Evaluating newborn screening programmes based on dried blood spots: future challenges

Carol Dezateux

Department of Epidemiology and Public Health, Institute of Child Health, London, UK

A UK national programme to screen all newborn infants for phenylketonuria was introduced in 1969, followed in 1981 by a similar programme for congenital hypothyroidism. Decisions to start these national programmes were informed by evidence from observational studies rather than randomised controlled trials. Subsequently, outcome for affected children has been assessed through national disease registers, from which inferences about the effectiveness of screening have been made. Both programmes are based on a single blood specimen, collected from each infant at the end of the first week of life, and stored as dried spots on a filter paper or ‘Guthrie’ card. This infrastructure has made it relatively easy for routine screening for other conditions to be introduced at a district or regional level, resulting in inconsistent policies and inequitable access to effective screening services. This variation in screening practices reflects uncertainty and the lack of a national framework to guide the introduction and evaluation of new screening initiatives, rather than geographical variations in disease prevalence or severity. More recently, developments in tandem mass spectrometry have made it technically possible to screen for several inborn errors of metabolism in a single analytical step. However, for each of these conditions, evidence is required that the benefits of screening outweigh the harms. How should that evidence be obtained? Ideally policy decisions about new screening initiatives should be informed by evidence from randomised controlled trials but for most of the conditions for which newborn screening is proposed, large trials would be needed. Prioritising which conditions should be formally evaluated, and developing a framework to support their evaluation, poses an important challenge to the public health, clinical and scientific community. In this chapter, issues underlying the evaluation of newborn screening programmes will be discussed in relation to medium chain acyl CoA dehydrogenase deficiency, a recessively inherited disorder of fatty acid oxidation.

The UK national newborn screening programmes for phenylketonuria and congenital hypothyroidism are based on a single sample of capillary blood collected by heelprick from all infants between 6 and 14 days of age. These samples are usually stored dried on special filter papers,
Screening

previously referred to as Guthrie cards, after the originator of the main screening test for phenylketonuria. These programmes are considered to be successful, achieving levels of population coverage, test performance and treatment effectiveness which compare favourably with other national screening programmes. A single blood sample is common to screening for both conditions and there is potential for the residual sample to be tested for other conditions. In 1995, Streetly et al. estimated that about 16% of all infants born in the UK each year were routinely screened for cystic fibrosis, 12% for homocystinuria, 9% for haemoglobinopathies, 9% for galactosaemia, and 3% for tyrosinaemia. Such local variation is inequitable and emphasises the uncertainty that exists at the margins of screening policies and the susceptibility to local pressures. If national screening policy is to be based on evidence of effectiveness and cost-effectiveness, mechanisms are required to ensure that innovations in screening are introduced only within the context of an evaluative study, ideally a randomised controlled trial. In their absence, the opportunistic and piecemeal extension of the existing national newborn screening programmes is likely to continue. Prioritising which conditions should be formally evaluated and developing a national framework to support their evaluation poses an important challenge to the public health, clinical and scientific community.

In the UK, two commissioned systematic reviews of newborn screening for inborn errors of metabolism have been published. Further commissioned reviews of cystic fibrosis and haemoglobinopathy screening are due to be reported (updated information can be obtained from the NHS R&D HTA website: http://www.soton.ac.uk/~hta/). Important issues for the organisation and delivery of newborn screening services have been identified through local audits and, more recently, through a national audit of the screening programmes for phenylketonuria and congenital hypothyroidism. It is, therefore, timely to consider the future directions of newborn screening based on dried blood spots and the research required to inform and strengthen the scientific basis of these developments. In this chapter, some principles underlying the evaluation of proposed newborn screening programmes are discussed, using the example of screening for medium chain acyl CoA dehydrogenase deficiency.

Medium chain acyl CoA dehydrogenase deficiency

Over the last two decades, a number of disorders of fatty acid oxidation have been recognised, the most common of which is medium chain acyl CoA dehydrogenase (MCAD) deficiency. MCAD deficiency is a recessively inherited metabolic disorder that reduces the ability to
maintain a normal blood sugar during episodes of metabolic stress\textsuperscript{13}. During intercurrent infections and illnesses, affected children may develop profound hypoglycaemia, encephalopathy and hepatic dysfunction. Affected children are usually asymptomatic initially, although neonatal symptoms have been reported\textsuperscript{14}. MCAD deficiency may not be suspected until late infancy or early childhood, when children present with an acute and frequently fatal episode of metabolic decompensation. These episodes are most frequent in the first 2 years of life and rare after 5 years of age. A spectrum of clinical severity is recognised and some affected individuals are not diagnosed until later childhood when they present with mild episodic hypoglycaemia or, in adult life, with symptoms of muscle weakness and fatigue\textsuperscript{4,5}.

Of those children presenting clinically, about one-quarter will die and about one-third of survivors will have irreversible neurological damage\textsuperscript{4,5}. In the largest published case series based on 120 cases from the US, 19 of the 23 children who died had been previously well and had died during their first illness\textsuperscript{15}. Failure to diagnose metabolic disease in those who die is a recognised problem, reflecting a low index of clinical suspicion for rare conditions as well as difficulties in making a biochemical diagnosis in an acutely sick child. It is estimated that 20–30\% of cases may go undiagnosed clinically, either because they die without a diagnosis being considered or made, or because they remain asymptomatic and well throughout early childhood. However, this percentage, and the proportion at either extreme of the clinical spectrum, has not been directly measured.

A single point mutation of adenine to guanine at position 985 (termed A985G mutation) in the MCAD gene sequence has been identified which is thought to account for almost 90\% of mutations\textsuperscript{16,17}. Among those clinically diagnosed with MCAD deficiency, 81\% are estimated to be homozygous and 18\% heterozygous for this common mutation. Although this mutation is thought to have originated from the northwestern European population\textsuperscript{16}, this genotype–phenotype relation has not been confirmed in Scotland\textsuperscript{18}. Based on A985G mutation prevalence studies and assuming random mating, it has been suggested that the birth prevalence of MCAD deficiency in the UK may be as high as 10 per 100,000\textsuperscript{19}. This is about twice the cumulative incidence to 16 years of age of clinically diagnosed MCAD deficiency reported from a national surveillance study\textsuperscript{20}. This disparity suggests that cases are underdiagnosed clinically.

Once a diagnosis has been made, treatment during intercurrent illnesses or periods of anorexia consists of ensuring adequate calorie intake, either by mouth or intravenously. Clinical case series suggest that outcome is favourable following such treatment, although its effectiveness in preventing subsequent episodes has not been proven\textsuperscript{15,21,22}. In the US,
affected children are also given L-carnitine supplements. This is not currently used in the UK, but evidence to support either practice is lacking.

**The rationale for screening**

As major sequelae are often sustained during the presenting illness rather than at a later stage, the role of earlier diagnosis before symptoms have developed has received increasing attention. Although prenatal diagnosis has been reported, this is not considered an appropriate strategy, as affected children are believed to be normal provided severe metabolic decompensation can be avoided.

**The screening test**

The screening test for MCAD deficiency is based on detection and quantification of acyl carnitines using tandem mass spectrometry. DNA analysis to detect A985G mutations can be undertaken on dried blood spots, but would not be suitable for primary screening. By linking two mass spectrometers together, tandem mass spectrometry allows the separation and analysis of complex samples such as dried blood spots to proceed simultaneously, so that a number of compounds can be identified in a single analytical step. More than 20 inborn errors of metabolism can be detected with this technique, including phenylketonuria, tyrosinaemia type I, MCAD deficiency, other disorders of fatty acid oxidation and organic acidaemias. However, at present, tandem mass spectrometry cannot be used to screen for congenital hypothyroidism and disorders such as cystic fibrosis, sickle cell disease and congenital adrenal hyperplasia, for which screening has been proposed or is currently undertaken.

A presumptive positive screening test for MCAD deficiency is based on the finding of raised concentrations of octanoyl (C8) carnitine. In one study from the US using isotope-dilution tandem mass spectrometry, octanoyl carnitine concentrations were raised (> 0.3 μM) in all (n = 62) cases of MCAD deficiency. The authors reported that symptom status, carnitine supplementation or genotype (A985G homozygote or compound heterozygote) did not influence octanoyl carnitine concentrations but numbers were small. Octanoyl carnitine levels were > 0.3 μM in retrieved Guthrie card samples obtained for 8 clinically diagnosed subjects, and below this level among unaffected controls and normal neonates. In one UK report based on electrospray tandem mass...
spectrometry, octanoyl carnitine concentrations > 0.38 μM were reported in 35 children with proven MCAD deficiency and > 1.5 μM in the retrieved neonatal blood spots. It has been suggested that all clinically ascertained cases may be identified with electrospray tandem mass spectrometry.

A presumptive positive screening test for MCAD deficiency may be confirmed by measuring blood spot cis-4-decenoic acid, white cell/fibroblast tritium release, or through a phenylpropionic acid load. Plasma non-esterified fatty acid:3-hydroxy-butyrate ratios and urinary organic acids may be measured during an acute episode or after a provocative fast.

**Current experience of screening for MCAD deficiency**

Screening programmes have been established in the US and Saudi Arabia. Published data from Pittsburgh suggest a birth prevalence for MCAD deficiency of 11 per 100,000, based on the identification of 9 affected infants among 80,371 newborns screened with tandem mass spectrometry. In this report, quantitative thresholds were not specified but infants were identified through a ‘characteristic pattern’ of elevated octanoyl, hexanoyl and decanoyl carnitines. The ability to distinguish between simple heterozygotes (carriers of the A985G mutation who are not themselves at risk of metabolic decompensation) and compound heterozygotes (affected individuals who have two different MCAD mutations and are at risk) is clearly an important criterion for a screening test. In the Pittsburgh study, the methods used to confirm the presumptive positive screening result were not stated, although the results of DNA analyses were reported. Four were homozygous and five heterozygous for A985G, a higher proportion of heterozygotes than the authors predicted from clinical studies. One possible explanation for this is that A985G homozygotes may be relatively more common among clinically diagnosed cases than among those detected through screening. The authors considered it unlikely that these cases were simple heterozygotes, as raised octanoyl carnitine concentrations had not been reported previously among a small sample of obligate simple heterozygotes.

Although it is assumed that the true prevalence of MCAD deficiency is higher than that determined from clinically diagnosed cases, estimates of prevalence based on ascertainment through screening with tandem mass spectrometry are not yet available for the UK. Existing UK prevalence figures for MCAD deficiency are based on A985G carrier status determined in about 10,000 newborn blood spot specimens from Trent and three predominantly Caucasian English counties.
suggest that MCAD deficiency may affect 10 out of every 100–200,000 children born in the UK.

Only limited information regarding outcome following screening is available from the Pittsburgh screening programme\(^{27}\). Two of the nine children identified through screening died subsequently: one (homozygote) developed a metabolic crisis after immunisation and one (heterozygote) during an intercurrent illness. The numbers involved are too small to allow any inferences to be drawn from these data. The effectiveness of treatment following early diagnosis through screening may also be reduced if clinical episodes occur before a screening test is performed or a result is available. Up to one-third of affected newborns may develop symptoms in the first 3 days of life, before a screening result would be available\(^{14}\). In the UK national surveillance study\(^{20}\), 3 of 45 children without a family history of MCAD deficiency presented before 17 days of age (R Pollitt, personal communication).

**What information is needed by decision makers?**

What information is needed to determine whether screening for MCAD deficiency should be introduced in the UK? The objective of a screening programme for MCAD deficiency would be to prevent death and major neurological handicap through early diagnosis and treatment. Current experience of population-based screening for MCAD deficiency is very limited. What priority should be given to obtaining better information about effectiveness and what information is needed to determine policy? At a public health level, MCAD deficiency is a rare disorder. Clinically, it is a condition with potentially major consequences for affected children and their families. In health technology terms, MCAD deficiency is the tip of the tandem mass spectrometry iceberg, and technologies for other conditions, including congenital adrenal hyperplasia, have already emerged\(^{28}\).

Some annual figures for a UK screening programme for MCAD deficiency can be estimated based on current data. If the birth prevalence of MCAD deficiency is assumed to be 10 per 100,000, this suggests that each year around 70 affected children are born in the UK, 49 (70%) of whom would be currently diagnosed clinically. Of the 49 clinically diagnosed cases, 12 (25%) would die while 12 of the 37 survivors would have irreversible neurological damage. Assuming a similar proportion of deaths among the 21 cases not diagnosed clinically, an additional 5 children would die, leaving 16 who are asymptomatic and presumed to be neurologically normal. If the true prevalence of MCAD deficiency were lower, say 5 per 100,000, these figures would be halved.
If screening combined with early treatment could reduce the proportion who die or sustain neurological damage by 50%, then, in the UK each year, screening might prevent 8–9 deaths and 6 children from becoming severely neurologically impaired. The benefits and harms to the 16 affected children who might never have developed symptoms have not been established. Potential consequences of diagnosis for this asymptomatic group include anxiety about the risk of hypoglycaemia during early childhood and genetic information for the parents which may have relevance for future children. The disbenefits to the false positives who are recalled for further testing include the potential anxiety and confusion caused to unaffected families, the risk of subsequent misdiagnosis among a truly unaffected child, and the health service resources required to recall, evaluate and diagnose all cases.

The magnitude of these outcomes depends on screening test performance and the prevalence of disease in the newborn population. If favourable assumptions about test performance are made (for example, a sensitivity of 99% and 0.5% false positive rate), 69 affected children would be correctly identified in the UK each year and 3500 unaffected children would require follow up for a presumptive positive screening result. This gives an odds of being affected given a positive result of 1:51 for a birth prevalence of 10 per 100,000. The number of children recalled for further investigation may be less than this: presumptive positive results can be investigated initially by using other dried spots from the same Guthrie card or by obtaining a second Guthrie card through the midwife or health visitor (currently a repeat sample is requested for about 1% of children in the UK). While DNA analysis could be used to select who should be recalled, at present this strategy would not allow the important question of phenotype–genotype associations within a screen-detected population to be assessed. The contribution of cis-4-decenolic acid and other in vitro assays suitable for dried blood spots to improving test specificity requires further investigation. It has been suggested that the recall and retest rates for analyses carried out by tandem mass spectrometry are ‘close to zero’. Further data are required from UK populations to confirm these views and to identify the determinants of the concentrations of individual acyl carnitine species in the neonatal period.

Distinct from the issue of recall policies is the issue of establishing or excluding a diagnosis of MCAD deficiency in a young baby who is well and without a family history. One recognised problem for families of newborn infants with a false positive screening diagnosis is residual anxiety despite re-assurance that the screening result has not been confirmed. It is, therefore, extremely important that a definitive diagnosis can be provided as quickly and unequivocally as possible. The investigations and criteria that can be used to make a diagnosis of...
MCAD deficiency in an asymptomatic presumptive positive infant are similar, but not identical, to those used to make a diagnosis in clinically symptomatic infants who are frequently investigated during an episode of metabolic decompensation. Diagnostic algorithms are required to avoid a potential risk of misdiagnosis and to ensure that a diagnosis may be confidently excluded in an asymptomatic presumptive positive infant. Without this there is a risk of misdiagnosis or of the screening result becoming a surrogate diagnostic result.

Current data are clearly insufficient for policy. This was effectively the conclusion of the teams responsible for both UK systematic reviews of screening for inborn errors of metabolism. Both concluded that more UK-based data were required to inform screening policies based on tandem mass spectrometry and identified a need for primary research on screening for MCAD deficiency and some other inborn errors of metabolism. The current debate relates to the nature of that primary research.

Information on likely benefits and possible harms is required to a precision that is meaningful in public health terms. In policy terms, this means identifying the level of benefit (in terms of mortality/disability) which would lead to a decision to introduce screening. An indication of the level of harm, and its relation to benefit, which would lead to a decision not to introduce screening is also required and is equally difficult to define. Given that a newborn infant will be screened for several conditions, the marginal disbenefits of each new screening programme may assume greater relevance. For most newborns screened, the probability of a false positive result will almost certainly exceed that of a true positive result. This issue has been discussed by Russell in relation to cervical screening, but applies equally to newborn screening where the probabilities of false positive diagnoses arising from each screening programme are likely to be additive.

Evaluating screening

The role of observational studies

Observational data can contribute to the assessment of potential screening, providing information about prevalence and test performance. Analysis of Guthrie specimens from large numbers of proven cases can be used to confirm current estimates of test sensitivity. However, these estimates will only apply to cases ascertained clinically. As the predictive value of a positive test will depend on pre-test probability of disease, geographically representative estimates of disease prevalence based on tandem mass spectrometry are required.
Evaluating newborn screening programmes based on dried blood spots

Observational study designs can also be used to determine the contribution of screen-detectable metabolic disorders such as MCAD deficiency to death in infancy and early childhood. This is currently being investigated through a UK collaborative case control study of death in the first 2 years of life. Finally, the role of further secondary data analyses designed to synthesise data from observational studies deserves mention. While the UK systematic reviews included some secondary economic analyses, further analyses based on decision trees and using existing observational data may be useful to assess cost-effectiveness and to identify key areas of uncertainty before a large and potentially expensive randomised controlled trial is undertaken.

The rationale for randomised controlled trials

More difficult is the issue of obtaining unbiased evidence that treatment is more effective following early detection than it is following clinical diagnosis, and that screening is not associated with significant risks to those screened. Is a randomised controlled trial of screening necessary or can observational studies of the outcomes of screening in 'pilot' regions provide this information? The rationale for a randomised controlled trial lies in its ability to overcome the biases inherent in observational studies of screening, whether they compare outcome before and after introduction of screening or outcome in different geographical areas with different screening strategies. Furthermore, because screening may be associated with unrecognised harms, which are difficult to assess within an observational study, a randomised trial is the best way to ensure that the potential risks of screening are minimised.

Bias due to differential methods used to ascertain cases is likely in an observational study of screening, as cases identified through screening will differ to those identified clinically. For example, an observational study of screening for MCAD deficiency may well show a reduction in mortality in screened periods or regions simply through identifying children with milder forms of the disease. To avoid ascertainment bias, the probability of ascertainment should be equal between the screened and an unscreened population at the time outcome is assessed. This approach was used in the Wisconsin trial of newborn screening for cystic fibrosis.

Outcome may be influenced by factors other than screening, such as variations in medical care over time or between different regions. The availability of tandem mass spectrometry as a diagnostic tool is likely to facilitate rapid diagnosis of a range of inborn errors of metabolism in symptomatic children or those with a relevant family history. The comparison of interest for a screening programme will be with a
concurrent unscreened control group that has access to this technology. Geographically based comparisons will be difficult to interpret due to regional variation in disease prevalence as well as recognised variations within the UK in access to specialist paediatric metabolic services.

The key limitation of observational evaluative studies lies in their inability to quantify benefit and harm in an unbiased manner. Screening and early treatment may reduce mortality but result in survival of children with neurological impairment. It is difficult to see how this outcome could be satisfactorily measured within an observational study as a comparable control group would be lacking. Other outcomes are also important, notably the psychological effects of the information given to parents of true or false positive infants about their apparently well baby.

Despite an increasing commitment to requiring evidence from high quality randomised controlled trials that a screening programme is effective in reducing mortality or morbidity, this kind of evidence is almost completely lacking for most newborn screening programmes. Neither congenital hypothyroidism nor phenylketonuria were assessed in this way before screening was introduced. For phenylketonuria, this was not because of lack of awareness about randomised controlled trials. Debates about how best to assess the effectiveness of early diagnosis and treatment for phenylketonuria when screening for this condition was first proposed bear an uncanny resemblance to the debates about screening for MCAD deficiency. Birch and Tizard questioned the evidence for the effectiveness of dietary treatment of phenylketonuria which they regarded as ‘not proven’ on the basis of observational studies which they pointed out were subject to selection bias. They proposed a randomised controlled trial of early treatment to assess the effectiveness of a low phenylalanine diet since ‘it may not only be ineffective but harmful as a treatment’. In the end, screening policy was determined on the basis of observational data comparing outcome for cases diagnosed through screening in Scotland, where screening had already been introduced, and England, where it had not. Almost 30 years later, data from the UK national register confirm that outcome for classical severe phenylketonuria is much improved, but there remain some outstanding issues about the benefits and possible social and financial costs of treating milder forms of phenylketonuria.

Almost certainly screening has altered the concept of a ‘case’ of phenylketonuria. Genetic variation is increasingly recognised: phenylalanine levels are a continuous distribution determined largely by the specific mutations present. The previous distinction between classical phenylketonuria (with little enzyme activity) and non phenylketonuria hyperphenylalaninaemia (with intermediate levels of enzyme activity) has become less meaningful as the genetic basis of phenylketonuria has been
clarified and experience with screening has increased. Observational studies do not help assess the benefits or otherwise of treatment in individuals with borderline or mild phenotypes yet such information is highly relevant when determining screening policies.

The scientific case for trials of newborn screening is strong. There are also ethical reasons why such trials should be carried out. Lumley has suggested that evaluation through randomised controlled trials is an 'ethical imperative' and cites Silverman's view of randomised trials as essentially 'risk minimising'. Objections to trials of screening usually arise because there are strongly held convictions that early treatment is effective and that the control group will be denied access to treatment, which will be beneficial. However screening, like most forms of medical treatment, is likely to produce marginal rather than dramatic benefits and the alternative possibility, that the intervention group may be offered something with unrecognised hazards, should also be considered. Potential hazards for MCAD deficiency include survival with neurological impairment, adverse effects of clinically-unwarranted treatment in asymptomatic individuals, the consequences of unwanted genetic information, misdiagnosis of false positives, anxiety among families of false positives.

How large a trial might be needed? Given the natural history of MCAD, mortality and/or neurological impairment as a combined measure and occurring in the first two years of life would be an appropriate primary outcome measure in a randomised trial. Provisional estimates of sample size based on this measure are summarised in Table 1, but cannot be considered definitive until more reliable data on birth prevalence are available. From this table, it can be seen that a trial would need to involve the entire UK population of births over a 5-year period in order to have sufficient power to detect a 50% reduction in the primary outcome of death and/or disability by 2 years of age. If a trial of this size and duration were to be undertaken, this would require that screening for MCAD deficiency using tandem mass spectrometry were only introduced into the UK as part of a randomised controlled trial.

There is no doubt that trials of newborn screening will pose a major challenge to the clinical, scientific and public health community and will

<table>
<thead>
<tr>
<th>% Reduction in outcome</th>
<th>Birth prevalence 5 per 100,000</th>
<th>Birth prevalence 10 per 100,000</th>
</tr>
</thead>
<tbody>
<tr>
<td>25% reduction</td>
<td>15.1 million</td>
<td>7.5 million</td>
</tr>
<tr>
<td>50% reduction</td>
<td>3.2 million</td>
<td>1.6 million</td>
</tr>
</tbody>
</table>

*For 90% power and 5% significance.
require the informed participation of parents on a national scale. Parents will require better information and will need to give informed consent to screening and its evaluation, which is not currently requested. Clinical uncertainty will be made more publicly explicit than is currently the case\textsuperscript{39}. The UK National Health Service and the existing infrastructure of the national newborn screening service could be developed to support such planned experiments and an investment in this aspect could increase the cost-effectiveness of newborn screening research. Lack of evidence for the effectiveness of early treatment for cystic fibrosis has been cited for over 15 years as the main reason for not introducing newborn screening\textsuperscript{41}. Were neonatal screening for cystic fibrosis to be proved effective as the early results of the Wisconsin newborn screening trial\textsuperscript{35} suggest, then affected children will have been denied an effective form of intervention for more than a decade because observational studies were not conclusive.

**Conclusion**

There are an increasing number of conditions for which neonatal screening is being proposed. Technological advances, such as tandem mass spectrometry, make it technically possible to test for several conditions in one analytical step. The priority given to evaluating these potential screening programmes will depend on the importance of the individual conditions, the performance of the proposed tests, as well as the anticipated benefits of screening. However, decisions to start new screening programmes should be informed by unbiased estimates of benefits and harms which cannot be derived from observational studies. These issues have been discussed in relation to screening for MCAD deficiency but apply to other proposed newborn screening programmes. The infrastructure established to support the current UK national newborn screening programmes for phenylketonuria and congenital hypothyroidism could be developed to allow new screening programmes to be introduced in the context of a formal trial. This would maximise the potential benefits of diagnostic and therapeutic advances to affected children and their families while minimising harm to the population being screened.

**References**

Evaluating newborn screening programmes based on dried blood spots

7 Pharoah POD, Madden MP. Audit of screening for congenital hypothyroidism. *Arch Dis Child* 1992; 67: 1073–6
8 Galloway A, Stevenson J. An audit of the organisation of neonatal screening for phenylketonuria and congenital hypothyroidism in the Northern Region. *Public Health* 1996; 110: 119–21
9 Cappuccio FP, Hickman M, Barker M. Performance is hard to monitor. *BMJ* 1996; 312: 182
16 Gregersen N, Winter V, Curtis D et al. Medium-chain acyl-CoA dehydrogenase (MCAD) deficiency: the prevalent mutation G985 (K304E) is subject to a strong founder effect from Northwestern Europe. *Hum Hered* 1993; 43: 342–50
20 Pollitt RJ, Leonard JV. Prospective surveillance study of medium chain acyl-CoA dehydrogenase deficiency in the UK. *Arch Dis Child* 1998; 79: 116–9
22 Wilson CJ, Champion MP, Collins JE, Clayton PT, Leonard JV. Outcome of medium chain acyl-CoA dehydrogenase deficiency after diagnosis. Submitted
33 Torgerson DJ, Donaldson C. Economic evaluations before clinical trials. Lancet 1996; 348: 687