Newborn screening for inborn errors of metabolism: a systematic review

Authors' objectives
The authors' stated objectives were: to undertake a systematic review of the data as a basis for the evaluation of newborn screening for inborn errors of metabolism; to prepare an objective summary of the evidence on the appropriateness and need for various existing and possible neonatal screening programmes for inborn errors of metabolism in relation to the natural history of these diseases; to identify gaps in existing knowledge and make recommendations for required primary research; and to make recommendations for the future development and organisation of neonatal screening.

Searching
MEDLINE was searched from 1966 to June 1996 for publications in any language; the keywords were provided. Other sources searched were EMBASE and various databases on BIDS (from 1980), the Science Citation Index (from 1981) and the Index to Scientific and Technical Proceedings (from 1982). Manual searches of reference lists of articles, textbooks, conference proceedings, Index Medicus and Current Contents were also conducted. The directors of all screening laboratories in the UK were contacted for any grey or unpublished literature.

Study selection
No inclusion criteria relating to the study design were specified.

Specific interventions included in the review
The studies were of neonatal screening for inborn errors of metabolism. The screening technologies considered included: automation of process, molecular (DNA) techniques, tandem mass spectrometry (MS), Gutrie bacterial inhibition assay, chromatography, fluorimetry, radioimmunoassay, enzymology, and immunoassay based on time-resolved fluorimetry.

Reference standard test against which the new test was compared
It was unclear whether diagnostic accuracy studies were included in the review. No inclusion criteria for the reference standard or method of establishing diagnosis in the included studies were reported.

Participants included in the review
Studies of newborn infants screened for inborn errors of metabolism were eligible for inclusion. Such errors of metabolism included: phenylketonuria (PKU), amino acidopathies, disorders of carbohydrate metabolism, disorders of organic acid metabolism, fatty acid oxidation defects, disorders of adrenal steroidogenesis, lipoprotein disorders, peroxisomal disorders, disorders of the urea cycle, respiratory chain or tricarboxylic acid cycle disorders, trace metal disorders, purine or pyrimidine disorders, and lysosomal disorders.

Outcomes assessed in the review
Any studies that contained data on population incidence, the effectiveness of screening, health outcomes, or screening and treatment costs for inborn errors of metabolism were included, as were all studies that described defined suitable screening technologies for blood samples. Different outcome measures relating to morbidity and mortality, specific to each type of inborn error of metabolism, were reported.

How were decisions on the relevance of primary studies made?
Two reviewers evaluated citations and abstracts of references identified. A copy of the full paper was obtained and assessed where there was insufficient information in the abstracts.

Assessment of study quality
The studies were assessed using a critical appraisal checklist that was reported in the review. A working group of two or three reviewers was convened to assess the papers selected.

**Data extraction**
The authors did not state how the data were extracted for the review, or how many reviewers performed the data extraction.

**Methods of synthesis**
How were the studies combined?
The studies were combined in a narrative description, and the evidence was summarised and assessed against World Health Organization criteria for screening programmes (see Other Publications of Related Interest no. 1).

How were differences between studies investigated?
Differences between the studies were investigated in the narrative and were discussed.

**Results of the review**
A total of 1,866 papers were identified, of which 407 were systematically selected for full critical appraisal.

Published evidence confirmed that universal newborn screening for PKU meets all of the screening criteria, and justifies the expense and infrastructure necessary for the collection and testing of neonatal blood spots. There was insufficient evidence in the literature to assess the cost-effectiveness of screening for any other inborn errors of metabolism. There was reasonable evidence to support the inclusion of biotinidase deficiency, congenital adrenal hyperplasia, medium-chain acyl coenzyme A dehydrogenase deficiency and glutaric aciduria type 1 in extended neonatal screening.

**Cost information**
PKU screening provides a positive net monetary benefit to society and justifies the collection of blood samples from neonatal infants. There was insufficient economic evidence to support a change from current methodology to tandem MS-based screening solely for PKU. More information is needed on the cost-effectiveness of extending screening to other disorders. There was insufficient evidence to assess the economic value of screening for any other inborn errors of metabolism.

**Authors’ conclusions**
Large-scale trials of screening for biotinidase, congenital adrenal hyperplasia, medium-chain acyl coenzyme A dehydrogenase deficiency and glutaric aciduria type 1 should be conducted, with careful evaluation to establish their clinical effectiveness and cost-effectiveness in practice. Screening for the latter two disorders would be dependent upon the use of tandem MS. The application of tandem MS to newborn screening requires further evaluation. The extension of neonatal screening programmes to other inborn errors of metabolisms is not currently justified.

**CRD commentary**
This review set out a comprehensive and appropriate question. The search strategy, criteria for inclusion, and screening criteria were adequately described. The evidence was narratively summarised against appropriate criteria. However, the reporting of study details, including the results, were limited. The results were derived largely from opinion-based interpretation using the critical appraisal checklist. The lack of data makes it impossible to judge the validity of the narrative summary. In addition to a review of the literature, further evidence were provided by a survey of the directors of all neonatal screening laboratories in the UK, and site visits to some laboratories in the USA and the UK. Given the limitations outlined, the authors' conclusions should be treated with caution.
Implications of the review for practice and research
Practice: Universal screening for PKU should be continued. National neonatal screening programmes for profound biotinidase deficiency and congenital adrenal hyperplasia would be justified on the evidence. Screening for medium-chain acyl coenzyme A dehydrogenase deficiency and glutaric aciduria type 1 should be seriously considered.

Research: Any new neonatal screening programmes should be the subject of ongoing evaluation to ensure cost-effectiveness. Tandem MS for neonatal screening for PKU, medium-chain acyl coenzyme A dehydrogenase deficiency and glutaric aciduria type 1 should be further evaluated by primary research.

Funding
NHS R&D Health Technology Assessment (HTA) Programme, project number 93/36/03.

Bibliographic details

Original Paper URL
http://www.hta.ac.uk/project.asp?PjtId=921

Other publications of related interest

Indexing Status
Subject indexing assigned by NLM

MeSH
Cost-Benefit Analysis; Incidence; Infant, Newborn; Mass Spectrometry /economics /instrumentation /methods; Metabolism, Inborn Errors /diagnosis /epidemiology /therapy; Neonatal Screening /economics /methods; Technology Assessment, Biomedical

AccessionNumber
11998008285

Date bibliographic record published
31/07/2005

Date abstract record published
31/07/2005

Record Status
This is a critical abstract of a systematic review that meets the criteria for inclusion on DARE. Each critical abstract contains a brief summary of the review methods, results and conclusions followed by a detailed critical assessment on the reliability of the review and the conclusions drawn.