

Clinical effectiveness and cost-effectiveness of neonatal screening for inborn errors of metabolism using tandem mass spectrometry: a systematic review

A Pandor,^{1*} J Eastham,² C Beverley,¹ J Chilcott¹
and S Paisley¹



¹ The School of Health and Related Research (SchARR),
The University of Sheffield, UK

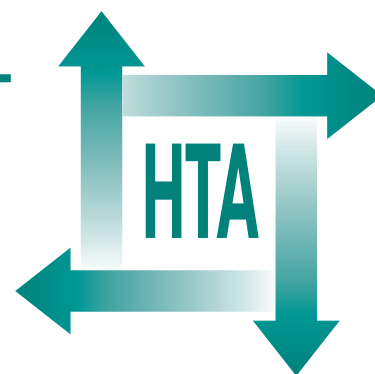
² Nuffield Institute for Health, University of Leeds, UK

* Corresponding author

Executive summary

Health Technology Assessment 2004; Vol. 8: No. 12

Health Technology Assessment
NHS R&D HTA Programme





Executive summary

Background

Inborn errors of metabolism are a rare group of genetic disorders that can have serious clinical consequences for an affected neonate or young infant. If undiagnosed and untreated, these disorders can cause irreversible mental retardation (ranging from mild to severe), physical disability, neurological damage and even fatality. Early detection (soon after birth) and an accurate diagnosis are very important for achieving a rapid and favourable patient outcome. Although the incidence of each specific metabolic disorder is rare, their collective importance is deemed to be of considerable public health significance.

Tandem mass spectrometry (MS) is seen as an important new technology for neonatal screening of inborn errors of metabolism. It has been demonstrated to be suitable for the reliable detection of phenylketonuria (PKU) and some other inborn errors of metabolism. This technology has the potential to screen for a range of metabolic disorders simultaneously.

The NHS R&D HTA programme commissioned two reviews of neonatal screening for inherited metabolic disorders (published in 1997). These reviews recommended further studies on the application of tandem MS to neonatal screening of inborn errors of metabolism. Both reviews were favourable to some introduction of screening for selected disorders, but with varying caveats. It was agreed that there should be no widespread introduction of the technology pending further evaluation, and the Child Health Subgroup of the National Screening Committee supported this position. Since the publication of these reports, no primary research has been undertaken in the UK. Recently, the HTA Diagnostic Technologies and Screening Panel felt that the technology was diffusing and that they needed answers more rapidly than would be obtained from proposed research. It was thought that a systematic review (to build on the two previous HTA reports) with economic modelling would be useful to bring the evidence base up to date and to identify the most urgent research needs.

Objectives

The aim of this review was to evaluate the clinical and cost-effectiveness of tandem MS-based neonatal screening for inborn errors of metabolism.

Methods

This review is an update of two previous HTA reports of neonatal screening for inborn errors of metabolism. These reports have been updated by a systematic review of recently published research (limited to studies published after 1995: the cut-off date in previous reviews) on neonatal screening of inherited metabolic diseases using tandem MS. This was supplemented by a search for economic literature and the application of a modelling exercise (based on the available evidence) to investigate the economics of using tandem MS within a neonatal screening programme in the UK.

Search strategy

New search strategies were developed based on scoping searches and strategies reported in the two previous systematic reviews. Fourteen electronic bibliographic databases covering biomedical, science, economic and grey literature were searched. The reference lists of relevant articles and abstracts of conference proceedings were checked. Eighteen health services research-related resources were also consulted via the Internet.

Inclusion and exclusion criteria

The titles, and abstracts where available, of all the articles identified by the literature searches were downloaded into a database. Any duplicates were removed. One reviewer assessed all citations and abstracts for relevance. Final decisions regarding the inclusion and exclusion criteria were based on full paper copies of manuscripts. Any uncertainties were resolved by discussion with another reviewer and/or clinical advisers.

Data extraction and quality assessment

All selected papers were read and critically appraised by a single reviewer, who extracted relevant information from the included studies directly into an evidence table. Any uncertainties ►

were resolved by discussion with another reviewer and/or clinical advisers. The overall strength and quality of evidence included in the review were graded according to the levels of evidence as defined in the previous systematic reviews of neonatal screening for inborn errors of metabolism.

Data synthesis

For the assessment of the clinical and cost-effectiveness evidence, formal meta-analysis techniques were not used because of the concerns over heterogeneity and poor study quality expressed in the previous systematic reviews. Instead, summary results were tabulated with detailed descriptive qualitative analyses.

Evaluation of the relative cost-effectiveness of adopting tandem MS for neonatal screening was undertaken using a probabilistic economic model. Data for the construction of the economic model were derived from previous HTA reports, other published sources and updated information on costs. A probabilistic approach was used to characterise the uncertainty within the model parameters. The model adopted an incremental net benefit approach. A value of information analysis was also undertaken to establish which elements of parameter uncertainty had the largest impact on model results.

Results

Number and quality of clinical effectiveness data

Fifteen studies were identified that evaluated the clinical effectiveness of neonatal screening for inborn errors of metabolism using tandem MS (including two abstracts and one paper that reported the same study but provided additional information or results on tandem MS-based newborn screening from different periods). Seven studies evaluated neonatal screening of amino acids and acylcarnitines using tandem MS. Four used a prospective cohort design with study durations from 2 to 3 years. Three studies did not report the study design. Of these seven studies, three were reported as abstracts and provided limited information. Four studies assessed newborn screening for medium-chain acyl-coenzyme A dehydrogenase (MCAD) deficiency using tandem MS. Three of these studies used a prospective cohort design with study durations from approximately 2 to 7 years, whereas a study from the UK, of approximately 3 years, used a retrospective cohort design. Two retrospective, analytical studies were also identified.

Assessment of clinical effectiveness

Evidence for neonatal screening of amino acids and acylcarnitines using tandem MS was primarily from observational data of large-scale prospective newborn screening programmes, from several centres outside the UK, namely, Australia, Germany and the USA. In general, newborn screening of dried blood spots for the amino acid and acylcarnitine group of disorders using tandem MS was shown to be rapid, highly sensitive (90–100%) and highly specific (99–100%). However, there was a lack of evidence regarding the false negatives and false positives for individual diagnosable disorders. The variation in the age of sampling and the heterogeneity in the choice of metabolite, as well as in thresholds used to define a positive result, limited direct comparison of the discriminative power of tandem MS between studies.

Evidence from several large-scale prospective newborn screening programmes in Australia, Germany and the USA, and retrospective data from the UK have, in general, illustrated high sensitivities and high specificities of neonatal screening for MCAD deficiency using tandem MS. However, direct comparison about diagnostic performance and outcome of tandem MS between studies was limited owing to use of various analytes and thresholds for detecting disease status and different criteria for the diagnosis of MCAD deficiency.

In most of the large-scale prospective screening studies conducted outside the UK, sampling for selected amino acids and acylcarnitines is usually performed within 72 hours after birth, whereas in the UK, blood samples are usually taken at about 6–14 days of age. The age at which screening is undertaken will affect the sensitivity and specificity of the screening process as concentrations of metabolites change over time, and this time-lapse may be detrimental for conditions that present acutely in infancy. Therefore, the evidence obtained in this review suggests that the collection of newborn blood-spots in the UK slightly earlier than the current 6–14 days may facilitate earlier detection and initiation of effective therapies. However, the earlier collection and reporting of results may influence test performance for other conditions and would underline the need for a good infrastructure for clinical follow-up, management and high-quality clinical services for identified cases and their families.

Evidence from the reviews of inborn errors of metabolism found that the UK screening



programme for PKU was well established and there was universal agreement that neonatal screening for PKU was justified. The average UK incidence of PKU (classical and atypical combined) is 11.0 cases per 100,000 live births. Early dietary interventions, including dietary treatment before or during pregnancy, are effective in reducing the severity of developmental delay, and neonatal screening using tandem MS is suitable for the reliable detection of PKU.

Of the many other disorders that can be detected by tandem MS, the best candidate condition for a new screening programme was MCAD deficiency, a disorder of fatty acid metabolism (expected UK birth prevalence/incidence ranges from 4.0 to 9.9 cases per 100,000 live births). The disorder is associated with increased morbidity and mortality. Treatment for MCAD deficiency is relatively simple with dietary management, thus preventing possible early death and neurological disability. After (early) diagnosis, current management makes death rare and improves outcome. Neonatal screening data from tandem MS-based studies show that this method is robust, highly sensitive and specific for MCAD deficiency. Without screening, an unknown number may remain asymptomatic and may never experience any ill effects.

For many other inborn errors of metabolism that can be detected by tandem MS, robust clinical evidence was limited (e.g. natural history of disease, UK incidence, and uncertainties regarding the effectiveness of treatments, and sensitivity and specificity of detection using tandem MS).

Number and quality of cost-effectiveness data

Since the two previous HTA reports, no published studies on the cost-effectiveness of screening for inborn errors of metabolism using tandem MS were identified.

Cost-effectiveness analysis using economic modelling

Cost-effectiveness analysis using economic modelling indicated that substituting the use of tandem MS for existing technologies for the screening of PKU alone could not be justified. Using tandem MS would incur additional costs with no measurable increase in health benefits. However, results from the economic modelling indicate that the addition of screening for MCAD deficiency as part of a neonatal screening programme for PKU using tandem MS would be cost-effective. Using an operational range of 50,000–60,000 specimens per system per year, the mean incremental cost for PKU

and MCAD deficiency screening combined using tandem MS from the model was –£23,312 (median –£14,810). This cost saving is associated with a mean incremental gain of 59 life-years for each cohort of 100,000 neonates screened.

Additional economic modelling for other conditions using probabilistic methods was undertaken, but the results are based on limited data; in particular, robust evidence on long-term outcomes, especially systematic differences in outcomes that could be attributed to screening. The evidence therefore does not support including other inherited metabolic diseases within a neonatal screening programme at present.

Conclusions

This systematic review evaluated the clinical and cost-effectiveness of neonatal screening for inborn errors of metabolism using tandem MS. The evidence appears to support the introduction of tandem MS into a UK neonatal screening programme for PKU and MCAD deficiency combined.

New technological approaches for automated processing coupled with the use of computer-assisted software would allow the analysis of hundreds of samples on a daily basis and minimise labour costs. Tandem MS has the potential for simultaneous multidisease screening using a single analytical technique. However, it is difficult to draw firm conclusions on extending the UK neonatal screening programme to all disorders detectable by tandem MS. Although the marginal cost of extending the programme to include other conditions may be relatively small, the application of this new technology to PKU and MCAD deficiency screening does not imply the wholesale inclusion of all disorders detectable by tandem MS. Robust evidence on the underlying incidence and outcomes for many of the disorders was lacking, particularly differences in long-term outcomes that could be attributed to therapies initiated as a consequence of presymptomatic detection using tandem MS.

Recommendations for research

- The economic evidence concerning the use of tandem MS for PKU and MCAD deficiency combined is very favourable. The key assumptions underlying this analysis that may constitute areas for further research

are: the future disability costs associated with symptomatic cases, and the underlying population incidence of the condition in England and Wales. However, the value of information analysis undertaken as part of the economic modelling of cost-effectiveness suggests that the overall value of obtaining this additional data is not high.

- More evidence is needed to establish the sensitivity and specificity of neonatal screening using tandem MS for other individual inborn errors of metabolism in the UK and to determine the underlying incidence of these conditions.
- Further research is needed to ascertain the natural history of some conditions and the potential economic impact of screening for

other metabolic disorders. It is suggested that the primary focus of this research could 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 tandem MS.

Publication

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 Technol Assess* 2004;**8**(12).

NHS R&D HTA Programme

The NHS R&D Health Technology Assessment (HTA) Programme was set up in 1993 to ensure that high-quality research information on the costs, effectiveness and broader impact of health technologies is produced in the most efficient way for those who use, manage and provide care in the NHS.

The research reported in this monograph was commissioned by the HTA Programme and was funded as project number 01/31/01. Technology assessment reports are completed in a limited time to inform decisions in key areas by bringing together evidence on the use of the technology concerned.

The views expressed in this publication are those of the authors and not necessarily those of the HTA Programme or the Department of Health. The editors wish to emphasise that funding and publication of this research by the NHS should not be taken as implicit support for any recommendations made by the authors.

Criteria for inclusion in the HTA monograph series

Reports are published in the HTA monograph series if (1) they have resulted from work commissioned for the HTA Programme, and (2) they are of a sufficiently high scientific quality as assessed by the referees and editors.

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

HTA Programme Director: Professor Tom Walley
Series Editors: Dr Ken Stein, Professor John Gabbay, Dr Ruairidh Milne,
Dr Chris Hyde and Dr Rob Riemsma
Managing Editors: Sally Bailey and Caroline Ciupek

The editors and publisher have tried to ensure the accuracy of this report but do not accept liability for damages or losses arising from material published in this report.

ISSN 1366-5278

© Queen's Printer and Controller of HMSO 2004

This monograph may be freely reproduced for the purposes of private research and study and may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising.

Applications for commercial reproduction should be addressed to HMSO, The Copyright Unit, St Clements House, 2-16 Colegate, Norwich, NR3 1BQ.

Published by Gray Publishing, Tunbridge Wells, Kent, on behalf of NCCHTA.
Printed on acid-free paper in the UK by St Edmundsbury Press Ltd, Bury St Edmunds, Suffolk.