethical, legal and social implications and writing the final report, under the guidance of the pilot project's management committee and relevant experts.

## 1.3 The report

The report is set out in eleven Chapters; Chapters 2-4 provide an overview of technology of MS/MS and its advantages and disadvantages; a description of each of the five diseases, their clinical features, and treatment; and the public health context of newborn screening and its evaluation. Chapter 5 provides a methodology for the systematic review. Chapter 6 includes a summary of the findings of health technology assessments in other countries carried out since 2004. Chapter 7 contains summaries of the extant evidence, grouped by disease in the following categories:

- Epidemiology
- Analytical validity of MS/MS as an assay
- Clinical validity of MS/MS as a test
- Clinical utility: comprising effectiveness of treatment, effectiveness and cost effectiveness of screening programmes

Chapter 8 contains a submission from Climb (a national charitable organisation focussed on inherited metabolic disease) representing the views of parents and families. Chapter 9 contains an overview of some of the key ethical, legal and social implications and Chapter 10, submitted by the Genetic Interest Group provides background on European policy on rare disorders. In the *Conclusions and Recommendations*, we bring together our main findings, discuss some of the implications and make recommendations for the research project and a wider audience.



## 2 Tandem mass spectrometry

Mass spectrometry is a generic analytical technique used for identifying and quantifying compounds by separating ions based on their molecular mass-to-charge ratio and measuring their intensities. A mass spectrometer has three main components: an ionisation source, an analyser, and a detector. Samples are introduced into the instrument and then ionised to generate charged molecules. These charged molecules are extracted into the analyser and separated by their mass-to-charge ratios and detected as they emerge. Results are presented as a mass spectrum, which is a graphical display showing each ion by its mass-to-charge ratio and its relative intensity. There are many variants of mass spectrometry and newborn screening is based on the use of electrospray tandem mass spectrometry, described further below.

In general, the tandem mass spectrometer (or MS/MS) analyser consists of three chambers (designated quadrupoles Q1, Q2, and Q3 respectively): Q1 and Q3 are mass spectrometers separated by a collision cell (Q2), which breaks down molecules into their constituent parts (see Figure 2.1). Different methods of sample introduction are available but electrospray atmospheric pressure ionisation is the method of choice for newborn screening because it uses an ionisation process that results in little or no fragmentation of the original molecules and sample injection can be automated. The principle advantages of the MS/MS configuration are that it can separate precursor ions (intact ionised molecules of the original sample, also known as parent ions) in Q1, fragment these in Q2, and quantify the resulting product ions (also known as daughter ions) in Q3. For each molecule of interest in the original sample, the results of the analysis in Q1 and Q3 are matched in 'precursor ion/product ion' pairs, identified by their specific mass-to-charge ratios. Quantification is based on comparison with internal standards or on the ratio of different metabolites. This is usually calculated by a sophisticated data system which can compare the levels of selected metabolites and, when used in newborn screening, flag abnormal results for attention.

Two MS/MS scan modes<sup>2,3</sup> are used in newborn screening:

- *Precursor ion scanning*: In this case, Q1 is programmed to allow all the precursor ions to enter Q2 but Q3 is programmed to only detect product ions with a specific mass-to-charge ratio. A spectrum of all precursor ions producing this product ion is obtained. This procedure is used for the analysis of acylcarnitines.
- *Neutral loss scanning*: This mode is used to detect all precursor/parent ions sharing a common neutral fragment lost after fragmentation in Q2. Thus, Q1 and Q3 are programmed in relation to a constant difference in mass, equal to the mass of the neutral fragment. This method is used to profile amino acids.

MS/MS can perform several specific analyses simultaneously by restricting measurements to specific ion-pairs rather than scanning as above. This technique, known as multiple reaction monitoring (MRM), allows analysis to be limited to the metabolites of interest while avoiding the detection of others and is used in most newborn screening programmes to restrict the number of conditions that are screened for. However, in some instances, the primary metabolite detected is not specific to one particular disease; in such cases multiple disorders are identified through detection of a single metabolite. For example elevated levels

of methionine may indicate homocystinuria, generalised amino acidemia (for example due to liver disease), parenteral nutrition, tyrosinaemia type 1, maternal B12 deficiency, or methionine adenosyl transferase (MAT) deficiency. In such cases, processes such as a full amino acid or acylcarnitine scan using MS/MS and/or second-tier tests for analytes that are more specific to a particular condition can be used to improve the specificity of the screening assay. These processes can be carried out on the initial blood spot sample. The detection of overlapping metabolic conditions is particularly relevant to screening for IVA, GA1 and homocystinuria and these are discussed in more detail in Chapter 3.

## 2.1 Tandem mass spectrometry: from sample to mass spectrum

A critical part of the analysis is sample preparation, which involves the extraction of amino acids and acylcarnitines from the dried blood spot. After the cards are punched, samples are mixed with methanol and internal standards before the solution is ready for the next step. These samples may be analysed directly or converted to their butyl esters prior to analysis (this latter process is known as derivatisation). The advantage of derivatisation is that it increases the strength of signal from some compounds. However, it can also lead to inaccurate measurement of some acylcarnitines and is also a long and laborious process. In the UK, MS/MS analysis is carried out on non-derivatised samples.

The final samples are then introduced into the MS/MS, ionised, and scanned. Signals corresponding to the selected range of mass-to-charge ratios are then detected and quantified by reference to internal standards. For any particular metabolite, the yield of product ions at the detector depends partly on the concentration of metabolite in the original sample extract but also on the degree of “ion suppression” by other components in the mixture (salts for example) and by instrumental settings, particularly those affecting the ion source and conditions in the collision cell. Internal standards are used in order to correct for these effects and thus allow the concentrations of the metabolites in question to be calculated. The internal standard must match the chemical properties of the analyte molecule as closely as possible. In general an isotopically-labelled version of the analyte is used. Thus for measuring phenylalanine the internal standard is usually [ $^2\text{H}_5$ ] phenylalanine. The substitution of the five hydrogen atoms on the aromatic ring by deuterium increases the molecular mass by 5. The ratio of the two peaks ( $m/z$  166 and  $m/z$  171 for underivatised samples) is proportional to the concentration of phenylalanine in the sample extract. For compounds where an isotopically labelled internal standard is not readily available a close chemical analogue may be used.

The metabolite concentrations and ratios are then compared to pre-determined cut-off values. Population cut-offs are usually determined following analysis of samples from a number of unaffected newborns to establish population means and ranges for the analyte of interest. These are compared to those from published reports and are often modified following accumulation of further data upon initiation of a pilot study or screening programme.

A further consideration when analysing samples by MS/MS is the impact of measuring isobaric compounds - *i.e.* different compounds that have the same molecular weight. Direct injection MS/MS does not distinguish between closely related compounds such as positional isomers. Thus leucine, isoleucine and *allo*-isoleucine all contribute to the peak at  $m/z$  188. Another amino acid, hydroxyproline, has the same molecular mass as (*i.e.* is isobaric with) the leucines despite having a different molecular formula. It will contribute somewhat to the

m/z 188 peak if it is present in blood in high concentration, as in the very rare condition of hydroxyprolinaemia. The three leucines can be resolved from each other (and hydroxyproline) by rapid short-column chromatography using MS/MS for detection. Positional isomerism and interference from isobaric compounds must also be considered during MS/MS analysis of acylcarnitines. Problems are rare but in many cases interference from isobaric compounds may be detected by comparing profiles from derivatised and non-derivatised extracts.

## **2.2 Tandem mass spectrometry: quality assurance**

Components of quality assurance of the test performance in the laboratory require ensuring standardisation of the various markers that are observed, the cut-off, internal standards and sample preparation process. In addition, laboratory-based analysis of external quality control materials and monitoring normal population data for specific analytes obtained by different laboratories provides a means of quality assurance. For example, quality assessment of the performance of the C8 assay for MCADD is monitored by circulating two groups of spiked blood spot specimens to laboratories; one is an in-house preparation and the other is obtained from the Centres for Disease Control and Prevention (CDC, USA). Comparing normal population data for the target analyte from different laboratories is also used as a means of quality assurance.

## **2.3 Using MS/MS in newborn screening: advantages and disadvantages**

### ***Advantages***

1. It enables the rapid detection of a large number of different analytes from a single sample. In addition, for some diseases (such as MCADD), there are no other available assays
2. It is versatile and can be programmed to detect as many, or as few, metabolites as required. Although the technology can potentially detect a very large number of different analytes relating to different conditions with ease and at minimal cost, this does not mean that all of these diseases should be screened. The latter is a matter for screening policy
3. The analysis can be performed on very small quantities of blood or urine, which means that it can be used for newborn screening using dried blood spots
4. The analysis does not require prior chromatographic separation because two mass spectrometers are used concurrently
5. The MS/MS time required for the analysis of each sample is around 2-3 minutes
6. The process can be automated, permitting a throughput of around 600 samples per 24 hours
7. The analysis does not depend upon commercially produced reagent kits and the cost per test is very modest

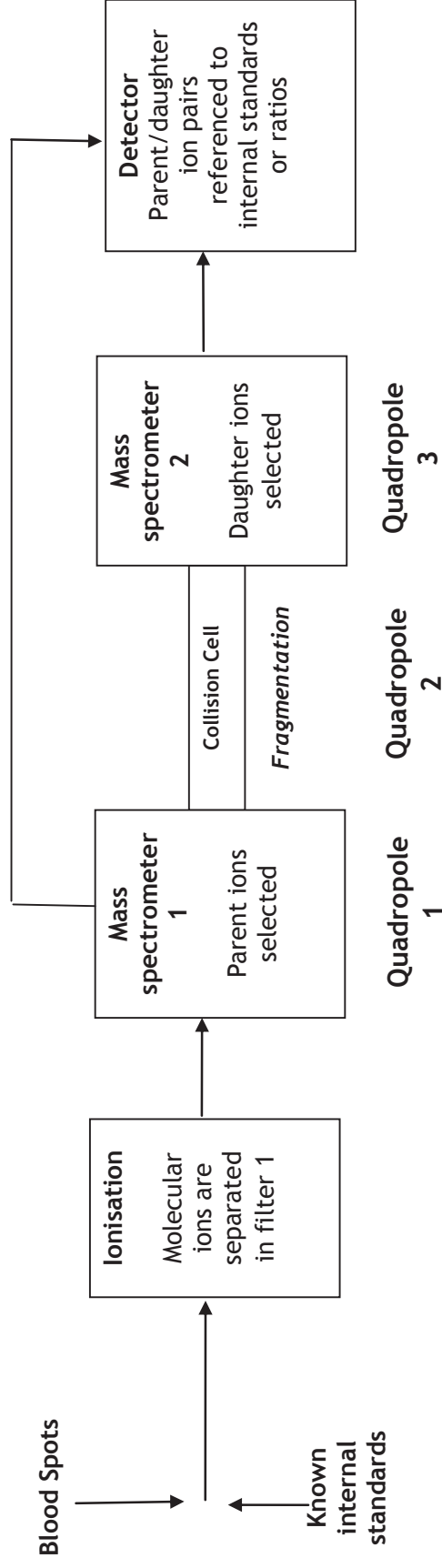
**Disadvantages**

1. In some instances, the metabolites detected are not specific to a particular disease and other diseases may be detected. This has implications for screening for IVA, GA1 and Homocystinuria and these are discussed in Chapter 3.
2. Technical and organisational care need to be taken when handling large numbers of samples to avoid analytical errors.
3. It is possible that the level of certain metabolites may not be significantly elevated in the neonatal period, even in the presence of an inherited metabolic disease, either because of insufficient protein ingestion or age of sampling (this is especially important in the context of premature births). Both this and the above point apply to any analytical technique.
4. Certain drugs and other substances, such as sodium valproate (an anticonvulsant), can interfere with the detection of some acylcarnitines.
5. The technique cannot be used to measure isobaric (*i.e.* having the same molecular weight) compounds, such as leucine, isoleucine and hydroxyproline separately. However, precise quantification can be made at the diagnostic phase.

**2.4 Conclusion**

MS/MS is a powerful technology that can be used for the rapid detection of a large number of IMDs, including some diseases that could not be previously detected by other methods. MS/MS has a number of advantages over other currently available analytical methods and is widely used internationally for newborn screening. However, it is imperative that the inherent capabilities of the technology should not be the primary drivers of screening policy, as there may be important reasons why certain conditions should not be screened for. In addition, the heavy technical requirements of the assay and the analytical process, from sample preparation through to interpretation of the results, mean that appropriate quality assurance mechanisms need to be established to ensure analytical validity and these analyses need to be performed in designated and accredited laboratory facilities.

Figure 2.1 Tandem mass spectrometry





### 3 Clinical and epidemiological overview of the five selected inherited metabolic diseases

This section of the report provides an introductory overview from the general literature of the five diseases selected to be evaluated in the research project for expanded newborn screening. The five conditions are given in Table 3.1.

**Table 3.1** Five conditions selected to be evaluated in the research project

<b>Amino acid disorders</b>	Maple syrup urine disease (MSUD) Homocystinuria (HCY)
<b>Organic acid disorders</b>	Isovaleric acidaemia (IVA) Glutaric acidaemia type 1 (GA1)
<b>Fatty acid oxidation defect</b>	Long-chain 3-hydroxyacyl CoA dehydrogenase deficiency (LCHADD)

It should be noted that estimates of incidence in this Chapter are subsequently revised in the systematic review.

#### 3.1 Introduction

Currently, there are well over 500 known inherited metabolic diseases (IMDs). This number is increasing as our knowledge of human metabolism increases and our ability to detect problems increases with the availability of new technologies. Whilst individual metabolic diseases are usually rare, the collective cumulative incidence is substantial: estimates of 1 in 2,500-5,000 live births are commonly quoted. However, new research has suggested that the incidence is much higher in certain parts of the UK, at around 1 in 1,000 live births<sup>4</sup>. More patients are surviving into adolescence and adulthood as a result of earlier detection (for example through expanded neonatal screening programmes) and improved treatment (such as specific replacement therapies, e.g. *Cerezyme* for Gaucher Disease).

#### *Definition of inherited metabolic diseases*

IMDs are classically defined as monogenic diseases resulting from deficient activity in a single enzyme in a pathway of intermediary metabolism. Clinical consequences arise from the accumulation of substances usually present in small amounts, deficiency of critical intermediate products or specific final products, or the toxic effects of products derived from alternative metabolic pathways. Although each disease is unique, there are some common features:

- They are genetic disorders, so their detection has implications for the person diagnosed and their family
- Certain disorders are particularly common in certain ethnic groups (such as Tay-Sachs disease in Ashkenazi Jews and galactosaemia in Irish travellers)
- The clinical consequences are almost invariably severe, and often require long-term

intervention by experts, factors which are compounded by their relative rarity

- Many have an asymptomatic period shortly after birth with subsequent metabolic decompensation or gradual chronic progression

### 3.2 Maple Syrup Urine Disease (MSUD)

Branched chain ketoaciduria, more commonly known as Maple Syrup Urine Disease (MSUD), is an autosomal recessive condition caused by defects in the branched chain 2-keto acid dehydrogenase (BCKAD) complex. It is a rare disease, with an estimated incidence of around 1 in 120,000 from pooled screening data in six European countries, although in the UK there are pockets of higher incidence in certain ethnic groups. Birth prevalence may be high in populations where there are high levels of consanguinity.

#### *Genetic, molecular and biochemical features*

The BCKAD complex carries out one of the key steps in the breakdown of branched chain amino acids (BCAA) - leucine, isoleucine and valine. Degradation of these amino acids begins with their conversion to branched chain 2-ketoacids (BCKA), which are subsequently broken down by the BCKAD enzyme complex. Deficiencies in the BCKAD complex lead to a blockage of this catabolic pathway and results in the toxic accumulation of branched chain amino acids as well as 2-ketoacids in tissues and body fluids.

BCKAD is a multienzyme complex made up of three catalytic components and two regulatory enzymes which are encoded by six genetic loci. Mutations causing MSUD are associated with genes encoding the catalytic components. As far as is known, mutations causing MSUD are always homozygous or compound heterozygous mutations in the same subunit. Correlations between the genotype and phenotype have as yet not been established, except in the case of E3 deficiency, which is a very rare form of MSUD that leads to additional deficiencies in pyruvate and alpha-ketoglutarate dehydrogenases.

#### *Clinical features*

The three commonest clinical phenotypes of MSUD are:

- Classical (severe)
- Intermediate\*
- Intermittent\*

\* also known as 'variant' forms

Two other forms of the disease are also recognised but extremely uncommon; these are:

- Thiamine responsive MSUD
- Lipoamide dehydrogenase (E3) deficiency

It is important to recognise that there is no clear distinction between the three commonest clinical phenotypes. As noted earlier, the disorder is genetically heterogeneous and these three types form a continuum of severity. Furthermore, differentiation between the classical and intermediate phenotypes has little significance in practical terms because the aims and methods of treatment are the same. Intermediate forms are more easily controlled whilst neonatal death is more likely in classical cases. Undiagnosed, intermediate forms are more likely to show prolonged survival but are likely to suffer very high morbidity.

The majority (75-80%) of patients have the classical (severe) form of the disease, which is characterised by presentation during the neonatal period with encephalopathy and cerebral oedema<sup>5</sup>. Symptoms usually appear at 2-4 days although breast fed infants may only become unwell in the second week of life. Due to the non-specific nature of presentation, there are significant variations in the time to diagnosis. Naughten *et al.*<sup>6</sup> (Ireland) described typical classical patients presenting at a median of 8 days and diagnosed on a median of day 15, while Morton *et al.*<sup>7</sup> (USA) diagnosed patients at a median of 7 days. It should be noted, however, that Morton works in an Amish community with a high incidence of certain IMDs and so his observation is very atypical. Untreated, this condition is progressive and fatal. Milder variants may present later<sup>8</sup> and those with intermittent forms may remain asymptomatic and biochemically normal at least throughout childhood.

It is the classic form of the disease that is most likely to be detected by MS/MS screening in the newborn period. Babies with intermittent MSUD and a minority of cases with intermediate MSUD have BCAA concentrations within the normal range in the newborn period and so MS/MS screening may not be able to detect all individuals with these types of MSUD. For these patients, diagnosis can only be made during attacks of metabolic decompensation, which may be induced by fasting or infection. The relatively late day of screening in the UK means that most classic cases would be symptomatic by the time the screening result was available. However, even when screened at 48-72 hours, most cases would probably still require ICU treatment. Further, the slightly later day of screening in the UK (day 5-7) means that there would be a slightly higher chance of picking up intermediate forms that would benefit from treatment.

Thiamine responsive forms of MSUD may be severe and still require dietary intervention as well as thiamine, so may behave classically. Cases have been missed on screening in the USA, but in the UK the later age of screening may mean that we are more likely to identify them. E3 deficiency seems too rare to be able to draw any conclusions but is not really treatable. The BCAAs may only be mildly raised so it may be missed by screening although there is no data in the literature to substantiate this.

**Table 3.2 Clinical classification of MSUD**

Clinical Phenotype	Age of onset	Clinical Features	Biochemical Features
Classic	Neonatal	Poor feeding, increased/decreased tone, ketoacidosis, and seizures	Elevated allo-isoleucine, BCAA and BCKA. Enzyme activity 0-2% of normal
Intermediate	Variable	Failure to thrive, often no ketoacidosis, developmental delay	Elevated allo-isoleucine, BCAA and BCKA. Enzyme activity 3-30% of normal
Intermittent	Variable	Normal early development. Episodes can be fatal	Normal BCAA when asymptomatic. Enzyme activity 5-20% of normal
Thiamine-responsive	Variable	Similar to intermediate	BCAA and BCKA levels respond to thiamine therapy. Enzyme activity 2-4% of normal
Lipoamide dehydrogenase (E3) deficiency	Variable	Usually no neonatal symptoms. Failure to thrive, developmental delay, progressive deterioration	Elevated $\alpha$ -ketoglutarate and pyruvate

**Screening, diagnosis and treatment**

Historically, screening for MSUD could be carried out by measuring the amount of leucine in dried blood spots using a bacterial inhibition assay (BIA) or by chromatographic methods. In 1993, approximately one-third of UK newborns were being screened using paper or thin-layer chromatography. Screening of the same samples is now possible using tandem mass spectrometry (MS/MS), which quantifies the amount of leucine, isoleucine, *allo*-isoleucine and valine. However the measurement obtained is not precise for each of the amino acids. The MS/MS procedure is unable to distinguish between leucine, isoleucine, *allo*-isoleucine and hydroxyproline as they have the same mass; hence a combined peak is obtained for these four analytes. Elevated levels of the “leucines” and valine are suggestive of MSUD. Second-tier tests measuring levels of *allo*-isoleucine in the blood spots can further confirm initial findings. One difficulty is there do not seem to be clear data on how quickly *allo*-isoleucine levels rise after birth, but screening on day five would make it likely that classical and intermediate forms would be detected. Confirmatory laboratory studies include quantitative analysis of the individual amino acids using column chromatography and quantification of organic acids in urine. In addition, measurement of enzyme activity may give an indication of disease severity.

### *Overlapping conditions*

A number of conditions may be mistaken for MSUD, including galactosaemia and hydroxyprolinaemia. Hydroxyprolinaemia is a very rare condition that would be clearly differentiated by column chromatography of amino acids eluted from the dried blood spot. Similarly conditions affecting the liver, including galactosaemia, may give rise to a secondary generalised amino acidemia in the newborn period and so may be picked up on the screen. It should be noted that most cases of galactosaemia are already diagnosed as a by-product of screening for phenylketonuria (PKU) and, further, that a full and properly interpreted MS/MS amino acid scan would differentiate between these conditions and MSUD. Diagnosis of galactosaemia would be an advantage to patients.

### *Treatment*

Treatment of MSUD depends on the clinical condition and leucine levels. Those diagnosed pre-symptomatically may be treated with enteral formula. Symptomatic infants require rescue therapies in intensive care units with intravenous therapy, extracorporeal detoxification and management of cerebral oedema. Subsequent management will be life-long and involves the institution of a special diet and rapid intervention in patients experiencing metabolic crises. The diet is synthetic and contains the minimum required levels of the relevant amino acids. Many patients who receive treatment develop normally, although vigilance is required to detect the onset of a metabolic crisis, which can be induced by stressors such as infection and fasting. For people with intermittent forms of the disease the diet can be more relaxed, requiring only the avoidance of excessive protein intake and an emergency regimen at times of intercurrent illness. All patients will require care in major metabolic centres including careful monitoring of their blood spot amino acid levels.

For those surviving the initial presenting episode, outcome appears to be related to duration of toxin exposure and, in particular, the time that plasma leucine exceeds a certain level<sup>9</sup>. In particular, Kaplan *et al.*<sup>10</sup> noted that for classical MSUD patients, those diagnosed at a mean of 3.5 days had normal neuro-developmental outcome whereas those diagnosed at a mean of 10 days had an abnormal outcome. As well as learning difficulties, adverse outcomes may include spasticity, particularly spastic quadriplegia and cortical visual impairment. In addition, diagnosis by screening is particularly important for the intermediate form, in which individuals may not experience a crisis and therefore have delayed diagnosis in the absence of screening.



### **MSUD 'at a glance'**

MSUD is an autosomal recessive disease caused by the defective catabolism and toxic accumulation of the amino acids leucine, isoleucine, *allo*-isoleucine and valine.

Incidence in Europe is estimated to be 1 in 120,000 live births but is higher in certain ethnic groups.

The clinical phenotypes form a continuum of severity and they are genetically heterogeneous.

The commonest form is the classical (severe), which presents with catastrophic illness very early in the newborn period.

The classical (severe) form can be detected by MS/MS screening but other (albeit rarer) forms, especially the intermittent form, are less likely to be detected.

Treatment is based on a special diet and the rapid detection and treatment of metabolic crises.

**Box 3.1 Dietary management of the patient with MSUD**

(author Hazel Rogozinski)

Dietetic management of MSUD aims to keep plasma BCAA within recommended ranges: Leucine 200-400/500  $\mu\text{mol/l}$ ; Isoleucine 100-200  $\mu\text{mol/l}$ ; Valine 100-300 $\mu\text{mol/l}$ . This is achieved by dietary manipulation as follows:

**Severe restriction of natural protein:** The amount given is determined by the plasma BCAA levels. It is given as prescribed, measured amounts of normal infant formula, breast milk or weighed leucine exchanges (the weight of natural food which provides 50mg leucine, or approximately 0.5g protein).

In classical MSUD, high protein foods such as meat, fish or egg will need to be avoided, as will many foods containing moderate amounts of protein, such as bread and ordinary flour products.

**Supplementation of amino acids free from BCAA:** This is referred to as the Protein Substitute, and may be given as a drink, paste or gel (depending on the age of the child), several times a day. Most are supplemented with vitamins and minerals.

**Including very low protein natural foods in the diet:** This will include some vegetables, many fruits, as well as pure fats and sugars.

**Including prescribed low protein foods for energy, bulk and variety:** A wide variety is available, including bread, flour, cereals, pasta, biscuits, milk replacements and energy supplements.

**Vitamin and mineral supplementation if not sufficient from the protein substitute**

Plasma leucine is usually elevated much more than isoleucine and valine. The diet is therefore related to leucine intake. The quantity of leucine given is adjusted according to plasma leucine levels. Most children with well controlled classical MSUD have total leucine intakes around 400-600mg/day (equivalent to 4-6g natural protein daily). If plasma levels of isoleucine and/or valine fall too low, supplementation will be required.

**Management of illness**

Plasma BCAAs can rise rapidly during illness. Patients are given a personalised 'Emergency Regimen' to use at the first sign of illness. This consists of 2 to 3 hourly glucose drinks to be taken day and night. The usual diet can be gradually introduced as the patient improves. If the drinks are not tolerated or the child is vomiting, hospital admission will be needed for intravenous glucose and careful monitoring of plasma BCAAs.

### **Box 3.1**    *continued*

#### **Routine monitoring**

This involves regular, lifelong blood tests for BCAA (from twice weekly in infancy, weekly for early childhood and fortnightly/ monthly thereafter as needed).

Patients will also be seen regularly in multidisciplinary metabolic clinics. Growth and nutritional status will be assessed regularly. They will have frequent contact with the metabolic status team, especially during periods of instability or illness.

#### **Problems in dietary management**

The MSUD diet is largely synthetic and complex in nature, which often leads to problems for patients and families. The following are commonly noted:

- The complexity of the diet can make it hard to understand; this problem is compounded if parents speak limited or no English
- It can be time consuming. Cooking separate meals for the child, weighing foods, organising prescriptions, taking blood and attending clinic appointments all contribute to this
- Foods and products can be unfamiliar and new skills need to be learned to produce acceptable results
- It can be socially isolating. Children feel different from their peers; they may not like to eat their food or take the protein substitute in front of their friends. Eating out is very difficult; holidays, sleepovers and parties, therefore, involve a lot of forward planning
- The necessary level of supervision by the parents is often high to prevent young children from eating high protein foods, and to ensure that the correct diet is eaten
- Parents can be reluctant to allow the child to go on residential school trips or to encourage increasing independence as the child grows, due to their concerns that the child may eat the wrong food or become metabolically unstable
- Feeding difficulties are common. The Protein Substitute is unpalatable and may be rejected by the child
- Management of illness is especially worrying as leucine levels can rise rapidly and parents need to ensure the correct treatment is given urgently

#### **Long term management**

Regular, careful monitoring of nutritional adequacy of the diet and biochemical monitoring is very important to ensure optimal growth and development. Any deficiencies will need to be rectified. Children with MSUD usually show normal growth if the diet has been satisfactory. However, nutritional deficiencies may be related to poor intake including rejection of the protein substitute. This can often be overcome with support and advice from the dietician and other members of the metabolic team but occasionally gastrostomy or nasogastric feeding may be necessary to ensure the child receives full nutrition.

### 3.3 Homocystinuria due to cystathionine $\beta$ -synthase deficiency

The most common cause of homocystinuria is a defect in the enzyme cystathionine  $\beta$ -synthase (CBS); this is referred to as “classical” homocystinuria. The overall incidence in the UK is reported to be around 1 in 100,000 live births, although it is considerably higher in certain ethnic groups, such as those of Irish ancestry. The other forms are less common than classical homocystinuria.

#### *Genetic, molecular and biochemical features*

The mode of inheritance of classical homocystinuria is autosomal recessive. A number of disease causing mutations in the CBS gene have been identified, the majority of which are sporadic mutations. These mutations lead to absent or reduced enzyme activity, causing the toxic accumulation of homocysteine and methionine in the blood and tissues. Those mutations associated with residual enzyme activity lead to a milder clinical phenotype.

Classical homocystinuria is associated with a number of clinical and pathological abnormalities. Infants are usually normal at birth and the diagnosis is not usually made until the first 2-3 years of life. Myopia followed by dislocation of the lens, osteoporosis, thinning and lengthening of the long bones, mental retardation and thromboembolism affecting larger and small arteries and veins are the commonest clinical features. Thus, clinical diagnosis is usually only made after irreversible damage has occurred. Without treatment, 25% of patients will die before the age of 30, usually as a result of arterial thromboembolism. There is a great deal of clinical heterogeneity, with some patients displaying all clinical symptoms whilst others display very few or none. The concentration of plasma total homocysteine can be measured to assess the clinical severity of disease and can be monitored to determine the response to treatment.

Homocystinuric patients can be sub-divided into two important biochemical phenotypes:

- Pyridoxine responsive (screen undetectable by usually employed techniques)
- Pyridoxine non-responsive (screen detectable)

In the UK approximately 50% of patients with classical homocystinuria are classified as pyridoxine responsive; these patients usually have milder symptoms and disease progression is slower and slowed further by oral pyridoxine (Vitamin B6) supplementation.

#### *Screening, diagnosis and treatment*

Screening for homocystinuria is based on quantitation of methionine levels either by bacterial inhibition assay, chromatography or tandem mass spectrometry. As these techniques are based on analysis of methionine levels, they are unable to identify milder pyridoxine responsive patients, who do not show markedly elevated levels of methionine. Methionine levels vary in the first few days of life and this variability can influence the final result; usually sampling before 24 hours of age is not recommended. Follow-up testing involves a full MS/MS amino acid scan in order to rule out a generalised rise in blood amino acid concentrations as a result of liver damage or parenteral nutrition. Confirmatory testing for classical homocystinuria involves the measurement of plasma and urine homocysteine using a standard amino acid analyser.

### Overlapping diseases

Liver disease (for example due to tyrosinaemia type I), parenteral nutrition, and methionine adenosyl transferase (MAT) deficiency can give rise to positive results on the initial screen. These can be rapidly resolved by second-tier testing and clinical contact. Raised homocysteine concentrations are also seen in some rarer inborn errors of metabolism (MTHFR deficiency and defects of vitamin B12 metabolism) and in maternal B12 deficiency but these would not be detected by screening as they are associated with low, rather than high, methionine concentrations.

### Treatment

Treatment strategies for homocystinuria were set out by Walter *et al.*<sup>11</sup> based on experience at the Willink Biochemical Genetics Unit in Manchester. Their comparison of patients detected in infancy through their screening programme and those presenting clinically, found that a normal outcome (IQ, and other abnormalities) was only achievable in those treated from infancy. For patients who are non-pyridoxine responsive (screen detectable) treatment involves dietary restriction of methionine and supplementation with betaine, vitamin B12 and folic acid. The diet is analogous to that for PKU; it should be started in the first few months and be overseen by a specialist metabolic service. In contrast, the pyridoxine sensitive patients (not screen detectable) may remain asymptomatic if treated with large doses of pyridoxine.

#### Homocystinuria due to CBS deficiency 'at a glance'

Classical homocystinuria is an autosomal recessive disease caused by a defect in the catabolism of methionine, leading to the toxic accumulation of homocysteine.

Incidence in Europe is estimated to be 1 in 100,000 live births but is much higher in certain ethnic groups.

In the UK the commonest form is classical homocystinuria which causes skeletal, ocular, vascular and nervous system pathology; untreated, the prognosis is poor.

In the UK, approximately 50% of patients with classical homocystinuria are responsive to pyridoxine (Vitamin B6). These patients tend to have a milder clinical phenotype.

Presently, only the 50% of patients with pyridoxine non-responsive homocystinuria can be detected by screening; this includes MS/MS screening.

There is some overlap with other metabolic diseases but second-tier testing can rapidly resolve these issues.

Treatment is effective and is based on diet and/or supplementation with treating other complications as they arise.



**Box 3.2a Patient history classical homocystinuria**

Richard is a 12 year old boy with homocystinuria. With hindsight, his development was delayed from the age of one year. He walked at 18 months but did not have any words with meaning till aged 2 years and did not speak in short sentences till aged 4 years. He has gone to normal primary school has required increasing amounts of extra help. Psychometric assessment has shown a full scale IQ of 65; he will need to go to a special secondary school and it is unlikely that he will ever be able to live independently. At 4 years, he started wearing glasses for myopia and his optician noted dislocated lenses at 8 years of age. The lenses were extracted and he now wears aphakic spectacles. The ophthalmologist sent blood for amino acid analysis, which showed a high concentration of homocysteine diagnostic of classical homocystinuria. At the time of diagnosis, Richard was unusually tall (above 98th centile) with long fingers and mild lumbar scoliosis. Scans showed mild osteoporosis.

Unfortunately, Richard showed no biochemical response to a trial of pyridoxine. He was started on treatment with betaine and a low-methionine diet. The latter involves avoiding foods that contain large or medium amounts of protein (such as ordinary bread as well as meat, cheese etc). Patients can eat fruit and most vegetables and artificial low-protein food (such as special pasta) that are available on prescription; they must also take a 'protein substitute' that contains all the amino acids except methionine. Richard refused to take the betaine medication and the protein substitute and it is now given through a gastrostomy. The dietary protein restriction has been a greater problem, as Richard had acquired a liking for meat, cheese etc over his first 8 years. Thus, his plasma homocysteine concentrations remain above the target levels and his skeletal abnormalities have worsened. Moreover, there is a high risk of thromboembolism, such as a deep venous thrombosis or sagittal sinus thrombosis.

**Box 3.2b Dietary management of patients with homocystinuria (non-pyridoxine responsive)**

(author Hazel Rogozinski)

Dietetic management aims to prevent high levels of homocysteine and methionine, while providing adequate methionine for growth. This is achieved by dietary manipulation; the diet is similar to that for MSUD/ PKU.

Natural protein is severely restricted, and the diet is supplemented by a protein substitute free from methionine. Dietary methionine is provided by a prescribed number of exchanges (the weight of a food providing 20mg methionine). The exchanges are adjusted according to blood levels to achieve the desired range. Diet is monitored in a similar way as with MSUD.

**Problems in dietary management**

These are similar to the problems seen in MSUD. In addition, the following problems may be seen:

- Patients who are diagnosed after the first year of age will have been eating a normal diet and may find the new restrictions difficult. They may refuse the protein substitute and/ or medications and a gastrostomy may be needed to give these.
- Later diagnosed patients may have learning difficulties. This may limit their understanding and ability to manage diet and medications.
- Compliance may be an issue. Patients do not become immediately unwell if they miss medication or eat high protein foods and in practice this can sometimes lead to reduced/ poor compliance and control.

**Long term management**

As in MSUD, the metabolic team are closely involved with patients and families. The team aim to build good relationships with families so that they are able to offer support and advice as needed, and to identify and address problems as soon as possible.

Growth is usually good as long as nutrition is satisfactory.

Adults with homocystinuria will need ongoing support as the dietary management is lifelong.

### 3.4 Glutaric Aciduria 1

Glutaric aciduria 1 (GA1) is an autosomal recessive condition caused by deficiencies in the enzyme glutaryl-CoA dehydrogenase (GCDH). The estimated incidence in the UK is around 1 in 100,000 live births.

#### *Genetic, molecular and biochemical features*

GCDH is involved in the dehydrogenation and subsequent decarboxylation of glutaryl-CoA, which is an intermediate in the breakdown of the amino acids lysine, hydroxylysine, and tryptophan. Defective catabolism causes the toxic accumulation of glutaric acid, 3-hydroxyglutaric acid, glutaconic acid, and glutaryl carnitine. Over 150 disease causing mutations have been identified; of these the R402W mutation is the most prevalent among Caucasians. Most mutations, including the R402W mutation, are associated with undetectable GCDH activity and excretion of high amounts of glutaric acid. However, mutations that lead to varying levels of residual GCDH activity and low excretion of glutaric acid have also been reported. Consequently, patients with GA1 can be divided into two biochemically defined subgroups based on the levels of glutaric acid present in the urine: low excretors are those with less than 100mmol/mol creatine (33%) and high excretors are more than 100mmol/mol creatine (67%)<sup>12</sup>. Although these subgroups are clinically similar, confirmatory testing in low excretors requires more complex follow-up, with either determination of GCDH enzyme activity or by mutation analysis of the GCDH gene.

#### *Clinical features*

The clinical features and natural history of GA1 are now very well understood following the publication by Kolker *et al.* of an international cross-sectional observational study of 279 patients from 35 metabolic centres<sup>12</sup>. About 70% of patients (including both high and low excretors) have an encephalopathic crisis, which is most commonly at around 9 months, with 90% by age 2 years. These are usually precipitated about 1-3 days after onset of a non-specific intercurrent illness, gastrointestinal infection or pneumonia and lead to dystonia and dyskinesia as permanent sequelae but with relative preservation of the intellect. Only 6% of patients had no neurological abnormalities following encephalopathy. Of symptomatically diagnosed patients about 50% die before the age of 25 years. For those who have had an encephalopathic crisis, the average handicap score was 2.7 (representing moderate to severe handicap) and the morbidity score was 2 indicating problems in at least two areas (relating to loss of mobility, feeding problems, respiratory problems and seizures). There is also evidence for two clinically defined sub-groups with insidious or later onset and occasionally for a neonatal onset with non-specific symptoms including irritability and transient lactic acidosis. These infants go on to show delayed motor development. Very occasionally an individual might be asymptomatic.

### ***Screening, diagnosis and treatment***

Population screening for GA1 was not possible prior to developments in tandem mass spectrometry; organic acid analysis has been used for high-risk screening of Amish and Irish Travellers who have a high carrier frequency. Screening using MS/MS is based on quantifying glutarylcarnitine (C5DC) in dried blood spots. The levels of C5DC are estimated based upon an internal standard or in comparison with other acylcarnitines. As the levels of C5DC are prone to considerable variation in the newborn period, in some cases metabolite ratios are also assessed to increase accuracy. Because low excretors tend to have normal concentrations of glutarylcarnitines, they will not be detected by MS/MS screening.

### ***Overlapping conditions***

Glutaric aciduria type 2 (GA2; multiple acyl-CoA dehydrogenase deficiency) would be suspected if a full acylcarnitines scan of the dried blood spot showed that other acylcarnitines were also increased. Urinary organic acid profiles are diagnostic. Medium chain acyl-CoA dehydrogenase deficiency (MCADD) can also give an elevated C5DC signal (due to an isobaric hydroxymonocarboxylic acid) but, as in the current MCADD screen, would be diagnosed on the basis of clearly increased octanoylcarnitine so that there should be no chance of confusion.

### ***Treatment***

Kolker *et al.*<sup>12</sup> found that timely diagnosis and treatment were crucial for a good outcome. Treatment protocols incorporate low-protein, lysine-restricted diets with supplementation of carnitine. High doses of riboflavin are effective in some patients. An emergency regimen with oral or intravenous glucose is provided for times of intercurrent illness in order to avoid metabolic crisis. In general complications were prevented if the treatment was started before the onset of encephalopathic crises but had only a limited effect in patients who already had symptoms. The risk of an encephalopathic attack (and the need for special diet) diminishes greatly after early childhood. Expert guidelines on the management of GA1 were published in 2007<sup>13</sup>.

#### **Glutaric aciduria type 1 'at a glance'**

GA1 is an autosomal recessive disease caused by defective catabolism of certain amino acids, leading to the toxic accumulation of glutaric acid and related compounds.

Incidence in Europe is estimated to be 1 in 100,000 live births.

The clinical phenotype is heterogeneous but common manifestations include macrocephaly as well as neurological and metabolic crises.

Some patients with GA1 are classified as low excretors; these patients cannot be detected by screening.

Overlap with GA2 is known but second-tier testing can rapidly resolve this issue.

Treatment is based on a special diet and the rapid treatment of intercurrent illness to prevent crises.

**Box 3.3a Case history glutaric aciduria type 1**

Elizabeth is a 10 year old girl with glutaric aciduria type 1. She was well until she had an upper respiratory tract infection at the age of 9 months. During this, she became lethargic, lost her ability to sit or to hold up her head and had episodes of extending her arms and legs. Magnetic resonance imaging of the brain showed changes in the basal ganglia and frontotemporal atrophy. Urine organic acid analysis showed glutaric acid and 3-hydroxyglutaric acid, establishing the diagnosis. Subsequently, Elizabeth developed an extrapyramidal movement disorder with dystonia, painful spasms and constant athetoid movements of the limbs and mouth. Muscle relaxants have led to some improvement in the spasms but there has been very little recovery of motor skills. Elizabeth is completely dependent and has to be nursed semi-recumbent in a specially moulded chair. She has developed severe scoliosis and has recurrent chest infections. Intellectual development is relatively preserved, though it is hard to assess due to communication difficulties. She gained weight poorly due to her constant movements and swallowing difficulties and she is now predominantly tube-fed through a gastrostomy. Following diagnosis, Elizabeth's metabolic disorder has been managed with carnitine and a regular glucose intake during infections. There have been no further episodes of neurological deterioration. Nevertheless, her life expectancy is markedly reduced: 40% of symptomatic patients die by 20 years of age, often due to aspiration pneumonia and respiratory failure.

**Box 3.3b Dietary management of patients with GA1**

(author Anita MacDonald)

The aim of dietary management is to reduce the risk of encephalopathic crisis by reducing flux of precursor amino acids into the brain and so reducing production of cerebral glutaric acid and 3-hydroxyglutaric acid by restricting protein. In addition, preventing protein catabolism and aggressive emergency management during situations likely to cause metabolic decompensation such as infections, immunisations and surgery are also important. Treatment does not reverse the neurological damage in symptomatic patients but might prevent further encephalopathic crises or progression of neurological damage. The aims are achieved by adherence to a low natural protein diet; a protein substitute free of lysine and low in tryptophan; an energy intake that meets estimated average requirements; vitamin and mineral supplementation and any other nutrients that may become deficient as a result of the dietary intervention. In all dietary therapy it is important to maintain normal growth and nutritional status.

### **Box 3.3b** *continued*

#### **Protein restriction**

The amounts of dietary protein intake vary with age and recommendations on the optimum levels have been made<sup>13</sup>. Natural protein consists of 2-9% lysine and 0.6-2% tryptophan and a low protein diet rather than low lysine diet is given in the UK. In infancy, natural protein requirements are met by breast milk or standard infant formula.

#### **Energy intake**

Normal energy requirements should be met to ensure that essential amino acids are not degraded to provide energy or nitrogen. Energy requirements are achieved through using a combination of natural foods, energy supplements, and low protein special foods.

#### **Role of enteral feeding**

Long term nocturnal tube feeding is used in some children for 2 main reasons:

- Feeding problems, usually caused by neurological difficulties following encephalopathic crisis
- To deliver emergency feeds during illness

#### **Supplementation with other nutrients**

A low protein diet must be supplemented with a comprehensive vitamin and mineral supplement, and other nutrients such as omega-3 fatty acids (intake is likely to be sub-optimal without supplementation).

#### **Illness management**

During stress (acute infections, high temperatures, trauma and poor feeding), emergency feeds are based on glucose polymer and supplemented with lysine-free, low tryptophan protein substitute. Natural protein intake is temporarily reduced or stopped according to the illness severity. Ten percent IV glucose together with enteral lysine-free, low tryptophan protein substitute is given if oral glucose polymer is not tolerated. Natural protein is introduced after 24-48 hours.

#### **Monitoring**

Regular monitoring is an essential component of management. Anthropometric measurements, medical and dietary histories, and blood nutritional markers are all used to evaluate adequacy of the diet.



### 3.5 Isovaleric acidaemia (IVA)

Isovaleric acidaemia (IVA) is caused by a deficiency in one of the enzymes (isovaleryl-CoA dehydrogenase, IVD) involved in the catabolism of the amino acid, leucine. It is an autosomal recessive disease, with an estimated incidence of around 1 in 100,000 with higher incidence in some locations and ethnic groups.

#### *Genetic, molecular and biochemical features*

Loss of function of the enzyme leads to the toxic build up of metabolites including isovaleric acid and its glycine and carnitine derivatives. Over 25 mutations in the IVD gene have been associated with disease, a number of which lead to complete lack of the enzyme. Although a firm phenotype-genotype correlation has not been identified, recent research suggests that the 932C>T mutation in the IVD gene may be associated with a milder phenotype<sup>14</sup>.

#### *Clinical features*

The disease has a spectrum of clinical phenotypes which might include acute neonatal presentations, acute presentations at a later age and chronic intermittent presentations<sup>14</sup>. The acute neonatal presentation is characteristically in the first two weeks after birth. Infants are initially well, then develop vomiting and lethargy, progressing to coma. Patients may also present with similar symptoms at a later age, usually precipitated by an infection. Other patients present with chronic symptoms, failure to thrive and/or developmental delay, usually within the first year. There are a few reports of acute and chronic presentations occurring in the same family, suggesting the involvement of additional or non-genetic factors in the clinical phenotype. All children are prone to intermittent acute episodes of decompensation with minor illness. In one summary of presentations with 37 patients from different publications, 28 presented in the first two weeks, 7 between 2 weeks and 1 year, and the remaining 2 after 1 year of life; 16 patients were deceased, and of the 21 who were still alive, 7 had mild to moderate learning disability<sup>15</sup>.

Newborn screening has identified individuals with partial as well as complete IVD deficiency. In one study, nearly half of the mutant IVD alleles from infants diagnosed by newborn screening contained a common mis-sense mutation, 932C>T<sup>16</sup>. All of the newborns with this mutation remained asymptomatic with mild or no dietary protein restriction when followed for up to 5 years. This represents a new, milder phenotype that has been unmasked by newborn screening. It is unclear at present whether this mutation will turn out to be common in the UK or whether it is present mainly in German-derived populations.

#### *Screening, diagnosis and treatment*

Population screening for IVA has only been possible since the development of MS/MS and is based on the detection of isovalerylcarnitine (C5) in dried blood spots. If IVA is suspected as a result of elevated C5 levels, a full acylcarnitine scan is carried out on the MS/MS. If other acylcarnitines are elevated, the diagnosis may be GA2. The diagnosis of IVA is confirmed by urine organic acid analysis (which will also differentiate between IVA and GA2). It is thus apparent that MS/MS will detect more cases of GA2. This will be advantageous for some with milder forms, but more severe forms will almost certainly have presented in the first week of life.

### *Overlapping conditions*

IVA screening will detect some cases of GA2 (see above) and 2-methylbutyryl-CoA dehydrogenase deficiency, a very rare condition, probably harmless, found mainly in Hmong Indians. Differential diagnosis for all these conditions is readily made by urinary organic acid analysis. False positive results due to pivaloyl-containing antibiotics have been described, but these drugs are not available in the UK.

### *Treatment*

Treatment can be divided into acute and long-term management. Prompt intervention is needed during intercurrent illnesses, to prevent metabolic crises (which can lead to permanent neurological damage). This intervention involves giving high carbohydrate drinks, to minimise protein breakdown; intravenous management is needed if the drinks are vomited or the patient deteriorates. Long-term management involves dietary protein restriction (to minimise isovaleryl-CoA formation) and treatment with carnitine and glycine (to promote excretion of isovalerate).

It is still uncertain whether subjects carrying the 932C>T mutation are at risk of clinical manifestations or simply express a clinically insignificant biochemical phenotype. It has been suggested that these patients should be observed clinically, particularly when exposed to metabolic stressors such as surgery or febrile illness and low dose carnitine supplementations should be considered if the plasma level is reduced.

#### **Isovaleric Acidaemia 'at a glance'**

IVA is an autosomal recessive disease caused by defective catabolism of the amino acid leucine, with toxic accumulation of isovaleric acid and its glycine and carnitine derivatives.

Incidence in Europe is estimated to be 1 in 100,000 live births.

Patients may present acutely with a metabolic crisis usually in the neonatal period, or chronically with a failure to thrive or developmental delay. All patients are at risk of crises during metabolic illnesses with high mortality and morbidity.

Screening for this disease has only been possible since the development of MS/MS.

Overlap with certain rare diseases is known but second-tier testing can rapidly resolve these issues.

Treatment involves a special diet and the rapid detection and treatment of crises.

**Box 3.4a Case history isovaleric acidaemia**

Bilal is the first child of consanguineous parents. He was born at term and had no neonatal problems except poor feeding and poor weight-gain on standard infant formula. At 2 months of age, he presented with vomiting and dehydration. Cow's milk protein intolerance was suspected and he was discharged on a pre-digested formula. He presented again 3 weeks later with vomiting, a reduced level of consciousness, dehydration and tachypnoea. Investigations showed a normal blood glucose concentration but ketoacidosis and moderate hyperammonaemia. Bilal was rehydrated and the acidosis and hyperammonaemia were corrected. Urine organic acid analysis subsequently revealed markedly increased concentrations of isovaleryl glycine, leading to the diagnosis.

Bilal was treated with a low-protein diet (to reduce the formation of isovaleric acid) and with carnitine and glycine (to promote isovaleric acid excretion). During infections, he has been given high carbohydrate drinks 3 hourly, day and night, to minimise catabolism and acute rises in isovaleric acid production. He has been admitted for intravenous glucose on four occasions when he vomited these drinks. He has had no serious illnesses since diagnosis. Unfortunately, at the age of 8 years he has learning difficulties (IQ 71); the clinical course suggests that this was probably caused by damage prior to diagnosis.

**Box 3.4b Dietary management of patients with IVA**  
(author Anita MacDonald)

The aims of diet therapy in IVA include: limiting the intake of leucine to minimise the formation of isovaleric acid, prevention of protein catabolism, supplementation with any nutrients that may become deficient as a result of the dietary intervention, maintaining normal nutritional status, encouraging normal feeding development and use of emergency regimen during illness.

These aims are achieved by adherence to a moderate protein restricted diet and ensuring normal energy requirements are achieved. Children also require a nutritious diet, meeting all vitamin and minerals, dietary fibre, and essential fatty acid requirements for age. Enteral feeding, supplementation with other nutrients and monitoring are similar to GA1.

### **Box 3.4b** *continued*

#### **Protein restriction**

In IVA, a moderate protein restriction is usually only required. Although protein requirements per kg/body weight decrease with age, protein restriction in early infancy is approximately 2g/kg/day decreasing to 1.5-2g/kg/per day in childhood. In infancy, natural protein requirements are met by breast milk or standard infant formula. In childhood, ideally, high biological value protein containing foods are used to make up part of the protein allowance, although there may be some self-restriction of protein containing foods. Variability of protein restriction has been reported and individual protein requirements should be prescribed. Some patients need minimal protein restriction; and others regulate their own dietary intake without an 'official' restriction. Leucine does not accumulate in plasma so, unfortunately, it is not possible to use this as a measure of metabolic control. However, low plasma concentrations of leucine may occur if protein restriction is too rigorous. A special amino-acid mix free of leucine is rarely advocated.

#### **Energy intake**

Normal energy requirements should be met to ensure that essential amino acids are not degraded to provide energy or nitrogen. Inadequate energy intake could lead to potential metabolic imbalance and decreased growth rate. Energy requirements are achieved through using a combination of natural foods, energy supplements, and low protein special foods. Energy intakes may be low due to restriction of protein containing foods which are often high in calories; anorexia; feeding difficulties associated with neurological dysfunction; and poor feeding experience: e.g. limited variety of foods or delayed exposure to oral feeds or solid foods.

#### **Emergency regimen**

Metabolic decompensation and encephalopathy in children with IVA can be triggered by protein catabolism associated with intercurrent illness, prolonged fasting e.g. due to surgery, or metabolic instability. Prompt intervention is essential with regular high carbohydrate protein-free feeds based on glucose polymer,  $\pm$ fat, administered as small frequent oral feeds (2-3 hourly) both day and night. Alternatively, emergency feeds may be administered as an enteral feed over 24 hours. If there is clinical deterioration or feed intolerance the patient is given 10% intravenous glucose in hospital.

### 3.6 Long-chain 3-hydroxyacyl CoA dehydrogenase deficiency (LCHADD)

LCHAD deficiency is one of a family of defects known as mitochondrial tri-functional protein (MTP) defects. These deficiencies impair the use of fatty acids as a source of respiratory energy. LCHAD deficiency is an autosomal recessive disease with an estimated incidence in the UK of around 1 in 100,000.

#### *Genetic, molecular and biochemical features*

The mitochondrial tri-functional protein (MTP) is a multi-enzyme complex made up of three components - long chain enoyl CoA hydratase (LCEH), long chain 3-hydroxyacyl CoA dehydrogenase (LCHAD) and long chain ketothiolase (LKAT). The  $\alpha$ -subunits of this enzyme complex harbour the LCEH and LCHAD activities, whilst the  $\beta$ -subunit contains the LKAT activity. Molecular defects can affect either the functioning of individual enzymes or the whole complex. Therefore MTP defects are classified as general MTP deficiency, isolated LCHAD deficiency or isolated LKAT deficiency (depending on the specific defect).

Isolated LCHAD deficiency is the more common form of MTP deficiency. The defects cause the toxic accumulation of long chain acyl-CoA esters and the inability to synthesise ketone bodies, which are a source of energy for organs such as the heart and brain.

**Table 3.3 LCHAD deficiency**

Disorder	Biochemical phenotype	Clinical phenotype
Isolated LCHAD deficiency	LCHAD primarily affected, partial deficiency in enolase & thiolase enzymes	Myopathy, cardiomyopathy and liver disease
General tri-functional protein deficiency	All three MTP enzymes deficient	Similar to LCHADD, can be more severe and present earlier with severe cardiac involvement
Isolated LKAT deficiency	LKAT primarily affected	

#### *Clinical features*

MTP deficiency is a clinically heterogeneous condition affecting many organ systems, including the heart, central nervous system, liver and eyes. A few patients (15%) present in the first month, usually with hypoglycaemia. Most patients present between 6 weeks and 6 months of age with hypoglycaemic encephalopathy, cardiomyopathy and raised liver enzymes. A number of patients die following these acute presentations, even if they are managed aggressively. A few patients with mild mutations present later in childhood or as adults with chronic symptoms, such as weakness, but this is much rarer than for other fatty acid oxidation disorders (such as VLCAD deficiency). Acute decompensation can usually be avoided once treatment is started. Prolonged exercise may, however, cause rhabdomyolysis. Most patients develop progressive retinopathy and peripheral neuropathy as adolescents or adults.

### ***Screening, diagnosis and treatment***

Screening for LCHADD only became possible with developments in mass spectrometry and is based on analysis of 3-hydroxyhexadecanoylcarnitine (C16-OH-acylcarnitine) levels. C16-OH-acylcarnitine is raised in all MTP defects, including isolated LCHAD and LKAT deficiencies and generalised MTP deficiency. Confirmation of an MTP defect is based on the analysis of urine organic acids and blood acylcarnitines. If desired, the precise biochemical phenotype can be defined by measuring the enzyme activity in cultured skin fibroblasts or lymphocytes. This information does not, however, alter management, although it may alter prognosis. Molecular genetic analysis starts with testing for the 1528G>C mutation.

### ***Treatment and prognosis***

Treatment includes maintaining a high carbohydrate diet intake during infections in order to prevent metabolic crises. It is also necessary to avoid prolonged fasting, to restrict long chain fat in the diet and to substitute medium chain fat. This is effective in preventing episodes of hypoglycaemia and sudden death and it usually prevents cardiomyopathy. It may also reduce the frequency of muscle problems and slow the progression of retinopathy although it does not completely prevent either the retinopathy or the peripheral neuropathy<sup>17;18</sup>. Unfortunately, most patients with complete MTP deficiency die within the first year despite treatment but this form of the disorder is very rare.

#### **LCHADD and other MTP defects ‘at a glance’**

MTP deficiency is an autosomal recessive disease caused by defective catabolism of certain fatty acids, with toxic accumulation of long-chain fatty acid esters.

LCHADD is the most common form.

Incidence in the UK is estimated to be 1 in 100,000 live births.

Most patients present at the age of a few months with hypoglycaemic encephalopathy, cardiomyopathy and liver disease. There is a high risk of metabolic crises, which can occur with fasting, infection or after prolonged physical exertion. Retinopathy and peripheral neuropathy occur as long term complications.

Screening for this disease has only been possible since the development of MS/MS.

Treatment involves a special diet and maintaining a regular carbohydrate intake during illnesses to prevent acute crises.



**Box 3.5a Case history LCHAD deficiency**

Sophie was born by a Caesarean section at 32 weeks gestation to non-consanguineous parents. Her mother had suffered a variant of pre-eclampsia known as HELLP syndrome (hypertension, elevated liver enzymes and low platelets). Foetal LCHAD deficiency is a rare cause of HELLP syndrome and the mother recovered promptly following delivery. Sophie had mild respiratory distress due to her prematurity and required extra oxygen and feeding through a nasogastric tube for 4 days. She was readmitted aged 3 months, extremely unwell with lethargy and breathlessness. Investigations revealed hypoglycaemia, acidosis and heart failure due to hypertrophic cardiomyopathy. Analysis of blood acylcarnitines and urine organic acids suggested LCHAD deficiency; the diagnosis was subsequently confirmed when she was shown to be homozygous for the common 1528G>C mutation. She was ventilated and treated with diuretics and inotropes but continued to deteriorate and nearly died. She improved after 48 hours of extra-corporeal membrane oxidation (ECMO) and introduction of an MCT-based infant formula. In this milk substitute, the normal long-chain fats are replaced by medium chain triglycerides (MCT), which bypass the metabolic block. She has remained on a minimal long-chain-fat diet since then with MCT supplements. To minimise fasting, she has a continuous overnight feed through a gastrostomy.

Sophie is now 12 years old. She is growing and developing normally. During infections, she has been given high carbohydrate drinks 3 hourly, day and night, and she has been admitted for intravenous glucose on 6 occasions when she vomited these drinks. She has only had one serious illness since infancy: after going sledging in the snow, she suffered acute muscle damage, leading to severe muscle pain, dark urine (myoglobinuria) and renal impairment for 3 days. Ophthalmological examination has shown a mild pigmentary retinopathy but she has no visual symptoms at present. Unfortunately it is likely that her vision will gradually deteriorate over the next 10-20 years and she may also develop a peripheral neuropathy, leading to weakness around her ankle in adulthood.

### **Box 3.5b Dietary management of patients with LCHADD**

(author Anita MacDonald)

The aims of diet therapy in LCHADD include: minimising fat lipolysis, reducing accumulation of toxic intermediates, ensuring normal growth (by achieving full growth potential but avoiding overweight) and helping maximise exercise tolerance. These are achieved by 1) adhering to a limited long-chain fatty acid intake, 2) avoiding prolonged fasting by regular daytime feeding and, usually, overnight tube feeds 3) replacing long chain dietary fat with medium chain triglyceride (MCT) oil, and 5) during illness, administering emergency, high carbohydrate feeds throughout 24 hours. Children also require a nutritious diet meeting all vitamin and minerals, dietary fibre, and essential fatty acid requirements for age.

**Low fat diet:** In LCHADD, there is evidence that long chain fat (LCT) should be restricted to no more than 10% of total energy intake, and in practical terms, long chain dietary fat is restricted to as low as possible. However, the degree of fat restriction varies according to the severity of the condition and age of presentation.

**Medium chain triglyceride oil:** Replacing long chain fat with medium chain fat is particularly beneficial for patients with LCHADD. Approximately 20% of energy intake from MCT is given to patients with LCHADD, although the optimal dose is still to be determined.

**MCT based infant formula:** Babies are given a nutritionally complete infant formula that is high in carbohydrate, low in long chain fats, and supplemented with MCT oil (Monogen; SHS International).

**Regular feeds:** Children with LCHADD require feeds at frequent intervals days and night. Frequency of feeding should be individually assessed according to severity of disorder and metabolic stability. Children presenting with symptoms in the first 2 years of life, are usually given three to four hourly daytime feeds with continuous overnight tube feeds (either via a gastrostomy or nasogastric pump).

**Uncooked corn starch:** In children over the age of 2 years, uncooked corn starch may be given to provide a source of slow release glucose which may allow the interval between feeds to be extended.

**Essential fatty acid supplementation:** Restriction of dietary LCT to less than 10% of calories is associated with a high risk of essential fatty acid deficiency. Low plasma docosahexaenoic acid (DHA) levels have been reported in patients with LCHADD. Therefore, essential fatty acid supplementation (EFAS) is necessary if adequate amounts are not provided by the MCT based infant formula. EFAS may be given from a small dose of walnut oil. Some centres also give additional supplementation with DHA.

**Vitamin and mineral supplementation:** Supplementation with fat soluble vitamins (A, D and E) is necessary unless a child is taking adequate volumes of MCT based infant formula which is supplemented with all micronutrients. Overall nutritional intake should be monitored at regular intervals.

**Box 3.5b** *continued*

**Vitamin and mineral supplementation:** Supplementation with fat soluble vitamins (A, D and E) is necessary unless a child is taking adequate volumes of MCT based infant formula which is supplemented with all micronutrients. Overall nutritional intake should be monitored at regular intervals.

**Exercise**

Fatty acid oxidation is increased during times of increased energy use (e.g. prolonged exercise). Some researchers advocate giving medium chain triglycerides supplements before exercise in LCHADD patients. There is evidence that this improves exercise tolerance. Also giving an extra carbohydrate supplement pre-exercise may be helpful.

**Illness management**

Management of illness is very important. It is well recognized that poor food intake due to infection or loss of intake due to vomiting or diarrhoea results in rapid body fat breakdown which can quickly cause metabolic instability. An emergency regimen, consisting of age appropriate glucose polymer drinks/feeds should be started at the first sign of illness or poor appetite in order to try and prevent or limit body fat breakdown. If high carbohydrate feeds are not tolerated, intravenous fluids providing 10% dextrose are administered immediately.

**Practical issues**

- Feeding problems and feeding difficulties are common in these children. This may be due to reduced appetite due to regular high carbohydrate feeds, poor oral feeding experience prior to diagnosis, or poor metabolic control. They gradually resolve over time.
- Dislike of low fat diet may be a particular issue if children are surrounded by high fat foods.
- Accidental disconnection of overnight tube feeding equipment can lead to a child failing to receive adequate feed overnight.
- Dependence on overnight feeding may lead to less social integration with other children in activities away from the home. Dependence on daytime tube feeds can lead to additional one to one carer help being required within school.

**Long term management**

This involves careful management of growth, nutritional intake, and monitoring of any symptoms associated with lack of dietary energy supply. Biochemical measures that are useful in guiding adjustment of day to day dietary management are limited.

If high carbohydrate feeds are not tolerated, intravenous fluids providing 10% dextrose are administered immediately.

## 4 Evaluating screening and newborn screening: an overview

### 4.1 Defining the key elements of screening

Although screening using periodic health examinations and child health examinations has been practised since the middle of the 19<sup>th</sup> century, it was not until 1951 that screening was formally defined. The United States Commission on Chronic Illness defined screening as:

*‘The presumptive identification of unrecognised disease or defect by the application of tests, examinations or other procedures which can be applied rapidly. Screening tests sort out apparently well persons who probably have a disease from those who probably do not. A screening test is not intended to be diagnostic. Persons with positive or suspicious findings must be referred to their physicians for diagnosis and necessary treatment’*

More recently, the UK National Screening Committee (NSC) in 2000 defined screening as:

*‘A public health service in which members of a defined population who do not necessarily perceive they are at risk of or already affected by a disease or complications are asked a question or offered a test, to identify those individuals who are more likely to be helped than harmed by further tests or treatments to reduce the risk of a disease or its complications’*

These definitions provide five key insights:

1. Screening is an important public health activity
2. Screening is the testing of people who believe they are well in relation to the disease that the screening test relates to
3. Screening tests sort people into higher and lower risk categories but they do not provide certainty
4. The benefits of screening should outweigh potential harms
5. Evaluation must encompass the whole system or programme of events necessary to achieve risk reduction. Thus, screening is a *programme* and not merely a *test*

Purposes of screening include: a) to reduce the risk of future morbidity or mortality (for example, breast cancer screening) and/or b) to give valuable information about a condition, even when that risk cannot be changed (for example, antenatal Down’s syndrome screening which may influence future reproductive choice).

Although most aspects of screening outlined above directly apply to newborn screening, there are some important differences. The first is that the newborn child is unable to give consent to the test and so consent to screening should be sought from the parent(s). Informed consent is required to ensure that participation in screening programmes (or any other clinical activity) is based on an adequate explanation of the potential benefits and harms and thus ensures that any proposed intervention is lawful. Second, because inherited metabolic diseases are genetic conditions, their results have potential relevance for other

family members and future siblings (who have not given their consent for the test). This raises issues about what should be done with this information and how it should be dealt with. Third, because of the rarity, complexity and heterogeneity of inherited metabolic diseases, it is impossible to obtain evidence of the effectiveness of screening programmes based on randomised trials; this is because there are insufficient patients to generate the necessary statistical power to answer the question and there may also be ethical considerations about the appropriateness of randomisation in such diseases in the light of rapidly developing understanding and treatment options.

#### 4.2 Newborn blood spot screening in the UK

Newborn blood spot screening started in the UK in 1969, based on the pioneering work of Robert Guthrie on phenylketonuria (PKU). A key feature of his work was the use of filter paper cards to collect blood samples and transport them to laboratories. By providing a robust method of sample collection, processing and storage, this development provided a mechanism for developing newborn screening programmes. From the early 1970s, the development of immunoassays began to increase the scope of blood spot screening and the detection of congenital hypothyroidism using this technique is now the most widely practised screening test in the world.

Until recently only PKU and congenital hypothyroidism were screened for over the entire UK. More recently the programme was expanded to include sickle cell disorders, cystic fibrosis and MCADD. The addition of MCADD to the disease panel requires the introduction of tandem mass spectrometers into screening laboratories, as MCADD cannot be screened for in any other way. Tandem mass spectrometry (MS/MS) allows simultaneous and rapid analysis of a number of different biomarkers in dried blood spot samples and the potential to rapidly identify a number of different inherited metabolic diseases from the same sample. Internationally, screening using MS/MS has been introduced at different rates and for varying panels of IMDs. These differences are largely attributable to variations in the prevalence of specific diseases, medical practice, or funding priorities.

Countries also differ in the extent to which scientific evidence influences screening policy, with the UK adopting a more conservative stance on expanding newborn screening than some other similar countries. This stance is justified in the UK by the importance attached to ethical imperatives of population screening and particularly the need to ensure that the benefits of a screening programme at the population level outweigh any harm. For example, recent controversy about the benefits of PSA (prostate specific antigen) screening for prostate cancer has highlighted the need for caution, especially in the face of public pressure for such a programme. Thus, it is important that there are explicit criteria against which new screening programmes can be evaluated.

#### 4.3 Evaluating screening programmes

In 1968, Wilson and Jungner proposed a number of criteria for evaluating screening programmes, comprising a range of evidence about the condition to be screened for, the screening test, and subsequent diagnosis and treatment. However, the pace of technological change since their adoption by the World Health Organisation (WHO) means that they are perceived to be inadequate for evaluating certain types of screening, such as screening for inherited diseases<sup>19</sup>. This is because they do not fully acknowledge the inherent difficulties of studying the epidemiology of conditions that are extremely heterogeneous and also very

rare. Recognising these difficulties, the WHO and other organisations, such as the Netherlands Health Council and the UK's National Screening Committee (NSC), have proposed updated criteria for use in evaluating screening programmes in general whilst other organisations, such as the International Society for Neonatal Screening have proposed criteria for evaluating newborn screening in particular. The NSC criteria, introduced in 1998, are the current list against which new screening programmes in the UK are judged (see Appendix 1). The NSC criteria are arranged into four categories:

1. The condition
2. The test
3. The treatment
4. The screening programme

Although these criteria are more focused and comprehensive than Wilson and Jungner's, they still do not provide the level of detail required for evaluating expanded newborn screening programmes, which are based on the simultaneous detection of biomarkers.

#### **4.4 Evaluating biomarker-based tests in screening programmes**

In this report, we distinguish between the concept of an *assay* (defined here as a laboratory-based method used to detect a substance or 'target analyte' in a sample) and the concept of a *test* (defined as the use of an assay to detect a disorder in a particular population for a particular purpose). Thus, the same assay may be used in different populations for several different purposes; similarly, the same test may use different assays. These distinctions are important because each assay/test application must be evaluated separately, especially as the same assay may be used as a diagnostic, predictive or screening test. Adequate performance of a diagnostic test does not necessarily mean that it will be a good screening test. This is because the populations and settings in which the test is used, and thus the prevalence of the pre-clinical/latent phase and the target disorder, are likely to differ. This is especially true when tests are used for screening in well populations, where the disease is likely to be extremely rare, as opposed to diagnostic or high risk settings. Similarly, test properties, such as sensitivity and specificity, are often assumed to be fixed properties of a test. This assumption has been shown to be false, both theoretically and clinically. Both parameters may be influenced by case mix, disease severity and other risk factors for disease, especially if these are different in cases and controls when the test is evaluated or applied in different settings.

The UK NHS Genetic Testing Network has issued requirements for evaluating new biomarker-based tests, particularly suitable for genetic conditions, based on the American ACCE (Analytic Validity; Clinical Validity; Clinical Utility; and Ethical, Legal and Social Issues) Framework. Although there is some overlap between ACCE and NSC criteria, the ACCE criteria provide a more detailed approach for identifying and evaluating published evidence about new diagnostic, predictive, and screening tests. We have used these criteria to supplement the NSC criteria in this report. The ACCE criteria are as follows:



***The target disease and its natural history***

A detailed understanding of the epidemiology of the disease in the target population is required in order to develop and target screening interventions. Nevertheless, even where knowledge of inherited metabolic diseases is extensive, there may be differences in presentation and manifestations, or there may be borderline cases. There is substantial genetic heterogeneity in a number of inherited metabolic diseases and it is important to recognize the contribution of environmental factors, such as the impact of fasting or infection on the risk of clinical presentation in certain fatty acid oxidation disorders. Obtaining evidence about the natural history of inherited metabolic diseases is primarily hindered by their rarity. Thus it can be difficult to know with certainty what the natural history of each disease is and the extent of variation in presentation, clinical and metabolic features and outcomes.

Finally, it is imperative that, even though they may share some common features, each inherited metabolic disease should be regarded as being unique and should be evaluated in its own right. It is only when these differences are appreciated that the implications of increasing the number of disorders screened for can be truly appreciated (for example, when deciding on the age of sampling). Some of these issues are discussed in more detail below.

***Analytical validity - assay performance in the laboratory***

Analytical validity is the ability of an assay to accurately and reliably detect and quantify target analyte(s) of interest in a sample. If an assay does not perform well in the laboratory, then it is very unlikely that it will be clinically useful. Unfortunately, many assays are assumed to be highly specific when they may not be; it is critical that quantitative estimates with confidence intervals are obtained for performance parameters, such as analytical precision (reproducibility of multiple assays on the same sample) and accuracy (bias, the closeness of the average result to the true value). Evidence of analytical validity is usually obtained from laboratory based studies. The precision of the assay is usually calculated by replicate analysis of the same normal blood samples. Accuracy is determined by comparing the results of the new assay against an alternative assay, which is defined as the reference (or 'gold') standard. One problem for the evaluation of newborn screening is that for certain diseases (such as some fatty acid oxidation disorders), there are no available reference assays or internal reference standards and alternative measures are required, such as population distributions or the use of metabolite ratios.

***Clinical validity - test performance in patients***

Clinical validity is the ability of a test to detect or predict the disorder of interest in the target population. If an assay is shown to have good analytical validity, the next step is to assess how a test performs in patients. The application of a cut-off to the MS/MS assay converts it to a test, giving categoric results and allowing determination of sensitivity and specificity. A distinction needs to be made between performance of the screening test and the performance of subsequent diagnostic confirmatory testing (often referred to as the 'sieve' and 'sort' phases of screening respectively), as both are important but should not be conflated. As with analytical validity, a critical component of assessing clinical validity is the definition of an appropriate reference standard. This reference standard may be a set of clinical diagnostic criteria or the application of an alternative testing regimen in all those screened. The distinction between sieve and sort phases with respect to our target conditions is not always clear-cut. In many disorders, both initial and differential diagnoses

are based entirely on laboratory tests which can be carried out as part of the standard screening protocol.

One of the problems commonly encountered in the evaluation of newborn screening programmes is that the reference test is usually only applied to those with screen *positive* results and not to those with screen *negative* results. It is thus almost impossible to identify those negative results that should have been positive (*i.e.* false negatives). Some indication of the number of false negatives may be obtained by examining the incidence of the disease in screen negative patients or through retrospective testing of blood spots. However, such evidence needs to be interpreted with caution because (a) screen detected and clinically detected cases may not be directly comparable (see the discussion of length time bias below) and (b) retrospective testing of blood spots is only possible for certain conditions. Furthermore, retrospective testing will only detect analytical errors. False negatives also occur when the marker is not abnormal (above cut-off) at the time the sample is taken: e.g. some cases of intermediate MSUD. There is a limit to how long spots can be stored and still provide useful analytical results and this is dependent on the metabolite<sup>20</sup>.

It is also important to recognise that sensitivity and specificity are influenced by the age at screening, the choice of appropriate cut-off values, by the method for determining appropriate cut-off values, and by screening policy. For example, the age at screening is important because in some diseases, such as amino acid-based disorders, the concentration of analytes increases with increasing age and so they are more easily detectable when screening is later. However, for other disorders, such as some fatty acid oxidation disorders, the converse is true and earlier screening would be better. When there are only a few disorders being screened for this is less of a problem, but as the number of conditions included in screening panels increases, it becomes more problematic.

In some countries, screening policy prioritises very high sensitivity so that no cases are missed. To achieve this goal, low cut-off values are required to ensure that all cases are detected. However, such programmes have low positive predictive values and a large number of false positive results, increasing both the number of newborns requiring extensive post-screen evaluation and parental anxiety, even though this may be relatively short-lived. These investigations may be costly, time-consuming and potentially harmful or painful. Accordingly, other countries adopt a more conservative approach, obtaining much higher positive predictive values and fewer false positives, but at the risk of missing some cases. A further complicating factor is that predictive values are also influenced by prevalence; because there are substantial regional variations for the prevalence of many IMDs, it may not be possible to extrapolate results of screening test and programme performance between countries.

### ***Clinical utility - does screening improve outcomes at an affordable cost?***

Clinical utility is concerned with evaluating the risks and benefits of screening in routine practice. Unless screening leads to measurable benefits to patients with minimal harm (as a result of false positive or false negative test results) and at an affordable cost, screening is not appropriate. For newborn screening, it is critical to establish that without neonatal screening, clinical diagnosis would have been made at a later stage and when irreversible damage had already occurred. The implementation of screening may also help to avert lengthy and difficult diagnostic processes in clinically-presenting patients, which may be

an important indirect benefit of screening. Identification of carriers through screening may also lead to health gain by providing valuable information, which may influence future reproductive choices for example. It is generally accepted that screening is appropriate when there is no available treatment if it provides information that is deemed to be valuable to those concerned (as in the example of Down's syndrome screening mentioned above).

However, where treatments are available, there must be evidence that they are both effective and cost-effective. For many inherited metabolic diseases, treatment may only lead to partial recovery or may only result in a slowing of the disease process. Thus, for newborn screening programmes, it is important to show that early detection leads to better outcomes than in clinically presenting children. If a condition can be effectively treated without screening (*i.e.* once it is manifest clinically) then screening is not appropriate for that condition.

Obtaining evidence of effectiveness in the context of rare inherited metabolic disease is difficult, because of the risk of a number of biases:

1. As the screening test may identify cases that would not otherwise have been diagnosed (often because the possibility of the diagnosis was never considered) the case mix, and hence the outcome, may be different. It is most likely that milder cases might be missed in clinically detected comparison groups.
2. Lead time biases: in classical screening for adult disease, the survival of screen detected cases may appear to be longer than in clinically-presenting individuals simply because the 'survival stop watch' has been started at an earlier stage even if screening makes absolutely no difference to outcome. This is known as *lead time bias* and might be somewhat less important in neonatal screening because onset is deemed to be at birth and survival is measured by age at death.
3. Length time bias: again in adult programmes, screening is generally better at picking up long-lasting or slowly progressive conditions than rapidly progressing severe diseases. Length time bias means that patients with a better prognosis are pulled into the screen detected group, whilst patients with a poorer prognosis are pulled into the 'control' group. Thus, outcomes in screen detected cases appear to be better than those in clinically-detected cases even if screening makes absolutely no difference to outcome. As a result of the time of screening in early life, newborn screening should not be subject to this form of bias, except in cases of very early neonatal disease not detectable by screening.

For all these reasons screen detected cases are not directly comparable with clinically-detected cases. Evidence of clinical utility is primarily obtained from randomised controlled trials of screening interventions, health services research (including economic evaluation), and quality assurance of established programmes. Where randomised trials are inappropriate, impossible or unethical, then non-randomised alternatives must be used (such as prospective cohort outcome studies). However, such studies are prone to bias and they need to be conducted and interpreted with caution.

## Conclusions

- Screening is an important public health activity
- Screening is the testing of people who believe they are well in relation to the disease being considered
- Screening tests sort people into higher and lower risk categories but they do not provide certainty
- The benefits of screening should outweigh potential harms
- Evaluation must encompass the whole system or programme of events necessary to achieve risk reduction. Thus, screening is a programme and not merely a test
- There are various international criteria used to evaluate screening programmes; the UK National Screening Committee criteria provide the current standard against which expanded newborn screening will be judged
- The NSC criteria are not entirely suitable for the evaluation of biomarker-based screening programmes, such as newborn screening, which presents unique challenges
- By supplementing the NSC criteria with the ACCE criteria, a more comprehensive assessment of newborn screening is possible
- Evaluation must focus on each target disease and its natural history; analytical validity (assay performance); clinical validity (screening test performance); and clinical utility (effectiveness of interventions and effectiveness of the screening programme)

## 5 Review methods

### 5.1 Systematic review of clinical evidence

This systematic review was undertaken to examine the epidemiology, effectiveness and cost-effectiveness of screening for the five additional inherited metabolic diseases chosen as candidates for the expanded newborn screening research project.

### 5.2 Search strategy

A two-stage strategy was used to identify articles for the review, which was based on previous UK HTA reports to ensure comparability. The first stage identified all articles relevant to the evaluation of MS/MS screening programmes in all of the electronic resources listed in Box 5.1. The second stage identified articles about the natural history, epidemiology, analytical validity, clinical validity and clinical utility for each of the five chosen diseases in PubMed (MEDLINE).

Articles were eligible for inclusion if they were published between January 2002 (the cut-off date of the previous HTA review published in 2004) and June 2009. The searches were conducted during May and June 2009, with no language restrictions. The search terms are presented in Appendix 2.

#### Box 5.1 Electronic resources searched

Pub Med (MEDLINE)

EMBASE

CINAHL

CRD Databases (NHS DARE, EED, and HTA)

Web of Science (formerly Science Citation Index)

Bandolier

Canadian Co-ordinating Centre for Technology Assessment (CCOHTA)

INAHATA (International Network of Agencies for Health Technology Assessment Clearing House)

NCCHTA (National Co-ordinating Centre for Health Technology Assessment)

SIGN (Scottish Inter-Collegiate Guideline Network)

In addition, reference lists of identified articles and international HTA reports were scrutinised and experts in the field were contacted to identify other articles that may have been missed. Literature searching was conducted by two reviewers (Sowmiya Moorthie and Gurdeep Sagoo) and the initial screen of titles and abstracts was conducted by a single reviewer (Simon Sanderson). Where articles could not be rejected with certainty, full text articles were obtained. The final list of potential articles was also reviewed by an expert in the field (Prof Rodney Pollitt) to identify any important omissions or to exclude any unsuitable articles.

### 5.3 Review inclusion and exclusion criteria

The criteria for including studies in the review are listed in Box 5.2.

#### Box 5.2 Inclusion and exclusion criteria

##### *Inclusion criteria*

- Studies published after January 2002 [AND]
- Target population: neonates or newborn infants [AND]
- Target IMDs: Homocystinuria, IVA, GA1, MSUD, LCHADD (studies investigating other IMDs must have data on at least one of the five targets) [AND]
- Primary intervention: Screening by MS/MS in the neonatal or newborn period [OR]
- Secondary intervention: Screening by other methods (Guthrie, TLC etc) [OR]
- Outcomes [OR] incidence and/or birth prevalence, natural history, prognosis; analytical and clinical validity of MS/MS screening (including sensitivity, specificity, predictive values and ROC curves); effectiveness of treatment; effectiveness and cost-effectiveness of screening
- Study designs [OR] primarily randomised controlled trials and cohort studies, case-control, other non-randomised evaluations of treatment effectiveness, cross-sectional epidemiological studies [AND]
- Studies of natural history, prognosis or treatment with five or more subjects

##### *Exclusion criteria*

- Studies published before January 2002 [OR]
- Non-human studies [OR]
- Studies of natural history, prognosis or treatment with less than five subjects

#### 5.4 Data extraction strategy

An electronic, pre-piloted extraction form was used by independent reviewers to extract data. A random sample of 10% of articles was subjected to a second, independent extraction as a quality control measure. Disagreements were resolved in conference or by a third reviewer (Simon Sanderson). A flow chart for the yield of identified articles is shown in Figure 5.1.

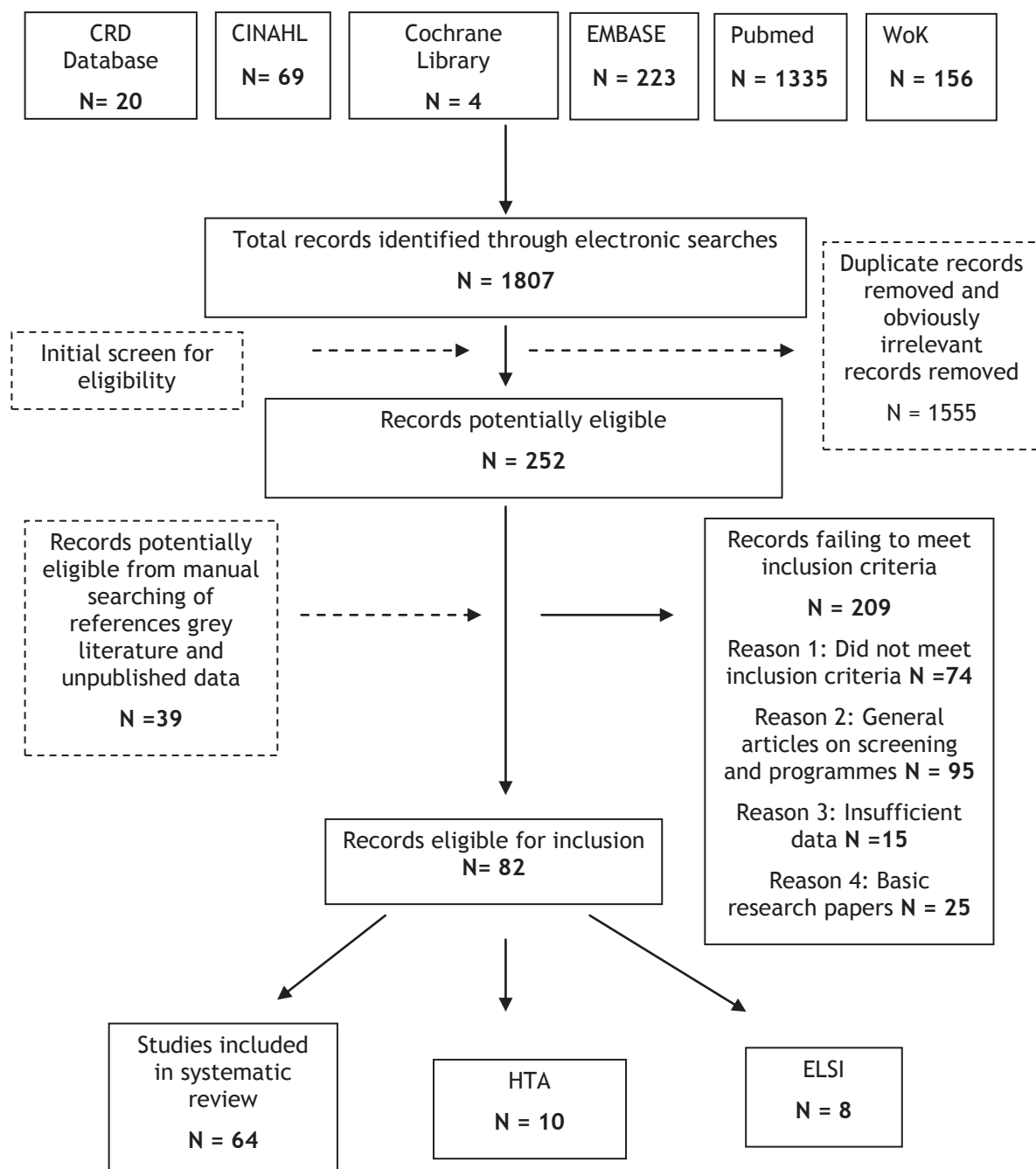
#### 5.5 Data synthesis

Substantial variation in the international prevalence of the selected diseases, the age at screening, and choice of population cut-offs, as well as different attitudes towards the trade-off between sensitivity and false positives precluded the use of meta-analysis. These factors also limit the direct applicability of international results to the United Kingdom; nonetheless, results from such studies provide context and useful information about epidemiology and natural history, as well as the likely range of results for assay and test performance. Thus, narrative synthesis was used to tabulate and summarise the results for each disease in five main sections in Chapter 7:

1. Epidemiology and natural history
2. Analytical validity of MS/MS as an assay
3. Clinical validity of screening using MS/MS
4. Effectiveness of screening programmes
5. Cost effectiveness of screening programmes



**Figure 5.1** Flow chart for the yield of all identified articles



## 6 Summary of findings from Health Technology Assessment reports identified by the systematic review

Ten health technology assessment (HTA) reports, published since 2002, were identified by the systematic review. These reports were from the UK, Canada (Ontario, Quebec and British Columbia), Argentina, Spain, the Netherlands, Finland, Denmark, and the United States. All reports except the Finnish HTA assessed the clinical benefit and the cost benefit of expanded newborn screening and many of them drew heavily on previous UK HTA reports published in 1997<sup>21;22</sup> and 2004. A comprehensive assessment of a number of inherited metabolic diseases was presented in six reports (UK, Netherlands, Spain, British Columbia, Denmark and United States). We present here brief summaries of these reports, their main findings and their recommendations.

### 6.1 United Kingdom (2004)<sup>1</sup>

The purpose of this report was to evaluate the clinical and cost-effectiveness of tandem mass spectrometry (MS/MS)-based neonatal screening for inherited metabolic diseases. It was an update of two previous HTA reports published in 1997<sup>21;22</sup>. Its conclusions were based on a systematic review of literature and an economic evaluation to investigate the economics of screening using tandem mass spectrometry.

Based on the literature review, the authors concluded that the evidence supported the introduction of MS/MS into a UK neonatal screening programme for PKU and MCADD deficiency. The authors felt there was a lack of robust evidence on the underlying incidence and outcomes for other disorders, particularly differences in long-term outcomes that could be attributed to therapies initiated as a consequence of presymptomatic detection using MS/MS. There was also limited evidence on screening test performance, the natural history of some diseases, and the potential economic impact of screening for other diseases. The report recommended that the primary focus of research should be on the long-term effectiveness of treatment strategies on adverse outcomes (disabilities and impairments) under conventional management and the potential impact of early diagnosis using MS/MS. As a result of the findings of this report and a subsequent pilot research programme, newborn screening in England and Wales now includes MCADD.

### **Summary UK HTA 2004**

- 23 conditions were evaluated including MSUD, homocystinuria, GA1, IVA and LCHADD
- The evidence appears to support the introduction of MS/MS into the UK for PKU and MCADD deficiency
- Other conditions should not yet be included because of the lack of robust evidence on the underlying incidence and outcomes for many of the disorders
- Further research was needed:
  - To establish the sensitivity and specificity of neonatal screening using MS/MS for other individual IMDs
  - To determine underlying incidence of these conditions in the UK
  - To ascertain the natural history and economic impact of screening for other disorders

## **6.2 Finland (2004)<sup>23</sup>**

The Finnish screening programme has important differences from a number of other countries, including the UK. For example, PKU is not screened for because of its rarity and screening for congenital hypothyroidism is based on the analysis of cord blood. Thus, its conclusions and recommendations should be seen in the light of the local context. The report evaluated evidence for screening for five additional diseases (PKU, MCADD, LCHADD, GA1 and congenital adrenal hyperplasia (CAH)), which were selected by an expert panel. However, expansion of the newborn screening programme was not recommended as it would have required a major change in the type of screening samples collected and the supporting infrastructure in place.

## **6.3 Netherlands (2005)<sup>24</sup>**

In 2005, The Netherlands Health Council evaluated the possibilities and limitations of neonatal screening for over thirty disorders, many of which are detectable by tandem mass spectrometry. The disorders were evaluated on the basis of screening criteria and separated into three categories based on the direct and indirect benefits of screening. The report also discussed ethical issues (such as informed consent) and briefly considered the economic implications of expanded screening. The committee recommended the inclusion of fifteen new disorders to the newborn screening programme; this expanded programme has been in place since January 2007.

**Summary Netherland Health Council report 2005**

- Substantial improvements have been made in the possibilities for both diagnosis and treatment of IMDs
- According to the present neonatal screening criteria, there are further disorders that qualify for neonatal screening
- In addition to PKU, congenital hypothyroidism (CH) and CAH, screening for the following 15 disorders was recommended:

Biotinidase deficiency	Cystic fibrosis
Galactosaemia	<b>Glutaric aciduria type 1</b>
HMG-CoA lyase deficiency	Holocarboxylase synthase deficiency
<b>Homocystinuria</b>	<b>Isovaleric acidaemia</b>
<b>LCHADD</b>	<b>Maple syrup urine disease</b>
3-MCC deficiency	MCADD deficiency
Sickle cell disease	Tyrosinaemia type 1
<b>VLCADD</b>	
(3-MCC, 3-methylcrotonyl-CoA carboxylase)	

**6.4 Spain (2007)<sup>25</sup>**

The conclusions of this report were based on an update of the UK HTA report of 2004 and an evaluation of newborn screening programmes in Spain's autonomous regions, as there were concerns about local variability in service provision. It concluded that screening using MS/MS should only be carried out for MCADD and PKU but additional studies were needed to establish the sensitivity and specificity of MS/MS for detection of other conditions.

#### Summary Spanish HTA 2007

- MS/MS is both a rapid and a highly sensitive and specific technology for detection of MCADD and phenylketonuria
- Doubts exist as to GAI and tyrosinaemia type I, and there is no evidence that would support the inclusion of the remaining IMDs
- Additional studies are needed to ascertain the sensitivity and specificity of tandem mass spectrometry in the detection of other IMDs, by assessing the long-term effectiveness of diagnostic strategies and conventional treatment, and the potential impact of early diagnosis using MS/MS
- Standardisation of existing screening programmes in Spain and the establishment of a case registry to enable active and regular follow-up of diagnosed patients was recommended

#### 6.5 Denmark (2008)<sup>26</sup>

This report examined the clinical effectiveness, cost-effectiveness and clinical utility of MS/MS screening for another 20 diseases. Based on a review of the literature and the results of a pilot research programme, the Danish Health Board recommended that the age at screening should be changed from day 4-6 to 42-72 hours and that MCADD and nineteen other conditions should also be screened for. The new programme was implemented in February 2009.

**Summary Danish Health Board report 2008**

- The working group proposed a consolidation and expansion of the screening program and a change in day of screening from day 4-6 to 42-72 hours.
- Conditions recommended to be included in the screening panel:

**Maple Syrup Urine Disease**

Argininosuccinic aciduria

HHH-syndrome

Methylmalonic acidaemia

Deficiencies of CPT 1 and II

Short chain acyl-CoA dehydrogenase deficiency

VLCADD

3-HMG-CoA lyase deficiency

Acetoacetyl-CoA thiolase deficiency

Citrullinaemia

Arginase deficiency

**Glutaric aciduria type 1**

Propionic acidaemia

MCADD

**LCHADD**

GA2

3-MCC deficiency

Galactosaemia

- Tyrosinaemia could be included following more work to reduce the false positive rate
- Although IVA was considered to meet virtually all requirements for screening it was not considered feasible for MS/MS screening due to the use of pivaloyl-containing antibiotics in Denmark which would interfere with the assay

(3-HMG, 3-hydroxy-3-methylglutaryl)

**6.6 Canada**

Three reports from three different health jurisdictions have been published in Canada since 2002 (Quebec, Ontario and British Columbia). The assessment carried out by Quebec was limited to the relevance of replacing current technologies by MS/MS for PKU and tyrosinaemia type I screening and including MCADD in their screening programme. The assessments carried out by Ontario and British Columbia considered a number of conditions which could also be detected by tandem mass spectrometry. Both reports updated information present in available HTA reviews carried out in the UK and reappraised the evidence in the light of local needs.

### ***Ontario 2003<sup>27</sup>***

This report evaluated seven conditions -CAH, MCADD, galactosaemia, PKU, GA1, CH and homocystinuria. Each of the conditions was weighted according to known incidence, available treatment, incremental cost to existing screening programme and their inclusion in international and other Canadian jurisdictions. The assessment showed that PKU and CH should continue to be screened for and that MCADD and CAH met most of the criteria for inclusion in the screening programme. Currently the screening programme in Ontario includes over 20 conditions, including GA1, homocystinuria, IVA, LCHAD deficiency and MSUD.

#### **Summary of report by Ontario Ministry of Health 2003**

- The assessment showed that screening for PKU and CH should be continued
- In addition, MCADD and CAH met most of the criteria for inclusion in a neonatal screening program
- MCADD can be screened with PKU by MS/MS while the test for CAH requires a different methodology
- An expanded neonatal program would require an enhanced infrastructure for result interpretation, reporting, care provision and counselling
- Important ethical and societal issues including informed consent need to be addressed

### ***British Columbia 2008<sup>28</sup>***

This report evaluated over 15 conditions that could be detected by MS/MS as well as some other conditions such as cystic fibrosis (CF), CAH and sickle cell disease. Recommendations were made following an update of the clinical evidence contained in reviews conducted by the UK and the United States and considerations of the harms and benefits of screening in the local context. A total of thirteen conditions were recommended for inclusion in an expanded screening program including homocystinuria, MSUD and IVA. LCHADD and GA1 were also among the conditions suggested for inclusion although they had already been part of the screening programme since October 2007. Full implementation of the expanded programme is expected by 2010.



### Summary of British Columbia HTA 2008

- The following conditions to be added to the newborn screening panel:

Cystic Fibrosis

Congenital Adrenal Hyperplasia

Sickle Cell disease

Glutaric aciduria type 1

VLCADD

LCHADD

Isovaleric acidaemia

Maple Syrup Urine disease

Propionic acidaemia

Homocystinuria

Tyrosinaemia

Citrullinaemia (argininosuccinate synthase deficiency)

Argininosuccinic aciduria

Methylmalonic acidaemia (Methylmalonyl-CoA mutase deficiency and Cobalamin A/B)

- Secondary targets (detected alongside other conditions but may not meet newborn screening criteria to be considered a primary target):

### Trifunctional protein deficiency (TFP)

2-Methylbutyryl-CoA dehydrogenase deficiency

Cobalamin C/D defects

## 6.7 The United States of America 2006<sup>29</sup>

Newborn screening programmes in the United States are governed by individual states and there were substantial variations in their newborn screening programmes. Some states mandated screening for as few as three conditions and others for as many as 43 conditions. This lack of uniformity led for calls to develop nationally recognised newborn screening system standards and policies. The US Department of Health and Human Services' Maternal and Child Health Bureau (MCHB) commissioned the American College of Medical Genetics (ACMG) to develop recommendations for the creation of a uniform panel of conditions to be screened for, along with model policies, procedures and minimum standards for state newborn screening programs. A total of 84 different conditions were considered and for each disease, information was collected on clinical characteristics (incidence, burden of disease), the analytical characteristics of the screening test, the diagnosis, treatment and management of the condition and the availability of health professionals experienced in this. The information was collected via questionnaires sent to health professionals and other interested parties in the USA and abroad.

Conditions were assigned to one of three categories (core panel, secondary targets, or not appropriate for newborn screening) based on the results of the questionnaire. The 2006 report recommended that screening for 29 core conditions should be mandatory and that diagnosis of a further 25 secondary conditions should also be reported. It also recommended that any abnormal results associated with clinically significant conditions, including the definitive identification of carrier status, should also be reported. Since 2006, those states that had not expanded their newborn screening programme have been doing so in order to reach a uniform panel.

**Summary of US newborn screening report 2006**

- A core panel of 29 conditions was recommended for mandatory screening in all states. These include:

Phenylketonuria	Maple Syrup Urine disease
Homocystinuria	Citrullinaemia
Argininosuccinic aciduria	Tyrosinaemia I(Tyr)
Isovaleric acidaemia	Glutaric acidaemia 1
3-HMG-CoA lyase deficiency	Propionic acidaemia
Holocarboxylase synthase deficiency MCD	$\beta$ -ketothiolase deficiency
Methylmalonic aciduria	MCADD
3-Methylcrotonyl-CoA carboxylase deficiency	VLCADD
Methylmalonic acidaemia (Cobalamin A/B)	LCHADD
Trifunctional protein deficiency	Carnitine uptake defect
Congenital hypothyroidism	Biopterin cofactors defects
Congenital adrenal hyperplasia	Galactosaemia
Hearing loss	Cystic fibrosis
Haemoglobinopathies	

- An additional 25 conditions were listed that could be identified in the course of screening for core panel conditions
- Recommended actions to overcome barriers to an optimal newborn screening included long-term data collection and surveillance, consideration of financial needs of programmes and standardisation of case definitions and reporting

## 6.8 Argentina (2005)<sup>30</sup>

This report concluded that although there was evidence for the effectiveness of screening for PKU, MCADD and GA1, the significance of this data for Argentina was uncertain. In addition, the cost effectiveness of screening would also be different as conventional screening is not extensively carried out. The authors concluded that more information about the various conditions and the impact of screening in the local setting was needed before further recommendations about newborn screening could be made.

### Summary of Argentine HTA 2005

- MS/MS sensitivity and specificity are adequate.
- More evidence is required to determine the natural history of these diseases and the economic impact of their screening in the Argentine setting.
- Major studies should focus on treatment and adverse effects compared to conventional screening.

## 6.9 Conclusions

Since 2002, many health jurisdictions have evaluated their newborn screening programmes with respect to MS/MS. All countries have based their evaluations on principles set out by Wilson and Jungner; however, these have been adapted to local contexts and interpreted differently. The variation in screening programmes is strongly influenced by the local epidemiology and burden of disease and the local structure and organisation of health care services.

Different countries have introduced MS/MS at different rates and for a varying number of conditions. Substantial expansions of newborn screening using tandem mass spectrometry were recommended by four health jurisdictions (British Columbia, Netherlands, Denmark, and USA). Certain conditions are included in all jurisdictions where newborn screening has expanded, including all five of the proposed disorders for the UK expanded newborn screening research project (MSUD, GA1, and LCHAD deficiency). The inclusion of IVA and homocystinuria to programmes has varied and this appears to be due to concerns over the analytical validity of the assay. The UK, Spain, and Argentina concluded that further research was needed to establish the performance of MS/MS as a screening test and to assess long-term effectiveness of strategies for diagnosis and treatment.

## 7 Evaluation of expanded screening programmes worldwide

This section of the report provides details of the systematic review undertaken to examine the epidemiology, effectiveness, clinical utility and cost-effectiveness of screening using tandem mass spectrometry.

### 7.1 Systematic review process and numbers of studies

#### *Quantity of research available*

This systematic review identified 1,807 additional publications on the epidemiology and natural history, clinical validity of MS/MS screening, effectiveness of interventions and of screening programmes for the five conditions of interest since 2002. Full copies of 252 papers were collected for further evaluation. A further 39 papers were identified by manual searching, grey literature, unpublished literature and by an expert in the field. Of these 291 papers, 82 were judged relevant to this review with 210 failing to meet our inclusion criteria (list of included and excluded articles available on request).

### 7.2 Epidemiology, natural history and outcome in screen detected cases

Sixteen centres published data on experience in newborn screening using MS/MS including one or all of these conditions since 2002. These are summarised in Table 7.1 and included 5 centres in Europe, 5 in North America, 2 in Australasia, 4 in East Asia and 1 in the Middle East. Nine centres included 4 or all 5 of our focus conditions, showing that there was broad agreement on the choice of these conditions as suitable for screening. Numbers of centres including each condition was MSUD (15 centres), GA1 (14 centres), homocystinuria (10 centres), IVA (11 centres) and LCHAD deficiency (8 centres).

#### *Incidence (birth prevalence) of disease in birth cohorts screened by MS/MS*

Each centre provided estimates of the birth prevalence of conditions for their cohorts. Tables 7.2-6 provide summaries by condition including information on age at screening and cut-offs applied. Table 7.7 provides overall summary data by condition and also incorporates estimated number of cases based on 708, 711 births in England and Wales (number of births from Office for National Statistics for 2008). When looking across all five conditions, the expected number of cases per year in England and Wales is between 16 and 23.

**Table 7.1 Identification of centres publishing studies including target IMDs**

Study Author	Country	HCY	MSUD	GA1	IVA	LCHADD
Europe						
Schulze <i>et al.</i> <sup>31</sup>	Germany	1	1	1	1	1
La Marca <i>et al.</i> <sup>32</sup>	Italy	1	1	1	1	1
Hoffman <i>et al.</i> <sup>33</sup>	Germany	0	1	1	1	0
Lund <i>et al.</i> <sup>34</sup>	Denmark	0	1	1	1	1
Vilarinho <i>et al.</i> <sup>35</sup>	Portugal	1	1	1	1	1
North America						
Frazier <i>et al.</i> <sup>36</sup>	North Carolina, USA	1	1	1	1	1
Matern <i>et al.</i> <sup>37</sup>	Mayo Clinic, USA	1	1	0	0	0
Comeau <i>et al.</i> <sup>38</sup>	New England, USA	1	1	0	1	0
Feuchtbaum <i>et al.</i> <sup>39</sup>	California, USA	0	1	1	1	1
Torres Sepulveda <i>et al.</i> <sup>40</sup>	Mexico	1	1	1	1	0
Australasia						
Wilcken <i>et al.</i> <sup>41</sup>	Australia	1	1	1	1	1
Boneh <i>et al.</i> <sup>42</sup>	Victoria State, Australia	0	0	1	0	0
Southeast Asia						
Yoon <i>et al.</i> <sup>43</sup>	S Korea	1	1	1	1	1
Abdul Rahman <i>et al.</i> <sup>44</sup>	Malaysia	0	1	0	0	0
Huang <i>et al.</i> <sup>45</sup>	Taiwan	0	1	1	0	0
Hsieh <i>et al.</i> <sup>46</sup>	Taiwan	0	0	1	0	0
Middle East						
Lindner <i>et al.</i> <sup>47</sup>	Qatar	1	1	1	0	0

Table 7.2 Homocystinuria prevalence in MS/MS screened cohorts

Study	Location	Period	Age at screening	Methionine cut-off $\mu\text{mol/L}$	Number of cases	Population screened	Birth prevalence 1 in	Per 100,000
Schulze <i>et al.</i> <sup>31</sup>	Germany	1998 - 2001	3 to 7 days (median 5)	65	0	250,000	NA	NA
La Marca <i>et al.</i> <sup>32</sup>	Italy	2002 onwards	48 to 72 hrs	46	0	160,000	NA	NA
Vilarinho <i>et al.</i> <sup>35</sup>	Portugal	2005-2009	3 to 6 days	50	1	316,243	316,243	0.32
Frazier <i>et al.</i> <sup>36</sup>	USA	1997 - 2005	Mean 39 hrs	96.6	2	944,078	472,039	0.21
Matern <i>et al.</i> <sup>37</sup>	USA	2004 - 2007	NS	100	1	260,936	260,936	0.38
Comeau <i>et al.</i> <sup>38</sup>	New England	1999 - 2003	At weight of 2500g	67	1	472,255	472,255	0.21
Wilcken <i>et al.</i> <sup>41</sup>	Australia	1998 - 2002	2 to 3 days	80.00	2	461,500	230,750	0.43
Wilson <i>et al.</i> <sup>48</sup>	Australia	2004 - 2006	2 to 3 days	80.00	1	270,000	270,000	0.37
Torres Sepulveda <i>et al.</i>	Mexico	2002 - 2004	24 to 48 hrs	132.75	1	42,264	42,264	2.37
Yoon <i>et al.</i> <sup>43</sup>	S Korea	2001 - 2004	95% before 72 hrs	87	0	79,149	NA	NA
Lindner <i>et al.</i> <sup>47</sup>	Qatar	2003 - 2006	36 to 72 hrs	NS	2	25,214	12,607	7.93
<b>Total worldwide</b>					<b>11</b>	<b>3, 281, 639</b>	<b>298, 331</b>	<b>0.34</b>
<b>Western populations</b>					<b>8</b>	<b>3, 135, 012</b>	<b>391, 877</b>	<b>0.26</b>

NS: Not stated



**Table 7.3** MSUD prevalence in MS/MS screened cohorts

Study	Location	Period	Age at screening	Cut-off μmol/L	Number of cases	Population screened	Birth prevalence 1 in	Per 100,000
Schulze <i>et al.</i> <sup>31</sup>	Germany	1998 - 2001	3 to 7 days (median 5)	Leu 490 Val 390	2	250,000	125,000	0.8
La Marca <i>et al.</i> <sup>32</sup>	Italy	2002 onwards	48 to 72 hrs	Val 210 Ile/ Leu 195	0	160,000	NA	NA
Hoffman <i>et al.</i> <sup>33</sup>	Germany	1999 and 2000	NS	NS	3	382,247	127,416	0.78
Lund <i>et al.</i> <sup>34</sup>	Denmark	2002 - 2005	5 to 10 days	NS	0	197,000	NA	NA
Vilarinho <i>et al.</i> <sup>35</sup>	Portugal	2005-2009	3 to 6 days	Leu 342 Val 350	3	316,243	105,141	0.95
Frazier <i>et al.</i> <sup>36</sup>	USA	1997 - 2005	Mean 39 hrs	Leu 450 Val 275	1	944,078	944,078	0.11
Matern <i>et al.</i> <sup>37</sup>	USA	2004 - 2007	NS	NS	0	204,281	NA	NA
Comeau <i>et al.</i> <sup>38</sup>	USA	1999 - 2003	at weight of 2500g	Leu 373	2	472,255	236,128	0.42
Feuchtbaum <i>et al.</i> <sup>39</sup>	USA	2002 -2003	NS	Leu 300 Leu/ Ala ratio > 1.75	2	353,894	176,947	0.57

Table 7.3 continued MSUD prevalence in MS/MS screened cohorts

Study	Location	Period	Age at screening	Cut-offs μmol/L	Number of cases	Population screened	Birth prevalence 1 in	Per 100,000
Wilcken <i>et al.</i> <sup>41</sup>	Australia	1998 - 2002	2 to 3 days	Leu/Ile 500	3	461,500	153,833	0.65
Wilson <i>et al.</i> <sup>48</sup>	Australia	2004 -2006	2 to 3 days	Leu/Ile 500	1	270,000	270,000	0.37
Torres Sepulveda <i>et al.</i> <sup>40</sup>	Mexico	2002 - 2004	24 to 48 hrs	Leu/Ile 450	0	42,264	NA	NA
Yoon <i>et al.</i> <sup>43</sup>	S Korea	2001 -2004	95% before 72 hrs	Leu/Ile 402	2	79,149	39,575	2.53
Abdul Rahman <i>et al.</i> <sup>44</sup>	Malaysia	2006 - 2008	1 to 7 days	NS	2	13,793	6,897	14.50
Huang <i>et al.</i> <sup>45</sup>	Taiwan	2001 -2004	48 hrs	Leu/Ile 171.66	2	199,992	99,996	1.00
Abdel-Hamid <i>et al.</i> <sup>49</sup>	Kuwait	2004 - 2006	First week	NS	1	1,158	1,158	86.36
Lindner <i>et al.</i> <sup>47</sup>	Qatar	2003 - 2006	36 to 72 hrs	NS	2	25,214	12,607	7.93
<b>Total worldwide</b>					<b>26</b>	<b>4,373, 068</b>	<b>168, 195</b>	<b>0.59</b>
<b>Western populations</b>					<b>17</b>	<b>3,807, 217</b>	<b>223, 954</b>	<b>0.45</b>

**Table 7.4** GA1 prevalence in MS/MS screened cohorts

Study	Location	Period	Age at screening	Cut-offs $\mu\text{mol/L}$	Number of cases	Population screened	Birth prevalence One in	Per 100,000
Schulze <i>et al.</i> <sup>31</sup>	Germany	1998 - 2001	3 to 7 days (median 5)	C5DC >0.14; Molar ratio C5DC/C8 (>1.8), C5DC/C16 (>0.06)	3	250,000	83,333	1.2
Kolker <i>et al.</i> <sup>50</sup>	Germany	1999 - 2005	NS	NS	38	3,807,600	100,200	1.0
La Marca <i>et al.</i> <sup>32</sup>	Italy	2002 onwards	48 to 72 hrs	0.018	1	160,000	160,000	0.63
Hoffman <i>et al.</i> <sup>33</sup>	Germany	1999 - 2000	NS	NS	2	382,247	191,124	0.52
Lund <i>et al.</i> <sup>34</sup>	Denmark	2002 - 2005	5 to 10 days	NS	0	197,000	NA	NA
Vilarinho <i>et al.</i> <sup>35</sup>	Portugal	2005-2009	3 to 6 days	C5DC >0.2	6	316,243	52,707	1.90
Frazier <i>et al.</i> <sup>36</sup>	USA	1997 - 2005	Mean 39 hrs	C5DC >0.38	5	944,078	188,816	0.53
Feuchtbaum <i>et al.</i> <sup>39</sup>	USA	2002 - 2003	NS	NS	1	353,894	353,894	0.28
Wilcken <i>et al.</i> <sup>41</sup>	Australia	1998 - 2002	2 to 3 days	NS	3	461,500	153,833	0.65
Wilson <i>et al.</i> <sup>48</sup>	Australia	2004 -2006	2 to 3 days	NS	4	270,000	67,500	1.48

Table 7.4 continued GA1 prevalence in MS/MS screened cohorts

Study	Location	Period	Age at screening	Cut-offs μmol/L	Number of cases	Population screened	Birth prevalence 1 in	Per 100,000
Boneh <i>et al.</i> <sup>42</sup>	Victoria	2001 - 2007	NS	NS	6	391,651	65,275	1.53
Torres Sepulveda <i>et al.</i> <sup>40</sup>	Mexico	2002 - 2004	24 to 48 hrs	NS	0	42,264	NA	NA
Yoon <i>et al.</i> <sup>43</sup>	S Korea	2001 - 2004	95% before 72 hrs	C5DC 0.2	2	79,179	39,590	2.53
Huang <i>et al.</i> <sup>45</sup>	Taiwan	2001 - 2004	48 hrs	0.015	2	199,992	99,996	1.00
Hsieh <i>et al.</i> <sup>46</sup>	Taiwan	2001 - 2006	3 days	C5DC 0.22	5	357,307	71,461	1.40
Lindner <i>et al.</i> <sup>47</sup>	Qatar	2003 -2006	36 to 72 hrs	NS	6	605,000	100,833	1.00
<b>Total worldwide</b>					<b>84</b>	<b>8,817,955</b>	<b>104,976</b>	<b>0.95</b>
<b>Western populations</b>					<b>69</b>	<b>7,534,213</b>	<b>109, 191</b>	<b>0.92</b>

**Table 7.5 IVA prevalence in MS/MS screened cohorts**

Study	Location	Period	Age at screening	Cut-offs µmol/L	Number of cases	Population screened	Birth prevalence 1 in	Per 100,000
Schulze <i>et al.</i> <sup>31</sup>	Germany	1998 - 2001	3 to 7 days (median 5)	C5 >2 or C5/acyl carnitine molar ratio >0.06	4	250,000	62,500	1.60
La Marca <i>et al.</i> <sup>32</sup>	Italy	2002 onwards	48 to 72 hrs	C5 levels; 0.56	1	160,000	160,000	0.63
Hoffman <i>et al.</i> <sup>33</sup>	Germany	1999 - 2000	NS	NS	2	382,247	191,124	0.52
Lund <i>et al.</i> <sup>34</sup>	Denmark	2002 - 2005	5 to 10 days	NS	1	197,000	197,000	0.51
Vilarinho <i>et al.</i> <sup>35</sup>	Portugal	2005-2009	3 to 6 days	C5 > 1	3	316,243	105,141	0.95
Frazier <i>et al.</i> <sup>36</sup>	USA	1997 -2005	Mean 39 hrs	C5 was >1.16	7	944,078	134,868	0.74
Comeau <i>et al.</i> <sup>38</sup>	USA	2002 -2003	1 to 3 days	C5 1.2	0	472,255	NA	NA
Feuchtbaum <i>et al.</i> <sup>39</sup>	USA	2002 - 2003	NS	C5 1.2	1	353,894	353,894	0.28
Wilcken <i>et al.</i> <sup>41</sup>	Australia	1998 - 2002	2 to 3 days	C5 >2	2	461,500	230,750	0.43
Wilson <i>et al.</i> <sup>48</sup>	Australia	2004 -2006	2 to 3 days	C5 >2	0	270,000	NA	NA
Torres Sepulveda <i>et al.</i> <sup>40</sup>	Mexico	2002 - 2004	24 to 48 hrs	C5 1.19	0	42,264	NA	NA
Yoon <i>et al.</i> <sup>43</sup>	S Korea	2001 - 2004	95% before 72 hrs	C5 0.396 to 2003 then C5 1.2	3	79,179	26,393	3.79
<b>Total worldwide</b>								
<b>Western populations</b>								
					<b>24</b>	<b>3,928,660</b>	<b>163,694</b>	<b>0.61</b>
					<b>21</b>	<b>3,263,328</b>	<b>155,396</b>	<b>0.64</b>

Table 7.6 LCHADD prevalence in MS/MS screened cohorts

Study	Location	Period	Age at screening	Cut-offs $\mu\text{mol/L}$	Number of cases	Population screened	Birth prevalence 1 in	Per 100k
Schulze <i>et al.</i> <sup>31</sup>	Germany	1998 - 2001	3 to 7 days (median 5)	C14OH 0.1 or/and C16:1OH >0.22; C16OH >0.20; C18:1OH >0.12 or C18OH >0.11	1	250,000	250,000	0.40
Sander <i>et al.</i> <sup>51</sup>	Germany	1999 - 2005	36h to 72 hrs	C14:1 >0.35; C14OH >0.2; C16OH >0.08; C18:1OH >0.06	11 <sup>a</sup>	1,200,000	109,000	0.92
La Marca <i>et al.</i> <sup>32</sup>	Italy	2002 onwards	48 to 72 hrs	C16OH 0.1	0	160,000	NA	NA
Lund <i>et al.</i> <sup>34</sup>	Denmark	2002-2005	5 to 10 days	NS	0	197,000	NA	NA
Vilarinho <i>et al.</i> <sup>35</sup>	Portugal	2005-2009	3 to 6 days	C16OH >0.10; C18OH > 0.06 C18:1OH >0.07; C16OH/C16 >0.04 Valine 350	3	316,243	105, 141	
Frazier <i>et al.</i> <sup>36</sup>	USA	1997-2005	Mean 39 hrs	C16:1OH >0.18; C18:1OH >0.14	3	944,078	314,693	0.32
Feuchtbaum <i>et al.</i> <sup>39</sup>	USA	2002-2003	NS	NS	1	353,894	353,894	0.28
Wilcken <i>et al.</i> <sup>41</sup>	Australia	1998 - 2002	2 to 3 days	C16OH >0.1	0	461,500	NA	NA
Wilson <i>et al.</i> <sup>48</sup>	Australia	2004 - 2006	2 to 3 days	C16OH >0.1	0	270,000	NA	NA
Yoon <i>et al.</i> <sup>43</sup>	S Korea	2001-2004	95% before 72 hrs	C16:1OH 0.12; C16OH 0.08; C18:1OH 0.07; C18OH 0.05	3	79,179	26,393	3.79
<b>Total worldwide</b>					22	4,231, 894	192, 359	0.52
<b>Western populations</b>					19	4,152, 715	218, 564	0.46
<b>Total worldwide excluding Sander study</b>					11	3,031, 894	275, 627	0.36
<b>Western populations excluding Sander study</b>					8	2,952, 715	369, 089	0.27

<sup>a</sup> 11 cases with proven MTP deficiency

**Table 7.7 Summary overall birth prevalence data**

Condition	European birth prevalence one per	European birth prevalence per 100,000	Estimated number of cases per 708,711 (in England and Wales)
MSUD	223,954	0.45	3.19
Homocystinuria	391,877	0.26	1.84
GA1	109,191	0.92	6.52
IVA	155,396	0.64	4.54
LCHADD	218,564	0.46	3.26
Overall		<b>2.73</b>	<b>19.35</b> (15.72 to 22.98)

***UK data on birth prevalence***

Data was available from two UK laboratories that continued to screen for selected amino acid disorders after the switch from chromatography to MS/MS. Results are given in Table 7.8.

**Table 7.8 UK lab data on incidence (birth prevalence) for homocystinuria and MSUD**

	Date	Newborns screened	Cases diagnosed	Birth prevalence 1 in
Homocystinuria				
Lab 1	2006 to 2008	128,003	1	128,000
Lab 2	2001 to 2009	450,000	3	150,000
MSUD				
Lab 1	2006 to 2008	128,003	2	64,000
Lab 2	2001 to 2009	450,000	3	150,000

These figures are somewhat higher than European figures in the literature. Both of these laboratories operate in areas with high immigrant populations and may therefore be expected to be high prevalence areas for inherited metabolic disorders.



### **Comparison of screened detected cases with birth prevalence of clinically detected cases**

Three studies looked at the same or very similar populations over similar time periods and applied methods to reduce the impact or allow for bias.

Hoffman *et al.*<sup>33</sup> presented data on clinically symptomatic cases (older than 7 days) ascertained with active surveillance through the German Paediatric Surveillance Unit (ESPED) in a birth cohort of 844, 575 conventionally screened children compared with a cohort of 382, 247 screened by MS/MS in two areas of Germany. In the surveillance cohort there were 39 cases in total, of which 6 presented in the first 7 days and so would not necessarily have been detected by the screening programme. These included 1 case of IVA. Numbers of our target disorders in the remainder of the surveillance cohort and the screened cohort are presented in Table 7.9.

**Table 7.9 Target disorders in surveillance and screen detected cohorts**

	Surveillance cohort	Newborn screening cohort (13 organic acidurias and 33 patients in total)
IVA	2	2
GA1	3	2
MSUD	3	3

The birth prevalence for organic acidurias (a group which included MSUD, IVA and GA1) based on diagnosis of screened children was 1 in 29,000 (55,000 to 17,000) compared with a frequency of 1 in 60,000 (1 in 110,000 to 1 in 36,000) for those diagnosed clinically. This is a ratio of approximately 2:1 (including correction for possible under-reporting using the capture/recapture method with data from two independent sources of surveillance). It was suggested that there would also be ascertainment bias due to some cases manifesting in the first 7 days (these were excluded from the analysis), and cases that might have escaped diagnosis due to fulminant death or omission of testing despite typical symptoms. However, there was no suggestion that any screen detected cases were not significant. The authors observe that *'considerable morbidity and mortality was observed for the ten conditions studied which may be diagnosed by MS/MS with no extra costs regarding the analytical procedure, only minor extra time required for the interpretation of additional results and little risk of harm'*. They argue that *'new screening concepts would allow identification of almost twice as many affected children compared with existing diagnostic procedures. Provided that the organisational, technical and ethical issues involved are addressed properly a health care problem of significant size could be beaten by the expansion of the MS/MS programme'*.

Wilcken *et al.*<sup>41</sup> compared the rates of detection of 31 IMDS (including diseases of urea cycle, amino acids (includes homocystinuria and MSUD), organic acids (includes IVA and GA1) and fatty-acid oxidation (includes LCHADD) among 362,000 newborns screened by tandem mass spectrometry over a 4 year period (1998 to 2002) born in New South Wales or the Australian Capital Territory with the rates in the 6 preceding 4 year birth cohorts in the same geographic area (1974 to 1998). During the 6 four year periods preceding implementation of MS/MS the rate of diagnosis for all conditions was 6.6 to 9.0 per 100,000 (adjusted to 6.9 to 9.5 cases per 100,000 when possible late diagnosis was taken into account). There was no trend toward increased rates during this period. After MS/MS was introduced the rate went up to 15.7 diagnoses per 100,000 (11.9 to 20.4) of which 48/57 newborns were diagnosed by MS/MS and a further 6 were diagnosed simultaneously with screening results.

The excess of cases in the screened cohort was confined to a small number of conditions, including medium chain and short chain acyl-CoA dehydrogenase deficiency (SCADD) and a small number of other conditions not included in our target group. The authors comment that some patients with these conditions could be at a small risk of never having symptoms and largely benign disorders. However, the authors argue that *‘with the possible exception of short chain acyl CoA dehydrogenase deficiency and 3-methylcrotonyl CoA carboxylase deficiency the other disorders can all lead to substantial morbidity and mortality’* and thus *‘the argument that MS/MS diagnoses cases that would not otherwise come to clinical attention should not be used as an argument against the programme’*.

Using the same Australian cases as a comparison, Wilson *et al.*<sup>48</sup> examined the New Zealand population through the New Zealand Paediatric Surveillance Unit. The study took place in a 3 year period from 2004 to 2006. During this time period there were 8 cases of treatable metabolic disease in the unscreened cohort (including 1 case of MSUD and 1 case of GA1), overall rate 8 per 175,000 or 1 in 22,000 compared with 45 cases in total (including 1 MSUD, 1 homocystinuria and 4 GA1) - overall rate 45 per 270,000 (or 1 in 6,000) in the Australian screened cohort.

Dionisi-Vici *et al.*<sup>52</sup> compare the birth prevalence of organic acidurias (including 1 case of IVA) in a cohort of clinically detected cases diagnosed and followed at the Bambino Gesù Children’s Hospital in Rome with a combined cohort of patients diagnosed by neonatal screening in Australia and Germany (including 7 cases of IVA). The incidence of IVA was much higher in the screened population,  $0.89 \pm 0.49$  per 100,000 vs.  $0.2 \pm 0.13$  in the clinically detected cohort. They ascribe this difference either to under-detection by conventional diagnosis or of over detection of mild cases by MS/MS. They consider that this might be due to a common mutation of the isovaleryl-CoA dehydrogenase gene (932C>T) associated with a mild phenotype or even asymptomatic IVA<sup>16</sup>. This is further supported by observation of 6 cases of IVA diagnosed by neonatal screening with a fully normal phenotype<sup>53</sup>.

Feuchtbaum *et al.*<sup>39</sup> comment on the issue of ‘missed expected cases’ in their study of the state mandated programme in California. They comment that most of the discrepancy is explained by MCADD (1 expected vs. 2 reported) because most would not yet have been diagnosed on the basis of clinical symptoms and SCADD (18 vs. 0 cases) because many cases are mild or would never be reported. They added that it is not known whether any of the 18 screen detected cases would have symptoms. **None of the diseases in our targeted group were included in this group of diseases that would be missed and/or mild.**

## Summary of epidemiology, natural history and outcome in screen detected cases

### *Issues with study design*

- Problems of bias due to different standards for diagnosis, evolving treatments, different populations
- Need to study:
  - Same geography different time
  - Different time same geography
  - Clinically ascertained cohorts should have some means of active surveillance

### *Findings from four studies comparing screened and clinical cohorts*

- Hoffman *et al.*<sup>33</sup> (Germany): organic acidurias - approximately double
- Wilcken *et al.*<sup>41</sup> (Australia): 31 IMDs - approximately double: much of this thought to be MCADD
- Wilson *et al.*<sup>48</sup> (New Zealand unscreened vs. Australia) screened - 3-4 times in screened
- Feuchtbaum *et al.*<sup>39</sup> (California) - much of the discrepancy is due to MCADD and SCADD

### *Conclusions from authors*

- Under-reporting
- Missed cases due to 'fulminant death' without diagnosis
- Excess cases due to MCADD, SCADD and a small number of other conditions that might not ever present - not including our target conditions

### 7.3 Analytical validity

Existing screening programmes utilizing MS/MS have gathered information on analytical validity initially through analysis of blood spots spiked with varying amounts of a specific analyte. Repeat analysis of these samples allows establishment of the precision and accuracy of the assay in detecting a target analyte. It also allows validation of the internal standards that are to be used to quantify the amounts of analyte in samples. Some centres have compared the results obtained by MS/MS with that of conventional assays such as the bacterial inhibition assay or HPLC where available. Retrospective analysis of samples from individuals with and without known abnormalities further establishes the analytical parameters of the assay. In such cases, the results of the analysis by MS/MS are confirmed through comparison with a gold standard method if one exists. However, it should be noted that for acylcarnitines there are no alternative assays.

### 7.4 Clinical validity - test performance in patients

#### *Reference standards*

A critical component of assessing clinical validity is the definition of an appropriate reference standard. This reference standard may be a set of clinical diagnostic criteria or the application of an alternative testing regimen in all those screened. The MS/MS screening test provides the 'sieve' initially sorting patients into positive (or 'flagged') patients and 'unflagged' patients. In most cases, in samples that are flagged or have borderline results, a repeat analysis of the blood spot is carried out by MS/MS. For some conditions, notably homocystinuria and MSUD, this may be followed by a second-tier test also done on the blood spot, after which screen positive patients would be recalled for further testing. In other conditions, notably GA1, initial flagging is followed by immediate clinical contact and further testing. Appropriate reference standards for final diagnosis include such elements as further biochemistry, biopsy and clinical assessment. In all cases the reference standards used by screening programmes are only applied to screen positive cases (examples in Box 7.1). The diagnosis of patients missed by screening in all studies thus relies on clinical presentation.

### Box 7.1 Examples of reference standard process to determine ‘true’ disease status

#### Baden-Wurttemberg, Germany

A sample was classified a true screen positive if the first and second tests were positive. If a distinct discrepancy between the first and second tests occurred a third test was done and the mean of the two corresponding values used. An experienced disease metabolic specialist decided if the flagged samples were abnormal or normal. In case of an abnormal ‘presumptive positive’ result a repeat dried blood spot specimen was obtained on recall or the patient was referred to a metabolic centre if avoidance of further delay was felt to be essential. Confirmatory investigations involved the following:

- Homocystinuria: enzyme activity in fibroblasts, mutational analysis and total homocysteine in plasma levels exceeding 20 $\mu$ M
- MSUD: analysis of plasma amino acids and presence of *allo*-isoleucine, enzyme activity in fibroblasts. For classical MSUD leucine in the range 500 to 5,000 $\mu$ M, and BCKD activity of less than 2%. For variant MSUD leucine in the range 40 to 4,000 $\mu$ M, and BCKD activity between 2 to 40%
- GA1: enzyme activity in fibroblasts and 3-hydroxyglutaric acid and glutaric acid in urine
- IVA: enzyme activity in fibroblasts and 3-hydroxyisovaleric acid and isovaleryl glycine in urine
- LCHADD: enzyme activity in fibroblasts/lymphocytes

Note this process was designed for a research study. In routine clinical use diagnosis can be confirmed in most cases (except in LCHADD and non-excretor GA1) without recourse to enzyme studies or mutation analysis.

#### New South Wales, Australia

Samples with all results inside the cut-off values were reported as “no further tests required”. If an analyte was marginally abnormal a repeat dried blood spot sample was collected. If follow-up tests were abnormal or if the original sample was significantly abnormal further investigation included a repeat blood spot sample and diagnostic work up, depending on the condition. The protocol for GA1 was altered to recommend immediate clinical and biochemical review, without request for a second sample.

### **Screening cut-offs**

Cut-offs set by newborn screening programmes varied and were usually based upon results obtained from screening a number of healthy newborns. They were usually set to optimise the balance between sensitivity and specificity, although the actual balance may vary from country to country. Cut-offs may be biased towards certain variants of disease as illustrated by Bhattacharya *et al.*<sup>54</sup>. The authors report on two cases of intermittent MSUD who participated in the New South Wales expanded screening programme but had normal results in the newborn period. They subsequently presented in early childhood with acute neurological problems and developmental delay. Review of the MS/MS screening result showed that the two patients had results that would have been below any cut-off level associated with a reasonable recall rate. These non-classical forms are in danger of being missed by the programme, (and of having worse outcomes than infants with more severe forms diagnosed at birth). The authors comment that *'ironically, newborn screening can identify classic MSUD, leading to a normal developmental outcome, but may not identify the 'milder' forms, in which developmental impairment is probable'*. They argue that this is not a reason not to screen but that it remains important to consider the diagnosis of MSUD in the appropriate clinical setting irrespective of the newborn screening result.

### **Test performance**

In this section we also look at test performance and particularly true positives, true negatives, false positive and false negatives and estimates of sensitivity, specificity and positive and negative predictive values. Twelve papers provided some information on test performance in relation to the individual conditions whilst a further five papers provided further information on relevant groups of conditions, which would have included our target conditions. Tables 7.10-14 present the data obtained on test performance from published reports of screening programmes. Many of the studies had very low positive predictive values. This maybe a consequence of the relatively young age of screening and hence low cut-offs. Furthermore, sensitivity and specificity were improved through the use of second-tier tests.

### *The use of second-tier tests*

One paper presented data on the performance of second-tier tests. Matern *et al.*<sup>37</sup>, presented data on the Mayo Clinic (USA) newborn screening program following the introduction of second-tier tests from 2004 to 2007. Second-tier tests were used in order to reduce false positive results for homocystinuria and MSUD. An initial result of methionine in excess of 100µmol/L is considered a positive screening result. Blood samples with methionine concentrations either below 9µmol/L or between 60 and 100µmol/L undergo a second-tier test (measuring homocysteine levels with a positive result for concentrations over 15µmol/L). The second-tier test was done for 516 samples with 515 screening negative and one true positive result. Two cases with elevated methionine and normal homocysteine were included in the 515 negative second-tier tests; these cases have been diagnosed with hypermethioninaemia. The authors state that analytical specificity of methionine for homocystinuria is relatively low and conservative cut-off values result in high false positive rates and unnecessary follow-up while reducing the risk of false negative results. The introduction of a second-tier test using homocysteine has reduced the amount of follow-up required to confirm positive screened cases as true positives.

With respect to MSUD, BCAAs are frequently elevated in newborns receiving total parental nutrition and approximately 0.1% of newborn screening samples in the series reported by Matern *et al.*<sup>37</sup> revealed elevations of BCAAs and in some cases require follow-up to rule out MSUD. Therefore, an LC-MS/MS-based assay was developed to quantify BCAAs, *allo*-ile, and OH-Pro. No patient with MSUD has thus far been identified prospectively by newborn screening. This second-tier test has however reduced the number of false positives for BCAA elevations.



**Table 7.10** Summary test performance for studies reporting data for homocystinuria

Study	Period	Day of screening	Cut-off Met $\mu\text{mol/L}$	Number screened	TN	TP	FN	FP	Sens. %	Spec. %	PPV	NPV
Schulze <i>et al.</i> <sup>31</sup>	1998-2001	3 to 7 (median 5)	65	250,000	249,992	0	0	8	NA	99.99	NA	100
<sup>a</sup> Matern <i>et al.</i> <sup>37</sup>	2004-2003	NS	60	260,936	260,935	1	0	0	100	100	100	100
<sup>b</sup> Frazier <i>et al.</i> <sup>36</sup>	1997-2005	Mean 39 hrs	Borderline 106 Diagnostic 300	239,415	239,407	1	0	7	100	99.99	12.5 (4)	100
Comeau <i>et al.</i> <sup>38</sup>	1999-2003	At weight of 2500g	67	472,255	472,046	1	0	209	100	99.96	0.5	100

<sup>a</sup> These data are from a study following the introduction of second-tier tests based on measurement of homocysteine. Of the 260,935 infants screened, 516 (0.29%) underwent second-tier testing and one positive screen case was confirmed with homocystinuria, 515 were negative on the second-tier screen.

<sup>b</sup> The study screened 944,078 births in total between 1997-2005 and two cases of homocystinuria were detected. These numbers were calculated from the detailed information provided for the period 2003-2004 based on 239,415 newborns screened.

**Note:** In consideration of the variation in this data it should be noted that Schultze sampled blood later in life than in the three USA studies. Frazier used a fairly high cut-off. Matern and Comeau used similar cut-offs and obtained somewhat similar initial positive rates but Matern carried out 516 “second-tier” tests and so had 0 false positives. Frazier has a 2-tier cut-off, babies with Met between 106 and 300 had a second blood sample taken. This example emphasises the importance of the full protocol for screening performance and of defining exactly what is meant by “positive”.

Table 7.11 Summary test performance for studies reporting data for MSUD

Study	Period	Day of screening	Cut-offs $\mu\text{mol/L}$	Population screened	TN	TP	FN	FP	Sens. %	Spec.%	PPV	NPV
Schulze <i>et al.</i> <sup>31</sup>	1998-2003	3 to 7 (median 5)	Leu 490 Val 390	250,000	249,975	2	0	23	100	99.99	8	100
<sup>a</sup> Matern <i>et al.</i> <sup>37</sup>	2004-2003	NS	NS	260,936	NS	NS	NS	FPR 0.09%	NS	NS	41	NS
Frazier <i>et al.</i> <sup>36</sup>	1997-2005	Mean 39 hours	Leu 450 Val 275	944,078	NS	1	1	0 <sup>b</sup>	NS	100	NS	100
Comeau <i>et al.</i> <sup>38</sup>	1999-2003	At weight of 2500g	Leu 373	472,255	472,075	2	0	178	100	99.96	1.1	100
Feuchtbaum <i>et al.</i> <sup>39</sup>	2002-2003	NS	Leu 300 Leu/Ala ratio >1.75	353,894	352,853	1	1	39	50	99.9	2.5 (0.14)	100
Wilcken <i>et al.</i> <sup>41</sup>	1998- 2002	2 to 3	Leu/Ile 500	461,500	NS	2	1	NS	66.7	NS	NS	NS
Huang <i>et al.</i> <sup>45</sup>	2001-2004	48 hrs	Leu/Ile 171.66	199,922	199,920	2	0	0	100	100	100	100
Abdel-Hamid <i>et al.</i> <sup>49</sup>	2004-2006	NS	NS	1,158	NS	1	0	NS	100	NS	NS	100
Lund <i>et al.</i> <sup>34</sup>	2002-2005	5-10	NS	170,000	169,977	1	0	22	100	99.99	4.3	100

<sup>a</sup> These data are from a study following the introduction of second-tier tests to quantify BCAA more accurately. No patient with MSUD has thus far been identified prospectively by newborn screening. This second-tier test has however reduced the number of false positives for BCAA elevations.

<sup>b</sup> The study screened 944 078 births in total between 1997-2005 and one case of MSUD was detected by the screening programme and one case of intermittent MSUD missed. The FP number was calculated from the detailed information provided for the period 2003-2004 based on 239,415 newborns screened.

Table 7.12 Summary test performance for studies reporting data for GA1

Study	Period	Day of screening	Cut-offs $\mu\text{mol/L}$	Population screened	TN	TP	FN	FP	Sens. %	Spec. %	PPV	NPV
Schulze <i>et al.</i> <sup>31</sup>	1998-2001	3 to 7 (median 5)	C5-DC 0.14	250,000	249,938	3	0	59	100	99.98	4.84	100
			C5-DC/C8 0.18									
			C5-DC/C16 0.6									
Frazier <i>et al.</i> <sup>36</sup>	1997-2005	Mean 39 hours	C5-DC 0.38	944,078	239,414 <sup>a</sup>	3	2	0 <sup>a</sup>	60	NS	100 <sup>a</sup>	99.99 <sup>a</sup>
Lindner <i>et al.</i> <sup>55</sup>	1998-2004	36h to 72h	C5-DC 0.17	605,000	604,788	6	0	206	100	99.96	2.83	100
Feuchtbauer <i>et al.</i> <sup>39</sup>	2002-2003	NS	NS	353,894	353,889	1	0	4	100	99.99	20	100
Huang <i>et al.</i> <sup>45</sup>	2001-2004	48 hrs	C5-DC 0.15	199,992	199,920	2	0	0	100	100	100	100
Hsieh <i>et al.</i> <sup>46</sup>	2001-2005	3	0.13	357,307	NS	2	0	38	recall rate 0.02	5%	100	100
	2005-2006		0.22									
Lund <i>et al.</i> <sup>34</sup>	2002-2005	5-10	NS	170,000	170,000	0	0	0	NA	100	NA	100

<sup>a</sup> The study screened 944 078 births in total between 1997-2005 and three cases of GA1 were detected by the screening programme; two cases were missed by the initial screening procedure, which involved a repeat screen. The procedure was changed and the FP number was calculated from the detailed information provided for the period 2003-2004 based on 239,415 newborns screened. The PPV for FAO disorders and GA1 was 60%

Table 7.13 Summary test performance for studies reporting data for IVA

Study	Period	Day of screening	Cut-offs $\mu\text{mol/L}$	Population screened	TN	TP	FN	FP	Sens. %	Spec. %	PPV	NPV
Schulze <i>et al.</i> <sup>31</sup>	1998-2001	3 to 7 (median 5)	C5 $2\mu\text{mol/L}$ C5/ acetyl carnitine molar ratio $>0.06$ .	250, 000	249, 961	4	0	33	100	99.99	10.81	100
Frazier <i>et al.</i> <sup>36</sup>	1997-2005	Mean 39 hours	C5 $>1.16\mu\text{mol/L}$	944, 078 (239,415 <sup>a</sup> )	239,413 <sup>a</sup>	7 (1 <sup>a</sup> )	0	1 <sup>a</sup>	100	100 <sup>a</sup>	50 <sup>a</sup>	100

<sup>a</sup> The study screened 944 078 births in total between 1997-2005 and seven cases of IVA were detected by the screening programme. The FP, specificity and PPV were calculated from the detailed information provided for the period 2003-2004 based on 239,415 newborns screened.

**Table 7.14 Summary test performance for studies reporting data for LCHADD**

Study	Period	Day of screening	Cut-offs μmol/L	Population screened	TN	TP	FN	FP	Sens. %	Spec. %	PPV	NPV
Schulze <i>et al.</i> <sup>31</sup>	1998-2001	3 to 7 (median 5)	C14OH 0.12 or/ and C16:1OH > 0.22 C16OH > 0.20 C18:1OH > 0.12 or C18OH > 0.11	250,000	249,988	1	0	10	100	100	9 (3.85)	100
<sup>a</sup> Frazier <i>et al.</i> <sup>36</sup>	1997-2005	Mean 39 hours	C16:1OH > 0.18 C18:1OH > 0.14	944,078	?	3	0	0	100	100	100	100
Feuchtbaum <i>et al.</i> <sup>39</sup>	2002-2003	NS	NS	353,894	353,893	0	1	0	NA	100	NA	99.99
Sander <i>et al.</i> <sup>51</sup>	2002-2005	36-72 h	C0 < 12 or > 60 C14:1OH > 0.35 C16OH 0.08 C18:1OH 0.06	1,200,000	1,199,979	11	0	10	100	99.99	52.4	100
Comeau <i>et al.</i> <sup>38</sup>	1999-2003	At weight of 2500g	C16OH 0.1	472,255	NS <sup>b</sup>	1	0	420 <sup>b</sup>	100%	NA	NA	100
Lund <i>et al.</i> <sup>34</sup>	2002-2005	5-10	NS	170,000	170,000	0	0	0	NA	100	NA	100

<sup>a</sup> The study screened 944 078 births in total between 1997-2005 and three cases of LCHADD were detected by the screening programme. The FP, specificity and PPV were calculated from the detailed information provided for the period 2003-2004. Overall positive predictive value for conditions with only one cut-off value was 60%

<sup>b</sup> The study by Comeau *et al.* identified one case of LCHADD amongst their MET19 panel which identified 28 positive cases. The FP number presented is not specifically for LCHADD but for the overall MET19 panel.

Five papers provided information on overall test performance for groups of disorders.

Torres-Sepulveda *et al.*<sup>40</sup> present the results of an expanded newborn screening programme conducted in Mexico, between 2002 and 2004. Samples were obtained at 24-48 hours after delivery from 42,264 newborns. Cut-offs were established following analysis of 500 samples, these were then revised after 10,000 samples were tested. Overall performance of MS/MS screening in this setting was:

TP	TN	FP	Recall rate	Cases still being followed
7	41,963	95	0.33%	42

Yoon *et al.*<sup>43</sup> present an updated summary report for the pilot MS/MS screening programme of newborns and children in high-risk groups for IMDs in South Korea. From 2001-2004, 79,179 newborns were screened within 72 hours of birth. Cut-offs were established following analysis of 6,000 full-term newborn blood spots. Samples flagged up for abnormal markers with the initial screen were reanalyzed to see if they confirmed the initial screen results. Confirmatory diagnosis involved repeating acylcarnitine profile, urine organic acid, plasma amino acid analysis, direct enzyme assay, or molecular testing. The overall performance of MS/MS screening in this setting was:

Sens.	Spec.	PPV	FPR	Recall Rate
96.67%	99.28%	6.38%	0.21%	0.05%

La Marca *et al.*<sup>32</sup> report on a six year screening programme in Italy where samples were analysed by LC-MS/MS. Samples from 160,000 newborns were analysed 48-72 hours after birth. Cut-offs were established during the pilot stage. The recall rate for whole programme was 1.35% in 2002 at start of pilot going up to 1.5% in 2005 after start of a project in Tuscany and then falling to 0.3% in 2008

Lindner *et al.*<sup>47</sup> report results from a screening programme in Qatar for the period 2003-2006. Samples from 25,214 newborns were analyzed 36-72 hours after birth. Overall recall rate for the programme was 1.8% which was reduced to 1% by 2006.

Feuchtbaum *et al.*<sup>39</sup> report on the overall performance of a pilot screening programme in California, USA. During the first five months of the programme the positive rate of screening (*i.e.* initially flagged specimens) was 0.49%, this was decreased to 0.07% following adjustments of cut-offs. During the entire 18 month pilot period, 0.20% of cases had an initial positive flag, of these 0.13% were reoffered follow-up services based on clinical review. Of these cases 72.7% were resolved normal, 7.8% were declined, 8.5% were lost and in 11% the diagnosis was resolved.

We were also able to gather data from Germany's national screening reports for the years 2004-2007. These were composed by the Deutschen Gesellschaft für Neugeborenen-Screening (DGNS) as well as the German screening laboratories and gave information on test performance and prevalence for four of our target conditions. The aggregated data for newborns greater than 32 weeks gestational age from fifteen laboratories is presented in Tables 7.15-18. The recommended timing of screening was 48-72 hours. The recall rate was fairly consistent over this four year period, except in the case of LCHADD. It is possible that changes in laboratory practice or screening algorithms had an influence on recall rate. Although the primary analyte measured was consistent in all laboratories, the secondary analytes varied, as did the cut-off values. The number of confirmed cases varied through the years and the positive predictive value was low for all conditions, however, this is to be expected given the rarity of these disorders.

Table 7.15 MSUD

Year	Primary screening (>36h)	Recall	Recall Rate %	PPV %	Patients Detected	Validity		Prevalence	False Negative	Primary analyte	Cut-off range (µM)	Secondary analytes
						Confirmed	Probable					
2004	728092	NS	0.03	3.26	4	4	0	1:145618	1	Leu/Ile	NS	Val, Ala, Leu/Ile:Phe, Leulle/Ala
2005	506864	84	0.01	NS	5	5	0	1: 139501	0	Leu/Ile	276-400	Val, Ala, Leu/Ile:Phe, Leulle/Ala
2006	670084	82	0.01	6.1	5	NS	NS	1:137161	0	Leu/Ile	263-400	Val, Ala, Leu/Ile:Phe, Leulle/Ala
2007	666071	91	0.01	2.2	2	2	NS	1: 343032	0	Leu/Ile	263-400	Val, Ala, Leu/Ile:Phe, Leulle/Ala



Table 7.16 GA1

Year	Primary screening (>36h)	Recall Rate %	PPV %	Patients Detected	Validity		Prevalence	False Negative	Primary analyte	Cut-off range (µM)	Secondary analytes
					Confirmed	Probable					
2004	728092	NS	0.06	6	6	0	1:121349	0	C5DC	NS	C5/Cn
2005	506864	151	0.02	6	6	0	1:116251	0	C5DC	0.12-0.51	C5/Cn
2006	670084	96	0.01	4	4	NS	1:171451	0	C5DC	0.12-0.66	C5DC/Cn
2007	666071	121	0.02	3	3	NS	1:228688	0	C5DC	0.12-0.666	C5DC/Cn

Table 7.17 IVA

Year	Primary screening (>36h)	Recall Rate %	PPV %	Patients Detected	Validity		Prevalence	False Negative	Primary analyte	Cut-off range (µM)	Secondary analytes
					Confirmed	Probable					
2004	728092	NS	0.02	10	9	1 (1*)	1:72809	0	C5	NS	C5/Cn
2005	506864	124	0.02	7	7	0	1:99643	0	C5	0.33-1	C5/Cn
2006	670084	91	0.01	10	10	NS	1:68580	0	C5	0.38-1	C5/Cn
2007	666071	86	0.01	4	4	NS	1:171516	0	C5	0.38-1	C5/Cn

Table 7.18 LCHADD

Year	Primary screening (>36h)	Recall	Recall Rate %	PPV %	Patients Detected	Validity		Prevalence	False Negative	Primary analyte	Cut-off range (µM)	Secondary analytes
						Confirmed	Probable					
2004	728092	NS	0.002	NS	4	2	2	1: 182023	0	C16OH	NS	C16:1OH, C18OH, C18:1OH, C16OH/C16
2005	506864	32	0.006	NS	3	3	0	1: 232501	0	C16OH	0.06-0.60	C16:1OH, C18OH, C18:1OH, C16OH/C16
2006	670084	19	0.003	21.05	5	5	NS	1:137161	0	C16OH	0.058-0.60	C16:1OH, C18OH, C18:1OH, C16OH/C16
2007	728092	NS	0.002	NS	4	2	2	1: 182023	0	C16OH	NS	C16:1OH, C18OH, C18:1OH, C16OH/C16

## Evidence on test performance from UK

Evidence was available from the two UK laboratories which continued to screen for selected amino acid disorders after the switch from chromatography to MS/MS.

**Table 7.19** Test performance in the UK

	Number screened	Screen positives	Eventual diagnosis
<b>Homocystinuria</b>			
Lab 1 2008 to 09	128,003	6	1 confirmed homocystinuria
			1 persistent hypermethioninaemia
			4 normal on follow up*
Lab 2	450,000	5	3 confirmed homocystinuria
			1 persistent hypermethioninaemia
			1 normal on follow-up*
<b>MSUD</b>			
Lab 1 2008 to 09	128,003	4	2
			1 very premature low birth weight baby, known as high risk for false positive
Lab 2	450,000	6	3 confirmed cases
			3 false positives

\*It is thought that two of these may have had liver disease.

Enquiry was made into the 'burden' of false positives in one of the services. The service has a very centralised clinical referral system and over this period all presumptive cases were dealt with at specialist inherited metabolic disease service. Presumptive positives were seen in the clinic on the next day. Blood and urine samples were collected for amino acids and liver enzymes (particularly important with increased methionine). Results were available the same or the next day and the parents could be told unequivocally whether their baby was affected or not.

**Was there evidence that the infants identified had significant disease or risk factors for disease?**

The reports were searched for evidence of clinical features of the screen detected cases that suggest the presence of pathology or evidence that the cases were deemed by metabolic disease specialists to require immediate preventive treatment and conversely for evidence that screening had uncovered diagnoses of uncertain significance of disease that metabolic specialists concluded was mild and did not need treatment. Only a few of the reports did provide this level of information for the five conditions studied. These are listed for each condition in relation to the reports published based on the various programmes (Table 7.20). Where no cases were detected (as noted in the Tables 7.2-6) there could be no description of the cases.

**Table 7.20 Evidence that infants identified had significant disease or risk factors for disease**

<b>Homocystinuria</b>	<p>Lindner <i>et al.</i><sup>47</sup> noted that treatment was commenced at days 35 and 38</p> <p>Yoon <i>et al.</i><sup>43</sup> Marked biochemical abnormalities noted</p>
<b>MSUD</b>	<p>Schulze <i>et al.</i><sup>31</sup> One classic form and one variant form. Classic MSUD symptomatic before result of screening and diagnosis made prior to screening. Treatment at 10 days still symptomatic (observed for 38 months). Variant MSUD asymptomatic at result of screening and treatment started at day 8 of life. Asymptomatic (observed for 24 months)</p> <p>Huang <i>et al.</i><sup>45</sup> Unwell day 9 respiratory distress and consciousness disturbance by day 19, treated</p> <p>Lindner <i>et al.</i><sup>47</sup> Treatment started at day 11 and 15 for these 2 cases</p> <p>Yoon <i>et al.</i><sup>43</sup> 2 patients with symptoms at day 10 and day 3 both had megalencephaly, 1 had developmental delay, both had mild motor delay and both were getting better under treatment</p>
<b>GA1</b>	<p>Schulze <i>et al.</i><sup>31</sup> The three confirmed cases of GA1 were asymptomatic when the screening result was obtained, treatment was started from 12-210 days and on follow-up which varied from 11-32 months, one patient had remained asymptomatic and two were symptomatic</p> <p>Hoffman <i>et al.</i><sup>33</sup> Both clinically asymptomatic</p> <p>Feuchtbaum <i>et al.</i><sup>39</sup> Diagnosis confirmed by clinical metabolic specialist</p> <p>Boneh <i>et al.</i><sup>42</sup> All patients were symptomatic and required treatment. Hospital admissions noted and some deficiencies in speed and motor activities</p> <p>Huang <i>et al.</i><sup>45</sup> Dietary therapy 42 days and 37 days. Both had mild hypotonia. The follow-up period was 2 years and 7 months for one patient and 1 year 10 months for the other</p> <p>Hsieh <i>et al.</i><sup>46</sup> All patients put on dietary therapy. The follow-up period varied from newly diagnosed to three years. All patients were reported to have normal development although initial symptoms such as macrocephaly, atrophy and frontotemporal atrophy were also reported</p> <p>Bijarnia <i>et al.</i><sup>56</sup> 6 patients - typical clinical and MRI findings in several, and mutation analysis or enzyme analysis on cultured skin fibroblasts in four cases. All of these patients had intercurrent illnesses however; six have remained well and had low scores on speech, motor and cognitive disabilities and one died aged 13 months</p> <p>Yoon <i>et al.</i><sup>57</sup> patient diagnosed at day 2, treated, no symptoms, well</p> <p>Schulze <i>et al.</i><sup>31</sup> Patients asymptomatic on diagnosis but all patients were treated starting days 7-11 and remained asymptomatic for mean 13.5 months</p>

IVA	<p>Hoffman <i>et al.</i><sup>33</sup> Not clinically symptomatic</p> <p>Frazier <i>et al.</i><sup>36</sup> 1 infant died on 7th day; no other information</p> <p>Schulze <i>et al.</i><sup>31</sup> 1 case symptomatic before the screening result was received and died one day later</p>
LCHADD	<p>Frazier <i>et al.</i><sup>36</sup> 1 case died from metabolic decompensation and/or prematurity before screening but test immediately after death was positive</p> <p>Sander <i>et al.</i><sup>51</sup> reports 11 screened detected cases: 1 patient had symptoms before screening result; 9 patients developed typical symptoms; only 1 patient developed normally to age 3. Symptoms included hypoglycaemia, cardiomyopathy, rhabdomyolysis in association with infections, episodes of metabolic decompensation</p>

The target diseases are heterogeneous, with severity related to a variety of factors (both known and unknown) and usually not capable of differentiation by the initial metabolites measured on the MS/MS screen. It is thus important to know how often MS/MS might pick up milder forms that might be asymptomatic or might even not require treatment. This will remain very difficult to assess although for some disorders e.g. MSUD, homocystinuria, GA1 and LCHADD, the biochemical phenotype expressed as metabolite concentration or residual enzyme activity might be used to help assess the severity of the clinical phenotype.

Fingerhut *et al.*<sup>58</sup> investigated whether classic and variant forms of MSUD can be discriminated by newborn screening so that the appropriate treatment can be provided whilst waiting for the result of a test positive. They concluded that an unambiguous discrimination is impossible by newborn screening and that a positive screen requires immediate referral to a specialised metabolic unit to confirm the diagnosis by quantitative determination of amino acids and to institute treatment. Measurement of plasma leucine concentrations then clearly differentiates between classic and variant forms with diagnosis being achieved between days 4 and 23. No newborn with variant MSUD required emergency management

Simon *et al.* (2006)<sup>8</sup> report on variants of MSUD, which are thought to represent about 20% of cases. Unlike classic forms with very low residual BCKD activity (0-2%), which present with catastrophic encephalopathic crisis and deep coma usually in second week of life, these patients (with residual activity of 3-40%) present differently and may have a milder clinical course and will be picked up by MS/MS. The authors studied 16 such patients aged 6-30 years (6 picked up by Guthrie screening). Of 6 patients picked up by screening, 4 were asymptomatic over a period of several years with no dietary restriction. These patients may represent a non-disease variant of MSUD; however, it was thought that they should still receive expectant management (e.g. during catabolic stress of illness); 4 patients were diagnosed clinically up to 60 days with poor feeding and other mild neurological problems and required treatments) 6 patients required treatment during acute encephalopathy accompanying catabolic stress; 2 patients had developmental delay and metabolic decompensation episodes. However BCKD activity does not predict the clinical course as the phenotype is modified by other genes responsible for amino acid transport and effects of BCAA and BCKA in the central nervous system (CNS). The authors concluded that intensity of treatment varies with BCAA concentrations but that those with variant forms would also benefit from newborn screening; each patient will need individual evaluation to determine severity and necessary intensity of treatment. **However, it is not clear whether the proportion of asymptomatic patients without treatment may rise as MS/MS screening becomes more established.**

It was notable that there were no suggestions in the literature that, once a diagnosis had been made after screening in any of our target conditions, there was any doubt about the necessity for some form of clinical management. Even where cases were asymptomatic some form of surveillance and 'expectant' management was the norm. Whilst recognising the need for medical supervision, the actual treatments were not necessarily too arduous (e.g. for MSUD this required protein restriction and strict surveillance during acute illness) and the specialist was usually able to titrate treatment against the biochemical and clinical profile.

## **Summary of clinical validity**

### ***Test performance***

- MS/MS as a screening test had high specificity and sensitivity in most settings
- For some conditions such as MSUD the test will never have optimal sensitivity due to the natural history of the disorder
- The use of second-tier tests as part of the screening protocol, for some conditions such as homocystinuria and MSUD improved the test performance
- In most settings very few or no false negatives were observed
- Cut-offs used by screening programmes vary and will need to be established using standardised procedures. Furthermore, cut-offs maybe biased towards certain variants of each condition
- False positive rates varied and rates of around 0.1% for parents actually recalled for further testing could commonly be achieved by adjustment of cut-offs and second-tier testing

### ***Problems with studies***

- Detailed information on false positive results were not obtainable for many studies
- Detailed information on the confirmed cases was not available to assess to what extent cut-offs influenced the cases that were identified

### ***Reference standards***

- Confirmatory tests were available for each condition and algorithms for which samples undergo confirmatory testing have been developed by a number of screening programmes
- There is a good idea of the correlation between biochemistry and phenotype for most of the conditions
- It was notable that there were no suggestions in the literature that, once a diagnosis had been made after screening in any of our target conditions, there was any doubt about the necessity for some form of clinical management



## 7.5 Clinical utility

The question at issue here is whether the screening programme provided benefit, primarily to the child tested but also including benefit to the family, and how this was balanced against possible harms. Possible benefits would include improvement in clinical outcome for the child, through better management and appropriate interventions when needed, benefits for siblings or other relatives who might receive a diagnosis for a hitherto undiagnosed clinical problem, reduction of stress and anxiety to parents and also the provision of information to parents that might enable them to take measures to avoid having further children with that disease. Conversely, harm could result from being 'diagnosed' unnecessarily, or when no treatment was available or through stress to parents because of a false positive diagnosis.

Because of the rarity, complexity and heterogeneity of these inherited metabolic diseases the gold standard for evidence of utility, meta-analysis of randomised control trials, is unlikely to be achieved in newborn screening. The next order of research evidence, the controlled observational studies, also has a set of problems that were summarised by Wilcken *et al.*<sup>59</sup> as follows:

**Power:** the population base to show a major change will be several millions

**Over diagnosis by screening:** more cases may be diagnosed by screening (especially true in MCADD) and in some instances milder disease and even non-disease (as in histidinaemia) may be hard to distinguish from the target affected population. Thus screen detected cases were more mildly affected and so would be expected to have a better clinical outcome

**Definition of a case:** there are difficulties of definition and even mutation analysis may be of little help because of incomplete knowledge of modifying genes. There is also a wide and graduated spectrum in any one disorder from severely affected individuals to functionally normal - biochemical definitions do at least signal an *in vivo* functional derangement

**Completeness of ascertainment:** especially where clinical and metabolic services are not centralised - systematic searching for missed cases may not be possible for controls. Cases of more insidious onset and non-specific manifestations, or where there is sudden or early death, may never be diagnosed

**Comparability of treatment:** this cannot always be taken into account

**Control groups:** could be contemporaneous from a different area or historical from the same area - in either the case difficulty is that there may have been less skilled and practised diagnoses, treatment or a different population demographic

For these reasons, we restricted our analysis of the literature to only those studies that were designed directly to compare screened and unscreened cohorts, over similar time periods and populations and centres, and where, if possible, some of the inherent biases had been accounted for.

Wilcken *et al.*<sup>60</sup> compared the outcomes for a cohort screened by MS/MS between 1998 and 2002 with an unscreened parallel cohort and one preceding 4 year cohort. These children from Australia were followed up to age 6. They analysed outcomes according to 3 groups, analysing patients with all conditions in the group separately from those with MCADD:

- Group 1 presenting in the first 5 days of life and including patients with some urea cycle and organic acidurias; half of group 1 patients died.
- Group 2 patients who presented later or were diagnosed by screening; this group included patients with GA1, IVA, and LCHADD. Overall, when excluding MCADD, a higher proportion of patients diagnosed clinically (generally after 5 days) died or had a significant intellectual or physical handicap when compared to the screened cohort (generally detected earlier).
- Group 3 patients with 'substantially benign' disorders for whom the authors thought that no benefit would be likely. This group included 3 methylcrotonyl-CoA carboxylase deficiency, 3-methylglutaconyl CoA hydratase deficiency and SCADD but none of our target conditions.

The authors concluded that screening by MS/MS provided a better outcome for patients at 6 years of age, with fewer deaths and fewer clinically significant disabilities. Most of the benefit from screening arises from Group 2 patients and it was arguable that Group 3 conditions should not be part of the programme.

Waisbren *et al.*<sup>61</sup> in their study in New England compared 50 screen detected patients in Massachusetts and Maine with 33 patients detected clinically from any New England state over a period from February 1999 to June 2002 (the total of screen detected cases was also supplemented by children diagnosed through a private laboratory that provides screening in New England). These groups included patients with two of our target conditions in the screened group (LCHADD and IVA) and two in the clinically detected group (LCHADD and GA1). Comparison of the two cohorts showed that hospitalisation rates for the overall groups of screen detected and clinically detected cases were similar. The group of clinically-detected cases showed a higher incidence of mental retardation with parents suffering greater stress and increased difficulty in meeting their child's needs than the screen detected cases and parents. This study could be criticised for potential bias due to the fact that no surveillance system was in place to ensure complete ascertainment of cases in the clinically detected group. The group thus might come from a larger affected group, of whom many may be alive but undiagnosed. This effect would bias the outcome towards a better outcome in screen detected cases.

Kölker *et al.*<sup>50</sup> in Germany studied outcomes in GA1 by comparing historical cohorts diagnosed from 1975 to 1998 with screen detected cohorts between 1999 and 2005. There were 62 patients in the historic cohort and 38 cases detected by MS/MS. Patients were followed up with a total cumulative follow-up of 685 years. The onset of prognostically relevant encephalopathic crises was markedly reduced in the screen detected cohort (11% of screened patients vs. 68% of the historical cohort). The same result was found in a small subgroup (9 patients) of historical patients diagnosed pre-symptomatically. Most of these children achieved motor milestones without or with only mild delay and dietary management did not affect growth and maturation. In contrast historical patients who already had symptoms when diagnosed had reduced life expectancy and a higher frequency of complications due to encephalopathic crises, variant disease forms with insidious onset (16%) and later

onset-types (8%). The authors discuss whether neonatal screening increases the selection of individuals with milder disease and note that the natural history of GA1 does not correlate with biochemistry or phenotype resulting in a similar *a priori risk* for all untreated patients suffering an encephalopathic crisis. Because of the findings of higher prevalence particularly between the historic cohorts and screened cohorts, they cannot exclude the identification by screening of milder disease as the number of missed cases in the historical cohort was unknown. They tackle the issue of ascertainment bias by assuming that all missed patients may do well, and estimate the maximal proportion of such patients from the difference in prevalence between the two historic cohorts. Even so, the absolute risk reduction of encephalopathic crisis is reduced from 0.67 to 0.57, meaning that there is still a beneficial effect from newborn screening and intensive management.

Again with respect to GA1, Strauss *et al.*<sup>62</sup> describe a study undertaken in Pennsylvania, USA on a cohort of 77 patients diagnosed between 1988 and 2000 and aged 0-44 years. During this period asymptomatic newborns from high-risk Amish families were screened for GA1 based on 3-hydroxyglutaric acid in urine. From 1994 onwards screening was state wide and based on blood spot analysis of C5DC using MS/MS. The 77 cases consisted of 37 Amish descent (20 identified by screening) and 40 of non-Amish descent (2 identified by screening). The majority were followed up prospectively over a 14 year period. The incidence of basal ganglia injury was 95% in clinically diagnosed Amish children and 35% in those Amish children who had been identified through newborn screening. The majority of the 40 non-Amish children were diagnosed after presenting with neurological disability and the incidence of basal ganglia injury in this group was 85%. Two children in this group were identified asymptotically and remained healthy. In the majority of infants head circumference is the only presenting sign of GA1; most patients (78%) are diagnosed after they develop striatal necrosis and their outcomes are poor. Only 12% of cases had no apparent motor impairment. Injuries acquired at an earlier age were more severe, detection of asymptomatic infants reduces risk of brain injury. Morbidity given for 20/22 screened children, mortality given for 22 children (21 alive). The authors state that without early diagnosis 80-90% of affected infants will come to an emergency room between 6 and 18 months of age with an evolving brain injury. As a further benefit the authors also comment on the finding that 13% of the non-Amish patients developed acute subdural haemorrhage after minor head trauma and, in two cases this was accompanied by retinal haemorrhages. Investigation of child abuse had preceded metabolic diagnosis in three of these children.

For IVA, Ensenauer *et al.*<sup>63</sup> analysed a cohort of 22 patients with IVA diagnosed in Germany in the preceding 30 years. Twenty seven percent were detected by newborn screening and 50% presented in the first weeks of life. In this cohort IQ was not related to the number of crises but did decrease with later diagnosis. They suggested that early diagnosis did prevent crises and was also related to better outcome - inferring that diagnosis by newborn screening might be beneficial.

Dionisi-Vici *et al.*<sup>52</sup> compared long-term outcomes of classical organic acidurias including IVA between a clinically identified cohort cared for at the children's hospital in Rome with newborn screening patients identified in German and Australian cohorts. Whilst overall better outcomes of screening were noted, including decreased early mortality, less severe symptoms at diagnosis and more favourable short-term neuro-developmental outcome it should be noted that there was only one case of IVA in the clinically detected cohort and the finding of a much higher birth prevalence of IVA in the screen detected cohort raises the possibility that screening was identifying milder or non-disease patients.

With respect to LCHADD deficiency, Spiekerkoetter and Sykut-Cegielska<sup>64</sup> provided results collected from seven European centres on the clinical presentation, treatment and outcome from patients identified before the screening era (25 patients) and since newborn screening was in place (13 patients). Three of the screened patients died compared with 14 of the clinically detected patients. Although death and long term complications such as retinopathy and peripheral neuropathy could not be prevented in the screened group, the risk of some complications was lower (e.g. retinopathy 2/13 vs. 14/25) and it was thought that overall morbidity was reduced in the screened group. A more recent publication by Spiekerkoetter *et al.*<sup>65</sup> expanded on the management of individuals with LCHADD deficiency, concluding that the toxic effects of long-chain 3-hydroxyacylcarnitines and CoA esters are such that long-chain fat intake has to be maximally restricted. Whilst this advice is current, it is acknowledged that the exact factors determining the development of irreversible neuropathic complications including retinopathy have not yet been elucidated and so it is impossible to differentiate treatment for different individuals.

Again, for LCHADD, Sander *et al.*<sup>51</sup> in their study of 11 patients identified by newborn screening in their laboratory in Hannover, Germany, note that all children other than those with complete MTP deficiency or isolated LKAT deficiency developed satisfactorily as compared with patients described in the literature who were diagnosed only after developing clinical symptoms. However metabolic crises did still develop in this group meaning that there were still possibilities to fine tune management. In contrast none of the patients with complete MTP deficiency survived longer than 13 months. The main advantage of screening for those patients was the avoidance of unnecessary diagnostic or therapeutic measures together with the early availability of information to guide future reproductive choices. In the future a better outcome for these patients seems possible.

Other authors provide shorter case-history approaches to identification of improved clinical outcomes from screen detected patients. These include Ibarra-Gonzalez *et al.*<sup>66</sup> (MSUD), Bijarnia *et al.*<sup>56</sup> (GA1) and Hoffman *et al.*<sup>67</sup> (MSUD) all of whom compare small numbers of pre-symptomatically detected versus clinically detected cases but without construction of comparative cohorts.

## **Summary of clinical utility**

### ***Possible types of benefit***

- Improved clinical outcome for child
- Benefits of relieving parental stress or anxiety
- Enabling reproductive choice for parents
- Economic benefits

### ***Problems with the studies***

- Lack of statistical power
- Over-diagnosis by screening
- Definition of what constitutes a case
- Completeness of ascertainment
- Comparability of treatment
- The lack of appropriate control groups

### ***Findings***

- Wilcken *et al.*<sup>60</sup> (Australia) 'Group 2' patients diagnosed by screening or presenting later (includes target conditions) - MCADD analysed separately - reduced risk of death, or significant intellectual or physical handicap
- Waisbren *et al.*<sup>61</sup> (New England) - hospitalisation similar but clinically detected had higher rates of mental retardation and higher parental stress
- Kolker *et al.*<sup>50</sup> (GA1) (Germany) - reduced mortality and encephalopathic crises; no or minimal delay in developmental milestones; no effect on growth or maturation
- Strauss *et al.*<sup>62</sup> (Pennsylvania) (GA1) Amish and other populations - reduced risk of brain injury in those diagnosed before symptoms present. Child abuse investigations noted in non-screened
- Spiekerkoetter and Sykut-Cegielska<sup>64</sup> (LCHADD) (7 European centres) - reduced mortality and some evidence of reduction in morbidity though details of management still to be worked out

## 7.6 Other evidence of benefit and possible harm

### *The danger of unnecessary treatment*

Arguments about clinical utility and the possibility of harm from overtreatment exist with concern expressed that newborn screening by tandem mass spectrometry will identify more and more asymptomatic infants with persistent biochemical disturbances that indicate likely enzyme deficiencies. Wilcken<sup>68</sup> takes up this concern and explores the problems that have arisen when it is not entirely clear which babies detected by newborn screening actually need treatment. She considers the disorders in two main categories: those that need active management and those that do not.

Wilcken considers that disorders **that need active management include:**

- Those associated with known pathology including 1) patients who already have symptoms at diagnosis (e.g. often classic organic acidurias) and 2) patients with disorders for which there is strong evidence that they will lead to preventable pathology (e.g. PKU, GA1, homocystinuria, MSUD, tyrosinaemia type 1 and most fatty acid oxidation disorders).
- Disorders that pose a risk of disease but with reduced penetrance. Examples are MCADD and some organic acidurias such as 3-ketothiolase deficiency. These patients need careful consideration and some will inevitably have treatment imposed when, with hindsight, there would never have been a need. MCADD is probably the best example in this category and in future it may be possible that risk may be further differentiated by knowledge of the mutations involved, but at present this is not possible and as MCADD management is not unduly burdensome, treatment is usually advised at present.

Wilcken considers that screening for conditions for which patients might not need management is more problematic. She includes a small number of disorders in this category and discusses the cases of 3-MCC deficiency and SCADD. However, it should be noted that **none of the targeted conditions for this review are included in this grouping.**

### *The effect on parental stress*

In the publication by Waisbren *et al.*<sup>61</sup> the effect of expanded newborn screening on child outcomes and parental stress is reported. A prospective study was undertaken involving a cohort of children with metabolic disorders identified from February 1999, through June 2002, and who were evaluated by December 2002. Children identified by expanded newborn screening in Massachusetts and Maine and a private screening laboratory in Pennsylvania were compared with those identified clinically from any New England state. In newborn screening programs, results are reported to primary care physicians. Depending on the results, a repeat filter paper blood specimen for follow-up testing is requested or an immediate referral to a metabolic unit is made for confirmatory assessment. Study participants were families of 50 affected children identified through expanded newborn screening (82% of eligible cases); 33 affected children identified clinically (97% of eligible cases); 94 screened children with false positive results (75% of eligible cases); and 81 screened children with normal results (63% of eligible cases).



Although not differentiated by disease the study found that:

- Mothers of cases identified by newborn screening experienced significantly lower overall stress than mothers of clinically diagnosed children
- Only one mother in screening group but 14 mothers in the clinically detected group scored in the clinical range for stress
- Mothers of children diagnosed through newborn screening were less likely to report a negative effect on reproductive plans (in particular, less likely to prefer not to have further children, less likely to request prenatal testing and only one mother in screened and two mothers in unscreened group said that they would terminate an affected pregnancy)

### ***The effect of false positives***

Waisbren *et al.* also reported on the effect of false-positive results. Parents reported a median age of 10 days when a repeat screen was requested and a median time of 7 days before learning the result. 55% correctly identified the reason for the repeat. 21% of parents were referred directly to the IMD specialist after the initial positive. These parents were 2.5-times more likely to report the correct reason for the follow up test and all reported that they were told the result of the test.

Mothers in the false positive group had significantly higher levels of stress, which was lower in those referred to a metabolic centre, and if they received information about the repeat result in person. It was suggested that there was a need for more education about newborn screening for parents prior to the birth of a child, and also education for primary care doctors and other health professionals since face-to-face contact seems to reduce stress.

Children in the false positive group were more likely to be hospitalised than children in negative group, but reasons for hospitalisation were similar. However, parents whose children were referred to a metabolic centre did not worry more than those who were not. It was suggested from these studies that parents of false positive children may have persistent altered perceptions of their child's health.

Similar results were reported by Gurian *et al.*<sup>69</sup> in the US in a comparison of 173 parents of children with false positive screening results for a biochemical genetic disorder in the expanded newborn screening panel compared with parents of 67 children with normal screening results. All were enrolled between 1999 and 2004. They conclude that false positive results may lead to increased parental stress and that this was especially true for parents who have not received adequate information about newborn screening. Features of this stress included sleep disturbance, maternal crying, shock and infant feeding problems. They noted similarities with other programmes that focused on stress induced by false positive results, which had noted that parents were over protective of child and more focused on physical symptoms. A study on cystic fibrosis screening showed that in parents of children with false positive results attachment relationships were insecure even at 12-18 months and high scores on parent-child dysfunctional interaction suggested that in these cases the parent child bond either is threatened or not established properly. Parents also interpreted the positive test as being a significant threat to the child's wellbeing and continue to fear that child might be developmentally delayed.

**Box 7.2 Possible harms**

- Over diagnosis; however, Wilcken<sup>68</sup> considers that all target conditions come into the category of conditions that will need active management because of links with known pathology or risk for disease but with reduced penetrance
- The effect on parental stress and effect on parent-child relationship for false positives as reported by Waisbren *et al.*<sup>61</sup> and Gurian *et al.*<sup>69</sup>; these effects were worse where parents had received less information, not received reported results in person and were not referred to a specialist centre

**7.2 Economic review (author Philip Shackley)**

Six studies examining the cost-effectiveness of screening for inborn metabolic disorders were included in the review. Of these, one study each was from Australia, Canada and Finland, while the other three were from the US.

Norman *et al.*<sup>70</sup> estimated the cost-effectiveness of replacing a programme of identifying IMDs based on single disease tests with MS/MS screening. The conditions considered were amino acidurias, urea cycle disorders, organic acidurias, MCADD and other fatty acid oxidation defects. The study was carried out from the perspective of the Australian health service with costs presented in 2002 Australian dollars (AUS\$). The additional (incremental) costs of screening of MS/MS compared to no screening were estimated to be AUS\$218,000 per 100,000 infants. The authors also estimated the total medical costs incurred during the first four years of life. These costs comprised emergency care costs, outpatient costs and inpatient costs, all of which were discounted at 6%. Adding these to the screening costs gave an overall incremental cost for MS/MS compared to no screening of AUS\$349,010. Combining this cost with the estimated incremental life years gained (LYG) from MS/MS screening of 32.378 years (discounted at 1.5%) resulted in an incremental cost-effectiveness ratio (ICER) for MS/MS screening of AUS\$10,799. The ICER compares favourably with existing standards for cost-effectiveness and was robust to sensitivity analysis.

Schoen *et al.*<sup>71</sup> were also concerned with estimating the cost-effectiveness of an MS/MS screening programme. The conditions considered here were MSUD, MCADD and other disorders of fatty acid oxidation, GA1, methylmalonic aciduria (MMA)/propionic acidemia (PPA) (included as example of organic acid disorders), urea cycle disorders, homocystinuria and PKU. The evaluation was carried out in a US setting, with no specific perspective for the analysis being stated. Costs were presented in US dollars (US\$), but the base year was not stated. Unlike Norman *et al.*<sup>70</sup>, the study considered lifetime treatment costs, which were discounted at 3%. Outcomes were measured in terms of quality adjusted life years (QALYs) gained, which were discounted, but the actual discount rate was not stated. Data for the QALY estimates were taken from four published studies of neurological defects in adults. The specific conditions were unruptured cerebral aneurysm, ruptured intracranial aneurysm, intractable temporal lobe epilepsy, and complex partial epilepsy. The relevance of health state values in adults for these conditions to health states associated with IMDs in children is open to question. In addition, numerous assumptions were made throughout the study, many of which were not justified.



Results were presented for three scenarios: a base case scenario, an unfavourable scenario (higher laboratory costs) and a favourable scenario. The incremental cost-effectiveness ratios (ICERs) for these scenarios were US\$5,827, US\$11,419 and US\$736 per QALY gained, respectively. The authors claim that these ICERs compare favourably with ICERs of other programmes, such as breast cancer screening and beta-interferon treatment for hepatitis C.

Autti-Ramo *et al.*<sup>72</sup> attempted to estimate the cost-effectiveness of MS/MS screening for multiple conditions compared to best and worst current practice. The conditions considered were CAH, MCADD, LCHADD, glutaric aciduria type 1 and PKU. The study adopted a modelling approach and was carried out in a Finnish setting. Costs were reported in 2002 Euros, and treatment costs up to the age of 16 years were included and discounted at 5%. Outcomes were measured in terms of undiscounted QALYs gained. A range of sensitivity analyses were performed, including testing different assumptions regarding disease incidence. Estimates of the ICER for MS/MS ranged from 5,500 Euros to 25,500 Euros per QALY gained. Preventing one severe handicap would reduce the ICER to a maximum of 18,000 Euros per QALY gained.

Carroll and Downs<sup>73</sup> were concerned with estimating the cost-effectiveness of each component of a multi-test newborn screening programme using MS/MS compared to no screening. The individual conditions considered were PKU, biotinidase deficiency, MSUD, galactosaemia, homocystinuria, CAH, and CH. The use of MS/MS to screen for multiple conditions, namely PKU, biotinidase deficiency, MSUD, homocystinuria and MCADD, was also compared with no screening. In an additional comparison, MS/MS for multiple conditions was compared directly with a panel of available conventional tests for the same conditions. The study was carried out in a US setting and adopted a societal perspective. Decision modelling was used, with costs being presented in 2004 US dollars (\$US) and outcomes being measured in QALYs. Lifetime treatment costs and QALYs were discounted at 3%. Comprehensive sensitivity analyses were performed. In the base case analysis, all except two of the screening strategies produced more QALYs and saved costs. The ICERs for CAH and galactosaemia relative to no testing were US\$20,357 and US\$ 94,000 per QALY gained respectively. At a conventionally accepted societal willingness to pay for a QALY threshold of \$US50,000, screening for galactosaemia would not be considered cost-effective. In the comparison of MS/MS for multiple conditions with a panel of conventional tests, MS/MS was found to dominate the conventional panel. The results were not sensitive to the costs of screening tests but were sensitive to the specificity of the tests and the cost of evaluating false positive results. This led the authors to suggest that false positive results must be minimised in order to realise the cost savings of the screening strategies.

Cipriano *et al.*<sup>74</sup> examined the cost-effectiveness of expanding the current newborn screening programme in Ontario, Canada, through the introduction of tandem mass spectroscopy screening. Twenty-one different diseases were considered, with the cost-effectiveness of each disease being assessed independently. In addition, the cost-effectiveness of screening for disease bundles was also estimated. The study used decision modelling to assess cost-effectiveness from a societal perspective. Costs and benefits were assessed over the patient's lifetime and were discounted at 3%. Costs were presented in 2004 Canadian dollars (CAN\$) and included not only treatment costs, but also non-medical costs such as costs of education and institutional care. The set up costs of the screening programme, including the acquisition of three mass spectrometers, were also included. Benefits were measured in terms of life years gained (LYG). The impact on the results of varying key parameters was

assessed via a series of one-way sensitivity analyses. In the base case analysis, the ICERs of screening for the individual diseases ranged from CAN\$221,719 per LYG for HMG-CoA lyase deficiency to CAN\$142,462,114 per LYG for glutaric acidemia type 2. To assess the cost-effectiveness of screening for disease bundles, the individual diseases were ranked in order of cost-effectiveness (from best to worst) and ICERs calculated for adding successive diseases to a bundle. The results suggest that the ICER can be optimised (at CAN\$65,373 per LYG) by screening for PKU plus the next nine most cost-effective diseases. In descending order of cost-effectiveness these are methylmalonic acidemia, HMG-CoA lyase deficiency, MSUD, propionic acidemia (PPA), VLCADD, carnitine transporter defect, GA1, IVA, and MCADD. Using a threshold of CAN\$100,000 per LYG, there was moderate evidence to support a programme of screening for PKU plus 14 additional diseases, including LCHADD at number 11. The addition of tyrosinaemia type 1 and homocystinuria resulted in an ICER of CAN\$331,200. The model results were found to be most sensitive to changes in the specificity of MS/MS. It was noted by our Steering Group, however, that this study was undertaken in Canada where there was an estimated prevalence of homocystinuria of 1 in 250,000, (less than the expected incidence in UK) and where lower cut-off values, (giving a lower specificity) may have been used.

Insinga *et al.*<sup>75</sup> attempted to estimate the cost-effectiveness of using MS/MS to screen for 14 fatty acid oxidation and organic acidemias in the state of Wisconsin in the USA. A sequential analysis was performed, with the cost-effectiveness of screening for MCADD assessed in the first instance. The study used decision modelling to assess cost-effectiveness from a societal perspective. Costs and benefits were assessed over the lifetime of a hypothetical cohort of 100,000 infants screened at birth and were discounted at 3%. Costs were presented in 2001 US dollars (US\$) and included costs of special education, health and social care. Benefits were measured in terms of QALYs, with the data for the QALY estimates being taken from previously published studies focusing on neurological impairment in teenagers and adults. One-way sensitivity analyses were performed. In the base case analysis, the authors used a series of conservative assumptions which they claimed biased the results against MS/MS screening. Using these assumptions, the ICER for screening for MCADD alone was US\$41,862 per QALY gained, which is below the accepted cost-effectiveness threshold value in the USA of \$50,000 per QALY. The adoption of what the authors suggested were more realistic assumptions yielded an ICER for MCADD screening of US\$6,008 per QALY gained. Adding in the costs and benefits of screening for the 13 additional conditions (including LCHADD, GA1 and IVA), resulted in an ICER of US\$15,252 per QALY gained. The results were robust to sensitivity analysis.

### ***Conclusions of the economic review***

On balance, the evidence from the six studies reviewed points to MS/MS being cost-effective when judged relative to nationally accepted norms for cost-effectiveness. In general, screening for multiple conditions is more efficient than screening for individual conditions, but there comes a point when adding more conditions ceases to be cost-effective. There would seem to be a good case for including MSUD, GA1, LCHADD and IVA in an expanded screening programme on cost-effectiveness grounds, but the case for homocystinuria may be more equivocal and would depend on the birth prevalence of the condition in the UK, as well as the determination of cut-off values to reduce the number of false positives.

## **7.8 Conclusions from the systematic review of screening using MS/MS**

### ***Birth prevalence***

The birth prevalence for the five target conditions is very small. Extrapolating from findings of programmes involving other Caucasian populations suggest around 19 new cases in England and Wales in any one year, although small numbers might mean that the actual number could be between 16 and 23 (these estimates are based on this systematic review and summarised in Table 7.7). However, UK data suggests that this number might be higher because of higher prevalence within immigrant populations where there are high levels of consanguinity.

Although about twice as many cases will be detected by MS/MS screening than would be diagnosed clinically for four of the target disorders this is not thought to represent milder cases where no treatment would be necessary, but is thought to be due to under-diagnosis and under-reporting. For IVA some diagnosis of milder cases may be taking place.

### ***Recommendations for the pilot***

The numbers of screen detected and clinically detected cases should be ascertained along with basic demographic details including age at diagnosis, sex, ethnic background of parents, history of consanguinity, and any family history of disease.

### ***Clinical validity of MS/MS screening tests***

#### ***Sensitivity of tests***

Laboratories are able to devise tests that are highly sensitive (mostly 100%) although in the case of MSUD variant or intermittent forms might occasionally be missed because of normal biochemistry in the newborn period. This is not thought to be an argument against screening, but for continuing clinical vigilance to diagnose later presenting or insidious onset forms. Further, MS/MS can also only detect pyridoxine non-responsive forms of homocystinuria and some forms of GA1 (low excretors) may not be picked up by newborn screening.

#### ***Specificity***

Laboratories have been able to adjust cut-offs and also develop second-tier laboratory testing for some conditions in order to minimise false positives, and, in particular, to reduce the recall rate for infants. False positive rates varied and rates of around 0.1% for babies actually recalled for a second test could commonly be achieved by adjustment of cut-offs and second-tier testing. This represents approximately 700 newborns who would be recalled each year in England and Wales for further testing. In some conditions, such as GA1, 'holding' treatment is commenced because of the urgency of the situation. For all patients a definitive diagnosis can be achieved in around 2 weeks at which point treatments for negatives can cease (or be tailored to any alternative diagnosis underlying the biochemical abnormality) and positives can go on to specialist treatment depending on precise and expert biochemical and clinical assessment.

An important aspect in determining the specificity and subsequently the positive predictive value of the test is an accurate definition of a screen positive result and a false positive result. As described in previous chapters screen positive results can arise as a result of transient elevations in levels of particular metabolites as a result of physiological variation in the analyte level, enzyme immaturity in preterm infants, parenteral nutrition or maternal factors such as the use of particular antibiotics. These are often considered false positives as these individuals do not have a disease. Screen positive results will also be obtained in those individuals who have the disease being screened and in addition, those with overlapping conditions. It was not certain from the literature if those individuals with overlapping conditions were considered as true positives or false positives. This is an important consideration when evaluating the test performance as the number of individuals assigned as false positives will have a bearing on the positive predictive value.

### *Recommendations for the pilot*

Laboratories participating in the pilot study must collectively devise and evaluate tests with respect to analytical and clinical validity, in particular to maximise test sensitivity whilst minimising false positives.

For each condition, this should result in flow-charts that show, for each condition:

- Initial cut-offs
- Cut-offs for any repeat testing/or for urgent assessment (depending on condition)
- Cut-offs for further sampling request and the further testing undertaking
- Diagnostic cut-offs
- The expected 'flow' of infants through the various branches
- Defined false positives
- It will be important to agree diagnostic criteria and independently assess compliance with the criteria when assigning true positive results

### *Clinical utility*

Comparison of clinical outcome for screened versus non-screened cohorts either for groups of conditions (including our target conditions) or for our conditions separately (notably GA1 and LCHADD) have shown better clinical outcomes for screened cases including less developmental delay, encephalopathic crises, psychological and motor disability. For parents, diagnosis following screening appears to lead to less stress and a more positive view of condition and options for antenatal testing with or without subsequent termination of pregnancy. Other positive outcomes have included diagnosis of other relatives with clinical symptoms of the condition.

### *Recommendations for the pilot*

Cases detected by newborn screening should be included on a register and followed up. Diagnosis should be recorded with details of underlying genetic and biochemical abnormalities and presenting clinical features. Patients should be followed up with details of treatment provided, centre of treatment and description of clinical progress including acute crises and outcomes in terms of morbidity and disability. Parallel active surveillance through the UK laboratories for cases of these conditions diagnosed clinically should be put in place and these cases should be followed up systematically and consistently.

### *Possible harm from false positives*

- These can cause stress
- In most cases the clinical diagnosis can be resolved within two or three days
- Despite eventual negative diagnosis there may be residual concern about the health of the child, over protectiveness and sometimes an adverse effect on the parent child relationship
- False positives will be limited to a fairly small group of people
- It is suggested that the harms from a false positive assessment can be alleviated by better education and information, personal communication and support from a specialist unit

### *Recommendations for the pilot*

The pilot study needs to look carefully at the numbers of false positives generated through screening and their pathway from flagged test to final negative diagnosis. It also needs to devise means of minimising stress caused by improving education in the antenatal period and around the time of screening and support in the event of a positive result. Educational support will be needed for parents, health professionals and the general public

### *Possible harm from over diagnosis*

This is not thought to be a substantive issue for any of the five target conditions, which are all thought to require active management once diagnosed. This is because either there is evidence of symptoms, or there is evidence that pathology will ensue because of the biochemical abnormalities present, or because of the risk for adverse events, even though the penetrance may not be 100%. At present, despite some heterogeneity of clinical course, there is no way of differentiating the occasional patient who might not progress to serious pathology on the grounds of biochemistry or genotype.

### *Recommendations for the pilot*

The pilot study should include clear documentation of biochemistry and clinical assessment of every case diagnosed. It should consider some form of independent external assessment to confirm whether treatment was strictly necessary for each case and record the reasons behind this.

### ***Economic analysis***

Addition of new target conditions will undoubtedly result in an increase cost to the screening programme but most experts think this is quite small. The evidence from the six studies reviewed point to MS/MS being cost-effective when judged relative to nationally accepted norms for cost-effectiveness. Screening for multiple conditions is found to be more efficient than screening for individual conditions and a good case for including MSUD, GA1, LCHADD and IVA in an expanded screening programme on cost-effectiveness grounds, but the case for homocystinuria may be more equivocal and would depend on the birth prevalence of the condition in the UK, as well as the determination of cut-off values to reduce the number of false positives.

### ***Recommendations for the pilot***

The pilot study needs to quantify the extra costs needed to expand existing newborn screening programmes to include these conditions. Costs falling on laboratories, specialist clinical, paediatric, community services and primary care need to be estimated.

### ***Organisational aspects***

High quality detection of infants with inherited metabolic disease and their subsequent assessment, diagnosis and management depends on a close relationship between specialist laboratory and clinical metabolic services. This can allow infants with abnormal results (whether diagnostic or borderline) to be referred for evaluation within hours of the completion of screening tests - a very important aspect of the programme when infants can become so rapidly unwell and suffer irreversible brain damage in the first few days of life.

Undoubtedly screening for a wider range of inherited metabolic conditions will initially create increased work for both laboratories and clinical services in further investigation of abnormal results and clinical diagnosis and treatment of patients. Our analysis again shows that the absolute numbers involved would be fairly small, but nevertheless, specialist IMD services throughout the UK are known to lack capacity and to be unequally distributed.

### ***Recommendations for the pilot***

The pilot programme should consider in detail the organisational requirements including the ways in which laboratory and clinical elements must be integrated in order to streamline the identification and follow up of those who screen positive.

### ***Capacity in specialist IMD service***

The review by Burton *et al.*<sup>76</sup> highlighted the problem of overall lack of capacity of specialist IMD services and patchiness of provision around the country. This, coupled with the inherent difficulty of making a diagnosis for a very rare condition with non-specific presentation (for example with hypoglycaemia or metabolic crisis in neonatal period, failure to thrive, or neurological or developmental problems to paediatric services) makes late diagnosis and the 'diagnostic odyssey' highly likely. This is an argument for screening (rather than against it) as it enables some work of the specialist services to be highly targeted on infants who have a higher *a priori* risk of disease. The infants can be rapidly pulled through the diagnostic processes and looked after by specialists (even if this is at a distance) rather than spending precious days without diagnosis or proper management.

### ***Recommendations for the pilot***

The pilot programme should consider a comparison of the route to diagnosis between screen detected and clinically detected cases. This could include clinically detected cases in areas not undertaking expanded screening, or different but equivalent conditions in the same population and also presenting in infancy.



## 8 Patient perspective

### 8.1 Introduction

As part of this Report, Climb, the umbrella group for inherited metabolic conditions was asked to provide the perspective of patients and families with direct experience of inherited metabolic disease. They were particularly asked whether they supported an expansion of newborn screening and their reasons behind this, and the factors that they felt would be important in providing a good screening service. Their submission is included in full in this Chapter. Key points include:

Climb is strongly in favour of an expansion of newborn screening to include other inherited metabolic conditions. Their reasons for this include:

- Prevention of death and substantial disability and improved quality of life for patients
- The need for a clinical diagnosis to explain diverse symptoms and to guide treatment, access to support, life planning and overall management - the feeling that knowledge empowers families
- A diagnosis as early as possible to prevent morbidity and aid family dynamics, communication and support
- Their view that, although there is a cost in screening, this is outweighed by the saving in caring that the family and state bear over the years
- Their view that, although the conditions are rare, the needs of people with rare disorders should count as much as those with common ones and that, cumulatively, there is a significant health burden from these conditions

Any expansion of the screening programme should include planning for the following:

- The systematic provision of clear, accessible information by professionals to parents prior to screening
- Good timely, communication, information, support and follow up to parents after a positive screening test and diagnosis. This should include involvement of the voluntary sector
- The likely need for increase in specialist metabolic nurses, linked to specialist centres and working in the community to provide advice on care



## 8.2 Submission on the patient and family perspective (provided by Steve Hannigan, Chief Executive at Climb)

### *The needs of the family*

The needs of the family revolve around a diagnosis and getting answers; understanding in layman's terms why this has happened, if it is likely to happen to anyone else in their family and if there is anyone else out there going through the same.

Following years of questions and of unexplained symptoms, deterioration, being labelled as 'troublesome or difficult by their GP', having unexplained extra needs that cannot be pinned down, to ask parents if they would have preferred their child to have been screened during the first few weeks of life, then the response would be a resounding 'yes'.

Having a clear diagnosis with a name, even though it may mean shortened life span or disability, is half of the problem of a metabolic condition and half the solution of learning how to cope and manage. Apart from being rare and being isolated by this rarity, if a condition does not have a name, the individual remains in the 'fog'. This is the period of wondering why and how they are like they are with no clarity on how to live their lives and what to expect. No-one can advise them.

There are families now in 2010 that are caring for their child and adjusting well to the loss of skills and continued deterioration. Due to a late diagnosis they have never quite found the right time to tell their child that they have an inherited metabolic disease. Informing a child that they have a metabolic disease and the issues surrounding their condition will always be difficult but for many families adjusting, planning, coping and seeking advice cannot prepare them enough to inform their loved one that their condition has a name and what it means to them.

A screening programme, for these families, in retrospect, would have given them more choice in preparing for the future, more choice in informing their child at an appropriate time before it is too late and empowered them to make their own decisions about the future. For those families who have two or three young children with a life-threatening condition and who face the loss of all their children one after the other, the choice of screening is obvious.

There are negative sides to screening and one is that there will always be families who say "*I don't want to know*" and would prefer to just get on with things. But for the vast majority of families with metabolic disease, knowledge is power. Knowledge gives time: time for planning, time for preparation and time to consider the future. It also allows time to consider what the condition means, what effects it will have on the affected child, what effects it will have on the parents, what effect it will have on the family dynamics and most of all what life-changes will have to be made. It also enables families, in many cases, to adjust to not working but caring and time to adjust to not having much time.

For those who are diagnosed through the extended mutation screening and newborn screening programmes it is all positive. Well, as positive as things can be after receiving a diagnosis of an inherited metabolic disease.

**Box 8.1 Patient history by Anne who lives in northern England**

In summer 2009 we gave birth to our third beautiful boy - Alex. We brought him home, he was breastfed and appeared to feed really well. He fed all night, and regularly during the day. At about 4pm on the Thursday, we noticed that Alex looked a little pale, and within minutes he had stopped breathing. The hospital staff managed to keep him alive for a further 2 days but advised us to withdraw treatment as Alex was 'brain dead'. They put the cause of death down to infection, as they couldn't find other answers. However, on the day we cremated him, a week later, we had a phone call to say he had been screened positive for MCADD. Of course, we then proceeded to spend hours on the internet to find out as much as we could about MCADD. Every day I question myself - why didn't I bottle feed? Why didn't I wake him more regularly to feed? We don't know of any other cases in our family. We are still awaiting advice from the geneticist as to whether our other two boys need further screening - they were tested at birth by heel prick and it has been checked that they are both clear. What is the future for our family? I find it so difficult, as we are a Christian family too, to comprehend that something so simple has caused us to lose our beautiful baby. The more I read, the more it appears that it is very rare that babies die from this disorder - but we are proof that it does happen. I am a medical professional and although know it was a disorder we tested for in the Guthrie test, I didn't know anything about it!

**Box 8.2 Patient history**

John is 35 years old, was well until 5 years ago, has a metabolic disease but was undiagnosed. John has worked since he was 15, he has a family, he has just gone blind and lost his job through this and has to now live with this disability and faces a much shortened life span, any knowledge that could have prepared him for this unknown journey would have helped his family. It would have also have helped the medical professionals who are trying to support him but who are also operating in an unknown zone.

***Why screen for inherited metabolic diseases?***

Why screen for metabolic disease when the incidence could be as low as 1:100,000, 1:200,000 or less? For those that have metabolic disease, for those who are alone with it, for those who will die before the age of 10 and for those families whose healthy baby has died before being a week old, is why screening for metabolic diseases is essential. There are thousands of known families affected by a child, teenager or adult with a metabolic disease. It is possible that when looked at collectively the total incidence of inherited metabolic diseases in newborn infants may well exceed 1:5000 in the UK.

Screening here is not about numbers and spreadsheets, it is about making a difference and enabling quality of life. It is about recognising that something must be done in the early days of a small life, before damage occurs and the family loses its happy, well child and its hopes

for the future. It is about all-round benefit and turning that family into an integral working part of the community and preventing them being a burden on society. Screening has a cost and this cost per person is minimal compared to the burden of disability that happens without screening. We know that hundreds of national studies on the burden of disability have been carried out and although the costs of extended screening may seem high for a possibly small amount of positives or diagnoses made, this is a minute amount compared to the input needed for a sick baby, a dying child, hospice and respite care, 24 hour carers, social services input and all the other services needed.

Screening benefit outweighs the potential harm of not screening no matter how small the numbers are. When considering why it is important to screen for these rare, unusual, complicated conditions and not conditions that affect the majority of people: it can be covered by 'Everyone Counts' (NHS Constitution).

### ***Information prior to screening***

Information prior to screening is critical. It should be shared in confidence, with understanding and the professional involved should be up-to-date regarding metabolic diseases. The MCADD project has shown that with the best will in the world, information-sharing remains haphazard in some areas and excellent in others. For the professionals lucky enough to work in an area where their workload permits reading, sharing, guidance and medical updates, they have the knowledge to share information about screening in a caring manner. For those whose who are already working to the limit, another leaflet on screening or a specific condition will be added to the already full drawer or cupboard.

Information leaflets for newborn screening are now at quite a high standard but the information assumes a high level of pre-knowledge of the reader. This is not the case. Metabolic diseases are not on everyone's reading list. No matter what information is given to families pre-screening it is not their priority to either read or understand what they have been given. Expectant parents do not need to be fully informed just adequately informed in order to understand what they are being asked to participate in whether it be the Newborn Screening programme or research projects.

### ***Information post screening***

It is not clear how much information new parents can retain and once the heel prick test has been undertaken by the midwife it is highly likely that the parents will not think of this again, unless they receive notification of a positive screen for inherited metabolic diseases. More information needs to be provided to the new parents before any newborn screening is undertaken and the midwife must satisfy themselves that the parents, especially the mother, have an adequate knowledge of what is being tested and the potential outcomes if a test is positive.

Another issue that raises itself during this period is "*do the midwives receive adequate training in inherited metabolic diseases?*" Comments we have received after the introduction of MCADD would seem to deny this. Any research project that is introduced into the screening programme must include training for midwives for without this training the very proud new parents will not get the right information to enable them to use their right of choice.

It is essential that before any screening is considered, long-term planning for those who will be interacting with the families at the raw edge of care and diagnosis are properly prepared, briefed and that national training is recommended. In 2006, in a large hospital where PKU is dealt with expertly, a nurse and midwife attended a briefing session by Climb and they commented that this is the first time they had understood properly what PKU meant for a family, even though they had been giving advice to families with PKU for several years.

As our communities are changing, the incidence of metabolic disease will continue to rise which means that any further introduction of screening must include leaflets or information that everyone can access.

Following the introduction of MCADD in England, it would be expected that screening follow-up would be as effective. In fact, there are several issues that need to be recognised:

- care after screening is sporadic
- some professionals cannot inform families due to time limitations
- occasionally families are left searching on the internet for support
- families are unsure of who they can ask questions following a screening diagnosis
- families assume that there is no-one who can help as the next appointment is 6 months away

One comment we regularly receive is *“They don’t care, we have our diagnosis and then we get to see someone in six months - who do we go to if something is wrong?”* There are pockets of excellent follow-up for many families but for some, especially those who are not used to asking questions, time is needed for the diagnosis to sink in and it may be months before they realise questions that need to be asked.

Voluntary sector networks have proven invaluable with newly diagnosed MCADD and recently diagnosed PKU families. They fill the gap between appointments and although cannot answer all the questions, they can provide support and day-to-day caring knowledge and experience about living with a condition. The gap that may exist with extended screening should be filled with those networks which are proven and recognised to be of value to the families and provide an inexpensive alternative to hospital/nurse/support. The added-value that networks provide ensures that families are not isolated, that there is somewhere to go to about sickness, caring, feeding, deterioration, family coping, education, social services, respite, benefits and general family health.

The implementation of screening programmes must ensure that a definite plan is introduced so that families are not left out in the cold with a diagnosis and a future appointment - this is a critical time whereby a small amount of care and information can make a difference to the rest of their lives. Leaving them on their own without a contact and without aftercare is unthinkable. Simple measures can include:

- Regular follow up
- Recommending quality groups and networks for support
- Ensuring that details are clear of what happens next

- Ensuring that if they have a crisis who do they contact
- Providing yearly national screening updates readily available for all
- Providing quality management to ensure that after-care is carried out

### ***Disease specific information for families***

Disease specific information for families is essential and whilst a lot of information is provided by the consultant, specialist nurse and/or dietician it is often left to the voluntary sector to continue to support the information and other needs of the family outside of the routine visits to the clinic.

The ability of Climb to give time to families who receive a diagnosis of an inherited metabolic disease is one of our strongest points. Family Advisers can often spend several hours on the telephone talking to the mother of a recently diagnosed child. Families contact us just to talk about their child's diagnoses and needs. For many families to hear that we can put them in touch with other families for mutual support either through a network of email contacts or through our managed internet forums is a real bonus.

This support enables families to better manage the effects on the family of the diagnosis and to talk things through with other families who have already been there.

### ***Specialist community nurses***

There is an increasing need for specialist metabolic nurses working in the community. Climb recently funded a two year research programme utilising a specialist nurse in the West Midlands to gauge the information needs of Black and ethnic minority (BME) families in that area. All the families who took part found it of great benefit to have a specialist nurse visit them who could talk, understand and advise on specific metabolic diseases in their own home. This remit would be well outside the role of the midwife and in the majority of areas a second language would be a priority when working with BME families.

There needs to be a national programme to introduce these specialist community nurses who would be hospital based but also able to liaise on a national basis and provide training to other medical professionals and families in their area.

### ***Family case study***

A family case study covering aspects of screening and diagnosis follows. It raises many issues.

**Box 8.3    Family case study**

When our son was diagnosed we lived in England (this was over 20 years ago) and he had been born 7 weeks premature. We spent many very difficult days, weeks and months at first because, although he was diagnosed with a urea cycle disorder, it was thought, at first, to be another one in the cycle, so effectively he was treated with the wrong medication.

**What was helpful:** having very supportive dieticians at the hospital in Europe to begin with and then later in the UK. Strangely, dietary help was less obvious in the US, where we went on to live for 5 years, but, on the other hand, there was a lot of support through family networks and lots of special diet foods already available over there.

Being able to talk to other parents about what worked for them - especially getting one's child to eat or take medicine.

Having an emergency kit of IV meds and details of protocol. We have still had to be very insistent in a lot of instances at the hospital but having the IV meds has, in most cases, helped to get treatment started more quickly. When the meds became available in capsule/tablet form it changed all of our lives.

**What has not helped:** medical professionals not wanting to spend time to listen to what needs to be done and doing it!

**Staff training, training and more training!** It's a shock to staff and careworkers to be told they have a new client who has such an unusual diet and that this client is NOT ALLERGIC TO PROTEIN (!) but needs to be monitored.

**Contact with doctors and specialists:** this has improved considerably in the past few years. Historically, I think, in the US, one was able to speak to a doctor or professor much more easily. In Europe and in England they were often unavailable or perhaps their staff were just over-protective. Thankfully, it seems that nowadays one is able to speak to someone on the phone or by email fairly quickly.

**Newborn Screening:** Yes please - the sooner the better.

**Climb has always been there for us and that has probably been the greatest source of support, information and understanding.** It's truly remarkable how much Climb (*i.e.* all its staff, members and volunteers) has been able to influence and involve so many people. Thank you!

Note. If Newborn Screening had been available for this family they would not have had months of stress with their new child. The misdiagnosis would have been avoided and the child would have been on the correct medication from the outset.

### 8.3 Conclusion

The concept of metabolic screening for the diagnosis and treatment of inherited metabolic diseases has evolved very slowly in the UK and as new methodology for detection and improved treatment have become available, we have not kept up with these advances but have let the ball drop.

The diagnosis of metabolic disorders is challenging because of the nature of the metabolic disease, the wide range of associated symptoms that are often linked with some more common conditions especially by the GP and the low incidence factor. The consequent lack of experience within the medical profession especially GP's and paediatrics often gives more stress and concern to a parent when they know their child is ill but the GP is not concerned. There are thousands of families affected by metabolic diseases in the UK and it is quite possible that when looked at collectively the level of incidence is higher than 1:5000. Although many children die of an inherited metabolic disease *before* a diagnosis can be determined it is fairly certain that inherited metabolic diseases represent a significant number of the morbidity and mortality rates in the newborn population.

The most opportune time to diagnose an inherited metabolic disease is during the first few weeks of life. Early diagnosis enables appropriate treatment, support and dietary management without which tragic outcomes are all too common. Public awareness of metabolic diseases in the UK is, at best, limited and even after the very low level announcements of the introduction of MCADD into the Newborn Screening programme the level of awareness is exceptionally poor.

In the US and Canada they have introduced expanded screening in most states to over 20 inherited metabolic diseases yet we, in the UK, lead the world in many areas of medicine but when it comes to inherited metabolic diseases we are sadly lacking in the areas of screening and diagnosis.

The early introduction of PKU and now MCADD screening exemplifies the benefits of early diagnosis of a metabolic disease to patients, their families and society as a whole. The benefits of finding and treating these patients far outweigh the costs of screening the entire newborn population.

Expanded newborn screening is fully supported by Climb and its service users. The use of MS/MS to screen for many inherited metabolic disease from a single blood spot is a benefit we cannot ignore and we must introduce this into the Newborn Screening programme with the minimum of fuss and certainly without the 7+ years of trials and testing had with MCADD.



## 9 Ethical, legal, and social implications of expanding newborn screening

### 9.1 Introduction

This chapter provides an overview of the ethical, legal and social issues (ELSI) identified as part of the systematic review and includes issues identified in the wider literature on screening. The systematic review process yielded eight references in which there was explicit mention of an ethical, social or legal issue arising as a result of screening. These references were analysed and results summarised. In addition a general literature search was carried out.

The ethical, legal and social issues arising in the context of newborn screening for rare metabolic diseases can be classified in a number of ways: firstly they may be a consequence of incomplete knowledge about the natural history of each disease, whether in terms of the incidence and prevalence of each condition, its severity, progression and response to treatment. They may also arise because conventional measures of acceptability (namely that of balancing risks against benefits) fail to translate well when research is applied at population level.

The appropriate framework for scrutinising the ethical basis of newborn screening programmes has increasingly been questioned by some commentators<sup>77-79</sup>. Tension arises because the ethical principles used to assess the benefits of interventions at the level of the individual (such as autonomy and non-maleficence), sit uneasily when applied to the population. Thus public health ethics, some argue, implies far more than a need to satisfy the requirements for participants to give an informed consent, but also demands that issues of liberty, self-determination and distributive justice are properly addressed. These issues are explored later in this chapter.

#### *Issues arising*

A number of clusters of issues were identified, from the nature of the conditions which form the subject of this review, those arising from the newborn screening process itself, the nature of the test, and governance issues around the storage, disposal and secondary use of samples<sup>80</sup>. In addition, a broader set of issues grounded in public health and the relative rights of individuals and the state in liberal pluralist societies - such as the requirement for distributive justice or public justification were also noted.

### 9.2 Issues arising from the nature of the conditions

A clear assessment of the benefits and burdens associated with newborn screening has been hampered by the characteristics of the conditions under review. Like many rare single gene disorders, lack of information about the natural history of the condition and lack of robust epidemiological information may make it difficult to predict the course of disease at the level of the individual, to the extent that the justification for intervention may be unclear. Within the US, arguably the availability of screening technologies, coupled with nationally agreed recommendations for screening a panel up to 84 conditions is said to have driven changes to the application of ethical principles to newborn screening<sup>81</sup>. However, as



preceding chapters of this review have demonstrated, increasing clarity about the natural history of each condition has allowed the balance of risks and benefits of offering expanded screening to be made more explicit.

Within the UK, for those conditions which form the subject of this report, the findings from previous chapters suggest that despite the heterogeneity of the conditions, the balance of the child's best interests almost always lie in testing and intervention. However some aspects require more examination:

### ***False positive results<sup>82;83</sup>***

One of the most problematic features of these conditions is the potential for the generation of false positive results. One concern is that if parents receive a 'false positive result' however swiftly resolved by subsequent tests, that this invokes an anxiety reaction. This might take the form of hypervigilance in safeguarding the child, or parents being convinced that their child is at risk despite subsequent negative tests<sup>84</sup>. Other manifestations of anxiety can include a general distrust of health professionals or institutionalised medicine, as has been reported in other newborn screening programmes<sup>85</sup>.

Whilst there is a need for ongoing research investigating the effect of false positive status on future health care use and parental anxiety, the available evidence suggests, in the case of these biochemical disorders that almost all of these cases can be resolved by further diagnostic tests so that parental anxiety can be mitigated - particularly if such tests are performed promptly within a structured care pathway. For example, one study found that the predicted burdens associated with false positive rates in the context of biochemical disorders were limited to the cost of diagnostic testing and follow-up<sup>69</sup>. Another found that although expanded newborn screening has potential to yield positive health outcomes for affected children, false positive results do place other children at risk of increased stress and parent-child dysfunction<sup>61</sup>.

Significantly, this parental stress can be partially mitigated by effective information strategies<sup>69</sup>. Arguably the potential for inducing increased parental anxiety through newborn screening might be distinguished from that observed on receipt of an abnormal screening test during pregnancy (where a sustained anxiety response may be observed even if the abnormal positive result is subsequently proved to be false by a normal diagnostic test)<sup>86</sup>. This might be because in the biochemical disorders under review, the interim result is typically confirmed to be either positive or negative within two weeks from the initial screening test, whereas in pregnancy there is typically a longer period of uncertainty which could be more likely to exacerbate an anxiety response<sup>87</sup>.

### ***The burden associated with diagnosis in cases of 'mild' disease***

In a minority of cases the variability of the natural history of the disease might be associated with uncertainties around the significance of a true positive result. In such cases it might be disproportionate to treat mild disease, if treatment involves significant burdens. For example, various mild forms of isovaleric acidemia (IVA) have been described following the advent of newborn screening, where those affected share a mis-sense mutation which results in a partial reduction in IVD activity. When followed for a period of 5 years, affected children were found to be asymptomatic and have not required significant dietary protein

restriction or carnitine supplementation<sup>16</sup>. However, the clinical consensus appears to be that even in these cases, a true positive result is genuinely beneficial because, at the very least, it aids care givers to be prepared if symptoms arise in times of metabolic crisis.

### ***Lack of gold standard trials***

The rarity of these conditions means that there is a lack of gold standard trials (*i.e.* randomised controlled clinical trials) and what research is available is often inadequately powered<sup>88</sup>. As we have seen, there is often a lack of descriptive natural history about the disease and the effect of treatment on preventing or ameliorating them. Moreover, since the diseases are autosomal recessive conditions, results may be confounded by founder effects in some populations or consanguinity which increases the reported incidence in some studies.

### ***Genetic carriers***

Unlike other screening tests (such as those for cystic fibrosis and sickle cell anaemia), the tests used for the metabolic diseases under review identify homozygotes, and do not allow the explicit diagnosis of carriers (although parents are, by implication necessarily carriers)<sup>89</sup>. Thus the sensitivities associated with the diagnosis of carriers, particularly those noted in certain ethnic groups, are unlikely to arise, but may be relevant to some population groups<sup>90</sup>.

## **9.3 Issues arising from the newborn screening process**

The ethical evaluation of population screening of newborns has tended to adopt the prevailing ethical principles current in Euro-America. Thus issues of autonomy, informed consent and privacy are concerns that arise frequently in the literature. Indeed uncertainties about the natural history of conditions being screened for tie in with worries about the practicability of obtaining a properly informed consent.

### ***Consent***

#### ***The form of consent (informed choice or informed consent)***

The requirement for an informed consent or informed choice is a topic that dominates much of the literature on newborn screening. Opinions vary as to the suitability of an informed consent or informed choice model and the extent to which a valid consent is possible within an opt-out system. One concern is that too high a threshold for consent (in terms of quantity of information communicated) might adversely affect participation<sup>91</sup>. Concerns have also been expressed that obtaining consent would be impractical, costly, or inappropriate (because it did not increase patient understanding of the issues involved) or that it is obtained at the expense of increasing patient anxiety<sup>92</sup>. In particular, there remains a lack of consensus about the degree of detail required to be given to participants, and the extent to which the requirement for a valid consent necessarily requires that participants understand all relevant issues. This is even more problematic where the aim of the screening test was to identify (or exclude) a range of conditions having very different presentation, symptoms, and treatments. A further point of concern is the systematic application of screening tests within the extended family raises issues around choice, and in particular how the right of parents to refuse newborn screening tests should be weighed against the child's best interests<sup>84</sup>.

### *Obtaining consent within an existing newborn screening programme*

Concerns about these requirements have been exacerbated by the experience of newborn screening programmes in other jurisdictions including the US, where there has been considerable criticism of the lack of an evidence base supporting the recent expansion of mandatory newborn screening. Indeed a recent commentary has advocated the revision of the consent processes associated with such programmes calling for the inclusion of preference-sensitive decision making process for receiving reproductive risk information<sup>93</sup>.

In other jurisdictions, such as the UK, offering testing within a screening programme seems likely to influence the balance of ethical issues involved by providing a context for ongoing diagnostic tests enabling a firm diagnosis to be made, and ensure availability of ongoing screening, and targeted preventative measures or treatment. Where the risks and benefits of screening are finely balanced, an alternative might be to adopt staged testing regimes as a means of gaining informed consent such as that used where carrier testing is offered to those at high risk.

### *Appropriate management of results*

A set of issues arise in relation to the need to keep test results confidential and the potential harms associated with their disclosure. Health professionals are familiar with the obligation to keep patient identifiable information confidential, and systems are already in place which protect against disclosure. It is important that sufficient numbers of adequately trained personnel are available to provide appropriate information to parents and secure meaningful consent prior to blood being taken for screening, and as noted above that there is sufficient capacity to deal with increased false positive cases that are likely to arise.

### *Feedback of results*

For each of the conditions under review, there are clinical benefits associated with the identification of those at risk. Research into parental attitudes suggests that even in cases where preventative treatment is not available at the time of disclosure, parents favour feedback. For example, in their survey of a Swedish birth cohort<sup>94</sup>, Helgesson *et al.* conclude that 74.3% of those surveyed were in favour of disclosure in the absence of a treatment being available.

### *Confidentiality*

The confidentiality of samples taken for newborn screening, and their subsequent use for research outside the terms of the original consent, are a potential source of concern. The requirement that data taken for one purpose are not processed for another without consent is made explicit in the UK Data Protection Act 1998 which establishes a statutory requirement for fair processing of data in accordance with data processing principles<sup>95</sup>. In addition the use of the blood spot samples which are generated by the existing newborn screening programme is regulated by codes of practice prepared by the Programme Centre in partnership with the UK Newborn Screening Laboratory Network<sup>96</sup> as well as legislation such as the Human Tissue Act<sup>97</sup>.

*Incidental findings such as misattributed paternity*

Like other genetic tests, newborn screening has the potential to reveal unintended findings which could be potentially harmful. These include the possibility of misattributed paternity, which has the potential to threaten existing social relationships. This is likely to be more harmful where assessing the carrier status of both parents might be an integral part of rolling out the test findings to other family members (through cascade testing). The conditions forming the subject of this review are unlikely to be cascaded out in this way in the UK, except perhaps where consanguinity has a role in increasing incidence of a condition.

*Implications for family members*

A positive test result in an affected child may also be used for purposes of reproductive choice in subsequent pregnancies. For some of these conditions, this might enable prenatal testing or even pre-implantation genetic diagnosis (where known mutations causing serious disease have been identified). Thus whilst knowledge of pathological mutations could allow screening tests to be offered in a targeted manner to other family members (especially siblings of affected children) who might be at risk of disease, for most of the conditions under review, mutation analysis tends to be used as an adjunct to clinical management rather than for the purposes of reproductive choice.

*Characterising 'benefits' in terms of direct benefits for the child and those for the wider extended family<sup>98</sup>.*

The potential benefits of testing and diagnosis may extend beyond any benefit potentially available to the newborn child to include better management of symptoms. Screening might also generate incidental results that relate to diseases for which there is no prospect of benefit, and it might be unethical to withhold results of that screening from parents. This practice has been criticised as the basis for a mandatory screening programme<sup>81</sup> since it offers the prospect of societal benefits through the findings allowing biomedical research on the condition, at the expense of individual participants. In the conditions under review, the tests may reveal such diseases, as well as potentially identifying mild forms of these conditions that have hitherto remained undiagnosed.

**9.4 Issues arising from the test itself***Timing of the heel prick test*

Newborn screening within the UK provides for samples to be taken by heel prick between 5-8 days after birth<sup>99</sup>. Since tandem mass spectrometry measures analyte concentration, the selection of this time period for sample collection has an impact on the conditions that can be detected using this methodology. Thus in the case of classical MSUD, an affected baby may already be symptomatic by the time the screening results become available, suggesting that the added value of the screening intervention is limited to confirming the diagnosis. As noted in Chapter 3, most cases of classical MSUD will require ICU treatment within 48-72 hours of birth. In the fatty acid oxidation conditions the reverse is the case, with analyte levels falling after birth. Some compromise is inevitable since in practice, were newborn screening to be extended to include the five conditions under review, the timing of sample collection has to fit with existing newborn screening strategies even though it might not represent the timing of choice for each condition. In addition, differences between the

timing of tests in other jurisdictions mean that testing protocols cannot easily be compared between countries that have a different collection strategy.

### ***False positives or negatives and timing of the heel prick test***

The use of 'day of screening' as a benchmark for testing, allows the possibility of false positives derived from low birth weight babies or metabolically stressed infants whose needs might be unmet<sup>86</sup>. Any protocol needs to take account of factors such as these which have potential to confound test results.

### ***Existing methodology is not representative of vulnerable groups***

Evidence from Down's syndrome cohorts suggests that those who decide not to proceed with testing are often poorly represented in studies. Thus it may be difficult to show a causal connection between the test and anxiety (as no comparative measure exists for the untested group and meaningful comparisons might be difficult). Although only a small minority (less than 1%) of the population refuse newborn heel prick testing, it is important that this group is not lost to follow up<sup>100</sup>. Furthermore, high rates of attrition of cohorts undergoing testing may mean that results are biased in favour of the group who complete all stages of testing<sup>86</sup>.

## **9.5 Issues relating to the governance of the material**

In other contexts, the retention of samples and data which could yield predictive information about a child might raise concerns about the security of the holding or the uses to which it might be put. Like other samples of human material, there are some general concerns arising from the potential of human blood spot material to yield information beyond the purposes for which it was originally taken, with or without consent. Some commentators are very suspicious of the regulatory environment which allows continued retention: others bemoan the lack of a systematic process for greater access to such samples. For completeness, these issues are reviewed here. If these samples and data are to be held as part of the existing newborn screening programme, a code of practice applies to the retention of newborn samples. This was drafted by an expert committee in 2005 and has not yet been updated<sup>96</sup>. This provides for information regarding research uses of blood spots to be published in UK National Screening Programme Centre (UKNSPC) Annual Reports, but the report for 2006/7 does not contain this information. Standard operating procedures also govern feedback of normal test results to parents which expressly exclude from their scope provision for feedback of screen positive results or reporting results to GP's<sup>101</sup>. Provision is therefore made for:-

### ***Security of storage and maintaining privacy and confidentiality***

Samples should be held in a secure and stable environment so that the quality of the samples is not compromised if used for retesting or research (subject to consent having been given). Codes of practice provide for de-linking of the physical card and personal identifiable information (including NHS number) but authorised persons may re-link sample and data provided that this can be justified.

***Disposal / unlinking of samples***

Within the UK, current protocol provides for a minimum retention period of five years. The retention of forensic samples has been the subject of litigation under Human Rights legislation and UK policy in this area is evolving<sup>102</sup>. What seems clear is that the retention of samples should be proportionate to the likely benefits.

***Use for secondary purposes, either with or without the knowledge and/or consent of the parents***

Anonymised blood spots can be used for health monitoring and research (for example to determine the spread of infections such as HIV and rubella within the population). This type of research can be carried out without consent (subject to satisfactory research ethics approval). Existing literature provided to mothers during pregnancy and at the time of taking heel blood samples provides an opportunity for mothers to opt out of further research. However there is no existing mechanism for women to consent to testing but opt out of ongoing retention of samples<sup>103</sup>.

***The effect of the Human Tissue Act 2004***

The development of a pilot project for the five conditions that have been described will need to take account of the requirements of the Human Tissue Act<sup>97</sup>. This legislation has profoundly changed the regulatory and governance structure from that which prevailed at the time the MCADD pilot was completed, and as a consequence, the methodology used in future pilot studies will need to be reassessed. In particular, the requirement for appropriate consent to be obtained for the use and retention of tissue (which includes retained blood spots) for research will influence the design of any proposed pilot project.

**9.6 Broader implications of public health ethics*****Conformity to other ethical frameworks***

The Programme Appraisal Criteria approved by the NSC are one example of an ethical framework by which an expansion of screening could be measured. There are other frameworks which have been proposed, against which newborn screening programmes have been considered in other contexts<sup>79</sup>:

***Meeting the criteria set out in an alternative ethics framework for public health***

Nancy Kass<sup>78</sup> has argued that public health interventions should be measured against six criteria. These are set out in Table 9.1 below, together with an assessment of the extent to which the proposed extension of the newborn screening programme might satisfy those criteria.



**Table 9.1      Criteria**

Stated Goal	Requirements	Measure of success of the proposed expansion of screening to include five metabolic conditions
1. What are the public health goals of the proposed program?	The goals of any proposed programme should be expressed in terms of public health improvement ( <i>i.e.</i> reduction in morbidity and mortality)	Reasonable assessments have been made of the likely reduction in morbidity and mortality that would result from introduction of an expanded screening programme in the UK.
2. How effective is the program in achieving its stated goals?	There should be evidence that such goals are achievable	The infrastructure for introducing the programme is already in place. Some details need to be clarified through a pilot study.
3. What are the known or potential burdens of the programme?	Identification of potential burdens or harms. These might include risks to privacy and confidentiality, risks to liberty and self-determination and risks to justice	Many of the risks and burdens have already been quantified. A pilot study is needed to quantify additional harms.
4. Can burdens be minimised? Are there alternative approaches?	The minimisation of those burdens and consideration of alternative approaches	For the conditions under review, burdens can be minimised by ensuring that robust procedures for obtaining an informed consent are in place, ensuring that protocols aim to minimise the number of false positive results, establishing protocol for processing positive results through second-tier testing and that parents are supported through this process. The alternative to expansion of screening is to retain the status quo (and test and treat affected individuals when they present with symptoms to health services).
5. Is the programme implemented fairly?	The programme should be implemented fairly (in accordance with principles of distributive justice)	If an expanded program is offered through existing provision of newborn screening, inequalities should not arise in relation to the initial screening test. However there might be geographical variation in accessing confirmatory testing and support from a tertiary centre.
6. How can the benefits and burdens of a programme be fairly balanced?	Whether the anticipated benefits of a programme outweigh the burdens which have been identified (via procedural justice)	The burdens associated with expanded screening will be widely distributed across the population of those screened. Expansion will require additional information to be provided to participants, a revised consent to be taken, which might in turn generate some additional anxiety both amongst the population of those screened, and more particularly, those parents who receive a false positive result. For the majority of parents however, screening will be beneficial in that the prospect of rare metabolic disease will be ruled out. The benefits of testing that accrue through prompt diagnosis and treatment will only apply to a small group of affected individuals.

In summary, the introduction of a limited programme for the conditions in this report would seem to satisfy the six criteria identified by Kass. A pilot study could help to clarify the extent to which offering a wider programme of testing might generate increased individual risks to liberty and self determination (such as the risks of coercion implicit in a universal programme) which have been identified by other commentators<sup>79</sup>.

### *Establishing justificatory conditions for screening*

Other writers have argued that screening should satisfy a public health ethics framework comprised of five 'justificatory conditions', namely effectiveness, proportionality, necessity, least infringement and public justification<sup>77</sup>. The fifth criterion, that of public justification, it is argued requires a degree of public accountability such that the rationale for programme setting and choice is made transparent.

### *Voluntariness and the role of the state in offering or mandating testing*

In some jurisdictions, most notably some states in the US, some screening programmes are mandatory. There has been wide debate about the extent to which mandatory testing can be justified where the benefits and burdens of testing remain inconclusive. In the UK, the right of parents to opt out of newborn screening has been respected, even where testing might be in the best interests of the infant. Some commentators have argued that a parent's right to refuse testing is supported by the Human Rights legislation which entrenches the right to respect for private and family life<sup>84;104</sup>.

Whether tests are voluntary or mandatory has implications for how data and samples emanating from the tests are governed. For example, in the US since screening is mandated by federal law, third party uses of data and samples are less restricted than in the UK where the terms for retention and storage are dictated by the terms of the original consent and code of practice<sup>105</sup>.

### *Assessing the burdens of screening: economic cost*

An additional argument which has been used in support of the expansion of newborn screening is that different disease profiles can be added to existing arrays at marginal additional cost, partly because the existing infrastructure for collection, analysis and reporting of samples is largely in place. This is the subject of a separate section of Chapter 7.

### *Beyond transparency to public engagement*

Finally, whilst the obligation for transparency requires publicity to assure 'accountability for reasonableness'<sup>106;107</sup> there would also seem to be a requirement for public engagement 'in order to engender and sustain public trust'<sup>77</sup>. If the decision is made to extend newborn screening, public engagement and public education is needed. Some work has already been done in this area<sup>108</sup>. Plans for a pilot project would take this work forward. Conversely, if the decision is taken not to proceed with expansion, the reasoning behind the policy decision should be made explicit.



## 10 Expanded newborn screening programme and EU policy towards rare diseases

This Chapter was contributed by GIG and RDUK. Authors: Alastair Kent, Melissa Hillier and David Brown.

### 10.1 Introduction

An expanded screening program based on pre-established screening technology would create opportunities to significantly improve the quality of life for affected individuals, and reflect a growing institutional and public awareness of the burden of rare diseases. While, by definition these diseases are individually rare, it is “*not unusual to have a rare disease*” and between 6-8% of the EU population will experience some form of rare disease in their lifetime<sup>109</sup>. The common problems presented by rare diseases are characterised by inefficiency and waste from misdiagnosis, delay, repeat consultation and inappropriate treatment, problems that could be in many cases alleviated by an expanded screening program<sup>109</sup>. These problems present a chronic challenge to the healthcare system as a whole and an acute disadvantage to individuals, for whom time is of the essence.

### 10.2 EU rare disease policy

EU policy with regard to rare diseases was initially addressed through the community action programme on Rare Diseases (1999-2003)<sup>110</sup>, which focused on providing information on the problems presented by rare diseases as a whole. Rare diseases have been recognised as a major public health issue in the Second Community Programme of Community Action in the Field of Health (2008-2013)<sup>111</sup>. This document identified non-communicable diseases as 77% of the EU burden of disease, and asserted the need to increase “Healthy Life Years” (HLY), in contrast to those years effected by disability.

The Communication from the Commission to the EU Parliament, Council and Economic and Social Committee and the Committee of the Regions on Rare Diseases<sup>112</sup>, which focused on the opportunity to address rare diseases on an EU level, encouraged national efforts at addressing rare diseases, co-operation on an EU level, and increasing visibility and recognition of rare diseases. This was followed by the adoption by all member states of the Council Recommendation on Action in the Field of Rare Diseases<sup>113</sup>, which focused on the adoption of coherent national strategies towards rare diseases.

The piloting of additional screening protocols would be a clear, visible and measurable movement on the part of the UK towards tackling rare diseases as a public health concern in line with their signing of the Council Recommendation on an Action in the Field of Rare Diseases. The five conditions are well below the 5 in 10,000 prevalence threshold for recognition as Rare Diseases. A pilot study examining the feasibility of extending screening to these conditions would contribute to any emerging strategy fostering research on rare diseases and their incorporation into the UK’s healthcare provision<sup>113</sup>.

### 10.3 Metabolic conditions and screening

The five conditions are screened for via MS/MS, a technology already in operation for MCADD screening. Given this clear commonality it is reasonable and beneficial to take inherited metabolic disorders with a common screening procedure together, and to examine their effect on public health cumulatively. This is in line with a trend in EU policy communications to address rare diseases as a whole and yet in a manner that allows for specific variation.

While primary prevention is impossible for these diseases, the symptoms can be mitigated significantly by early diagnosis and nutritional treatment, with a statistically significant increase in disability-free life years. While these diseases are rare, the burden of the disease is significantly lower if detected prior to the emergence of symptoms. Metabolic disorders affect multiple aspects of a patient's mental and bodily development over time and if undiagnosed will produce a 'polyhandicap' or multiple disability<sup>109</sup>. As has been demonstrated in the measurable differences in IQ, neurological development, and chances of survival, discovery via screening prior to clinical symptoms is the best means of diminishing the effects of metabolic disorders.

It is important to note that while other EU countries have trialled and retained these additional screening protocols, prevalence of these disorders differs between countries and this limits our ability to generalise such a result to the UK. Conversely, given higher incidence of rare inheritable diseases in ethnically diverse populations with higher levels of consanguinity, screening may be of greater importance in a multicultural Britain than indicated by other pilots.

The disparity in prognosis between early and late diagnosis is a common concern in rare disease policy and screening for these disorders would indicate a positive trend towards addressing this problem<sup>112</sup>. Given the issue presented by rarity and scale, a full national approach presents the best opportunity for catching cases early and treating them effectively. Early diagnosis also allows for more rapid mobilisation and implementation of expertise that may not be immediately available due to the rarity of the condition<sup>113</sup>. Screening will also identify individuals whose condition may not become symptomatic until permanent damage or disability has occurred.

## 10.4 Patient stress and empowerment

Early identification and consequent contact between parents and specialised support centres and charities has been associated with lower stress and therefore increases patient empowerment (see Chapter 8). Patient empowerment is described by the World Health Organisation as a “*pre-requisite for health*”, and is a recurring theme in the EU Rare Disease literature<sup>112</sup>. Rare diseases in particular, are: “*Highly painful in terms of psychosocial burden: the suffering of rare disease patients and their families is aggravated by psychological despair, the lack of therapeutic hope, and the absence of practical support in everyday life.*”<sup>109</sup>

Screening programmes offer the opportunity for both therapeutic hope and practical support, and thus will help to mitigate despair and disempowerment for patients and their families.

While there is potential for false positive results in screening to cause unnecessary stress, this can be mitigated through the provision of advice and a clear healthcare pathway. Where the UK has been reticent with increased neonatal screening due to concerns over privacy, the metabolic disorders under discussion do not produce the same cascade results as other inheritable diseases or else are already addressed by pre-existing blood-spot screening protocols. Thus the forms of screening under consideration empower patients and their relatives by informing them of treatment whilst presenting minimal additional intrusion into patient privacy.

At present no clear general procedure for quality assurance for a potential expanded newborn screening program exists and a pilot study would address this omission. Specifically, MS/MS has a clear and established method of quality assurance. Transparent quality assurance of diagnostic procedure is a clear theme in the EU recommendation and would be well served by a pilot study.

## 11 Expanded newborn screening: discussion and recommendations

### 11.1 Introduction

The purpose of this chapter is:

- To discuss the current NSC criteria in relation to the five conditions, including the relevance and shortcomings of the criteria, and other criteria that may aid in decision making
- To assess which of the criteria are met by each of the five conditions
- Where fulfilment of criteria is uncertain, to identify and outline the gaps in our knowledge
- To make recommendations for questions that could be answered by a pilot programme
- To set out some of the important factors that the National Screening Committee might wish to consider before recommending a pilot programme

In their 'Public Health Classic' published in the Bulletin of the World Health Organisation Andermann *et al.*<sup>19</sup> comment on screening in the genomic age and note that '*an ever-widening gap between what is technologically possible and the services available is creating pressure to introduce or expand screening programmes often before adequate safeguards and regulatory frameworks are in place*'. Whereas in the US during the last decade there has been rapid expansion in newborn screening following a report commissioned by HRSA/MCHB and undertaken by the American College of Medical Genetics<sup>29</sup>, the UK has adopted a more conservative and stepwise approach insisting on individual condition evaluation against screening criteria.

The successful completion, following favourable Health Technology Assessment, of a pilot study for MCADD, and subsequent implementation in England as part of the newborn blood spot screening programme has established that methodology and instrumentation are now sufficiently robust for large scale routine use. In September 2007, a joint meeting of the UK National Screening Laboratory Network and the National Metabolic Biochemistry Network (MetBioNet) was held to consider the lessons learned from the pilot, to hear about progress in international screening programmes and to discuss an approach to expanded screening. In particular, it was noted that expanding the scope of MS/MS was not the same as starting a new programme from scratch as the infrastructure for sample collection, analysis and reporting are already in place. The meeting generally agreed that expanded screening was:

- clinically useful
- widely available internationally
- safe
- requiring only marginal resources
- popular with patients and parents
- politically acceptable

It was clear that many clinicians would like to see the scope of screening expanded. In general the desire was to include those conditions that would be vigorously treated once they were diagnosed. As Rodney Pollitt commented in his opening address, these conditions would fulfil the categorisation of *'if only we had known earlier - in that the earlier in life the diagnosis was made, the more effective such treatment is'*. As a result of this meeting, a list of potential candidates was drawn up, including the five target conditions of this study.

Subsequent approaches to the NSC to consider an expanded programme were met with the response that screening should not be technology led, but that it should be possible to demonstrate that each condition individually fulfilled screening criteria. Various 'vignettes' were prepared in response to this requirement. However, because of the non-quantifiable nature of the screening criteria, the subjective nature of many of them, and their overall lack of suitability for rare genetic conditions, it has never been clear how judgments against the criteria would be made.

The requirement that each condition fulfil screening criteria on its own merits contrasts with the line taken in the US. Here it was argued that *'change in the technological landscape in the form of multiplex platforms that enable screening for multiple disorders from a single specimen had the potential to reduce costs per condition tested and could lead to expansion if these technologies could be integrated safely and effectively into newborn screening'*. Inherited disorders in the newborn were largely considered as a group, and even some untreatable conditions were eventually included on the grounds that screening might provide 'benefit to society' through reducing the cost of diagnosis and by generating research benefit as the natural history of conditions could be followed from an early stage at population level<sup>29</sup>. The recommendation of a nationally agreed panel of conditions meant that by November 2008 almost all states had adopted the ACMG panel of 29 core conditions and a proportion of 25 secondary conditions. Whilst such an approach seems unlikely to be acceptable in UK, and indeed has since been criticised<sup>81</sup>, nevertheless, some of the screening criteria and methods used to assess and rank conditions can assist in our discussion of the screening criteria.

However, it is not for this report to establish whether or not the criteria are fulfilled but simply to provide evidence that the NSC may use to support their decision making and, where possible, to make suggestions to guide the NSC in the process of interpreting the criteria against this evidence. We will also point out where there are gaps in evidence and suggest where these may be filled from a pilot programme. Finally, we recognise that, even when criteria are met, there may be logistical, social or ethical reasons that preclude service development and that the eventual decision will be influenced by broader political and economic considerations.

## 11.2 Screening criteria

The NSC screening criteria are set out in the Appendix 1 and will form the basis of this discussion. As an additional resource, the ACMG established a set of criteria against which experts reviewed a panel of conditions using both scientific evidence, expert knowledge and subjective opinion and provided some quantitative and qualitative guidance for scoring. In the following pages we set out each of the screening criteria identified by the NSC (where appropriate as supplemented by the ACMG) and discuss some of the issues arising in relation to the five target conditions.

The three main categories and criteria identified by the ACMG are summarised in Table 11.1 together with some guidance on the way conditions should be scored against the criteria.

**Table 11.1 ACMG criteria**

<b>Clinical characteristics of the condition</b>	Incidence; clinical identifiable signs and symptoms in the first 24 hours ( <i>i.e.</i> prior to screening test); burden of disease (natural history if not treated)
<b>The screening test: availability and characteristics</b>	Availability of a sensitive and specific test algorithm; ability to test on neonatal blood spot; test is based on a platform for high throughput; low cost of test; multiple testing of analytes for same condition (to increase specificity); ability to identify other conditions (secondary targets); multiplex testing (for multiple conditions)
<b>Diagnosis, follow-up, treatment and management</b>	Availability of treatment; cost of treatment; potential efficacy of existing treatment; individual benefits of early intervention (benefit to child); family and societal benefits of early identification; prevention of mortality; availability of diagnostic confirmation; availability of acute management; simplicity of therapy

As a helpful starting point, the ACMG noted explicitly that low scores in a particular area would not mean that screening for that condition should never be conducted but that they could be radically overruled, for example, by new advances in testing and treatment.

## ***The Condition***

### **1. *The condition should be an important public health problem***

The importance, or otherwise, of a condition as a 'public health problem' is usually considered to relate to:

- a) the number of people affected
- b) the severity of the condition

Together these constitute the 'burden of disease' within the population.

The ACMG illuminates this somewhat by noting that, in terms of public health importance, the more common the condition, the higher the justification for screening. They describe the burden of disease as an important criterion because it favours screening for conditions that constitute a greater burden. However, they note that for individual conditions there may be a wide spectrum of severity and tests may not necessarily discriminate between the milder and more severe forms.

Each of the target disorders in our discussion, and most other inherited metabolic and other heritable diseases are relatively rare. However, in total the European Organisation for Rare Diseases (EURORDIS) estimates that there exist between 5,000 and 7,000 distinct rare diseases. Although each individual disease is rare, the sheer number of individual rare diseases results in between 6% and 8% of the population of the European Union being affected by a rare disease<sup>109</sup>. Many rare diseases have severe consequences for the child and family and a high proportion of the children will require expensive care from health services, education and social care services. This is particularly the case as treatments improve survival, but survival may be with disability, so that the burden of care extends for many years and often into adulthood. Genetic or heritable conditions are thought to be responsible in children for:

- 35% of deaths under one year, 10% between the ages of 1 and 5 years old and 12% between 5 and 15 years old
- 30% of all childhood admissions to hospital
- 1/3 of children with visual disability or blindness
- 2/3 of all disability<sup>114</sup>
- around half of all severe learning disability and 15% of patients presenting with moderate learning disability

Rates may be even higher in communities such as the West Midlands with large ethnic minority populations, especially those with high levels of consanguineous marriage<sup>114</sup>.



The only recently published estimate of IMD as a group comes from a study in the West Midlands and gave an overall birth prevalence of 1 in 784 for the years 1999-2003<sup>4</sup>. Extrapolating this to the entire UK population would suggest approximately 800 new diagnoses per year, of which one third would be diagnosed by the age of one year. Whilst these results may be a little high, because of the effect of consanguinity in ethnically diverse populations, it should also be noted that there is likely to be under-reporting due to death occurring before a diagnosis is made, or other reasons for failure to diagnose. Some children will die without treatment whereas others may survive but suffer irreparable consequences of the disease such as learning disability.

The overall burden of disease due to genetic conditions should, therefore, merit a public health response that examines the possibility of prevention. EURORDIS advocates that this should be a global approach - rather than a piecemeal policy for each disorder separately. In particular, they recommend that *'where antenatal and asymptomatic phase screening methods for rare disorders allow for early and effective medical coverage, they should be implemented because they can significantly improve quality of life'*<sup>109</sup>.

Public health approaches to prevention will, however, be likely to differ from standard health promotion and disease prevention programmes since inherited metabolic diseases will not be amenable to 'blockbuster' prevention programmes, such as those tackling neural tube defects, sudden infant death, childhood obesity or infectious diseases. Rather than looking at risk factors for common diseases or indicators of latent or early stages of common diseases, prevention in rare genetic conditions is usually based on advances in understanding of individual disease causation and pathophysiology (often at a molecular level) and the development of technologies for identifying relevant biomarkers. This can allow those at risk of the conditions (or in some cases parents at risk of having children with these diseases) to be identified and informed of prevention options, and those with the conditions to be diagnosed as early as possible so that they can begin treatment. The ideal scenario for the latter is to identify the disease and start treatment before the development of irreversible damage so that a child will remain symptom free and develop as normally as possible. Prevention in individual genetic disorders is thus largely achieved by an approach that targets individual conditions leading to a gradual accumulation of health benefit for the population.

The development of relevant technologies can sometimes pave the way for preventive programmes for groups of genetic conditions. For example, amniocentesis and karyotyping allows identification of chromosomal disorders such as Down's syndrome and other trisomies in the foetus so that termination of pregnancy can be considered; routine antenatal ultrasound screening allows identification of congenital anomalies such as neural tube defects and congenital heart disease. In a similar way the introduction of MS/MS into newborn screening allows the identification of a cluster of inherited metabolic disorders. We argue that, just as other technologies can be harnessed to achieve more than one health outcome, it is legitimate and indeed logical to view this cluster of inherited metabolic disorders as a group which constitutes an important public health problem. Our view is that this does not detract from any requirement for each individual condition to adhere to the prerequisites that there must be effective interventions as well as good test performance, few adverse effects on the unaffected population and adequate services to look after affected children.



**Table 11.2 Burden of disease: details for the target conditions**

	Birth prevalence	Mortality and morbidity
MSUD	1 in 224,000	Patients may die during acute episodes. Untreated the condition is progressively fatal Learning difficulties, spasticity, and cortical visual impairment
Homocystinuria	1 in 392,000	Mortality 25% by age 30 Skeletal, ocular, vascular and nervous system pathology
GA1	1 in 109,000	Acute encephalopathic crises with high levels of mortality if untreated Children severely disabled with dystonia and dyskinesia will have difficulties in communication, feeding and swallowing.
IVA	1 in 155,396	Mortality in around 50% Severe neurological symptoms and learning disability
LCHADD	1 in 218,564	Mortality in more than half clinically presenting patients Cardiac, CNS, liver and ocular pathology, progressive encephalopathy
<b>Total</b>	<b>1 in 37,000</b>	

The total number of new cases in England and Wales each year would be on average 19 (range 16-23).

- The epidemiology and natural history of the conditions 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***

This criterion attempts to answer the question: ‘do we understand the disease well enough to feel confident that interventions will alter (hopefully improve) the expected natural history of the condition?’ It is expected that this question will be answerable by classical epidemiological studies. However, the rarity of inherited metabolic conditions means that the methods in standard epidemiological research, which rely on a population based assessment of disease comparing populations with and without disease, will be unlikely to provide sufficient statistical power to provide classical evidence on incidence, causation, risk factors and natural history.

Whilst lacking the strength of epidemiological evidence using large numbers, in inherited metabolic disease we have a group of conditions whose aetiology has been studied in detail in individuals with precise understanding of the underlying abnormal biochemistry (abnormal levels of metabolites, enzyme activity and underlying genetic defect) and where clinical

history has been followed in relation to evolving biochemistry.

Thus, we can use the argument of a detailed understanding of the aetiology and pathophysiology of disease and its clinical evolution in known cases to answer the question about natural history. In some cases (for example GA1) fairly large international studies (around 300 patients) have provided evidence of natural history and related this to different genotypes, biochemistry, early clinical events, stage at diagnosis and treatment. This sort of study, however, can raise new questions and decision makers should guard against allowing such detailed knowledge to create an unreasonable set of standards for this criterion that would not be applied to more common diseases. For example, with MSUD, clinicians would be confident that raised serum levels of leucine cause neurotoxicity and the clinical phenotype. However, for many of the conditions, detailed study of small groups of patients shows variable clinical outcomes that may be attributed to variations in genotype, relevant enzyme levels, a range of largely unknown other genetic factors, and the degree of exposure to environmental stressors such as infection, prolonged fasting or extreme physical exertion. Some such factors may indeed be broadly similar to those that influence the speed with which an individual with cervical abnormalities progresses to cancer and the ultimate aggressiveness or response to treatment. However, in the latter condition, prediction of average risk or progress, on the basis of study of large populations has enabled screening programmes to provide overall statistics for the population and to progress without the need to question the natural history of each condition on the basis of the precise genotype, phenotype and environment of each individual.

The question at issue thus becomes: do we understand the underlying pathology and expected natural history well enough to recommend treatment that we believe to be beneficial?

**Table 11.3 Relationship of abnormal biochemical phenotype to pathology**

	Pathology
MSUD	Related to toxic accumulation of branched chain amino acids and alpha ketoacids due to deficiencies in enzyme complex (BCKAD)
Homocystinuria	Pathology due to toxic accumulation of homocysteine in the blood and tissues
GA1	Toxic accumulation of glutaric acid, 3-hydroxyglutaric acid, glutatonic acid and glutaryl carnitine. Neurological damage due to encephalopathy
IVA	Acute and chronic forms related to toxic build up of IVA and its glycine and carnitine derivatives and depletion of intramitochondrial coenzyme A during crises
LCHADD	Toxic accumulation of long chain acyl-CoA esters and inability to synthesise ketone bodies which are a source of energy for organs such as heart and brain

The issue of ‘latent period’ is interpreted by ACMG in relation to whether or not the condition would be identifiable in the newborn period through routine clinical evaluation and, in particular, before the result of screening was available. They caution that recognition of a particular phenotype would differ according to typical health care provider or specialist and also that there would be phenotypic variability in presentation. In IMD, therefore, a continuum often exists between those that are never detectable in the newborn period and those where clinical manifestation would always be detectable.

**Table 11.4 Age of clinical presentation**

	Age at presentation
MSUD (Classic)	Symptoms days 2-4 (later for breast fed infants) Clinical presentation and diagnosis usually more than 1 week
Homocystinuria	Clinically normal at birth and diagnosis usually not made until age 2-3 years
GA1	Onset with encephalopathic episodes between 3-36 months. Occasionally non-specific clinical features before this time
IVA	Acute presentation characteristically in first two weeks of life
LCHADD	A few infants present in first few days but majority present later

**3. *All the cost effective primary prevention interventions should have been implemented as far as practicable.***

Since many of the conditions are recessive or sporadic they will appear to occur randomly in the population. Although, in general, measures such as improved detection of carriers in high risk communities (such as ethnic minorities at risk for Tay Sachs Disease), could be considered, no particular high risk communities are identified for any of our target conditions. For these conditions and for other IMDs, there should also be follow up of relatives of cases of IMD in large consanguineous families (with offers of prenatal testing), and counselling offered to parents of affected children. None of these options will have a major effect on overall infant mortality or morbidity, but they should be in place as part of high quality genetic services and, for some populations and conditions as part of antenatal (and possibly pre-conceptual) screening.

**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.***

Carriers of the mutation would not be directly identified by the screening test, but, by implication, the parents would be identified as carriers if their child was diagnosed with one of the conditions. In reality, however, this makes very little difference to parents as it is assumed that, for these conditions, the child would ultimately be diagnosed clinically, at which point the parents’ carrier status would be revealed anyway. Lessons for a newborn

screening programme could, however, be learned from experience in the neonatal cystic fibrosis and sickle cell and thalassaemia screening programmes. Although a recent HTA review (monograph 1357) was principally concerned with identification of carrier status in the newborn, issues were also described related to determining which of the parents was affected and whether other children might also be carriers, for the purposes of eventually giving them information that might be useful when planning pregnancies<sup>115</sup>.

### **The Test**

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

Neonatal screening for the target conditions is based on screening for elevated levels of particular analytes by MS/MS. This is performed in laboratories all over the world and the methodology has been shown to be simple and safe. Furthermore, it is based on existing infrastructure for sample collection and analysis. The validity of the test is measured by the sensitivity, specificity and positive predictive value. Data on this are available from both published and unpublished reports of international screening programmes. These data indicate that the individual tests for our target conditions are highly sensitive and specific in most cases. There are some instances, as discussed in Chapter 7, where the absence of abnormal biochemistry in the newborn period with certain sub-types of disease can impact on the sensitivity of the assay.

The positive predictive value of the test in most cases is very low and variability in the false positive rate across screening programmes can be observed. This is probably a reflection of the varying ages at sampling and different screening algorithms in use, which have an impact on the false positive rate and subsequently the positive predictive value. A number of factors contribute to the false positive rate including the fact that the screening test employed is not disease specific, but can rather be considered as specific to a “disease bundle” much in the same way as screening for PKU and MCADD. The development of testing algorithms, based on the quantitation of multiple analytes relevant to one condition or second-tier tests has allowed improvements in test performance and resolution of screen positives into appropriate disease or non-disease categories.

The ACMG considered the ability of a test to detect multiple analytes relevant to one condition as a positive feature, as it allowed further discrimination of true positives and false positives. In addition, value was also given to the incidental detection of disorders, which were classified as “secondary targets”. The ability to detect these conditions was felt to add value to the primary target, as they were clinically significant and could lead to tangible benefits for the newborn and the family.

#### **6. *The distribution of test values in the population should be known and a suitable cut-off level defined and agreed.***

Information from existing screening programmes can give some indication of test values in Caucasian populations. However, these data cannot be directly extrapolated to the UK as differences in age at screening and method of assay calibration will have an impact on the values. Definition of a suitable cut-off level for the UK will require analysis of samples from unaffected newborns to establish population means and ranges for the analyte of interest. These can be modified following accumulation of further data upon initiation of a pilot study or screening programme.

**7.      *The test should be acceptable to the population.***

Judging the acceptance of the test is very subjective and is influenced by a number of factors including parents' understanding of the screening process and the impact of false positive results. In countries where screening programmes for the five target conditions have been put in place, they have been widely accepted and this is likely to be the case in the UK.

**8.      *There should be an agreed policy on the further diagnostic investigation of individuals with a positive test and on the choices available to those individuals.***

Further diagnostic investigations of individuals with a positive test result involve analysis of a repeat blood sample and further analysis using urine and/or plasma specimens. For some individuals definitive diagnosis is based on measurement of enzyme activity. This might be complemented by mutational analysis, although the latter would not be required in the majority of cases. (It may, however, be included as part of a pilot study). These confirmatory assays are well known to diagnostic laboratories and the pathways that will be used for this process will be similar to those existing for PKU and MCADD.

**9.      *If the test is for mutations the criteria used to select subset of mutations to be covered by screening should be clearly set out.***

The tests in use are biochemical tests and not tests for mutations.

***The treatment***

**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.***

For newborn screening it is critical to show that treatment established once the diagnosis is made can improve the outcomes in terms of mortality, morbidity and disability. Treatments should be both effective and cost-effective. As these treatments will not, in general, provide a cure for the disease, effectiveness may be viewed in terms of partial recovery and slowing the disease processes and reduction in morbidity and disability. It is particularly important to show that early detection (via screening), if possible before symptoms and clinical signs have appeared, will lead to better outcomes than outcomes for children who are diagnosed clinically. Where conditions (such as MSUD) can present before the screening test is taken (or before the result is known) this does not necessarily preclude screening as it is important to show that the screening test can still lead to a beneficial earlier diagnosis.

In addition the ACMG also considered:

- Availability of treatment
- Cost of treatment
- That treatment might only affect a subset of individuals or may not be equally effective in all patients

- Availability of healthcare professionals with expertise in acute management
- Simplicity of therapy (this determines whether infants are likely to be able to be managed locally or whether sub-specialist care will be required)

One set of evidence involved our understanding of the pathophysiology of the disease and, in particular, the likely damage that will occur by exposure to toxic substances as a result of the impaired metabolic processes, and the ability of treatments to reduce those toxic effects.

**Table 11.5 Treatment and evidence of benefit**

Condition	Treatment regimes agreed	Benefit of early treatment
MSUD	Dietary restriction of amino acids	Yes - outcome relates to duration of leucine exposure
Homocystinuria	Dietary restriction of methionine and supplementation of betaine, B12 and folic acid	Yes - clinical manifestations can be avoided but only in those patients treated from soon after birth
GA1	Low protein lysine restricted diet with supplementation of carnitine. Guidelines available	Whereas outcome is poor in patients diagnosed after acute encephalopathy, most patients remain clinically asymptomatic if diagnosed and treated before these crises 'value of early diagnosis and treatment is not disputed'
IVA	Amino acid restriction and carnitine supplementation (No formal guidelines published)	Rapid detection and intervention in metabolic crises is critical and mis-management lead to permanent neurological damage and learning disability
LCHADD	Fasting avoidance and limiting long chain fatty acid intake plus vitamin and mineral supplementation The use of an agreed emergency regimen to family, primary and secondary care	Identification and early treatment to avoid metabolic compensation is effective in reducing risk of death and long-term complications and morbidity in all except those with complete MTP deficiency or isolated LKAT deficiency



There are already a considerable number of patients under treatment for these disorders and authoritative guidelines are increasingly available. Of these disorders LCHAD deficiency is probably the most complicated to treat. To provide support across the country it is recommended that there should be NHS provision of a formal panel of expert dietitians able to give telephone advice to their colleagues.

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

Agreed treatments are described in Chapter 3 and above for all target conditions, including details of the dietary management, the ways in which this can be titrated against plasma levels of the relevant analytes and also 'Emergency Regimens' that are instituted at times of metabolic stress. Regular and careful monitoring of nutritional adequacy of the diet, biochemical monitoring of the response and the provision of support and advice, (which may at times include alternative methods of feeding such as through gastrostomy or nasogastric feeding) are important aspects of supervision by specialised dietitians.

For most patients with our target conditions there is no substantive doubt that patients with abnormal biochemistry should be offered treatments. However, for two conditions, there are suggestions that there are milder variants that might not warrant a full-scale dietary regimen. Some patients with MSUD have some residual enzyme activity, although they will be picked up by MS/MS, they may have a milder clinical course and some may even be asymptomatic. The degree of dietary management required for each patient cannot be determined by the level of residual enzyme activity as the phenotype is modified by other genes. However the intensity of treatment can be modified according to plasma branched-chain amino acid concentrations. Each patient thus needs individual evaluation to determine severity of disease and necessary intensity of treatment. Some can be managed without dietary treatment but should still receive an emergency regimen for times of metabolic stress.

For the more recently recognised, milder phenotype of IVA (including all of the newborns with a particular recurring mis-sense mutation in the relevant mutant IVD allele) that can be found by newborn screening, there is a recommendation that these individuals should be observed clinically, particularly when exposed to metabolic stressors and low dose carnitine considered if plasma level reduced.

**12. *Clinical management of the condition and patient outcomes should be optimized in all healthcare providers prior to participation in a screening programme.***

Throughout the UK, there are specialised services capable of evaluating and managing patients with inherited metabolic disorders once they are suspected or identified and referred to the service. These services are quite closely linked with each other (e.g. through national and international professional groups to ensure best practice). However, it has to be admitted that there is some lack of capacity and, in some cases, services may lack the full multi-disciplinary team. In these circumstances agreed hub and spoke models of care, particularly for dietetic support, will need to be in place. For MCADD a detailed handbook on dietary management has been developed, a model that could be used in other conditions. In particular, services may not have been as active as they might be in providing outreach to district hospitals, in raising awareness about inherited metabolic conditions and setting

up systems for referral and joint care. This in turn exacerbates the problem of diagnosis in these conditions, which occurs because of the non-specific way in which they present; diagnosis may be delayed for days, months or even years for some of the more insidious onset conditions. Throughout all this time the child is not receiving optimal care.

The process of diagnosis does not begin with specialists, but more likely by district hospital paediatric services (for acute presenting) and community paediatric services (for those presenting with developmental delay *etc.*). It is thus crucially dependent on them having the necessary suspicion of disease and for the necessary testing protocols being in place. Paradoxically, therefore, although the specialised services may not be optimised in terms of overall capacity, it could be argued that this makes a screening programme more, rather than less important, by simplifying the diagnostic process for some of the infants and ensuring that positive screening results are immediately flagged up by the specialised services.

### ***The screening programme***

- 13. *There should be evidence from high-quality Randomised Controlled Trials that the screening programme is effective in reducing mortality or morbidity. Where screening is aimed solely at providing information to allow the person being screened to make an 'informed choice' (e.g. Down's syndrome, cystic fibrosis carrier screening), there must be evidence from high quality trials that the test accurately measures risk. The information that is provided about the test and its outcome must be of value and readily understood by the individual being screened.***

The main purpose of the test is to improve morbidity and mortality in the infant screened and so the second requirement in this criterion relating to 'solely providing information' does not apply. As noted in Chapter 4, it is impossible to obtain evidence of effectiveness of screening programmes from randomized trials because of the rarity, complexity and heterogeneity of the conditions. Not only will there be insufficient patients to generate the necessary statistical power, but also there would be significant ethical considerations in allocating patients to a non-screening group in the light of rapid advances in dietary and other aspects of management.

In Chapter 7 we provide evidence from the 'next best' study designs (observational cohorts) in which screened and clinically detected cohorts are compared in nearby geographic areas with similar services or sequentially with groups closely related in time before and after a screening programme is put in place. Even such comparisons are prone to biases, which some of the studies have attempted to minimise.



**Table 11.6 Evidence of effectiveness of screening programmes**

Wilcken <i>et al.</i> <sup>41;60</sup> Australia	General group including target conditions	Patients diagnosed by screening have reduced mortality and less significant intellectual or physical handicap than those presenting clinically
Waisbren <i>et al.</i> <sup>61;116</sup> New England	General group including target conditions	Hospitalisation levels similar but clinically detected group had higher rates of mental retardation and higher parental stress
Spiekerkoetter <i>et al.</i> <sup>65</sup> European centres	LCHADD	Screen detected patients had reduced mortality and some evidence of reduction in morbidity
Kolker <i>et al.</i> <sup>50</sup> Germany	GA1	Reduced mortality and encephalopathic crises in screen detected cases. Screen detected cases had no delay in developmental milestones and no effect on growth or maturation
Strauss <i>et al.</i> <sup>62</sup> Pennsylvania Amish and other populations	GA1	Reduced risk of brain injury in those diagnosed before symptoms present

**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.***

This criterion requires that we consider the range of relevant stakeholders in a potential expanded newborn screening programme and seek evidence about the clinical, social and ethical acceptability of otherwise of the programme. Relevant stakeholders might include:

- Patients, parents and other family members affected by these conditions
- Voluntary organisations and relevant patient groups concerned with these disorders (including umbrella organisations)
- Health professionals and others involved in provision of other newborn screening programmes
- Laboratory scientists involved in provision of diagnostic IMD services who would potentially be involved in the provision of a screening programme
- Clinicians involved in provision of specialised IMD services
- Dieticians involved in provision of specialised IMD services
- Providers of maternity services who would be involved in explaining expanded programme and gaining consent for testing

- Practitioners involved in community child health services who might be involved in assisting parents concerned about positive results, obtaining repeat tests, and communicating diagnosis to parents and providing or assisting in providing initial clinical support
- General practitioners involved in long-term care of the family

In the UK much of the pressure to expand the newborn screening programme has come from patient and voluntary organisations (see Chapter 8) and from professionals closely associated with management of these conditions (see documentation of meeting held by the NSC with professionals in June 2008). We have less evidence from providers in maternity services, community paediatrics and primary care services; this represents information that could usefully be obtained from the pilot programme.

### **What sort of clinical, social and ethical acceptability would be included in these conditions?**

Clinical issues with a wider group of professionals might include considerations of:

- Whether all professionals involved in the programme or affected by it have the necessary educational, information, time and other resources to provide this screening programme or feel 'out of their depth' in this area
- Whether a disproportionate amount of clinical time is taken up by the programme in relation to the perceived outcomes - e.g. providing education and support for those administering screening tests, organising diagnostic work-up and initial management for screen positives, supporting eventual false positives, providing care for diagnosed children and their families. It may be perceived that the expanded programme is 'more bother than it is worth'
- How the programme has affected relationships between specialised services and the periphery (e.g. the programme might enhance awareness and communication between specialists and the periphery across the range of IMDs)
- Whether health professionals consider that the services is appropriately prioritised when compared to other areas of need
- Whether this has a good or bad effect on diagnosis of other IMDs (e.g. more awareness, or complacency that diseases have been tested for)

Social aspects might include:

- Whether the need to provide information and gain informed consent from parents prior to testing provokes unacceptable stress and further medicalisation of pregnancy and the newborn period
- Whether the number of false positives and the level and length of uncertainty created for them and their families is acceptable
- Whether professionals and other stakeholders feel that this programme fits in with the cultural norms and expectations of the community. This might be particularly important to explore in a range of ethnic and faith communities
- Whether health communities such as PCTs agree that expanded newborn screening would

be a service to prioritise

- Enhanced public awareness of these metabolic disorders and rare disorders in general

Ethical issues are discussed in detail in Chapter 9 and include:

- Difficulties in achieving informed consent made worse by the number of conditions added, their complexity and individual variability
- Implications for family members and how these would be handled, particularly with regard to further reproductive choice and possible cascade testing following positive diagnosis
- False positive results - the problem of doing harm to some individuals
- Lack of gold standard results from RCT before implementing the expanded programme
- Possible identification of disease carriers (should not happen with the protocols envisaged)
- Finding misattributed paternity (fully covered in blood spot programme)
- Governance of material including security of storage, disposal of samples and use for secondary purposes

In considering ethical, legal and social issues with regard to expanding the programme to include a further five conditions, the question should be asked whether these conditions raise any new issues that have not already been addressed with regard to PKU and MCADD and other neonatal screening programmes. Alternatively, issues arising from an expansion of the programme may just be a question of scale, complexity, the relative balance of benefit to harm and the balance of resource input to outcome. In particular, the rarity of these disorders compared with MCADD and PKU might mean a smaller return for resource input, and a lower positive predictive value of positive screens, meaning that the balance of false and true positives was less favourable. Similarly, it is not known whether the addition of a further five conditions might be five times as burdensome, or, with appropriate support, just add a little to the work of each health professional. All of these questions could be addressed through a pilot study and through the continuing information coming from the MCADD programme.

**15. *The benefit from the screening programme should outweigh the physical and psychological harm (caused by the test, diagnostic procedures and treatment).***

A number of benefits could flow from expanded newborn screening. The principle benefits of the screening programme are those to the individual with the condition, who is diagnosed and treated earlier and with better outcomes. These health benefits also accrue for family members, who may also receive a disease diagnosis and be able to commence treatment, or by parents or other family members who might be able to receive counselling about risk to offspring and opportunities to avoid the birth of another affected child. In addition to health benefits are the benefits to families of not having to care for a very unwell or disabled child over long periods. It could be anticipated that there would be benefits to society of a reduced burden on medical, social and educational services. There may also be a widespread raised awareness of rare metabolic disorders, which, if properly managed could ensure that a wider group of patients is diagnosed and managed effectively and cared for with understanding in the community. For health services there may be streamlining of diagnostic processes, better

targeting of specialised services on those with diagnosed conditions, raised awareness of rare conditions with other paediatric and community services, more predictable care focused on long-term surveillance and care rather than crisis management. Finally there is the benefit of further research in the epidemiology, diagnosis and management of these conditions through the systematic study of patients diagnosed through a screening programme and clinically.

Possible harms are discussed in Chapter 9 and a background paper for the proposed expanded newborn screening pilot on parental views on expanded screening<sup>108</sup>. Harms are mainly accrued by the large majority (several hundred thousand each year) of parents who need to decide about expanded screening test and go through the testing process and the smaller number who have screen positive results. They include: anxiety caused by decision making in whether or not to have their infant screened; lack of information, confusion about the information or difficulty in understanding information about the test and its purpose; any follow-up testing; concern during the waiting period for results and uncertainty caused by lateness of results or any confusion in communication; heightened concern, hospital visits and clinical assessment for screen positives; depending on the condition, institution of dietary therapy and/or cessation of breast feeding or use of an emergency regime for screen positive infants whilst a definitive diagnosis is achieved; for a very small number of patients, possibly the identification of rare variants or uncertain biomarker levels with uncertainty of the utility of treatment and possibly unnecessary treatment.

The scale, range and quality of these benefits and possible harms cannot be described or quantified at present on the basis of published evidence, including reports of experience from the recent MCADD pilot study. However, a well constructed prospective evaluation of a pilot study could elucidate some of these areas. Moreover, it is likely that benefits could be enhanced through attention to design of the programme - for example, an awareness of the need to involve the wider health community in the programme and enhance relationships with specialist services might be built in to the project objectives. At the same time some of the harms associated with false positives could be minimised by better education of the public about the programme, personal communication of any positive results and expert support during diagnostic work-up; continuing support to parents of children with false positive results to ensure anxiety about the child's health is resolved. Most importantly, the number of false positives could be minimised by careful setting of laboratory cut-offs as detailed in Chapter 7.

- 16. *The opportunity cost of the screening programme (including testing, diagnosis and treatment, administration, training and quality assurance) should be economically balanced in relation to expenditure on medical care as a whole (i.e. value for money). Assessment against this criterion should have regard to evidence from cost benefit and/or cost effectiveness analyses and have regard to the effective use of available resource.***

The opportunity costs of an expanded screening programme include not only the price of the test but the costs of parental education, follow up of all positives to a definitive diagnosis, treatment of affected children, and ongoing data collection and evaluation. These inevitably draw scarce resources away from other public health programmes and needs. Whilst this might be quantified in general terms, at the level of individual clinicians and services it is inevitable that time spent on screening cannot be spent on other aspects of care. This, more qualitative aspect of service provision could be explored through the pilot by an in depth study of the day to day effect of dealing with the programme from maternity services, health visitors, general practitioners, screening coordinators and others.

Evidence was presented in Chapter 7 from six economic studies showing that MS/MS was cost-effective when judged relative to nationally accepted standards. In general, screening for multiple conditions is more efficient than screening for individual conditions. The evidence from these studies was supportive of all target conditions although somewhat equivocal about homocystinuria. However, it should be noted that in the 2004 UK HTA review, where a higher incidence of homocystinuria more relevant to the UK population was assumed, homocystinuria was ranked in 4th place in a table of priority for extended newborn screening using MS/MS<sup>1</sup>.

- 17. *All other options for managing the condition should have been considered (for example, improving treatment and providing other services), to ensure that no more cost-effective intervention could be introduced or current interventions increased within the resources available.***

For all of the target conditions there is general agreement that diagnosis based on clinical symptoms will always be difficult because of the non-specific presentations of all of these conditions and will always be associated with a poorer outcome. Clearly encouraging earlier clinical diagnosis would also be a possibility, but for most of these conditions, the child may die or suffer irreversible damage before the diagnosis is made. Although an alternative diagnostic method might be 'high-risk' testing, where MS/MS is run on specimens from infants presenting in relevant clinical ways, experts agree that results of testing would be more difficult to interpret because of a lack of standard cut-offs for infants at various ages who are unwell with varying underlying clinical conditions causing biochemical abnormality. The body of evidence from research, and clinical opinion is that screening is the only way to make a timely diagnosis.

- 18. *There should be a plan for managing and monitoring the screening programme and an agreed set of quality assurance standards.***

At present neither a detailed plan for managing and monitoring an expanded screening programme nor a set of quality assurance standards exists. However, one of the purposes of a pilot phase would be to develop such a plan. It is envisaged that the laboratory guide for newborn screening in the UK for MCADD would provide an excellent template to which additional conditions could be added. This guide covers:

- Protocol for testing; who should be tested, testing of siblings and late testing
- Specimen requirements and factors affecting screening results, particularly those associated with false negatives and false positives
- Laboratory analysis of relevant metabolites, expected internal quality control and performance criteria
- Laboratory protocol for full diagnostic work-up
- Protocol for clinical follow up of presumptive positives
- Reporting to child record departments with onward communication of negative results to parents by health visitor and communication of positive results by local trusts
- Published laboratory and clinical guidelines and standards

**19. Adequate staffing and facilities for testing, diagnosis, treatment, and programme management should be available prior to the commencement of the screening programme.**

At present, involved laboratory and clinical services consider the expanded newborn screening could be delivered as an extension of the current screening service and substantially within the current resource envelope of laboratory and clinical capacity and only marginal increases in administration. Indeed, it can be argued that the reduced burden of critically ill infants presenting clinically may benefit the services to some extent. However, this assertion should be tested as part of a pilot study.

**20. Evidence based information explaining the consequences of testing, investigation and treatment, should be made available to potential participants to assist them in making an informed choice.**

Resources similar to those provided for the MCADD programme would need to be available prior to commencement of a programme and could be developed as part of a pilot project.

**21. Public pressure for widening the eligibility criteria for reducing the screening interval, and for increasing the sensitivity of the testing process, should be anticipated. Decisions about these parameters should be scientifically justifiable to the public.**

The issues about widening eligibility criteria and screening interval are not relevant in the context of newborn screening. Pressure to increase the sensitivity of testing arises because of missed cases. The number of missed cases is likely to be extremely small. However, test performance must be monitored strictly to create the optimum balance of false negatives and false positives.

**22. If screening is for a mutation the programme should be acceptable to people identified as carriers and to other family members.**

Screening is not for a mutation.

### 11.3 Gaps in knowledge

Consideration of the criteria discussed above and the findings of the systematic review have shown that there are gaps in our knowledge, that ought to be filled before a decision on expanded newborn screening could be made. These have been discussed fully in the relevant chapters and are listed here as areas that could be pursued through a pilot programme.

#### **Recommendations for the pilot**

##### *Epidemiology*

The numbers of screen detected and clinically detected cases should be ascertained along with basic demographic details including age at diagnosis, sex, ethnic background of parents, history of consanguinity, and any family history of disease. *Comment: in a limited pilot study this will not provide definitive knowledge on the epidemiology of the conditions, but will*



*begin to set a basis for possible future epidemiological studies of these conditions within the UK.*

#### *Clinical validity of MS/MS screening tests*

Participating laboratories must collectively devise and evaluate their tests with respect to analytic validity and clinical validity - in particular to maximise test sensitivity whilst minimising false positives.

For each condition this should result in flow-charts that show, for each condition:

- initial cut-offs
- cut-offs for any repeat testing/ or for urgent assessment (depending on condition)
- cut-offs for further sampling request and the further testing undertaking
- diagnostic cut-offs
- expected 'flow' of infants through the various branches

#### *Clinical utility*

Cases detected by newborn screening should be included on a register and followed up. Diagnosis should be recorded with detail of underlying genetic and biochemical abnormalities and presenting clinical features. Patients should be followed up with details of treatment provided, centre of treatment and description of clinical progress including acute crises and outcomes in terms of morbidity and disability. Parallel active surveillance through the UK laboratories for cases of these conditions diagnosed clinically should be put in place and these cases should be followed up in a similar way.

#### *Possible harm from false positives*

The pilot study needs to look carefully at the numbers of false positives and their pathway from flagged test to final negative diagnosis. It also needs to devise means of minimising stress caused by improving education in antenatal period and around the time of screening and support in the event of a positive result. Educational support will be needed for parents, health professionals and the general public.

#### *Possible harm from over diagnosis*

This is not thought to be a major issue for the five target conditions. However, the pilot study should include clear documentation of biochemistry and clinical assessment of every case diagnosed. It should consider some form of external assessment to confirm whether treatment was strictly necessary for each case and record the reasons behind this.

#### *Economic analysis*

The pilot study needs to quantify the extra costs needed to expand the screening programme to include these conditions. Costs falling on laboratories, specialist clinical, paediatric, community services and primary care need to be estimated.



### *Availability of specialist care*

The pilot study needs to assess whether there is sufficient capacity in the specialist system to undertake the necessary diagnostic, clinical assessment and follow-up of patients identified through the screening system. A comparison of clinical input for screen detected versus clinically detected cases would be useful.

### *Organisational aspects*

The pilot programme should consider in detail the organisational requirements including the ways in which laboratory and clinical elements must be integrated in order to streamline the identification and follow up of those who screen positive.

### *Guidelines and treatment protocols*

The pilot programme should result in a set of guidelines for laboratory and clinical assessment of screen positive patients and for clinical management and follow-up of patients, where possible depending on initial genetic, biochemical and clinical profile.

### *Clinical outcome*

The pilot study should begin the process of tracking health and other outcomes for patients and families and should lead to the development of a system and agreed outcome measures.

### *Wider benefits and harms*

The pilot study should provide evidence on wider benefits and harms including those to parents and extended family of cases, false positives and their families, health services, researchers and society in general.

Where possible the pilot programme should develop and publish operational protocols and resources that will maximise benefit and minimise harm.

### *Clinical, social and ethical acceptability*

The pilot programme should work through a group of stakeholders to collect evidence on the clinical, social and ethical acceptability of expanded screening.

### *Opportunity cost of screening*

The pilot programme should study qualitative aspects of the opportunity cost by working with relevant health professionals involved in the whole range of the programme.

### *Managing, monitoring and setting quality assurance standards*

A detailed plan, resources for running and monitoring the programme and set of quality assurance standards should be developed as part of a pilot programme.

### *Resources for parents and public*

Evidence based information about the conditions, advantages and disadvantages of testing, process of testing, consequences of the test result and expected follow-up should be developed as part of a pilot project. The MCADD resources could be used as a template.

## **11.4 Conclusion**

There is much evidence to support expanded newborn screening within the UK as a means of preventing death and disability from a target range of inherited metabolic conditions through expansion of the current set of tests. The conditions are all treatable and there is evidence that outcomes are better if infants are diagnosed early and treatment is commenced before any symptoms occur. The strength of evidence supports the view that an expanded 'bundle' of conditions is more cost-effective than restricting testing to only one or two tests, as at present. However, whilst there will be significant benefit to a small number of infants and their families, as well as, in total, to the wider health, social and educational services, the possible harms applied to the entire population of newborns arising from this expanded screening programme must be considered. The most significant harms arise from the diagnostic work-up for those eventually found to be disease free and include those directly impacting on the family (such as anxiety) and 'knock-on' effects of the extra work on laboratory, specialist and general paediatric and primary care services. The evidence suggests that such impacts will be manageable but nevertheless there is a need for them to be investigated and quantified.

### *Newborn screening for MSUD, homocystinuria, GA1, IVA and LCHADD deficiency*

Through this review we have collected and synthesised all the available national and international literature and concluded that the evidence for expanded newborn screening for the five target conditions is favourable.

#### **Recommendation 1**

The NSC should study this report in detail and determine whether or not it agrees with the conclusions. If not, it should set out:

- Any factual points of disagreement, where possible indicating how such disagreements can be reconciled
- The areas where it considers the evidence is not strong enough to support a favourable case for expanded screening in general. To the extent that this evidence is unavailable, it should describe what it considers would constitute suitable evidence, and whether it is reasonably foreseeable that this could be obtained in the UK in the short to medium term

### *An expanded national newborn screening programme*

Whether or not the NSC newborn screening programme should be expanded to include each of these five conditions requires the NSC to weigh up the conditions against their screening criteria. This review has provided interpretation and discussion against each of the criteria. It has concluded that none of the criteria are unfulfilled but that the criteria in very rare genetic conditions may need to be judged differently; there will be trade-offs between

criteria; and judgements about ‘fulfilment’ or otherwise are subjective.

### **Recommendation 2**

The NSC should consider the conditions against each of its screening criteria and decide:

- whether each criterion is met or not
- where there is insufficient evidence for a given criterion/condition, what it would consider to be sufficient evidence that could be collected within the UK in the short to medium term

### ***Developing a pilot programme***

Following its conclusions that expanded newborn screening would improve health outcomes without causing undue harm and that this could best be undertaken through a national newborn screening programme, the review group of this study has concluded that the next step should be a large scale pilot study. The aim of such a study would be to place expanded newborn screening into experimental practice on a sufficiently large scale to allow some of the unanswered questions relevant to the programme to be answered. Central questions include whether or not laboratories can develop tests with optimal performance for screening, the actual cost to laboratory and clinical services, the impact on these services, and public and professional acceptability. A more complete set of questions that relate also to the effectiveness of screening programmes is outlined above. The NIHR CLAHRC in the Sheffield region has provided funding for a pilot programme and it is believed that many of these questions could be addressed within the funding available.

### **Recommendation 3**

The NSC should:

- recommend that a pilot programme should be undertaken to address gaps in our knowledge relevant to the expansion of newborn screening in the UK
- ask the NIHR funded CLAHRC project to conduct the pilot programme
- set out a mechanism for agreeing further evidence requirement and a process for obtaining and judging this evidence
- agree to receive a report from the pilot programme on completion

## References

- (1) Pandor A, Eastham J, Beverley C, Chilcott J, Paisley S. Clinical effectiveness and cost-effectiveness of neonatal screening for inborn errors of metabolism using tandem mass spectrometry: a systematic review. *Health Technology Assessment* 2004; 8(12):iii-NaN.
- (2) Ashcroft AE. An Introduction to Mass Spectrometry (accessed 05.01.2009). <http://www.astbury.leeds.ac.uk/facil/MStut/mstutorial.html>
- (3) Wilcken B, Wiley V. Newborn screening. *Pathology* 2008; 40(2):104-115.
- (4) Sanderson S. The incidence of inherited metabolic disorders in the West Midlands, UK. 2006.
- (5) Chuang D, Wynn R, Shih VE. Maple Syrup Urine Disease (Branched-Chain Ketoaciduria). In: Scriver C, Beaudat A, Sly W, Valle D, editors. *The Metabolic and Molecular Bases of Inherited Disease Vol. II*. 8th ed ed. New York: McGraw-Hill; 2001. 1971-2005.
- (6) Naughten ER, Yap S, Mayne PD. Newborn screening for homocystinuria: Irish and world experience. *Eur J Pediatr* 1998; 157 Suppl 2:S84-S87.
- (7) Morton DH, Strauss KA, Robinson DL, Puffenberger EG, Kelley RI. Diagnosis and treatment of maple syrup disease: a study of 36 patients. *Pediatrics* 2002; 109(6):999-1008.
- (8) Simon E, Flaschker N, Schadowaldt P, Langenbeck U, Wendel U. Variant maple syrup urine disease (MSUD)--the entire spectrum. *J Inherit Metab Dis* 2006; 29(6):716-724.
- (9) Hilliges C, Awiszus D, Wendel U. Intellectual performance of children with maple syrup urine disease. *Eur J Pediatr* 1993; 152(2):144-147.
- (10) Kaplan P, Mazur A, Field M, Berlin JA, Berry GT, Heidenreich R *et al*. Intellectual Outcome in Children with Maple Syrup Urine Disease. *Journal of Pediatrics* 1991; 119(1):46-50.
- (11) Walter JH, Wraith JE, White FJ, Bridge C, Till J. Strategies for the treatment of cystathionine beta-synthase deficiency: the experience of the Willink Biochemical Genetics Unit over the past 30 years. *Eur J Pediatr* 1998; 157 Suppl 2:S71-S76.
- (12) Kolker S, Garbade SF, Greenberg CR, Leonard JV, Saudubray JM, Ribes A *et al*. Natural history, outcome, and treatment efficacy in children and adults with glutaryl-CoA dehydrogenase deficiency. *Pediatric Research* 2006; 59(6):840-847.
- (13) Kolker S, Christensen E, Leonard JV, Greenberg CR, Burlina AB, Burlina AP *et al*. Guideline for the diagnosis and management of glutaryl-CoA dehydrogenase deficiency (glutaric aciduria type I). *Journal of Inherited Metabolic Disease* 2007; 30(1):5-22.
- (14) Vockley J, Ensenauer R. Isovaleric acidemia: New aspects of genetic and phenotypic heterogeneity. *American Journal of Medical Genetics Part C-Seminars in Medical Genetics* 2006; 142C(2):95-103.
- (15) Tanaka K. Isovaleric acidemia: personal history, clinical survey and study of the molecular basis. *Prog Clin Biol Res* 1990; 321:273-290.
- (16) Ensenauer R, Vockley J, Willard JM, Huey JC, Sass JO, Edland SD *et al*. A common mutation is associated with a mild, potentially asymptomatic phenotype in patients with isovaleric acidemia diagnosed by newborn screening. *Am J Hum Genet* 2004; 75(6):1136-1142.

- (17) Gillingham MB, Connor WE, Matern D, Rinaldo P, Burlingame T, Meeuws K *et al.* Optimal dietary therapy of long-chain 3-hydroxyacyl-CoA dehydrogenase deficiency. *Mol Genet Metab* 2003; 79(2):114-123.
- (18) Gillingham MB, Weleber RG, Neuringer M, Connor WE, Mills M, van CS *et al.* Effect of optimal dietary therapy upon visual function in children with long-chain 3-hydroxyacyl CoA dehydrogenase and trifunctional protein deficiency. *Mol Genet Metab* 2005; 86(1-2):124-133.
- (19) Andermann A. Revisiting Wilson and Jungner in the genomic age: a review of screening criteria over the past 40 years. 2008.
- (20) Fingerhut R. Stability of acylcarnitines and free carnitine in dried blood samples: implications for retrospective diagnosis of inborn errors of metabolism and neonatal screening for carnitine transporter deficiency. 2009.
- (21) Pollitt RJ, Green A, McCabe CJ, Booth A, Cooper NJ, Leonard JV *et al.* Neonatal screening for inborn errors of metabolism: cost, yield and outcome. *Health Technol Assess* 1997; 1(7):i-202.
- (22) Seymour CA, Thomason MJ, Chalmers RA, Addison GM, Bain MD, Cockburn F *et al.* Newborn screening for inborn errors of metabolism: a systematic review. *Health Technol Assess* 1997; 1(11):i-95.
- (23) Autti-Ramo I, Laajalahti L, Koskinen H, Sintonen H, Makela M. [Screening for rare metabolic disease in newborn infants] Vastasyntyneiden harvinaisten aineenvaihduntatautien seulonta FinOHTAn raportti 22. *STAKES / FinOHTA* 2004.
- (24) Health Council of the Netherlands. Neonatal Screening. *The Hague: Health Council of the Netherlands* 2005; publication no. 2005/11E.
- (25) Paz Valinas L, Atienza Merino G. Clinical effectiveness of neonatal screening for inborn errors of metabolism using tandem mass spectrometry. Systematic Review. *Santiago de ComPostela: Galician Agency for Health Technology Assessment (AVALIA-T)* 2007.
- (26) Danish National Board of Health. Biokemisk screening for medfødt sygdom hos nyfødte - Principper, erfaringer, anbefalinger. *Rapport fra en arbejdsgruppe nedsat af Sundhedsstyrelsen* 2008.
- (27) Medical Advisory Secretariat. Neonatal screening of inborn errors of metabolism using tandem mass spectrometry: an evidence-based analysis. *Ontario Health Technology Assessment Series* 2003; 3(3).
- (28) Vallance H, Green C, Stockler S, Sirrs S, Davidson G, Lillquist Y *et al.* What is the evidence that expanding the current newborn screening panel will provide clinical benefit to newborns? *Health Technology Assessment Report prepared for The Newborn Screening Advisory Committee* 2008.
- (29) Watson MS., Lloyd-Puryear MA, Mann MY, Rinaldo P, Howell R. Newborn Screening: Toward a Uniform Screening Panel and System. *Genetics in Medicine* 2006; Suppl 1:1S-252S.
- (30) Ferrante D, Garcia MS, Glujovsky D, Lopez A, Regueiro A. Efficacy of expanded newborn screening in postnatal testing. *Ciudad de Buenos Aires: Institute for Clinical Effectiveness and Health Policy (IECS)* 2005.
- (31) Schulze A, Lindner M, Kohlmüller D, Olgemöller K, Mayatepek E, Hoffmann GF. Expanded Newborn Screening for Inborn Errors of Metabolism by Electrospray Ionization-Tandem Mass Spectrometry: Results, Outcome, and Implications. *Pediatrics* 2003; 111(6):1399-1406.

- (32) La Marca G, Malvagia S, Casetta B, Pasquini E, Donati MA, Zammarchi E. 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.
- (33) Hoffmann GF, von KR, Klose D, Lindner M, Schulze A, Muntau AC *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.
- (34) Lund AM. Neonatal metabolic screening in Denmark. 101.Jahrestagung der Deutschen Gesellschaft für Kinder- und Jugendmedizin, Bremen.Abstract K058.02V Available at <http://www.mh-hannover.de/tagungen/abs/dgkj2005/web/abstract/0190.htm>; accessed 23/09/2009.
- (35) Vilarinho L, Rocha H, Sousa C, Marcao A, Fonseca H, Bogas M *et al.* Four years of expanded newborn screening in Portugal with MS/MS. *J Inherit Metab Dis* 2010; in press.
- (36) Frazier D, Millington D, McCandless S, Koeberl D, Weavil S, Chaing S *et al.* The tandem mass spectrometry newborn screening experience in North Carolina: 1997-2005. *Journal of Inherited Metabolic Disease* 2006; 29(1):76-85.
- (37) Matern D, Tortorelli S, Oglesbee D, Gavrilov D, Rinaldo P. 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):585-592.
- (38) Comeau AM, Larson C, Eaton RB. Integration of new genetic diseases into statewide newborn screening: New England experience. *American Journal of Medical Genetics Part C-Seminars in Medical Genetics* 2004; 125C(1):35-41.
- (39) Feuchtbaum L, Lorey F, Faulkner L, Sherwin J, Currier R, Bhandal A *et al.* California's experience implementing a pilot newborn supplemental screening program using tandem mass spectrometry. *Pediatrics* 2006; 117(5 Pt 2):S261-S269.
- (40) Torres-Sepulveda MR, Martinez-de Villarreal LE, Esmer C, Gonzalez-Alanis R, Ruiz-Herrera C, Sanchez-Pena A *et al.* [Expand newborn screening using tandem mass spectrometry: two years' experience in Nuevo Leon, Mexico]. *Salud Publica Mex* 2008; 50(3):200-206.
- (41) Wilcken B, Wiley V, Hammond J, Carpenter K. Screening Newborns for Inborn Errors of Metabolism by Tandem Mass Spectrometry. *N Engl J Med* 2003; 348(23):2304-2312.
- (42) Boneh A, Beauchamp M, Humphrey M, Watkins J, Peters H, Yapliito-Lee J. Newborn screening for glutaric aciduria type I in Victoria: Treatment and outcome. *Molecular Genetics and Metabolism* 2008; 94(3):287-291.
- (43) Yoon H-R, Lee KR, Kang S, Lee DH, Yoo H-W, Min W-K *et al.* Screening of newborns and high-risk group of children for inborn metabolic disorders using tandem mass spectrometry in South Korea: A three-year report. *Clinica Chimica Acta* 2005; 354(1-2):167-180.
- (44) Abdul Rahman S, Mohd Yunus Z, Yew Sing C, Omar A, Othman NA, Shaharudin AS. Pilot project on neonatal screening of inborn errors of metabolism (IEM) using tandem mass spectrometry in Malaysia. *J Inherit Metab Dis* 2007; 30:4.
- (45) Huang H-P, Chu K-L, Chien Y-H, Wei M-L, Wu S-T, Wang S-F *et al.* Tandem mass neonatal screening in Taiwan - Report from one center. *Journal of the Formosan Medical Association* 2006; 105(11):882-886.



- (46) Hsieh CT, Hwu WL, Huang YT, Huang AC, Wang SF, Hu MH *et al.* Early detection of glutaric aciduria type I by newborn screening in Taiwan. *Journal of the Formosan Medical Association* 2008; 107(2):139-144.
- (47) Lindner M, Abdoh G, Fang-Hoffmann J, Shabeck N, Al Sayrafi M, Al Janahi M *et al.* Implementation of extended neonatal screening and a metabolic unit in the State of Qatar: developing and optimizing strategies in cooperation with the Neonatal Screening Center in Heidelberg. *J Inherit Metab Dis* 2007; 30(4):522-529.
- (48) Wilson C, Kerruish NJ, Wilcken B, Wiltshire E, Webster D. The failure to diagnose inborn errors of metabolism in New Zealand: The case for expanded newborn screening. *New Zealand Medical Journal* 2007; 120(1262).
- (49) Abdel-Hamid M, Tisocki K, Sharaf L, Ramadan D. Development, validation and application of tandem mass spectrometry for screening of inborn metabolic disorders in Kuwaiti infants. *Medical Principles and Practice* 2007; 16(3):215-221.
- (50) Kolker S, Garbade SF, Boy N, Maier EM, Meissner T, Muhlhausen C *et al.* Decline of acute encephalopathic crises in children with Glutaryl-CoA dehydrogenase deficiency identified by newborn screening in Germany. *Pediatric Research* 2007; 62(3):357-363.
- (51) Sander J, Sander S, Steuerwald U, Janzen N, Peter M, Wanders RJ *et al.* Neonatal screening for defects of the mitochondrial trifunctional protein. *Mol Genet Metab* 2005; 85(2):108-114.
- (52) Dionisi-Vici C, Deodato F, Roschinger W, Rhead W, Wilcken B. "Classical" organic acidurias, propionic aciduria, methylmalonic aciduria and isovaleric aciduria: Long-term outcome and effects of expanded newborn screening using tandem mass spectrometry. *Journal of Inherited Metabolic Disease* 2006; 29(2):383-389.
- (53) Knerr I, Weigel C, Rauh C, Trefz FK, Scheible D, Haase C *et al.* Evaluation and follow-up care of children with benign metabolic phenotypes of inborn errors of metabolism, fortuitously detected in newborn screening programs using tandem mass spectrometry. *J Inherit Metab Dis* 2005; 28(Supplement 1):12.
- (54) Bhattacharya K, Khalili V, Wiley V, Carpenter K, Wilcken B. Newborn screening may fail to identify intermediate forms of maple syrup urine disease. *J Inherit Metab Dis* 2006; 29(4):586.
- (55) Lindner M, Kolker S, Schulze A, Christensen E, Greenberg CR, Hoffmann GF. Neonatal screening for glutaryl-CoA dehydrogenase deficiency. *J Inherit Metab Dis* 2004; 27(6):851-859.
- (56) Bijarnia S, Wiley V, Carpenter K, Christodoulou J, Ellaway C, Wilcken B. Glutaric aciduria type I: outcome following detection by newborn screening. *Journal of Inherited Metabolic Disease* 2008; 31(4):503-507.
- (57) Yoon HR, Lee KR, Kim H, Kang S, Ha Y, Lee DH. Tandem mass spectrometric analysis for disorders in amino, organic and fatty acid metabolism: two year experience in South Korea. *Southeast Asian J Trop Med Public Health* 2003; 34 Suppl 3:115-120.
- (58) Fingerhut R, Simon E, Maier EM, Hennermann JB, Wendel U. Maple syrup urine disease: newborn screening fails to discriminate between classic and variant forms. *Clin Chem* 2008; 54(10):1739-1741.



- (59) Wilcken B. Mini-Symposium: Newborn screening for inborn errors of metabolism - Clinical effectiveness. *Journal of Inherited Metabolic Disease* 2006; 29(2-3):366-369.
- (60) Wilcken B, Haas M, Joy P, Wiley V, Bowling F, Carpenter K *et al.* Expanded newborn screening: outcome in screened and unscreened patients at age 6 years. *Pediatrics* 2009; 124(2):e241-e248.
- (61) Waisbren SE, Albers S, Amato S, Ampola M, Brewster TG, Demmer L *et al.* Effect of expanded newborn screening for biochemical genetic disorders on child outcomes and parental stress. *JAMA: Journal of the American Medical Association* 2003; 290(19):2564-2573.
- (62) Strauss KA, Puffenberger EG, Robinson DL, Morton DH. Type I glutaric aciduria, part 1: natural history of 77 patients. *Am J Med Genet C Semin Med Genet* 2003; 121C(1):38-52.
- (63) Ensenauer R, Grunert S, Willard JM. Natural History of Isovaleric Acidemia. *Journal of Inherited Metabolic Disease* 2003; 26(Suppl 2).
- (64) Spiekerkoetter U, Sykut-Cegielska J. Prognosis and Treatment of LCHAD deficiency. *Milupa Metabolics* 2007.
- (65) Spiekerkoetter U, Lindner M, Santer R, Grotzke M, Baumgartner MR, Boehles H *et al.* Management and outcome in 75 individuals with long-chain fatty acid oxidation defects: results from a workshop. *J Inherit Metab Dis* 2009; 32(4):488-497.
- (66) Ibarra-Gonzalez I, Fernandez-Lainez C, Belmont-Martinez L, Vela-Amieva M. [Increased mortality and disability in a cohort of Mexican children with maple syrup urine disease]. *Gac Med Mex* 2007; 143(3):197-201.
- (67) Hoffmann B, Helbling C, Schadewaldt P, Wendel U. Impact of longitudinal plasma leucine levels on the intellectual outcome in patients with classic MSUD. *Pediatr Res* 2006; 59(1):17-20.
- (68) Wilcken B. The consequences of extended newborn screening programmes: Do we know who needs treatment? 2008.
- (69) Gurian EA, Kinnamon DD, Henry JJ, Waisbren SE. Expanded newborn screening for biochemical disorders: the effect of a false-positive result. *Pediatrics* 2006; 117(6):1915-1921.
- (70) Norman R, Haas M, Chaplin M, Joy P, Wilcken B. Economic evaluation of tandem mass spectrometry newborn screening in Australia. *Pediatrics* 2009; 123(2):451-457.
- (71) Schoen EJ, Baker JC, Colby CJ, To TT. Cost-benefit analysis of universal tandem mass spectrometry for newborn screening. *Pediatrics* 2002; 110(4):781-787.
- (72) Autti-Ramo I, Makela M, Sintonen H, Koskinen H, Laajalahti L, Halila R *et al.* Expanding screening for rare metabolic disease in the newborn: An analysis of costs, effect and ethical consequences for decision-making in Finland. *Acta Paediatrica, International Journal of Paediatrics* 2005; 94(8):1126-1136.
- (73) Carroll AE, Downs SM. Comprehensive cost-utility analysis of newborn screening strategies. *Pediatrics* 2006; 117(5 Part 2):S287-S295.
- (74) Cipriano LE, Rupar CA, Zaric GS. The Cost-Effectiveness of Expanding Newborn Screening for up to 21 Inherited Metabolic Disorders Using Tandem Mass Spectrometry: Results from a Decision-Analytic Model. *Value in Health* 2007; 10(2):83-98.

- (75) Insinga RP, Laessig RH, Hoffman GL. Newborn screening with tandem mass spectrometry: examining its cost-effectiveness in the Wisconsin Newborn Screening Panel. *Journal of Pediatrics* 2002; 141(4):524-532.
- (76) Burton H, Sanderson S, Shortland G, Lee P. Needs assessment and review of services for people with inherited metabolic disease in the United Kingdom. *J Inherit Metab Dis* 2006; 29(5):667-676.
- (77) Childress JF, Faden RR, Gaare RD, Gostin O, Kahn J, Bonnie RJ *et al.* Public Health Ethics: Mapping the Terrain. *Journal of Law, Medicine and Ethics* 2002; 30:170-178
- (78) Kass NE. An Ethics Framework for Public Health. *American Journal of Public Health* 2001; 91(11):1776-1782.
- (79) Miller FA. The complex promise of newborn screening. *Indian Journal of Medical Ethics* 2009; VI(3):142-148.
- (80) Potter BK, Avard D, Entwistle V, Kennedy C, Chakraborty P, McGuire M *et al.* Ethical, Legal, and Social Issues in Health Technology Assessment for Prenatal/Preconceptual and Newborn Screening: A Workshop Report. *Public Health Genomics* 2008; 12:4-10.
- (81) The President's Council on Bioethics. The Changing Moral Focus of Newborn Screening: An Ethical Analysis by the President's Council on Bioethics. *The President's Council of Bioethics* 2008.
- (82) Sorenson JR, Levy HL, Mangione TW, Sepe SJ. Parental Response to Repeat Testing of Infants with 'False-Positive' Results in a Newborn Screening Program. *Pediatrics* 1984; 73(2):183-187.
- (83) Svenger T, Thelin T. Four-year old children with alpha 1-antitrypsin deficiency. Clinical follow-up and parental attitudes towards neonatal screening. *Acta Paediatr Scand* 1981; 70:171-177.
- (84) Newson A. Should parental refusals of newborn screening be respected? *Camb Q Healthc Ethics* 2006; 15(2):135-146.
- (85) St-Jaques S, Grenier S, Forest J, Rousseau F, Légaré F. Decisional needs assessment regarding Down syndrome prenatal testing: a systematic review of the perceptions of women, their partners and health professionals. *Prenat Diagn* 2008; 28:1183-1203.
- (86) Green JM, Hewison J, Bekker HL, Bryant LD, Cuckle HS. Psychosocial aspects of genetic screening of pregnant women and newborns: a systematic review. *Health Technol Assess* 2004; 8(33):iii, ix-iii, 109.
- (87) Boivin J, Lancaster D. Medical waiting periods: imminence, emotions and coping. *Future Medicine* 2010; 6(1):59-69.
- (88) Wilcken B. Ethical issues in newborn screening and the impact of new technologies. *Eur J Paediatr* 2003; 162 Suppl 1:S62-S66.
- (89) Oliver S, Dezateux C, Kavanagh J, Lempert T, Stewart R. Disclosing to parents newborn carrier status identified by routine blood spot screening. *Cochrane Database Syst Rev* 2004;(4):CD003859.
- (90) Sharma A, Phadke S, Agarwal S. Sociocultural and ethical dilemmas of genetic counselling: a suggested working approach. *J Genet Couns* 1994; 3(1):81-83.

- (91) Liebl B, Nennstiel-Ratzel U, von KR, Fingerhut R, Olgemoller B, Zapf A *et al.* Very high compliance in an expanded MS-MS-based newborn screening program despite written parental consent. *Prev Med* 2002; 34(2):127-131.
- (92) Laing IA, McIntosh N. Practicalities of consent. *Lancet* 2004; 364(9435):659.
- (93) Bombard Y, Miller FA, Hayeems RZ, Avar D, Knoppers BM, Cornel MC *et al.* The Expansion of Newborn Screening: Is Reproductive Benefit An Appropriate Pursuit. *Nature Reviews Genetics* 2009; 10(October 2009).
- (94) Helgesson G, Swartling U. Views on data use, confidentiality and consent in a predictive screening involving children. *J Med Ethics* 2008; 34:206-209.
- (95) Data Protection Act. UK Parliament Schedule 1 1998.
- (96) UK Newborn Screening Programme Centre. Code of Practice for the Retention and Storage of Residual Blood Spots. *The Programme Centre* 2005.
- (97) UK Parliament. Human Tissue Act. *Chapter 30* 2004.
- (98) Stewart F. Ethics of newborn screening. *Current Paediatrics* 2006; 16(3):216-220.
- (99) UK Newborn Screening Programme Centre. Standards and Guidelines for Newborn Blood Spot Screening. *The Programme Centre* 2008.
- (100) UK Newborn Screening Programme Centre. Why we need a newborn screening programme. <http://newbornbloodspot.screening.nhs.uk/public> date accessed 25.01.2010.
- (101) UK Newborn Screening Programme Centre. Newborn blood spot screening, standard operating procedures for child health records departments. NHS Antenatal and Newborn Screening Programmes 2009.
- (102) *S. and Marper v The United Kingdom (2008)*. ECHR <http://www.bailii.org/eu/cases/ECHR/2008/1581.html> date accessed 22.01.2010
- (103) Genewatch UK. Keeping blood spots from newborn babies [http://www.genewatch.org/uploads/f03c6d66a9b354535738483c1c3d49e4/Blood\\_spots\\_from\\_newborn\\_babies.doc](http://www.genewatch.org/uploads/f03c6d66a9b354535738483c1c3d49e4/Blood_spots_from_newborn_babies.doc) 2006.
- (104) UK Parliament. Human Rights Act. Chapter 42, Article 8 1998.
- (105) Health Insurance Portability and Accountability Act. Public Law 104-191 104th Congress 1996.
- (106) Daniels N. Accountability for reasonableness. *BMJ* 2000;1300-1301.
- (107) Daniels N, Sabin J. Moving forward on rationing: an ethical perspective. *BMJ* 2008; 337:904-905.
- (108) Moody L, Pottinger E, Choudhry K. Parental views on the provision of informed consent for expanded newborn screening. *Applied Research Centre Health and Lifestyle Interventions, Coventry University* 2009.
- (109) European Organisation for Rare Diseases (EURORDIS). Rare diseases: understanding this public health priority. [http://www.eurordis.org/IMG/pdf/princeps\\_document-EN.pdf](http://www.eurordis.org/IMG/pdf/princeps_document-EN.pdf) 2005; Retrieved 22 December 2009

- (110) Decision No 1295/1999/EC of the European Parliament and of the Council of 29 April 1999. Community action programme on rare diseases. *Official Journal of the European Communities* 1999; [http://eurlex.europa.eu/pri/en/oj/dat/1999/L\\_155/L\\_15519990622en00010005.pdf](http://eurlex.europa.eu/pri/en/oj/dat/1999/L_155/L_15519990622en00010005.pdf).
- (111) Decision No 1350/2007/EC of the European Parliament and of the Council of 23 October 2007. *Programme of Community Action in the Field of Health (2008-13)*. *Official Journal of the European Communities* 2007; <http://eurlex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2007:301:0003:0013:EN:PDF>
- (112) Commission of the European Communities. Communication from the Commission to the European Parliament, The Council, The European Economic and Social Committee and the Committee of the Regions on Rare Diseases: Europe's Challenges. *COM(2008) 679 final* 2008; [http://ec.europa.eu/health/ph\\_threats/non\\_com/docs/rare\\_com\\_en.pdf](http://ec.europa.eu/health/ph_threats/non_com/docs/rare_com_en.pdf).
- (113) The Council of the European Union. Council Recommendation of 8 June 2009 on an action in the field of rare diseases . *Official Journal of the European Communities* 2009; <http://eurlex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:C:2009:151:0007:0010:EN:PDF>
- (114) Bunday S, Alam H. A five-year prospective study of the health of children in different ethnic groups, with particular reference to the effect of inbreeding. *Eur J Hum Genet* 1993; 1(3):206-219.
- (115) Kai J. Communication of carrier status information following universal newborn screening for sickle cell disorders and cystic fibrosis: qualitative study of experience and practice. 2009.
- (116) Waisbren SE, Read CY, Ampola M, Brewster TG, Demmer L, Greenstein R *et al*. Newborn screening compared to clinical identification of biochemical genetic disorders. *J Inherit Metab Dis* 2002; 25(7):599-600.

## **Appendix 1     UK National Screening Committee criteria (updated June 2009)**

### **The Condition**

1. The condition should be an important health problem
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.
3. All the cost-effective primary prevention interventions should have been implemented as far as practicable.
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.

### **The Test**

5. There should be a simple, safe, precise and validated screening test.
6. The distribution of test values in the target population should be known and a suitable cut-off level defined and agreed.
7. The test should be acceptable to the population.
8. There should be an agreed policy on the further diagnostic investigation of individuals with a positive test result and on the choices available to those individuals.
9. If the test is for mutations the criteria used to select the subset of mutations to be covered by screening, if all possible mutations are not being tested, should be clearly set out.

### **The Treatment**

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.
11. There should be agreed evidence based policies covering which individuals should be offered treatment and the appropriate treatment to be offered.
12. Clinical management of the condition and patient outcomes should be optimised in all health care providers prior to participation in a screening programme.

## The Screening Programme

13. There should be evidence from high quality Randomised Controlled Trials that the screening programme is effective in reducing mortality or morbidity. Where screening is aimed solely at providing information to allow the person being screened to make an “informed choice” (e.g. Down’s syndrome, cystic fibrosis carrier screening), there must be evidence from high quality trials that the test accurately measures risk. The information that is provided about the test and its outcome must be of value and readily understood by the individual being screened.
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.
15. The benefit from the screening programme should outweigh the physical and psychological harm (caused by the test, diagnostic procedures and treatment).
16. The opportunity cost of the screening programme (including testing, diagnosis and treatment, administration, training and quality assurance) should be economically balanced in relation to expenditure on medical care as a whole (*i.e.* value for money). Assessment against this criterion should have regard to evidence from cost benefit and/or cost effectiveness analyses and have regard to the effective use of available resource.
17. All other options for managing the condition should have been considered (e.g. improving treatment, providing other services), to ensure that no more cost effective intervention could be introduced or current interventions increased within the resources available.
18. There should be a plan for managing and monitoring the screening programme and an agreed set of quality assurance standards.
19. Adequate staffing and facilities for testing, diagnosis, treatment and programme management should be available prior to the commencement of the screening programme.
20. Evidence-based information, explaining the consequences of testing, investigation and treatment, should be made available to potential participants to assist them in making an informed choice.
21. Public pressure for widening the eligibility criteria for reducing the screening interval, and for increasing the sensitivity of the testing process, should be anticipated. Decisions about these parameters should be scientifically justifiable to the public.
22. If screening is for a mutation the programme should be acceptable to people identified as carriers and to other family members.

## Appendix 2 Search strategies used in the major electronic bibliographic databases

**Stage 1** Identification of all articles relevant to the evaluation of MS/MS screening programmes

<p><b>CDSR and CCTR</b> Limited 2002 to 2009 year of publication</p> <p>#1 NEONATAL SCREENING*:ME  #2 (NEONAT* NEAR/L SCREEN*)  #3 (NEWBORN* NEAR/L SCREEN*)  #4 (#1 OR #2 OR #3)  #5 MASS-SCREEN*:ME  #6 INFANT-NEWBORN*:ME  #7 (#5 AND #6)  #8 (#4 OR #7)  #9 METABOLISM-INBORN ERRORS*:ME  #10 ((INBORN NEAR/L ERROR*) NEAR METABOLISM)  #11 (#9 OR #10)  #12 (#8 AND #11)</p>	<p><b>CINAHL Database search VIA NHS Evidence website</b> <a href="http://www.library.nhs.uk/booksandjournals/Default.aspx">http://www.library.nhs.uk/booksandjournals/Default.aspx</a>  Limited 2002 to 2009 year of publication</p> <p>1 exp health screening/  2 exp infant, newborn/  3 1 AND 2  4 neonat* ADJ2 screen*  5 newborn* ADJ2 screen*  6 3 OR 4 OR 5  7 exp metabolism, inborn errors/  8 inborn ADJ2 error*  9 7 OR 8  10 6 AND 9  11 spectrum analysis/  12 mass ADJ2 spect*  13 ms ADJ2 spect*  14 tandem ADJ2 mass  15 11 OR 12 OR 13 OR 14  16 10 AND 15</p>
---	--



<p><b>CRD database search (NHS DARE, EED, HTA) via website <a href="http://www.crd.york.ac.uk/crdweb/">http://www.crd.york.ac.uk/crdweb/</a></b></p> <p>Limited 2002 to 2009 year of publication</p> <p>((neonat* AND screen*) OR (newborn AND screen*)) AND ((mass AND spect*) OR (ms AND spect*) OR (tandem AND spect*))</p>	<p><b>EMBASE Database search VIA NHS Evidence website <a href="http://www.library.nhs.uk/booksandjournals/Default.aspx">http://www.library.nhs.uk/booksandjournals/Default.aspx</a></b></p> <p>Limited 2002 to current year of publication</p> <ol style="list-style-type: none"> <li>1 exp newborn screening/</li> <li>2 neonat* ADJ2 screen*</li> <li>3 newborn* ADJ2 screen*</li> <li>4 exp mass screening/</li> <li>5 exp newborn/</li> <li>6 #4 AND #5</li> <li>7 #1 OR #2 OR #3 OR #6</li> <li>8 exp inborn-error-of-metabolism</li> <li>9 inborn ADJ2 error ADJ2 metabolism</li> <li>10 #8 OR #9</li> <li>11 #7 AND #10</li> <li>12 exp mass spectrometry/</li> <li>13 mass ADJ2 spect*</li> <li>14 ms ADJ2 spect*</li> <li>15 tandem ADJ2 mass</li> <li>16 #12 OR #13 OR #14 OR #15</li> <li>17 #11 AND #16</li> </ol>
<p><b>Medline via Pubmed</b></p> <p>Limited 2002 to current year of publication</p> <ol style="list-style-type: none"> <li>1 Neonatal screening</li> <li>2 Neonat* screen*</li> <li>3 Newborn* screen*</li> <li>4 Mass screening</li> <li>5 Infant, newborn</li> <li>6 #4 AND #5</li> <li>7 #1 OR #2 OR #3 OR #6</li> <li>8 Metabolism, inborn errors</li> <li>9 Inborn error*</li> <li>10 #8 OR #9</li> <li>11 #7 AND #10</li> <li>12 Spectrum analysis, mass</li> <li>13 Mass spect*</li> <li>14 MS spect*</li> <li>15 Tandem mass</li> <li>16 #12 OR #13 OR #14 OR #15</li> <li>17 #11 AND #16</li> </ol>	<p><b>Web of Knowledge Search Terms</b></p> <p>Topic=(neonat* OR newborn*) AND Topic=(screen*) AND Topic=(inborn error*) AND Topic=(metabolism) AND Topic=(mass OR MS OR tandem) AND Topic=(spect*)</p> <p>Timespan=All Years. Databases=SCI-EXPANDED, SSCI, A&amp;HCI, CPCI-S.</p>

**Stage 2**      search terms used to identify articles about the natural history, epidemiology, analytical validity, clinical validity and clinical utility for each of the five chosen diseases in PubMed (MEDLINE) Limited to PubMed MEDLINE 01/01/2002 to 21/05/2009

<b>Glutaric acidemia</b> <ol style="list-style-type: none"> <li>1. glutaryl coa</li> <li>2. glutaricaciduria</li> <li>3. gcdh</li> <li>4. ga 1</li> <li>5. glutaric aciduria</li> <li>6. glutaric acidemia</li> <li>7. glutaric acidemia</li> <li>8. #1 OR #2 OR #3 OR #4 OR #5 OR #6 OR #7</li> </ol>	<b>Homocystinuria</b> <ol style="list-style-type: none"> <li>1. homocystinuria</li> <li>2. hypermethioninaemia</li> <li>3. hypermethioninemia</li> <li>4. (cystathionine OR cbs) AND deficient*</li> <li>5. #1 OR #2 OR #3 OR #4</li> </ol>
<b>Isovaleric acidemia</b> <ol style="list-style-type: none"> <li>1. isovaleric acidemia</li> <li>2. isovaleric acidemia</li> <li>3. isovaleric aciduria</li> <li>4. ivd deficient*</li> <li>5. isovaleric acid AND dehydrogenase deficient*</li> <li>6. isovaleryl AND dehydrogenase deficient*</li> <li>7. isovalericacidemia</li> <li>8. isovalericacidemia</li> <li>9. #1 OR #2 OR #3 OR #4 OR #5 OR #6 OR #7 OR #8</li> </ol>	<b>Maple syrup urine disease</b> <ol style="list-style-type: none"> <li>1. maple syrup urine disease</li> <li>2. msud</li> <li>3. branched chain ketoaciduria</li> <li>4. keto acid decarboxylase deficient*</li> <li>5. #1 OR #2 OR #3 OR #4</li> </ol>

<p><b>Long-chain hydroxyacyl-CoA dehydrogenase deficiency</b></p> <ol style="list-style-type: none"> <li>1. trifunctional protein deficient*</li> <li>2. 3-hydroxyacyl coa dehydrogenase</li> <li>3. multienzyme complexes</li> <li>4. long chain AND dehydrogenase deficient*</li> <li>5. lchad</li> <li>6. hadh deficient*</li> <li>7. hydroxyacyl AND dehydrogenase</li> <li>8. long chain</li> <li>9. #7 AND #8</li> <li>10. hydroxydicarboxylic aciduria</li> <li>11. hydroxydicarboxylic aciduria</li> <li>12. #1 OR #2 OR #3 OR #4 OR #5 OR #6 OR #9 OR #10 OR #11</li> </ol>	<p><b>Epidemiology terms</b></p> <ol style="list-style-type: none"> <li>1. epidemiology</li> <li>2. morbidity</li> <li>3. mortality</li> <li>4. survival analysis</li> <li>5. disease susceptibility</li> <li>6. disease progression</li> <li>7. natural history</li> <li>8. epidemiolog*</li> <li>9. genetic heterogeneity</li> <li>10. incidence</li> <li>11. prevalence</li> <li>12. #1 OR #2 OR #3 OR #4 OR #5 OR #6 OR #9 OR #10 OR #11</li> </ol>
<p><b>Diagnosis terms</b></p> <ol style="list-style-type: none"> <li>1. sensitivity AND specificity</li> <li>2. diagnosis</li> <li>3. false positive reactions</li> <li>4. predictive value of tests</li> <li>5. false negative reactions</li> <li>6. false positive</li> <li>7. false negative</li> <li>8. diagnostic use</li> <li>9. specificity</li> <li>10. sensitivity</li> <li>11. #1 OR #2 OR #3 OR #4 OR #5 OR #6 OR #7 OR #8 OR #9 OR #10</li> </ol>	<p><b>Treatment terms</b></p> <ol style="list-style-type: none"> <li>1. research design</li> <li>2. clinical trials</li> <li>3. comparative study OR placebos</li> <li>4. treatment outcome</li> <li>5. double-blind method OR single-blind method</li> <li>6. (single OR double OR triple) AND blind</li> <li>7. random*[tiab]</li> <li>8. controlled clinical trial[pt]</li> <li>9. randomized controlled trial[pt]</li> <li>10. practice guideline[pt]</li> <li>11. clinical trial[pt]</li> <li>12. clinical trial[tiab]</li> <li>13. placebo*[tiab]</li> <li>14. clinical protocols OR feasibility studies OR pilot projects</li> <li>15. epidemiologic research design</li> <li>16. control*[tiab] AND trial*[tiab]</li> <li>17. #1 OR #2 OR #3 OR #4 OR #5 OR #6 OR #7 OR #8 OR #9 OR #10 OR #11 OR #12 OR #13 OR #14 OR #15 OR #16</li> </ol>

<b>Screening terms</b>	
------------------------	--

- |  |  |
|--|--|
| <ol style="list-style-type: none"><li>1. neonatal screening</li><li>2. neonat* screen*</li><li>3. newborn* screen*</li><li>4. mass screening</li><li>5. infant, newborn</li><li>6. genetic screening</li><li>7. #4 AND #5 AND #6</li><li>8. #1 OR #2 OR #3 OR #7</li><li>9. metabolism, inborn errors</li><li>10. inborn error*</li><li>11. #9 OR #10</li><li>12. #8 AND #11</li></ol> |  |
|--|--|





**PHG Foundation**  
2 Worts Causeway  
Cambridge CB1 8RN

Tel: +44 (0) 1223 740200  
Fax: +44 (0) 1223 740892

ISBN 978-1-907198-03-8

Registered charity no. 1118664  
Company no. 5823194