EVALUATION OF NEWBORN SCREENING: STUDIES IN CYSTIC FIBROSIS AND DISORDERS DETECTABLE BY TANDEM MASS SPECTROMETRY

Bridget Wilcken AM, MBChB, FRACP, FRCPA (hon)

A thesis submitted for the degree of Doctor of Medicine in the University of Sydney

Acknowledgements

The work presented in this thesis consists of papers published over a period spanning a quarter of a century. This of course could not be an individual effort, but rather the result of working closely with many colleagues all of that time. I am grateful to so many people. It is impossible to acknowledge all those many who have contributed and have been supportive – some played a major role in the work described here, and some were only involved in small but important parts.

I should like to thank particularly those who were in charge of the laboratory operations in newborn screening, the late Andrew Brown, and subsequently my good colleague Veronica Wiley, and other newborn screeners – Ursula Bayliss, Gina Sherry, Irene Favelle-Wilson and many more. Other scientists who have been particularly helpful in this work include Judith Hammond, Kevin Carpenter, now in charge of Biochemical Genetics, Ch'ng Sim, some of whose work on fatty acid oxidation is referenced here, and Bruce Bennetts who has been greatly helpful at all times about molecular genetics issues. All my clinician colleagues in the cystic fibrosis clinic at the Royal Alexandra Hospital for Children at Camperdown, and in its later life now at Westmead have been ever supportive and helpful, as have many others who were called upon for little bits of data, cheerfully supplied.

More recently I am really indebted to some of my research colleagues, especially Marion Haas and Pamela Joy, as well as the metabolic physicians and biochemical geneticists around Australia, all of whom have collaborated in the recent studies of tandem mass spectrometry.

Over the more than 25 years that this work proceeded my husband David Wilcken has had a lot to put up with. He has always been encouraging, and has helped in so many ways, not the least as editor of many papers. Many overseas colleagues have also been great sources of information and have provided advice and useful data over the years.

So it is a team effort on many fronts, and I feel most grateful to have had the opportunity to work in this rewarding field.

Bridget Wilcken
December 2008

TABLE OF CONTENTS:

The papers W 1 to W 36 then follow in order.

INTRODUCTION

Inborn errors of metabolism are individually rare, but together comprise an important component of morbidity and mortality, mainly in the paediatric population. I have been involved in the diagnosis and management of inborn errors ever since I rejoined the medical work-force in 1973 – a period of 35 years. Newborn screening to diagnose some of these inborn errors has undergone considerable development over that time, and I have had a major research interest in the implementation, and particularly in the evaluation of outcomes of two screening programmes: for cystic fibrosis and for the inborn errors detectable by tandem mass spectrometry. The work described here was undertaken in Sydney, initially at the former Oliver Latham Laboratories, a facility directly supervised by the NSW Health Department, and subsequently at the Royal Alexandra Hospital for Children (now called the Children's Hospital at Westmead).

The aim of newborn screening is to detect babies with serious, treatable disorders early enough to facilitate appropriate interventions and thus avoid or ameliorate adverse outcomes. Mass biochemical testing of newborn babies was pioneered in the 1960s with the introduction of screening for phenylketonuria, a rare inborn error of metabolism, using dried bloodspot samples[1]. The use of a dried blood sample was key to enabling the implementation of mass screening, because of the ease of collection, and of transport to a central laboratory. The next disorder introduced into screening programmes was congenital hypothyroidism[2], and a few more much rarer disorders were included in what was initially a very gradual progress of screening. In Australia, as in most of the developed countries, almost 100% of newborns are routinely tested.

Cystic fibrosis is the commonest potentially lethal inherited disorder in Caucasian populations, with a birth incidence of around 1: 2,500. It is due to mutations in the gene encoding the cystic fibrosis transmembrane conductance regulator which forms a chloride channel pore. Impairment affects particularly the pulmonary defence mechanism and the pancreas – exocrine pancreatic insufficiency is present in around 85% of older cystic fibrosis patients. When cystic fibrosis was first fully described clinically in 1938 by Dorothy Anderson, most known affected children died before the age of two years[3]. By 1980, with advances in treatment, and perhaps an expansion of what was included in the definition of classical cystic fibrosis, the mean age of survival was thought to be about 20 years, but up to 50% of patients remained undiagnosed at 12 months of age, many receiving sub-optimal or no treatment for several years. There has long been interest in newborn screening for cystic fibrosis, as it was felt that early treatment was very likely to be beneficial. Effective screening measuring blood immunoreactive trypsin was first described in 1979[4], but was at the time very controversial, and uptake around the world was extremely slow until the last few years.

In the 1990's, the technology of tandem mass spectrometry was modified to make it suitable for newborn screening[5], and this has greatly changed the pace of development. By the mid 1990s 30-40 rare disorders of aminoacid, organic acid, and fatty acid metabolism could be detected simultaneously using a single dried blood-spot assay[6]. In

contrast to the speed at which cystic fibrosis screening was taken up, this mutiplex testing has spread very rapidly, well ahead of evidence of efficacy and clinical effectiveness.

Newborn screening deals with rare disorders. Statistically significant measures of benefit require very large pilot studies of usually several million infants; there may need to be reliance on lesser orders of evidence. There were no actual trials of screening for phenylketonuria or for hypothyroidism, the two most widely undertaken screening programmes. The evidence of benefit appeared obvious, and the era was not one where clinical trials were much conducted. In newborn screening, there have been randomised controlled trials only for cystic fibrosis[7,8]. Now several studies (but not randomised controlled trials) are beginning to establish the benefit of tandem mass spectrometry screening for disorders of fatty acid and amino acid metabolism.

I have been involved in several areas of newborn screening, and my thesis deals with two of these: – cystic fibrosis and tandem mass spectrometry. It comprises 36 papers describing work carried out between 1981 and 2008. This is arranged in sections to describe the implementation of screening and the evaluation of outcomes for each of these areas, cystic fibrosis and tandem mass spectrometry, as well as papers reviewing important general aspects of newborn screening – a review of current newborn screening, the problems of evaluation of outcome where disorders are very rare, and the ethical issues involved in screening.

I believe my major contributions to knowledge and medical practice have been:

- Fostering the recognition that the clinical diagnosis of cystic fibrosis was considerably delayed, that newborn screening was feasible in the field, and that patients with pancreatic sufficiency were readily diagnosed by newborn screening;
- The evaluation and review of different screening strategies for cystic fibrosis;
- The demonstration that identification of cystic fibrosis by screening confers a significant health benefit by greatly reducing hospitalisation in the early years, and providing nutritional and pulmonary benefits later in life;
- Initiating screening by tandem mass spectrometry in New South Wales – the first publicly-funded state-wide screening for disorders of amino acid and fatty acid metabolism – and delineating some of the disorders which could and could not be reliably detected by this technology;
- Providing the first population-based audit of the rate of diagnosis of these disorders clinically compared with diagnosis by newborn screening;
- Providing the first clear demonstration of significant clinical benefits of tandem mass spectrometry screening for the commonest disorder of fatty acid oxidation, medium-chain acyl-CoA dehydrogenase deficiency.
- Promoting the proper use of newborn screening programmes by examining the rationale for screening, ethical aspects, and what disorders should and should not be included in newborn screening programmes.

My contributions to the field have been recognised by award of the Order of Australia in 2003, and the award in 2004 of the Guthrie Medal by the International Society for Neonatal Screening, both for services to newborn screening.

STUDIES IN NEWBORN SCREENING FOR CYSTIC FIBROSIS

1) Performance of the screening test for case-finding

The measurement of electrolytes in sweat and of albumin in meconium of cystic fibrosis patients were developed in the 1950s, and in 1964 the possibility of using the latter to screen the population of newborns for this disorder was first mooted[9] but specificity and sensitivity were poor. Other proposed tests included a test for faecal trypsin[10]. We investigated the feasibility of this latter method, and found that although we could achieve a high population uptake, the sensitivity was low [11].

The immunoreactive trypsin test: In late 1979 Crossley described a blood-spot method measuring blood immunoreactive trypsin (IRT) which appeared effective in identifying newborns with cystic fibrosis[4] who, surprisingly , had high circulating levels of this protein. We investigated the method in our screening laboratory, and in 1981 we started a 3-year pilot project, funded by the New South Wales Health Department, to screen all babies in New South Wales using the IRT method on dried blood spots. This was incorporated into the existing routine newborn screening programme which at the time was testing babies for phenylketonuria and congenital hypothyroidism. We had originally thought to screen only part of the state for cystic fibrosis, using the unscreened population as a control group, but the CF clinics did not favour this approach. This was some time ago, and I feel sure we would have made different decisions today. The prospective screening of 75,000 babies was described in 1983 **[W1]**. In this paper we demonstrated the distribution of IRT values in healthy babies, and in babies with cystic fibrosis, and found the test to have apparently high sensitivity. The protocol involved obtaining a second sample for IRT analysis from babies with IRT levels above the empirical cut-off level, and the results from this second sample gave a high specificity. This paper was not the first to report a field trial of cystic fibrosis screening using an IRT test, but it was the first large-scale study, and together with an English study[12], was the most complete. I was responsible for initiating this pilot screening programme, for supervising the work and analysing the data, and for preparation of the manuscript.

Retrospective analysis of delayed clinical diagnosis: At the same time I had also instigated a retrospective study of the delay in diagnosis of cystic fibrosis in New South Wales in the three years immediately before screening started. There had been only a few publications about age at diagnosis, largely from North America, but including one from Victoria[13]. We thought it important to document the apparent delay between onset of symptoms and clinical diagnosis in our state before screening, and to record the occurrence of symptoms at diagnosis or soon thereafter in the babies detected by newborn screening. Our data showed that the delay in clinical diagnosis was far greater than previously supposed, and that a majority of babies detected by screening already had some symptoms which warranted treatment. A colleague abstracted information from the notes of patients diagnosed clinically, and I was responsible for the preparation of the

manuscript for publication. This paper **[W2]** was abstracted for the Year Book of Pediatrics 1984.

Early justification of screening: In 1983 the USA Cystic Fibrosis Foundation Ad Hoc Committee Task Force on Neonatal Screening had published a critique of screening, exploring a number of hypothetical reasons why there might be more harm than good in such an enterprise[14], and others had also written similarly, on theoretical grounds[15,16]. I was invited to address these issues at a Canadian conference. The proceedings were published as a book. One chapter, **[W3]** attempts an evaluation in the light of experience we had gained in the first 5 years. This chapter includes the first data showing the decline of the IRT level with age of the baby, and also the results of a survey we had conducted on parental attitudes to screening, showing overwhelming support for screening from parents of both screened and unscreened babies. We later addressed the possible causes of false negative results. No-one had previously written on this, and I presented our data at an international conference, which was published in book form as part of the proceedings **[W4]**.

The screening detects patients with pancreatic sufficiency: Two papers, **[W5 and 6]**, report the work of colleagues who investigated pancreatic function in neonates identified by the screening programme by formal pancreatic stimulation tests. I was involved in the conception of the studies, I analysed the data on the results of the screening tests, and assisted in the preparation of the manuscripts. The work was important because it demonstrated not only that some babies with pancreatic sufficiency develop insufficiency with time, but also because the data unequivocally supported our earlier observation that the IRT test diagnosed not only pancreatic insufficient babies, but also babies with pancreatic sufficiency, something which earlier writers had denied would be the case.

A review of newborn screening for cystic fibrosis: In 1993, by invitation, I reviewed the situation of screening for cystic fibrosis **[W7].** This paper, published in a journal devoted to screening, reviewed all the literature to that time, and concluded screening test performance was satisfactory, and that there was good evidence of short-term benefit, with little if any evidence of harm. When this was published, two randomised controlled trials of screening were proceeding, but only very preliminary data had been published[8].

The use of a protocol incorporating DNA analysis: In 1989 the CF gene was cloned, and the common mutation ΔF508 was described[17]. This made a new approach to screening available[18, 19], and from the beginning of 1993 we made use of this approach. Initially we had used only IRT measurements, which necessitated calling for a second blood sample from approximately 0.6% of babies, because of the poor predictive value of IRT measurements at 3-5 days of age, when newborn screening tests are done. From 1993, we performed mutation analysis for the common mutation on blood samples in which the IRT level was in the highest 1% of values. Babies with two copies of the common mutation had cystic fibrosis, but babies with only one copy might either be compound heterozygotes or merely carriers, and a sweat test was needed to distinguish these. This

protocol did away with the need for second blood samples, and dramatically increased the specificity of the test.

We published the results of the first two years of the new protocol, and compared the two protocols, in a paper which detailed results from 1981-92, and 1993-95 **[W8]**. We recorded several new findings. As we had actively sought missed cases, we were able to give an accurate estimate of the birth prevalence in New South Wales of "classic" CF, taking a period of 1981-1984, before prenatal diagnosis was available, and when members of the cohort were at least 10 years old, and therefore likely to have been diagnosed clinically if missed by the screening tests. We were also able to document much more accurately the performance of the IRT/IRT protocol. While it was then too early to be sure about the performance of the DNA protocol, this is much more amenable to theoretical prediction, based on knowledge of the distribution of genotypes in this state. For this paper I did most of the data analysis, and was responsible for the preparation of the manuscript. This paper was abstracted for the Year Book of Pediatrics 1997. By this time most of Australian neonates were screened for cystic fibrosis, and we were invited to write a book chapter reviewing the use of DNA testing in newborn screening for cystic fibrosis **[W9]**.

During the analysis of these data we noted the over-representation of apparent carriers among the babies with one copy of the common mutation but normal sweat tests. This led to a study of the pancreatic function and molecular genetics in these babies, described in paper **W10**. We found that indeed some of the babies had two CF mutations, and thus had mild CF, of as yet unknown clinical significance, and that over half had decreased pancreatic function. This is a novel contribution of considerable interest. I was involved in developing the concept of this investigation, in carrying out parts of it and in the writing of the manuscript.

Strategies for newborn screening for cystic fibrosis: There has been recent interest in the different strategies in use for cystic fibrosis screening. All use IRT assays, and most but not all incorporate some DNA analysis. We have not investigated other strategies than those reported in our Journal of Pediatrics paper of 1995, but I have recently reviewed all the other strategies currently in use **[W11 ,12]**.

2) The clinical outcome of early detection of cystic fibrosis by newborn screening

The only justification of a screening programme is if it is associated with clinical benefit, without too much harm. We were interested in the outcome of screening from the start, but as described had not been able to have a trial with a contemporaneous control group – we were left with the careful use of historical controls. Groups in Wisconsin and in the United Kingdom were both able to start randomised controlled trials in 1985, the only such trials so far ever conducted of any newborn screening[7,8].

Study of early morbidity: Early on we investigated morbidity, as indicated by days of hospitalisation, in the first two years of life, (excluding the immediate neonatal period), in patients diagnosed in the first two years of screening, using as controls those who were

born in the three years immediately before screening began. Although this study used historical controls, there were several lines of evidence indicating that that these two groups were comparable, and we carefully examined and recorded management changes which had taken place during the period, but which affected only a minority of the patients. In the event, we found a major difference between the two groups, with a substantial reduction in hospital days for the screened patients, from an average of 27 days for the clinically diagnosed to 3 days for the screened. There was no trend with time, the change having occurred suddenly, coincident with the start of screening. Patients who had presented with meconium ileus and so were diagnosed early, but were severely affected, formed a type of internal control. These patients had an intermediate number of days of hospitalisation, 16, in each cohort. I was responsible for the concept of the study, data analysis and writing of the paper **[W13]**, which was the first study with a control group to show clear short-term benefit from early diagnosis.

Growth and lung function: Later we were able to extend the study of the outcome of newborn screening for CF using the same cohorts - those born in the first three years of screening and the three years immediately before this – to look at growth and lung function at ages 1, 5, and 10 years. Again this study had the problems of being a historical cohort study, but we have carefully addressed these issues. We believe that the groups are comparable and that the results are valid and show a consistent advantage for the screened group. A strength of this study is the large number in each cohort, 41 and 57 at 10 years, and the degree of apparent advantage in respiratory status at 10 years for the screened patients. The work for this study was performed by a colleague, but the original concept was mine, and I was involved in the preparation of the manuscript **[W14]**. The importance of this work has been emphasised, rather than detracted from, by the prior publication of outcome data from one of the only two randomised controlled trials. This was a ten year follow-up, but had only 4 and 9 patients in the two arms at 10 years, and had a great deal of missing data, which compromised the conclusions drawn[20]. A further study in which I was not involved showed that the benefit to our same screened cohort extended to age 15 years[21].

The work presented in the papers described here has been influential in the developed world in demonstrating the feasibility of screening for cystic fibrosis in the newborn, and for exploring the apparent advantages and disadvantages. The papers have been widely cited in the literature and two have been abstracted for the Year Book of Pediatrics. As a result of these studies I have been invited to present work at conferences in Australia, North America and Europe, to the World Health Organisation, and to the Centre for Diseases Control in Atlanta, Georgia. The feasibility and outcomes documented in the results of the studies presented here, together with the findings of others, have had a significant impact on public health policy, despite surprising delays **[W15]**. All babies born in Australasia are now tested for cystic fibrosis, and in the last few years screening has been uniformly recommended in the United States, has been implemented throughout the United Kingdom, and is widespread in Europe and becoming so in Canada.

STUDIES IN NEWBORN SCREENING BY TANDEM MASS SPECTROMETRY FOR DISORDERS OF AMINO ACID AND FATTY ACID METABOLISM

1) Performance of the MSMS test in diagnosis

Although tandem mass spectrometry (MS/MS) was developed over 25 years ago, it could only be applied to mass newborn screening in the mid-1990's after the development of electrospray ionisation and associated automation of sample handling. This, together with automated data analysis, allowed a sufficiently fast assay to handle numbers needed for population screening[5]. The technology introduced a huge change to newborn screening: it enabled a single assay on a 3mm dried blood spot to detect markers for over 30 rare disorders of amino acid, organic acid and fatty acid metabolism simultaneously, in about two minutes per sample. These groups of disorders were serious, often potentially fatal disorders with recognised treatment which was thought to improve the outcome. In newborn screening carried out previously, each disorder had needed a separate assay, and extremely rare disorders were therefore not even considered for newborn screening.

Development of screening in New South Wales: MSMS applied to newborn screening was put into practice first by a private laboratory in the United States (although much preliminary work had been done elsewhere) and their first results were reported at a conference in 1996, and later elsewhere[6]. At that 1996 meeting I decided that we needed to embrace this technology in New South Wales. We had expertise in the diagnosis and management of the inborn errors then detectable by MSMS screening, and there was much evidence that presymptomatic diagnosis would be likely to be of benefit in many of the disorders. There was at this time no formal evidence of benefit from trials. We managed to obtain the necessary equipment, and started screening in 1998. Our screening programme was the first in Australia, and was in fact the first state-wide publicly funded programme in the world. After our initial report of the first two years of screening **[W16]** we reported on the potential to detect several rare disorders, and also provided data showing that certain others could not be reliably detected. We documented the first detection of the carnitine transporter defect **[W 17]** and carnitine palmitoyltransferase type I deficiency **[W18]**, and later the inability to detect non-ketotic hyperglycinaemia, the congenital lactic acidoses, and mild maple-syrup urine disease **[W19 – 21]**. These latter negative findings were as important as positive ones in exploring the potential of this new technology. By 2003 we were able to report more fully.

Rates of clinical detection versus detection by screening: In 2003 we published a major report in the New England Journal of Medicine **[W22].** In this we compared the rates of clinical detection of 31 disorders in New South Wales over six 4-year periods from 1974 to 1998 with detection rates during four years of screening, 1998 to 2002. We showed that the detection rate had remained steady until newborn screening began when, over-all, it had doubled, although the increased diagnosis rate was confined to certain disorders. The detection rate for medium-chain acyl-CoA dehydrogenase (MCAD) deficiency, the commonest disorder of fatty acid oxidation, and for fatty acid oxidation disorders overall had approximately tripled. This paper was important for several reasons. This was the

first paper showing the extent of "extra" diagnoses made by newborn screening, and in addition, we had data on diagnoses in New South Wales going back for 24 years, which we believed to be complete (because only our facility could confirm the diagnoses of these disorders). There had then been only one previous report of similar data on an unscreened population[22], and none that were able to compare this to screened groups. This was also the first intimation of how hard it would be to show significant benefit from observational studies, or even randomised controlled trials, because of the likely non-comparability of groups.

Biochemical and molecular features in MCAD deficiency: We conducted specific studies of screening in New South Wales evaluating the *diagnosis* of MCAD deficiency by tandem mass spectrometry, which included a description of the biochemical and molecular features of patients detected **[W 23, 24]**. These studies too were important, as they were the first to compare systematically the biochemical and molecular features of the disorder, and they were additionally valuable, as we had closely defined what we accepted as MCAD deficiency. While this may seem obvious, it is not so, as most inborn errors of metabolism occupy a spectrum from clinically affected cases, through mildly affected, to clinically unaffected cases with only polymorphisms detected in the gene in question. In **W 24** we showed some correlation between functional (biochemical) disturbance and the spectrum of mutations in the ACADM gene, associated with MCAD deficiency.

2) Evaluation of the outcome of MSMS screening

Problems in evaluation: There are obvious difficulties in evaluating outcomes of interventions for rare diseases, especially when there is a preconceived idea of benefit, and I outlined these with particular reference to newborn screening in an invited presentation in Cambridge in 2000.**[W 25]**. Following on from our early publications about tandem mass spectrometry screening I was invited to chair a symposium on this screening in Paris in 2005 **[W 26]** and we also collaborated in gathering data for this symposium on outcomes for other disorders **[W 27]**.

Australia-wide study of MSMS screening. Outcome for MCAD deficiency: In 2003 we started an Australia-wide evaluation of MSMS screening, comparing morbidity, neuropsychological outcome and costs of a cohort of screened patients born in New South Wales and South Australia between 1998 and 2002, with two unscreened cohorts, one Australia-wide historical cohort (born 1994 to 1998), and one contemporaneous, born in the remainder of Australia from 1998 to 2002. I was the chief investigator of this NH&MRC - funded study. Firstly we concentrated on MCAD deficiency, the disorder most frequently found by MSMS newborn screening, after phenylketonuria. This is a fatty acid oxidation disorder which causes hypoketotic hypoglycaemia during catabolic stress. In published series of clinically diagnosed patients, approximately a quarter had died, usually before the diagnosis was made, and around 20% of the survivors suffered brain-damage, apparently secondary to the hypoglycaemia[23]. We assessed the occurrence of a severe event of metabolic decompensation or death by age 4 years, and we conducted neuropsychological studies on survivors. I directed the whole study, and authored the

Lancet paper which described the medical outcome **[W28]**. This was the first paper to show clearly a statistically significant advantage to the screened patients when compared to unscreened cohorts, on a population basis, at 2 and 4 years of age, as was commented on in the same issue[24]. A paper from Germany had preceded this[25] but with the comparison (unscreened) group taken from a small retrospective study conducted in another country[26], so that the populations were not necessarily comparable. In our Australia-wide study we also investigated the neuropsychological outcome in our patients, finding, surprisingly, no major overall difference, and we also evaluated costs. These studies were led by others in our research group, and detailed reports of the economic evaluation and the neuropsychological study were published separately **[W29, 30]**.

Ongoing work: I have ongoing work on the outcomes of screening by MSMS. In the Australia-wide study of screening, we have examined the economics overall of MSMS screening [27], and will soon be able to report on clinical outcomes at the age of 6 years of the whole spectrum of disorders currently screened for by MSMS. In NSW, we recently evaluated screening for another rare inborn error of metabolism, glutaric aciduria type I **[W31]**. Without early diagnosis, this disorder usually results in devastating disability following a "metabolic stroke", and we showed (although not the first to do so[28]) that early diagnosis by screening enables this to be avoided in most cases.

The work described here has been recognised world-wide, and we are considered to have a leading place in the field of tandem mass spectrometry screening for rare inborn errors of metabolism.

REVIEWS OF NEWBORN SCREENING

I have also been invited to contribute articles of a general nature about newborn screening, together with a book chapter. I believe these reviews round out the thrust of this thesis.

In 2003 I had addressed the ethical issues around newborn screening, particularly in relation to the new testing by MSMS **[W 32].** More recently I reviewed the advances in newborn screening, mainly by MSMS screening **[W33]**, and the very difficult subject of the detection of cases of mild disease identified during a newborn screening programme, and the definition of which cases actually need treatment **[W 34]**. This is a major issue for multiplex screening by MSMS, which is only now beginning to be accepted as a major problem world-wide. I have contributed a chapter to a standard text-book which focuses mainly on MSMS screening **[W35]**, and this year, a colleague and I reviewed newborn screening as a whole, particularly in reference to Australasia **[W 36].**

These and some other general papers I believe have influenced thinking around the world about which disorders ought to be included in newborn screening programmes. The whole of newborn screening has become a more mainstream field in the last few years, and I believe the work here has contributed to the understanding of the current and future challenges. That our work is well recognised is attested to by the number of invitations (several per year) to address international conferences on the subjects discussed here.

REFERENCES

1. Guthrie R, Susi A. A simple phenylalanine method for detecting phenylketonuria in large populations of newborn infants. Pediatrics 1963; 32:338-343.

2. Dussault JH, Coulombe P, Laberge C, Letarte J, Guyda H, Khoury K. Preliminary report on a mass screening program for neonatal hypothyroidism. Journal of Pediatrics 1975; 86(5):670-674.

3. Anderson DA. Cystic fibrosis of the pancreas and its relation to coeliac disease. Am J Dis Child 1938; 56:344

4. Crossley JR, Elliott RB, Smith PA. Dried blood spot screening for cystic fibrosis in the newborn. Lancet 1979; I: 472-4.

5. Millington DS, Kodo N, Norwood DL, Roe CR. Tandem mass spectrometry: a new method for acylcarnitine profiling with potential for neonatal screening for inborn errors of metabolism. J Inherit Metab Dis. 1990;13(3):321-4.

6. Naylor EW, Chace DH. Automated tandem mass spectrometry for mass newborn screening for disorders in fatty acid, organic acid, and amino acid metabolism. J Child Neurol. 1999;14 Suppl 1:S4-8.

7. Mischler E, Farrell P, Bruns T, et al. Neonatal screening for cystic fibrosis in Wisconsin.Wis Med J. 1989 Mar;88:14-8.

8: Chatfield S, Owen G, Ryley HC, et al. Neonatal screening for cystic fibrosis in Wales and the West Midlands: clinical assessment after five years of screening. Arch Dis Child. 1991;66:29-33.

9. di Sant'Agnesi PA. The Pancreas. In: Vaughan VC, McKay, RJ eds. Nelson Textbook of Pediatrics, WB Saunders Company, Philadelphia; 11th edition; 1980

10. Crossley JR, Berryman CC, Elliott RB. Cystic fibrosis screening in the newborn. Lancet 1977; ii: 1093-5.

11. Forrest DC, Wilcken B, Turner G. Screening for cystic fibrosis - evaluation of a stool trypsin method. Archives of Disease in Childhood 1981;56:1651-52

12. Heeley AF, Heeley ME, King DN, Kuzemko JA, Walsh MP. Screening for cystic fibrosis by dried blood spot trypsin assay. Arch Dis Child 1982: 57: 18-21.

13. Phelan PD. Screening for cystic fibrosis (commentary). Arch Dis Child 1982; 57: 779-80.

14. Ad Hoc Committee Task Force on Neonatal Screening. Cystic Fibrosis foundation. Neonatal screening for cystic fibrosis: position paper. Pediatr 1983; 72: 41-45.

15. Phelan PD. Screening for cystic fibrosis In: Meadow R, ed. Recent Advances in Pediatrics. Edinburgh, Churchill Livingstone, 1984; 103-20.

16. Farrell PM. Early diagnosis of cystic fibrosis: to screen or not to screen - an important question.Pediatr 1984; 73: 115-7.

17. Kerem BS, Rommens JM, Buchanan JA et al. Identification of the cystic fibrosis gene: genetic analysis. Science 1989; 245: 1073-80.

18. Ranieri E, Ryall RG, Morris CP et al. Neonatal screening strategy for cystic fibrosis using immunoreactive trypsinogen and direct gene analysis. BMJ 1991; 302: 1237-40.

19. Laroche D, Travert G. The application of PCR amplification and the polymorphic marker KM.19 to dried blood spots: comparison with deletion 508 for the conformation of the neonatal screening test for cystic fibrosis. Pediatr. Pulmonol. 1991; Suppl 7: 19-22.

20. Farrell PM, Kosorok MR, Laxova A, et al. Nutritional benefits of neonatal screening for cystic fibrosis. N Engl J Med 1997; 337: 963-9.

21. McKay KO, Waters DL, Gaskin KJ. The influence of newborn screening for cystic fibrosis on pulmonary outcomes in New South Wales. J Pediatr. 2005 Sep;147(3 Suppl):S47-50.

22. Applegarth DA, Toone JR, Lowry RB. Incidence of inborn errors of metabolism in British Columbia, 1969-1996.Pediatrics. 2000 Jan;105:e10.

23. Wilson CJ, Champion MP, Collins JE, Clayton PT, Leonard JV. Outcome of medium chain acyl-CoA dehydrogenase deficiency after diagnosis. rch Dis Child. 999;80:459-62.

24. Grosse SD, Dezateux C. Newborn screening for inherited metabolic disease.Lancet. 2007;369(9555):5-6.

25. Nennstiel-Ratzel U, Arenz S, Maier EM, et al. Reduced incidence of severe metabolic crisis or death in children with medium chain acyl-CoA dehydrogenase deficiency homozygous for c.985A>G identified by neonatal screening.Mol Genet Metab. 2005 Jun;85(2):157-9.

26: Pourfarzam M, Morris A, Appleton M, Craft A, Bartlett K. Neonatal screening for medium-chain acyl-CoA dehydrogenase deficiency. Lancet. 2001;358(9287):1063-4.

27. Norman R, Haas MR, Chaplin M, Joy P, Wilcken B. Economic evaluation of Tandem Mass Spectrometry Newborn Screening in Australia, *Pediatrics*, 2009; in press.

28. Kölker S, Garbade SF, Boy N, et al. Decline of acute encephalopathic crises in children with glutaryl–CoA dehydrogenase deficiency identified by newborn screening in Germany. Pediatr Res 2007; 6

PAPERS TO BE SUBMITTED FOR THE DEGREE OF DOCTOR OF MEDICINE: W 1-36

W1. Wilcken B, Brown ARD, Urwin R, Brown DA. Screening for cystic fibrosis by blood spot immunoreactive trypsin: Results in 75,000 neonates.. Journal of Pediatrics 1983; 102: 383-88

W2. Wilcken B, Towns SJ, Mellis CJ. Diagnostic delay in cystic fibrosis: lessons from newborn screening. Archives of Disease in Childhood 1983; 58:863-6

W3. Wilcken B. An evaluation of screening for cystic fibrosis. In: Genetics and Epithelial Cell Dysfunction in Cystic Fibrosis. M. Buchwald and J. O'Riordan, Eds., Alan R. Liss Inc. New York, 1987; 201-15. (also Prog Clin Biol Res. 1987;254:201-15).

W4. Wilcken B, Brown ARD. An analysis of false negative screening tests for cystic fibrosis. In: Current Trends in Infant Screening B.J. Schmidt, A.J. Diament and N.S. Loghin-Grosso, Eds., Excerpta Medical, Amsterdam 1989; 19-25.

W5. Waters DL, Dorney SFA, Gaskin KJ, Gruca MA, O'Halloran M, **Wilcken B**. Pancreatic function in infants with cystic fibrosis from a neonatal screening program New England Journal of Medicine 1990; 322:303-08

W6. Gaskin K, Waters D, Dorney S, Gruca M, O'Halloran M, **Wilcken B**. Assessment of pancreatic function in screened infants with cystic fibrosis. Pediatr. Pulmonol 1991; suppl 7: 69-71.

W7 Wilcken B. Newborn screening for cystic fibrosis. Its evolution, and a review of the current situation. Screening; 2: 43-62.

W8. Wilcken B, Wiley V, Sherry G, Bayliss U. Newborn Screening for Cystic Fibrosis: a comparison of two strategies for case detection in 1.2 million babies. J.Pediatr 1995; 127:965-72.

W9. Ranieri E, Lewis BD, Morris CP, **Wilcken B**. Neonatal Screening Using Combined Biochemical and DNA-based Techniques. In: Cystic Fibrosis, Current Topics volume 3. Dodge JA, Brock DJH, Widdicombe J Eds. Chichester, New York, John Wiley and Sons, 1995; 181-206.

W 10. Massie RJ, **Wilcken B**, Van Asperen P, Dorney S, Gruca M, Wiley V, Gaskin K. Pancreatic function and extended mutation analysis in DeltaF508 heterozygous infants with an elevated immunoreactive trypsinogen but normal sweat electrolyte levels. J Pediatr. 2000;137:214-20.

W 11. Wilcken B, Wiley V. Newborn screening methods for cystic fibrosis. Paediatr Respir Rev. 2003;4:272-7. Review.

W 12. Wilcken B. Newborn screening for cystic fibrosis: techniques and strategies. J Inherit Metab Dis. 2007 Aug;30(4):537-43. Epub 2007 May 12.

W13. **Wilcken B**, Chalmers G. Reduced morbidity in patients with cystic fibrosis detected by neonatal screening. Lancet 1985; ii:1319-21.

W 14. Waters DL, **Wilcken B**, Irwing L, Van Asperen P, Mellis C, Simpson JM, Brown J, Gaskin KJ. Clinical outcomes of newborn screening for cystic fibrosis. Arch Dis Child Fetal Neonatal Ed. 1999;80:F1-7.

W 15. **Wilcken B**, Gaskin K. More evidence to favour newborn screening for cystic fibrosis. Lancet. 2007 Apr 7;369(9568):1146-7.

W 16: Wiley V, Carpenter K, **Wilcken B**. Newborn screening with tandem mass spectrometry: 12 months' experience in NSW Australia. Acta Paediatr Suppl. 1999; 88(432):48-51.

W 17: Wilcken B, Wiley V, Sim KG, Carpenter K. Carnitine transporter defect diagnosed by newborn screening with electrospray tandem mass spectrometry.
J Pediatr. 2001; 138:581-4.

W 18. Sim KG, Wiley V, Carpenter K, **Wilcken B**. Carnitine palmitoyltransferase I deficiency in neonate identified by dried blood spot free carnitine and acylcarnitine profile.J Inherit Metab Dis. 2001;24:51-9.

W 19. Tan ES, Wiley V, Carpenter K, **Wilcken B**. Non-ketotic hyperglycinemia is usually not detectable by tandem mass spectrometry newborn screening. Mol Genet Metab. 2007; 90:446-8.

W 20. Sim KG, Carpenter K, Hammond J, Christodoulou J, **Wilcken B**.
Acylcarnitine profiles in fibroblasts from patients with respiratory chain defects can resemble those from patients with mitochondrial fatty acid beta-oxidation disorders. Metabolism. 2002;51:366-71.

W 21. 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: 586

W 22. Wilcken B, Wiley V, Hammond J, Carpenter K. Screening newborns for inborn errors of metabolism by tandem mass spectrometry. N Engl J Med. 2003; 348:2304-12.

W 23. Carpenter K, Wiley V, Sim KG, Heath D, **Wilcken B**. Evaluation of newborn screening for medium chain acyl-CoA dehydrogenase deficiency in 275 000 babies.
Arch Dis Child Fetal Neonatal Ed. 2001;85: F105-9.

W 24. Waddell L, Wiley V, Carpenter K, Bennetts B, Angel L, Andresen BS, **Wilcken B.**
Medium-chain acyl-CoA dehydrogenase deficiency: genotype-biochemical phenotype correlations. Mol Genet Metab. 2006 ;87 :32-9.

W 25. Wilcken B. Rare diseases and the assessment of intervention: what sorts of clinical trials can we use? J Inherit Metab Dis. 2001; 24:291-8.

W 26. Wilcken B. Mini-symposium: newborn screening for inborn errors of metabolism—clinical effectiveness. J Inherit Metab Dis. 2006; 29:366-9.

W 27. Dionisi-Vici C, Deodato F, Röschinger 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. J Inherit Metab Dis. 2006; 29: 383-9.

W 28. Wilcken B, Haas M, Joy P, Wiley V, Chaplin M, Black C, Fletcher J, McGill J, Boneh A. Outcome of neonatal screening for medium-chain acyl-CoA dehydrogenase deficiency in Australia: a cohort study. Lancet. 2007 Jan 6;369(9555):37-42.

W 29. Haas M, Chaplin M, Joy P, Wiley V, Black C, **Wilcken B**. Healthcare use and costs of medium-chain acyl-CoA dehydrogenase deficiency in Australia: screening versus no screening. J Pediatr. 2007 Aug;151(2):121-6, 126.

W 30. Joy P, Black C, Rocca A, Haas M, **Wilcken B.** Neuropsychological functioning in children with medium chain acyl-CoA dehydrogenase deficiency (MCADD): The impact of early diagnosis and screening on outcome. Child Neuropsychol. 2008 Apr 3:1-13.

W 31. Bijarnia S, Christodoulou J, Ellaway C, Wiley V, Carpenter K, **Wilcken B.** Glutaric aciduria type I: outcome following detection by newborn screening. J Inher Metab Dis 2008; In press.

W 32. Wilcken B. Ethical issues in newborn screening and the impact of new technologies. Eur J Pediatr. 2003; 162 Suppl 1:S62-6.

W 33. Wilcken B. Recent advances in newborn screening. J Inherit Metab Dis. 2007; 30:129-33.

W 34. Wilcken B. The consequences of extended newborn screening programmes: Do we know who needs treatment? J Inherit Metab Dis. 2008; 31: 173-7

W 35. Wilcken B. Newborn screening for inborn errors of metabolism In: Fernandez J, Saudubray J-M, van den Berghe G, Walter J. Inborn Metabolic Diseases. 4th edition,. Springer Medizin Verlag, Heidelberg, 2006. pp 49-57

W 36. Wilcken B. Wiley V. Newborn screening. Pathology. 2008; 40:104-15

PAPERS ON NEWBORN SCREENING FOR CYSTIC FIBROSIS

Cystic fibrosis screening by dried blood spot trypsin assay: Results in 75,000 newborn infants

Seventy-five thousand 5-day-old babies were screened for cystic fibrosis by blood spot immunoreactive trypsin (IRT) assay as part of a statewide screening program.

IRT was elevated in 433 babies; retesting revealed persistent elevation in 38. Sweat testing confirmed cystic fibrosis in 35 babies and was normal in two babies, whose IRT remained elevated at the time of the test. Sweat testing was refused by one mother. Of the 35 babies with cystic fibrosis, 13 had meconium ileus or an already diagnosed affected sibling, but the diagnosis was unsuspected in 22, although all but four had some symptoms suggestive of cystic fibrosis. Stool trypsin activity at the time of the diagnostic screen was normal in nine and reduced in seven of the babies with cystic fibrosis. One baby did not have elevated IRT, and the cystic fibrosis was missed by the screening test. In a retrospective study of blood spot samples from 36 newborn infants, who were later diagnosed as having cystic fibrosis, all had IRT levels greater than in matched controls. Our study confirms that elevated IRT is characteristic of newborn babies with cystic fibrosis, and shows that this test is very specific and sensitive when used as a newborn screening test. (J PEDIATR 102:383, 1983)

Bridget Wilcken M.B., Ch.B, A. R. D. Brown B.Sc, M.A.A.C.B., Ruth Urwin, and D. A. Brown, A.S.T.C. *Sydney, Australia*

SCREENING IN THE NEWBORN POPULATION for cystic fibrosis has been controversial for two reasons: (1) Described methods have employed meconium or feces and have had high false-positive and false-negative rates, in addition to many other drawbacks[1,2]; (2) More important, there is no general agreement about the benefit to the infant or the family of the very early diagnosis of cystic fibrosis.

In 1979 Crossley et al.[3] demonstrated that blood immunoreactive trypsin (IRT) concentrations are elevated in newborn babies subsequently shown to have cystic fibrosis, and that immunoreactive trypsin could be assayed in dried blood spot samples collected routinely for other neonatal screening tests. Further study indicated that the test could distinguish between patients with cystic fibrosis and age-matched controls, even when the patient had adequate pancreatic function.[4]

Small pilot studies showed that this new assay could be used for neonatal screening,[4-6] and there have been recent reports of the results of screening in several larger series of infants.[7-11]

We describe a newborn infant screening program for detecting cystic fibrosis in New South Wales, Australia, using a modification of Crossley's radioimmunoassay for IRT, and report the results of screening 75,000 babies.

METHODS

The samples used were the dried blood specimens received in this laboratory for current screening programs from all 5-day-old babies born in New South Wales. Similar samples collected in the past three years from 36 babies subsequently diagnosed as having cystic fibrosis were retrieved from our files, together with control samples for each from four apparently normal babies born in the same hospital on the same day or within one or two days.

IRT was assayed in blood contained on a 3 mm disk punched by an automatic multipunch machine from the dried blood spots. Purified ^{125}I trypsin, human trypsin antiserum, and trypsin standards were obtained from Elliott and Crossley's laboratory (Department of Pediat-

From Oliver Latham Laboratory, Health Commission of New South Wales.

Reprint address: Dr. Bridget Wilcken, Oliver Latham Laboratory, P.O. Box 53, North Ryde, 2113 N.S.W., Australia

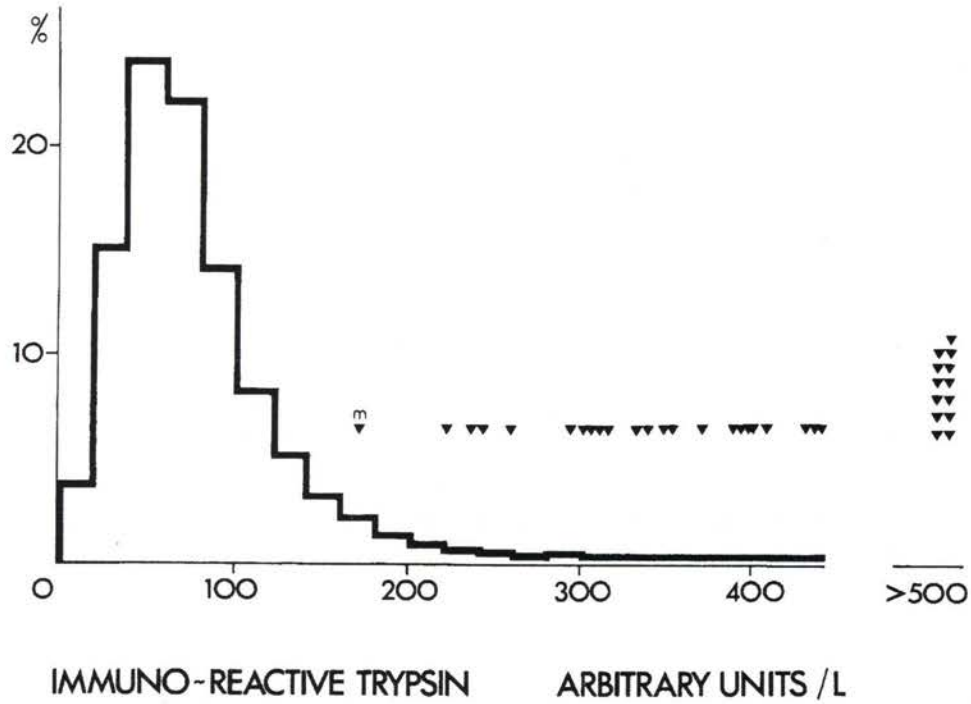


Fig. 1. Distribution of IRT values in dried blood samples from 10,244 neonates. ▼, IRT values in samples from neonates with cystic fibrosis detected by screening 75,000 babies. *m*, cystic fibrosis missed by screening.

rics, Auckland University, New Zealand). As there was no recognized standard human trypsin against which the preparation could be assayed, Crossley expressed the activity assayed as Arbitrary Units (AU), where one AU represents the immunoreactivity of 1 μg purified trypsin extract,[4] and we have adopted this terminology.

Batch size varied from approximately 250 to 500. When IRT values fell into the top 2% of the batch, or exceeded 200 Arbitary Units/1 (AU/1), the samples were reassayed in duplicate, using hand-punched 3 mm disks of blood (so that an evenly distributed portion of the blood spot could be selected). When the mean IRT value exceeded 220 AU/1, a second heel prick sample was requested; this was assayed in duplicate. If the IRT value was again above 220 AU/1, a sweat test was requested.

At the time of the sweat test, filter-paper stool samples were obtained from 16 of the babies and assayed for stool trypsin activity according to the method of Crossley et al.[12]

Details of the duration of pregnancy, birth weight, feeding, medical problems, and medication during the first week of life were sought from physicians of 110 consecutive babies whose initial IRT value was raised but whose repeat blood sample had a normal IRT value.

RESULTS

The distribution of IRT values in 10,000 neonates is shown in Fig. 1. The assay had a within-run coefficient of variation of 12.4%, calculated from 90 duplicate samples, with a range of values of 196 to 465 AU/1. The between-run coefficient of variation was 14.2% at both 80 and 280 AU/1 (n = 53) using prepared control samples, and was 12% at 452 AU/1 (n = 16) using patient samples.

Initial elevation of IRT value, above 220 AU/1, was seen in 0.58% of the total sample of 75,000 infants. Of 433 second samples received, 38 had persistently elevated IRT values, and in 35 a diagnosis of cystic fibrosis was confirmed by sweat testing. In two infants the sweat test was negative, although they continued to have elevated blood IRT at 8 months and 2 months, respectively. In the remaining baby whose IRT was persistently elevated, a sweat test was not conducted successfully, and repeat testing was refused by the mother. The stool trypsin activity in 16 of the babies with newly diagnosed cystic fibrosis was normal in nine and reduced in seven. The IRT values of the patients with cystic fibrosis are shown in Fig. 1.

Of the 35 babies with confirmed cystic fibrosis, the diagnosis would have been made early in 13, because of

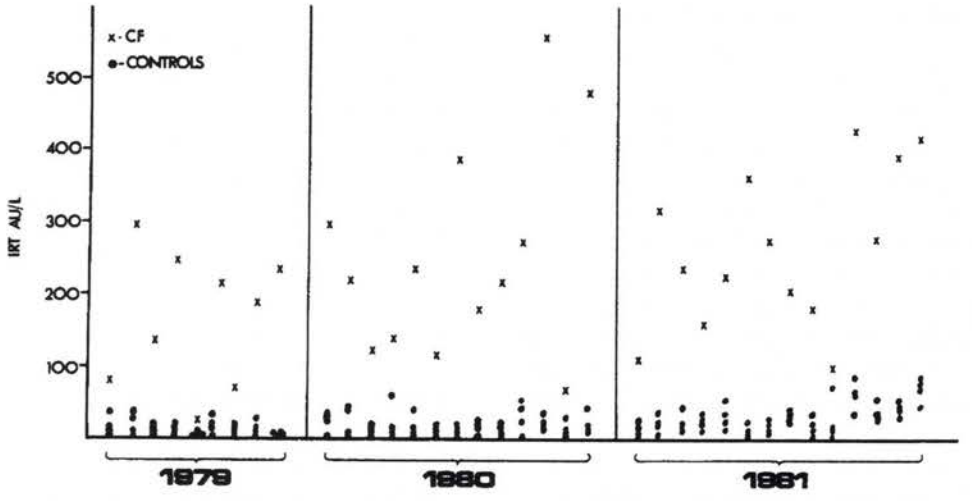


Fig. 2. IRT values in a retrospective study of neonatal blood samples from patients diagnosed as having cystic fibrosis *(X)* and from four controls for each patient (•). Control blood samples were from babies born in the same hospital within two days of each patient. All blood samples had been stored for a few days at the hospital of birth, under unknown conditions, and in the laboratory at room temperature for up to three and one-half years.

presence of meconium ileus[7] a sibling with previously diagnosed cystic fibrosis,[4] or both.[2] In the remaining 22 patients the diagnosis was quite unsuspected, and indeed four infants were entirely asymptomatic at the time of diagnosis (although one developed symptoms soon thereafter), whereas the other babies had some slight symptoms or signs consistent with cystic fibrosis. One baby had an older sibling with severe chest symptoms and previously undiagnosed cystic fibrosis.

One case is known to have been missed by screening. The baby had had meconium aspiration, and still required oxygen therapy when the routine blood sample was taken for screening. Her IRT level was in the top 2% of the batch, but subsequent duplicate assays were below the cutoff point, and no further action was taken. The baby developed bronchiolitis at 8 weeks, and sweat electrolyte concentrations were elevated. However, the blood IRT was then only 54 AU/1. Another baby with cystic fibrosis, who was born during this period, was never screened because of critical illness in the neonatal period.

Fig. 2 shows the results of a retrospective study of blood samples from newborn babies subsequently found to have cystic fibrosis. Although the absolute values of IRT may have declined with long storage of the sample, the level was in each instance higher than levels in four sample age-matched and hospital-matched controls.

Of 70 babies whose initial IRT was elevated, but whose second test was normal, 67 weighed more than 2,500 gm at birth and 58 had had no significant medical problem in the neonatal period. Among the total of 433 babies with initial elevation of IRT, three had Hirschsprung disease and one had gastroschisis.

The cost of the screening test in May 1982 was about $0.70 per infant screened, for technical and secretarial staff, consumables, and depreciation of equipment. The cost of the reagents accounted for about $0.15 per infant.

DISCUSSION

The results of our retrospective study confirm the findings of Crossley et al.[4] that IRT elevation is characteristic of newborn babies with cystic fibrosis. Combining the two studies, which were conducted comparably, every one of the 60 children with cystic fibrosis whose neonatal blood sample could be retrieved from files had had blood IRT levels in the neonatal period that were clearly distinguishable from appropriately matched controls. King et al.[13] had similar findings in six of seven neonates with cystic fibrosis. However, the recovery of IRT from dried blood spots appears to decrease in a variable way on storage, so it is not certain what percentage of these patients would have been identified by a screening program using the protocol we describe. Our study also confirms that blood IRT values are elevated even in those patients with cystic fibrosis whose stool chymotrypsin is normal, which indicates that the test, unlike those using meconium or stool, can detect cystic fibrosis even when pancreatic function is adequate.

The prospective study of 75,000 newborn infants demonstrates that the assay is effective, with a detection rate of 1:2,143, which is close to that expected in a predominantly

white population. The specificity of the assay performed on blood spots taken at 4 to 7 weeks (the repeat blood samples) was very high, but in the newborn period the assay was not so specific, and resampling was necessary in 0.58% of babies. In screening 14,000 babies, Heeley et al.[7] had a resampling rate of only 0.2% and a detection rate of 1:2800. Had we halved our resampling rate, we would not have detected nine of our cases, and our detection rate would have been 1:2,885, similar to Heeley's. It may well be that for this screening test a compromise must be reached between an acceptable resampling rate and an acceptable rate of false-negative tests. However, other factors also are operative, and Adriaenssens et al.,[8] with a resampling rate similar to ours, detected only five cases in 19,125 babies, with no known false-negatives.

Our detection rate is high compared with rates so far reported by other groups.[7-11] There is no evidence that the reagents we used are more sensitive than those currently available commercially, as no direct comparisons of sensitivity have yet been made. Lyon et al.,[11] using the same reagents, had a lower detection rate than Heeley,[7] who used the CIS trypsin radioimmunoassay kit (CIS(UK)Ltd), but the high incidence of births to Maori or part-Maori parents in Lyon's population makes this comparison uncertain.

False-positive results from the 5-day blood sample seem inevitable, and the reason for elevated IRT in babies without cystic fibrosis is not apparent. Our study, and those of Crossley et al.[4] and Kirby et al.[6] have not indicated any association with prematurity, conduct of labor, or medication. Collecting a second routine heel prick blood sample from all babies at 4 weeks of age, to test in a more highly specific way for cystic fibrosis, cannot be recommended because of increased cost and decreased coverage that would ensue. Our coverage with the newborn assay is close to 100% of all births in our population.

There is no agreement at present on whether diagnosis in the newborn period improves the outcome in cystic fibrosis, and early studies purporting to show this have had serious drawbacks. However, evidence for medical benefit in early diagnosis is becoming more convincing. Orenstein et al.[14] examined data from 16 sibling pairs when each patient was 7 years of age, and found that clinical scores and x-ray examination scores were higher and residual lung volumes lower in the younger siblings in whom the disease was diagnosed earlier. A larger sibling study has been carried out by Corey and Levison,[15] who studied 44 families with more than one patient with cystic fibrosis, and found that, with correction for age and sex, order of diagnosis within the family was the most significant discriminating variable. The first diagnosed in a family was less likely to maintain normal lung function for two years. The study of Kraemer et al.[16] of survival in 204 patients also supports the notion that early diagnosis is related to a better prognosis, and Reid's observations on lung remodeling in cystic fibrosis in relation to hypoxia, and its probable reversibility in the first three months of life, appear to be compelling reasons favoring early diagnosis.[17]

Certainly, knowledge of the diagnosis of cystic fibrosis alters the therapeutic approach of most pediatricians to the treatment of bronchopulmonary infections in infants, and this in itself is an argument for screening, because the diagnosis of cystic fibrosis in infants not having meconium ileus is often difficult. Of the new patients seen in the cystic fibrosis clinic at the Royal Alexandra Hospital for Children in Sydney in 1979 and 1980, the mean age at diagnosis in 24 patients who did not have meconium ileus or a sibling with cystic fibrosis was 22 months, but symptoms had been present for a mean of 17 months. It is reassuring that almost all of the patients in whom we detected cystic fibrosis had some respiratory or gastrointestinal tract symptoms from an early age, even though the diagnosis was not being considered. The screening program does not appear to be detecting the disease in patients who otherwise would remain symptom free until late childhood or adulthood.

Neonatal screening may be somewhat beneficial in relation to genetic counseling, but full benefit cannot be realized in the absence of a reliable method for prenatal diagnosis.[18]

Our study demonstrates that blood spot IRT assay performed in the neonatal period is an excellent screening test for detection of cystic fibrosis. Sensitivity and specificity are satisfactory, but more experience is required before these can become optimal. We believe that the available evidence for benefit in early diagnosis is sufficiently strong that large-scale and prolonged trials of screening should be undertaken in several centers.

REFERENCES

1. Stephan U: Neonatal screening for cystic fibrosis. *In* Bickel H, Guthrie R, Hammersen G, editors: Neonatal screening for inborn errors of metabolism. New York, 1980, Springer, pp 81-88.
2. Forrest DC, Wilcken B, Turner G: Screening for cystic fibrosis by a stool trypsin method. Arch Dis Child **56**:151, 1981.
3. Crossley JR, Elliott RB, Smith PA: Dried blood spot screening for cystic fibrosis in the newborn. Lancet **1**:472, 1979.
4. Crossley JR, Smith PA, Edgar BW, Gluckman PD, Elliott RB: Neonatal screening for cystic fibrosis, using immunoreactive trypsin assay in dried blood spot. Clin Chim Acta **113**:111, 1981.
5. Dominici R, Monaco F, Morano S, Antonozzi I: Immunoreactive trypsin on dried blood spots as a possible neonatal test for cystic fibrosis. La Ricerca Clin Lab **10**:511, 1980.
6. Kirby LT, Applegarth DA, Davidson AGF, Wong LTK,

Hardwick DF: Use of a dried blood spot in immunoreactive trypsin assay for detection of cystic fibrosis in infants. Clin Chem **27**:678, 1981.

7. Heeley AF, Heeley ME, King DN, Kuzemko JA, Walsh MP: Screening for cystic fibrosis by dried blood spot trypsin assay. Arch Dis Child **57**:18, 1982.
8. Adriaenssens K, Janssens H, VanSoom H: Two tier screen for cystic fibrosis. Lancet **1**:833, 1981.
9. Travert G, Mustin O, Fernandez Y: Dosage radioimmunologique de la trypsine dans le sang séché. Nouv Presse Med **10**:2093, 1982.
10. Kirby LT, Davidson AGF, Applegarth DA, Wong LTK, Frohlich M, Hardwick DF: Use of an immunoreactive trypsin assay to screen newborns for cystic fibrosis (in press).
11. Lyon ICT, Smith PA, Crossley JR: Neonatal screening for cystic fibrosis in New Zealand (in press).
12. Crossley JR, Berryman CC, Elliott RB: Cystic fibrosis screening in the newborn. Lancet **2**:1093, 1977.
13. King DN, Heeley AF, Walsh MP, Kuzemko JA: Sensitive trypsin assay for dried blood specimens as a screening procedure for early detection of cystic fibrosis. Lancet **2**:1217, 1979.
14. Orenstein DM, Boat TF, Stern RC, Tucker AS, Charnock EL, Matthews LW, Doershuk CF: The effect of early diagnosis and treatment in cystic fibrosis. Am J Dis Child **131**:973, 1977.
15. Corey M, Levison H: Pulmonary function in siblings with cystic fibrosis. CF Club Abstract 35, 1978.
16. Kraemer R, Hadorn B, Rossi E: Classification at time of diagnosis and subsequent survival in children with cystic fibrosis. Helv Paediatr Acta **32**:107, 1977.
17. Reid LM: The pulmonary circulation: Remodelling in health and disease. Am Rev Respir Dis **119**:531, 1979.
18. Nadler HL, Rembelski P, Mesirow KH: Prenatal detection of cystic fibrosis. Lancet **2**:1226, 1981.

Archives of Disease in Childhood, 1983, **58**, 863–866

Diagnostic delay in cystic fibrosis: lessons from newborn screening

B WILCKEN, S J TOWNS, AND C M MELLIS

Oliver Latham Laboratory, New South Wales Department of Health, and the Department of Respiratory Medicine, Royal Alexandra Hospital for Children, Camperdown, Australia

SUMMARY Newborn screening for cystic fibrosis (CF) by dried blood spot immunoreactive trypsin (IRT) assay is now feasible, but the benefits are disputed. We have studied the symptoms and signs at diagnosis in 48 babies detected during a newborn screening programme, and also the delay between presentation with symptoms and diagnosis in all 33 babies diagnosed at our CF clinic in the two years before screening began. Eleven of the 48 screened babies had meconium ileus, 16 had gastrointestinal symptoms only, and 14 had both respiratory and gastrointestinal symptoms at the time of diagnosis. Five of the remaining 7 babies developed clear cut symptoms or signs soon after diagnosis. Thus, 96% (46 of 48) of the babies had symptoms by 3 months of age. Of the 33 infants diagnosed clinically in the two years immediately before screening, 13 (39%) were over 12 months of age at diagnosis. Moreover, the mean delay between presentation with symptoms and diagnosis of CF in these infants was 2·6 years. Our data show that the delay between onset of symptoms and diagnosis is far greater than previously supposed and that most babies detected by our screening programme already had symptoms that warranted treatment at the time of their diagnosis.

While it has been taken for granted that the earlier a child with symptoms of cystic fibrosis (CF) is treated the better the prognosis,[1 2] there is no agreement on whether very early diagnosis by newborn screening followed by prophylactic treatment confers any extra benefit.[3 4] Early diagnosis by screening would obviously facilitate prompt, aggressive treatment of respiratory or gastrointestinal symptoms, or both, as soon as they arise. The delay between the onset of symptoms and clinical diagnosis of CF has been examined in 33 new patients presenting at a large CF clinic during 1980 and 1981. The natural history of the disorder during the first weeks of life has also been studied in 48 patients diagnosed by newborn screening in 1981–3.

Patients and methods

Patients diagnosed clinically in the two years before newborn screening started. All 33 patients with CF diagnosed at this hospital in the two years before neonatal screening was started were included. The parents of the children were interviewed on one or more occasions by one of us (SJT). Case notes were consulted and the following information recorded: the age of the patient at diagnosis; age of initial presentation to medical attention; the mode of presentation (meconium ileus, sibling of a CF patient, or with symptoms, namely, failure to thrive, gastrointestinal, respiratory, and combined gastrointestinal plus respiratory symptoms); the number of visits to a medical practitioner; and the number of hospital admissions, before diagnosis. The babies were grouped according to age at diagnosis as follows: group 1—before 2 months of age; group 2—between 2 months and 12 months; group 3—later than 12 months of age.

Patients diagnosed by newborn screening. One hundred and twenty thousand infants were screened for CF by blood spot immunoreactive trypsin (IRT) assay[5] as part of a state wide screening programme, and 50 babies with CF are known to have been born during this period. Forty eight babies had a positive screening test. One baby was accidentally not screened because of severe illness (small bowel atresia) and another was missed by the screen but diagnosed at 8 weeks with severe failure to thrive and bronchiolitis. The first 35 infants detected with cystic fibrosis have been reported previously, together with details of the methods.[6]

Results

Patients diagnosed clinically.

Group 1 comprised 11 patients diagnosed before two months of age. Five of these had meconium ileus, one had meconium plug syndrome, and three had a close relative with CF. Although a definitive diagnosis by sweat test was not made in these babies until up to 8 weeks of age, the possibility of CF had been discussed with the parents during the first week of life. The remaining two patients in this group had symptoms—one had two admissions to hospital with pneumonia before diagnosis at 6 weeks, and one failed to thrive from the age of two weeks and had three medical consultations before the diagnosis was established at 7 weeks.

Group 2 comprised 9 patients diagnosed between 2 and 12 months of age. The average age of presentation was 1·9 months, and the average age at diagnosis was 5·7 months. The commonest mode of presentation in this group was failure to thrive, usually with obvious gastrointestinal symptoms. Details of these 9 patients are shown in Table 1.

Group 3 comprised 13 patients (39% of the total) in whom a diagnosis of CF was not made until after 12 months of age. Most had respiratory symptoms, usually in association with gastrointestinal symptoms. The patient details are summarised in Table 2. The mean delay between presentation with symptoms and diagnosis was 2·6 years in this group.

In all of these 33 patients whose newborn screening blood sample could be retrieved from our records and tested, the IRT value was raised above that of control samples.[6]

Babies diagnosed during the newborn screening period. Forty eight babies had a positive screening test. Notification of a raised IRT value on the second blood sample and request for sweat test was made usually between 3 and 6 weeks of age, and by 8 weeks in all but two babies. The mean age at diagnosis was 37 days.

Table 1 *Patients aged over 2 months and under 12 months at the time of diagnosis*

Case No	*Age at diagnosis (mths)*	*Age at first presentation*	*Symptoms at presentation**	*Visits to doctor before diagnosis (No)*	*Hospital admissions before diagnosis (No)*
1	5·5	3 mths	Respiratory	2	1
2	3	5 dys	FTT/GIT respiratory	5	1
3	8	6 mths	FTT/GIT	5	1
4	11	3½ mths	FTT/GIT, respiratory	10	—
5	6	birth†	FTT/GIT	3	2
6	4	3 wks	FTT/GIT	9	2
7	7	3 mths	FTT/GIT, respiratory	3	1
8	3	3 wks	FTT/GIT	2	2
9	4	3 wks	GIT	3	1

*First listed is major mode of presentation.
†Congenital hydrocephalus, Erb's palsy noted at birth—2 previous admissions for VP shunting.
FTT = failure to thrive; GIT = gastrointestinal.

Table 2 *Patients aged 12 months and over at the time of diagnosis*

Case No	*Age at diagnosis (yrs)*	*Age at first presentation*	*Symptoms at presentation*	*Visits to doctor before diagnosis (No)*	*Hospital admissions before diagnosis (No)*
1	3·5	1·5 yrs	Rectal prolapse	5	—
2	1·3	1 yr	Rectal prolapse, resp	6	1
3	2·3	9 mths	GIT resp	12	2
4	2·5	4 mths	FTT, resp	14	—
5	1·3	Birth	FTT	10	—
6	1·1	3 mths	FTT/GIT, resp	14	1
7	4	5 mths	Resp	13	1
8	3	6 wks	Resp	26	3
9	2·5	3 mths	GIT, resp	6	4
10	3·5	2 mths	GIT, resp	10	8
11	4·2	1 yr	Relative with CF, GIT, resp	30	2
12	8	3 mths	Sibling with CF, GIT, resp	72	2
13	2·5	3 mths	Sibling with CF, GIT, resp	4	—

FTT = failure to thrive; GIT = gastrointestinal; Resp = respiratory.

Table 3 *Symptoms in 48 babies diagnosed by screening*

Symptoms	*At diagnosis (3–8 weeks) (No)*	*At 3 months (No)*
Previous meconium ileus	11	11
Gastrointestinal	16	19
Gastrointestinal, respiratory	14	14
Respiratory	0	2*
Asymptomatic	7	2

*Includes an asymptomatic baby with *Staphylococcus aureus* and *Pseudomonas aeruginosa* in pharyngeal aspirates (see text).

Eleven of the 48 babies presented in the neonatal period with meconium ileus. Thirty of the remaining 37 had gastrointestinal symptoms at the time of diagnosis. These symptoms ranged from severe failure to thrive, with weight loss and oedema to increased diarrhoea, with fat globules in the stools. Fourteen of these 30 babies also had pulmonary symptoms, although these were mild in 7. One baby with no pulmonary symptoms had an abnormal chest radiograph with right upper lobe collapse. No baby had pulmonary symptoms alone at this time. Only 7 babies had no symptoms at diagnosis, but four of these developed symptoms—bronchitis in one, and a fall off in weight gain in three—soon afterwards. A fifth asymptomatic baby, upon his first visit to the clinic at two months, had *Staphylococcus aureus* and *Pseudomonas aeruginosa* cultured from a pharyngeal aspirate. (Table 3) Among those 30 babies with symptoms, the diagnosis of CF was only being considered in four, three of whom had siblings with known CF.

Discussion

Our report of the age pattern at diagnosis of CF in patients before the screening programme started is similar to findings in other Australian states and in other countries. In Victoria, between 1955 and 1978, 32% of all CF patients were over the age of 12 months at diagnosis.[4] In 1979–82 in Victoria, 43% of patients were diagnosed under 3 months of age, but 35% were older than 12 months (personal communication). In South Australia, between 1977 and 1981, 40% of patients were under 3 months of age and 43% older than 12 months at diagnosis (personal communication). Gitzelmann reports that in Switzerland, where clinical awareness is traditionally high, approximately 30% of cases remain undiagnosed at 12 months.[2] A figure of 35% has been reported for East Germany[7] and 40% for Canada.[8] One comprehensive report from the United States shows, however, that at least 52% of patients remained undiagnosed at 12 months of age.[9]

During the period of screening in New South Wales, it is likely that most of the cases were identified and only one child is known to have been missed by the screen. The incidence of CF detected during this period was 1/2400, which is close to the expected incidence.[10] At the time of diagnosis, 85% of the screened babies already had symptoms or signs of CF, and by 3 months of age this figure was 96%. These findings indicate that the usual delay between onset of symptoms and diagnosis is far greater than that suggested by retrospective enquiry.

Now that screening for CF is feasible, with acceptably low false positive and false negative rates,[6] it is important that controlled studies be undertaken to investigate its value. In the meantime, our study has important implications for screening—particularly in light of the high incidence of CF symptoms at such an early age. It is not clear at present when to start treatment of CF. Nevertheless, once there are respiratory or gastrointestinal symptoms, or signs of the disease, appropriate treatment must be offered. Our investigation shows that without screening there are appreciable delays between onset of symptoms and diagnosis. The data from the group of screened babies suggest that these delays are far more substantial than previously suspected; almost all babies had symptoms at 3 months. Hence, screening for CF may be justified not only in relation to prophylactic treatment, the benefit of which is disputed, but also to facilitate early and aggressive treatment of mild symptoms, which is not at present being achieved.

We thank Dr J Brown and the many other paediatricians who supplied information about their patients, Mr A R D Brown and staff members at the Oliver Latham Laboratory who performed the screening assays, and Mrs E Pluck and Mrs S Morven for secretarial assistance.

References

1 Nadler HL, Girimaji JSR, Taussig LM. Cystic fibrosis. In: Stanbury JB, Wyngaarden JB, Fredrickson DS, eds. *The metabolic basis of inherited disease*. 4th edition. New York: McGraw-Hill, 1978:1683–710.

2 Gitzelmann R. Why we should not screen our newborn for cystic fibrosis. *Helv Paediatr Acta* 1981;**36**:493–4.

3 Dodge JA, Ryley HC. Screening for cystic fibrosis. *Arch Dis Child* 1982;**57**:774–80.

4 Phelan PD. Screening for cystic fibrosis (Commentary). *Arch Dis Child* 1982;**57**:779–80.

5 Crossley JR, Elliott RB, Smith PA. Dried-blood spot screening for cystic fibrosis in the newborn. *Lancet* 1979;**i**:472–4.

6 Wilcken B, Brown ARD, Urwin R, Brown DA. Cystic fibrosis screening by dried blood spot trypsin assay: results in 75,000 newborn infants. *J Pediatr* 1983;**102**: 383–7.

[7] Gottschalk B. Die mukoviszidose in der Deutschen Demoncratischen Republik. *Deutsche Gesundheitwesen* 1972;**27**:2292–5.
[8] Nesbitt CJ, Gregg MB, Olmsted N, Pogue RE, Warwick WJ, Wood RE. 1976 report on survival studies of patients with cystic fibrosis. Atlanta Georgia; Cystic Fibrosis Foundation, 1978.
[9] Warwick WJ. Undiagnosed patients with cystic fibrosis. *J Chronic Dis* 1980;**33**:685–96.
[10] Allan JL, Robbie M, Phelan PD, Danks DM. The incidence and presentation of cystic fibrosis in Victoria, 1955–1978. *Aust Paediatr J* 1980;**16**:270–3.

Correspondence to Dr B Wilcken, Oliver Latham Laboratories, Macquarie Hospital, PO Box 169, North Ryde NSW 2113, Australia.

Received 30 June 1983

Fifty years ago

Proceedings of the Sixth Annual General Meeting of the British Paediatric Association

Dr Alan Moncrieff (London) suggested that in view of the recent persecution of Jewish paediatricians in Germany members should, as far as possible, find temporary posts for some of them in this country until things became more settled there.

Archives of Disease in Childhood 1933; **8**: 361.

(A future editor of the *Archives* and inaugural professor of child health at the Institute of Child Health, London, speaks up in the year of his election to the British Paediatric Association. 1933 was pretty quick off the mark on this topic. Philip Evans).

Studies in the anaemias of infancy and early childhood Part IV. The Haemolytic (erythronoclastic) anaemias of the neonatal period with special reference to erythroblastosis of the newborn

LEONARD G PARSONS, J C HAWKSLEY AND ROBERT GITTINS (*Birmingham*)

Haemolysis occurs constantly in intra- and extrauterine life but normally is kept within bounds by some mechanism. Hampson has suggested that haemolysis may be controlled in intrauterine life by something produced by the mother and passed on by her to the fetus, which in extrauterine life is elaborated by the infant itself; for at birth a considerable degree of haemolysis occurs owing to alterations in the oxygen tension of its surroundings, which is prevented from becoming excessive by the anti-haemolytic agent. If this factor is absent or insufficient, excessive haemolysis with the production of anaemia and perhaps severe jaundice may occur; this may however, be prevented if the anti-haemolytic factor is given by administration of blood serum. This is the basis of the treatment of icterus gravis by injections of blood serum, a treatment which we owe to Hampson and which has proved a great success . . . Such an hypothesis would furnish an adequate explanation for the fact that jaundice is more frequent and severe in preterm than term children; also of the development of haemolytic anaemia in the preterm and of its prevention by blood transfusion . . . It . . . fails adequately to explain hydrops fetalis and those cases of icterus gravis which develop spastic diplegia in later life and those which show necrotic changes in the liver.

Archives of Disease in Childhood 1933; **8**: 159–183.

(Rhesus incompatibility not then recognised. The history of an error; the Hampson hypothesis was false and the treatment by injection of maternal serum directly wrong. Philip Evans).

GENETICS AND EPITHELIAL CELL DYSFUNCTION IN CYSTIC FIBROSIS

Proceedings of a Symposium
Held in Kimberly, Ontario, Canada,
November 12–15, 1986

Editors

John R. Riordan
Department of Biochemistry
The Hospital for Sick Children
Toronto, Ontario, Canada

Manuel Buchwald
Department of Genetics
The Hospital for Sick Children
Toronto, Ontario, Canada

ALAN R. LISS, INC. • NEW YORK

Genetics and Epithelial Cell Dysfunction
in Cystic Fibrosis, pages 201–215

AN EVALUATION OF SCREENING FOR CYSTIC FIBROSIS

Bridget Wilcken

New South Wales Department of Health, Oliver Latham Laboratory, P.O. Box 53, North Ryde, Sydney, N.S.W. 2113. Australia

Many patients with cystic fibrosis now live into the third and fourth decades, whereas forty years ago death in early childhood was the rule. The relative importance of the different factors which have effected this great improvement has not been established, but appropriate treatment of symptoms and early referral to a specialised clinic seem to be mandatory (Schiotz et al 1981). As clinical diagnosis of cystic fibrosis is often delayed, there has been much interest in developing a reliable newborn screening test, but until recently methods have been unreliable (Phelan 1984). In 1979, Elliott, Crossley and their colleagues in New Zealand found that newborns with CF had elevated immunoreactive trypsin (IRT) levels in the blood, and they developed a radio-immunoassay of IRT which could be applied to dried blood samples already being collected for newborn screening for phenylketonuria and hypothyroidism (Crossley et al 1979). Further investigation showed that this assay was adequately specific and sensitive for newborn screening (Crossley et al, 1981).

However, although the earliest possible treatment of symptoms in a patient with CF seems desirable, there has been no clearcut evidence that very early diagnosis by newborn screening and prophylactic treatment would confer any extra benefit, and with the advent of a suitable screening test came doubts about the wisdom of the introduction of national screening programmes (Ad Hoc Committee Task Force 1983, Gitzelman 1981, Holtzman 1984) as well as strong arguments in favour of screening (Dodge and Riley 1982).

Several regional or national screening programmes have in fact been started (Naylor 1985). This paper describes the experience of screening for CF in the state of New South Wales, Australia for the past 5 years, since July 1981, summarises briefly some findings of other screening programmes, and demonstrates the apparant advantages and possible drawbacks of such screening.

THE NEW SOUTH WALES STUDY

New South Wales (N.S.W.) has a population of almost five and a half million. Cystic fibrosis is treated in four clinics, two based on each of the two children's hospitals, and two in other major children's units. The latter have become independent clinics only in the last 2 years. The referral to the two major clinics which operated during most of our study is on a geographical and physician - network basis. Over the past eight years, approximately sixty percent of all CF patients diagnosed in N.S.W. were referred to one centre, the Royal Alexandra Hospital for Children (RAHC), Camperdown, Sydney. Those parts of our study which refer to the performance and results of the screening test refer to all CF patients diagnosed over the period in N.S.W. Several subsidiary studies related to delay in diagnosis, evaluation of the effects of screening, and parental attitudes, have been based on only those patients attending the RAHC clinic.

Delay in Diagnosis

Diagnostic delay in cystic fibrosis has been well documented (Rosenstein and Langbaum 1984). We examined the age at diagnosis, and the elapsed interval between presentation with symptoms and diagnosis, in all children first presenting to the RAHC clinic in the two years immediately before screening began (Wilcken et al 1983 B). Excluding those with meconium ileus, or an already-diagnosed sibling, the mean age at diagnosis was 1.8 years, the mean delay after first consultation with a physician was 1.6 years, and the mean number of visits to a physician before diagnosis was 11. Thirty-nine percent of all CF children remained undiagnosed at 12 months. These figures were similar to those reported from other populations with good clinical awareness (Rosenstein and Langbaum 1984).

The Screening Programme

In July 1981 we started screening all babies born in N.S.W. for CF as part of the routine newborn screening programme (Wilcken et al 1983 A). A heel-prick dried blood sample was already collected from all newborns on about day 5, and delivered to our laboratory, for assays to detect phenylketonuria and hypothyroidism. These samples were in addition assayed for immunoreactive trypsin by radio-immunoassay using a modification of Crossley's method (Crossley et al 1981). The reagents were supplied by the Department of Paediatrics, Auckland University, New Zealand. There is no internationally accepted reference standard for immunoreactive trypsin, and thus our results are not numerically comparable to results obtained using either of the two commercial reagent kits commonly available (Hoechst and Sorin).

The Screening Protocol The distribution of IRT values in neonates is a continuous non-normal distribution. For defining a positive result a cut-off point was chosen empirically. Initially the level of cut-off was 220 AU/L, but we now use the mean +2.97 SD of log transformed data. IRT values in the top 2% of each batch of 300-500 samples are repeated in duplicate. Where the mean value exceeds the cut-off point a second dried blood sample is requested, and this is assayed in duplicate. If the IRT value of this second sample is also above the cut-off, a sweat test is requested. In N.S.W. this can usually be arranged within a few days, and occasionally even on the day of the second screening result.

Results Results in 400,000 babies tested are shown in table 1. The false negative rate was just under 4%, similar to the findings in other large series (Kuzemko 1986), but the false positive rate of 1:10,000 of the population screened is two to three times higher, and this is undoubtably due to the differing quality of reagents used. Thus, in our assay, the positive predictive value after the second test was 80%. The positive predictive of the initial test is of course quite low, and this assay can only be successfully performed in a two-stage system, requiring a resampling step. However, for newborn siblings of already diagnosed CF patients, who have an a priori risk of CF of 1:4, the positive predictive value is 98%. The incidence we found was 1:2564 (95% confidence intervals 1:2200-1:3040), almost exactly the same as that

TABLE 1 Neonatal screening in N.S.W. Australia from July 1981 to June 1986.

Babies tested	400,000
Elevated IRT level in second sample	188
CF confirmed	149
CF missed by screen (false negative)	6
CF never screened	1
Total CF population	156
Apparent incidence	1:2564

found in the neighbouring state of Victoria by a painstaking survey (Allan and Phelan, 1985). If the geographical origin of mothers is taken into account, for babies of Northern European stock the incidence was 1:2200, for Mediterranean European stock 1:3500, and for other babies, 1:12,000 (the latter mainly from middle eastern and far eastern countries). There were 95 male babies and 61 females. The male to female ratio of 1.56 is not significantly different from that of the whole cohort (1.05). However, male predominance has been mentioned in another survey (Sturgess et al 1985) and in offspring of uncles of CF patients (Pritchard et al 1983). The variation we found in sex ratio might well be real, and would have very interesting implications. The incidence of meconium ileus was 20%, (including 2 babies with neonatal volvulus). Four babies, all with meconium ileus, died during infancy. Babies such as these may be missed by population surveys not based on neonatal screening, and this could help to account for the much lower incidence reported for meconium ileus in some centres (Rosenstein and Langbaum 1984), but not in others (Phelan 1984).

Thus, of the CF cases among our 400,000 babies screened, fewer than 4% were missed by the screen, or not screened, and 33% had either meconium ileus, or a sibling with previously diagnosed CF. Sixty three percent had unsuspected CF and were detected.

False positive and false negative results Transient hypertrypsinaemia, causing a false positive result in the first test only, was associated with low birthweight (under 2000G), and with illness in general. The chance of a baby in an intensive care unit having an IRT level above the cut-off point was 3%, and was much higher in babies with

bowel atresias and renal failure. There was no association with age at sampling from 3-7 days, or with mode or timing of feeding. Persistent hypertrypsinaemia was associated with congenital infections, presumably because of pancreatitis, bowel atresias, and trisomy 13, and these causes accounted for 11 of our 39 false positives. The remaining babies were apparently healthy, and included three pairs of siblings.

Of the six false negative cases one had meconium ileus, and another, with an initial IRT that was just below the cut-off point, was a sibling, so that these cases were diagnosed early. One other baby had a low and one a borderline IRT value on the first test and no further action was taken. One had an elevated IRT in the first sample, but a repeat sample had a level just below the cut off. Finally one whose initial IRT level was later found to be elevated, was missed because of a laboratory error. CF was diagnosed at 2, 5, 13, and 16 months in these four children.

The rate of decline of IRT in the CF patients is shown in figure 1. High levels, greater than 2 multiples of the median of newborn values, persisted in most patients during the first 12 months, but the percentage of patients with

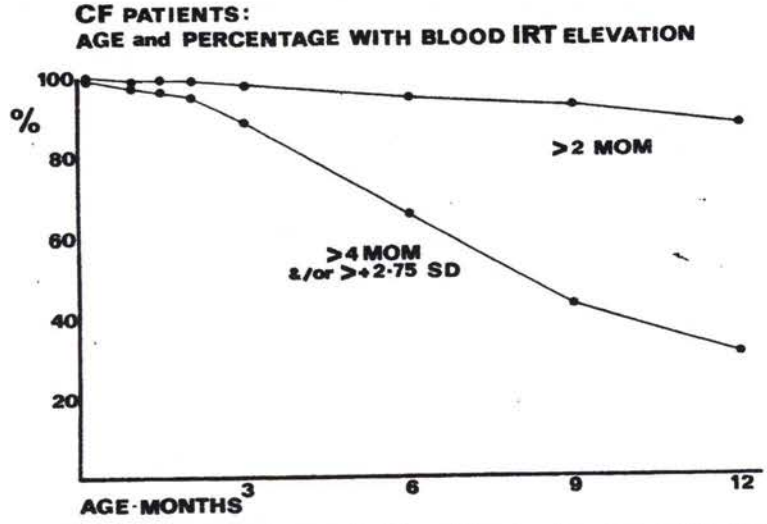


Figure 1. The percentage of CF patients by age with blood IRT levels above the neonatal cut-off level (lower line; MOM=multiples of median, see text) or below the cut-off level but greater than twice the median (upper line).

levels above the neonatal cut-off level declined rapidly after three months.

Symptoms in screened patients At diagnosis (mean time seven weeks) 92 of the first 114 babies we detected had symptoms related to cystic fibrosis. Twenty three had had meconium ileus, 41 had failure to thrive or other gastrointestinal symptoms, and 28 had some respiratory symptoms as well as the gastrointestinal symptoms. None had respiratory symptoms alone, and 22 were symptom-free. Only 8 of these 114 were symptom-free by three months. Amongst all the CF patients referred to the RAHC clinic, fifty so far have had some tests of pancreatic function at ages 3-12 months, carried out by Dr. Kevin Gaskin. Fat balance studies showed that nine of the fifty had normal faecal fat, and in these 9 pancreatic stimulation tests showed adequate pancreatic function with lipase secretion of >1% and colipase >2%, and up to 25% of normal.

EVALUATION OF THE EFFECTS OF SCREENING IN N.S.W.

We considered, and rejected on ethical grounds, the possibility of a controlled trial, and opted instead to use as groups for evaluation 1) All CF patients born during the first three years of screening (mid 1981-mid 1984) and initially referred to the CF clinic at the Royal Alexandra Hospital for Children, Sydney. 2) All CF patients born during the three years immediately before screening was instituted, and initially referred to the same clinic, and 3) All CF children born in the neighbouring state of Victoria in 1981-1984 and treated at the Royal Children's Hospital, Melbourne. These children were not diagnosed by screening.

Comparability of groups These groups are not strictly comparable because of a number of factors which could be operating. However, we analysed those points of difference between screened patients, group 1, and their historical controls, group 2, and found that there were few biases (Wilcken and Chalmers 1985). For both groups there seemed likely to be near-complete ascertainment, with any missing patients from either group likely to be mildly affected. There were no identifiable changes in hospital admissions policy, and only minor changes in management, shown not to have affected morbidity. Comparisons between group 1 and their contemporaneous controls, group 3, are more difficult.

There is likely to be greater under-ascertainment in group 3, as more children will remain undiagnosed because of their younger current age, and although management policies have been compared and seem similar, there are certain to be differences. There are climatic differences also, with colder, wetter winters in much of Victoria.

Nutrition Heights and weights of children in all three groups, excluding those presenting with meconium ileus or with a diagnosed affected sibling, were measured within two months of their first, second, third, and fourth birthdays, and the data adjusted to actual birth dates (Valencia Soutter-personal communication). Screened children were heavier at each age than either of the unscreened groups and were taller at 1, 2 and 3 years. However, the contemporaneous unscreened group from Victoria had similar heights to screened children at 4 years. These data are being prepared for publication.

The existence of subtle nutritional deficits in CF babies under 2 months of age diagnosed by screening has been documented carefully by the Colorado group (Reardon et al, 1984). Our data, and that from other screening programmes (Barlocco et al 1985) support the notion that early nutritional management following diagnosis by screening improves the nutritional status in the early years.

Hospitalisation as a measure of morbidity We studied data on hospital admissions during the first two years of life in screened and unscreened patients (Groups 1 and 2) (Wilcken and Chalmers, 1985). We were able to obtain full information on all admissions in all patients in these two groups. We included in our analysis all admissions for chest infection, intravenous antibiotics, failure to thrive, malabsorption and diarrhoea, and excluded all hospital days at birth, days related to treatment of meconium ileus, including ileostomy closure, and days for illnesses unrelated to cystic fibrosis. Admissions purely for sweat test and instruction, which were uncommon, were analysed separately. The results are shown in table 2. Whereas there is no difference in hospital days for patients with meconium ileus between the two groups, for those without, the unscreened patients had a mean of 27 hospital days in the first two years for CF related illness, compared with 3.9 days for screened patients ($p<0.01$ by Student's t-test, and Wilcoxon rank sum test). There was no trend with time. This big

TABLE 2 Hospital days in the first two years of life in unscreened and screened patients (Mean days)

Reason for admission	Unscreened patients	Screened patients	
Without MI	n = 48	n = 34	
CF related illness	27.3	3.9	p<0.01
Instruction	0.4	0.9	
Total	27.7	4.8	p<0.01
Meconium ileus	n = 8	n = 6	
CF related illness	11.5	15.8	N.S.

reduction in hospital days occurred suddenly, coincident with the onset of screening, as is shown in figure 2. Careful analysis showed that this could not be explained by the minor changes in management that occurred during the study period.

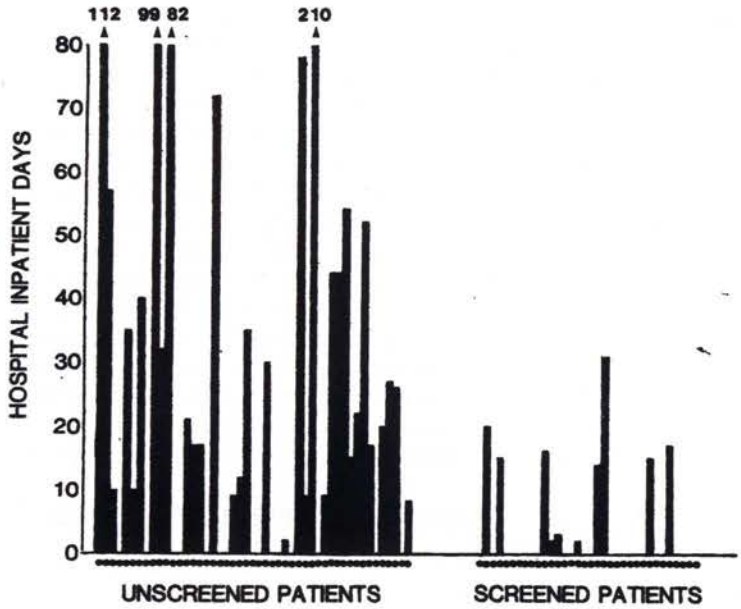


Figure 2. Hospital days in the first two years in 48 unscreened and 34 screened patients without meconium ileus. Individual patients are shown by closed circles. (Wilcken and Chalmers, 1985, Lancet ii 1319-1321, with permission).

Nor was the difference due to greater readiness to admit to hospital an undiagnosed child with a puzzling illness, since even if only those hospitalisation are included which occurred after the diagnosis was made, there is still a significant difference (screened, mean of 3.9 and unscreened 16.3 days). Twenty four of 34 screened patients had no hospitalisation in the first 2 years, compared with only 15 of 48 unscreened patients, and only 1 of 34 screened was hospitalised for over 21 days, compared with 20 of 48 unscreened ($p<0.0005$ for both).

Effects on early parent-child interaction It has been postulated that early identification of an asymptomatic child with CF may interfere with the parents' attatchment to their newborn (Ad Hoc Committee 1983, Holtzman 1984, Phelan 1984). Although we have shown that most children had some symptoms at diagnosis, or soon thereafter (Wilcken 1983 B), some were indeed asymptomatic. We invited 71 mothers from groups 1 and 2 into a study to explore the effect of screening on their relationships with their children (Boland 1986). Mothers were excluded if the baby had had meconium ileus, if there was more than one child with cystic fibrosis in the family, or if there was a chronic or life-threatening illness or any death in the nuclear family. Two mothers did not reply, five refused, and a further 6 were later excluded. From interviews with 58 mothers measurements were made of the psychological distance between mother and CF child, and of over-protectiveness and denial. This study is being prepared for publication. In summary, Boland found that increased diagnostic delay did not increase over-protectiveness; increased length of the asymptomatic phase did not increase denial; and absence of symptoms at diagnosis did not increase over-protectiveness. Overall, there was no significant difference between the caring attitudes of mothers of screened and unscreened children, and no evidence of a harmful effect on family interactions from early diagnosis.

Attitudes of physicians and parents All 12 physicians who currently treat children in all four major CF clinics in New South Wales were asked by postal questionnaire about their attitudes to screening for CF, and all replied. Eleven of the 12 agreed that newborn screening was generally useful, 10 that screened patients had better health in the early years, and that the cost was 'money well spent'; the remainder were unsure. The most enthusiastic support for newborn screening came from the five physicians with the

heaviest case-loads who had had extensive experience in the clinic before as well as during the screening period. Parents of CF children were also asked by postal questionnaire about their feelings on newborn screening. All families in groups 1 and 2 who still attended the clinic at the Royal Alexandra Hospital for Children, Sydney, were contacted, except those whose babies had had meconium ileus, two who spoke little English, and three who had recently moved. Sixty five out of seventy one sets of parents replied. All parents in the non-screened group and all but one from the screened group stated that they would prefer to know the diagnosis early, before any symptoms were apparent, and that if they had another child they would want the baby to have a screening test. The remaining parent, from the screened patient group, was unsure about both questions. Thus it appears that there is strong support from physicians who have actual experience of newborn screening for CF, and strong support from parents.

Sweat tests Numbers of sweat tests performed at the Royal Alexandra Hospital for Children, Sydney, a major centre for sweat testing, declined markedly after newborn screening was established, from 746 and 736 in the years 1979-80 and 1980-81 respectively, to 384 and 362 in 1984-85 and 1985-86. An even greater decline, from over 500 tests per annum before screening, to 167 in the most recent complete year, occurred in the other major sweat-test centre at the Prince of Wales Children's Hospital, Sydney. In these centres, no difficulty has been experienced with performing sweat-tests on 3-5 week old infants diagnosed by the screening test.

Cost benefit analysis It is difficult to think about cost and benefits without reminding oneself that the most cost-beneficial approach to CF is to offer no treatment at all. The patient would then die usually in the first few years and huge costs would have been averted. We have not done a formal cost-benefit analysis of screening, but it appears likely to be cost-effective in N.S.W. in the short term. The incremental screening cost of detecting each otherwise unsuspected case is A$2300 in 1986, and the cost of treatment in the first six months, assessed generously, is A$2500. Averted costs include A$9500 for hospitalisation in the first two years, A$1000 for a reduction in the number of sweat-tests, and an uncomputed amount for inappropriate investigations, treatment, and professional attention in the first six months. A similar suggestion of short term cost

benefit has been made by Dauphinais' group (Ganeshananthan et al, 1985, A).

Attendance at a CF clinic All N.S.W. patients but one identified by newborn screening are currently registered with one of the centralised CF clinics, and all but one were referred immediately the diagnosis was confirmed. We have no data on the extent of clinic attendance in the pre-screening period, but it was also high. In Basse-Normandie, France, there was no CF centre when screening was started in 1980, but screening provided the impetus for the formation of a central clinic - the first in this region of France. (G. Travert, personal communication.)

Evaluation of the Results of Screening by Other Groups

Several groups (Ganeshenanthan et al, 1985 B, Mastella et al 1983, Dankert-Roelse et al 1982) have compared the outcome in screened and unscreened patients with respect to a variety of measurements including clinical scores, X-ray scores, heights and weights, bacteriology, and survival. None of these studies has had a randomised controlled design and for some there is no evidence that management was comparable. Even so, all results suggest a benefit in the short term for babies diagnosed by screening particularly when treatment is started before 2 months of age. A carefully designed randomised controlled trial of screening was started in Wisconsin in 1985 (Hassemar et al 1985) and the results of this are eagerly awaited.

ARE THE CONTROVERSIES NOW RESOLVED?

The report and recommendations of the Ad Hoc Committee Task Force on Neonatal Screening of the Cystic Fibrosis Foundation (1983) has been perhaps the most influential paper of recent years. This report explored several possible problems related to mass screening for CF, and suggested that more research was required before screening could be endorsed. Several other papers raised similar doubts (Gitzelman, 1981; Phelan, 1984). How far have we come since then?

Effectiveness of presymptomatic treatment Our studies, and those of Reardon et al (1984) have shown unequivocally that the great majority of CF patients have symptoms or more

subtle signs of poor nutritional status and other problems before two months of age, so that for most, it is not strictly true to speak of presymptomatic treatment. The task force report (Ad Hoc Committee, 1983) implied, and others have suggested (Gitzelman, 1981) that better attention to these symptoms would lead to early diagnosis, but unfortunately there is abundant evidence that this approach has not been effective. We have shown decreased morbidity in the first two years, we and others have shown improved nutritional status in the screened children (Barlocco et al 1985), and others have demonstrated other clinical advantages (Ganeshananthan 1985). The weight of evidence is now very strong that early diagnosis by screening, and appropriate treatment (not yet fully defined) provides better health in the early years. There is as yet no evidence as to whether this advantage will persist later in life. Screening would also provide an admirable basis for defining optimal treatment, but this possible advantage of screening has not yet been explored.

Effects on early parent-child interaction Our study (Boland - in preparation) has failed to demonstrate any adverse effect of screening on mothers of CF children diagnosed. Nor is there any evidence of inappropriate 'labelling' of very mild cases, since all have developed some symptoms quite early. Families with a false-positive diagnosis undoubtably suffer. However, reports of severe long-lasting psychological problems among such families screened for other conditions have come from investigations without control groups (Bodegard et al 1984) and those findings tend to be denied by a much better designed study (Sorenson et al, 1984). More work is required in this area.

Usefulness in making a reproductive decision The advent of reliable prenatal diagnosis of cystic fibrosis for families with an affected child makes very early diagnosis more important. Phelan (1984) found that in Victoria, Australia, 7% of all CF children had been born into a family in which there was another living undiagnosed CF child. Overall, 10% of all patients were born to parents who knew they were CF carriers. The possibility of averting the birth of even only 7% of all CF babies seems extremely worthwhile numerically, and would certainly be cost-effective.

Characteristics of the test The IRT has now been shown to have high specificity and sensitivity when used in a two-

stage protocol (Hammond 1985; Kuzemko 1985; Travert 1986; Mastella et al 1983). Certainly it will diagnose some and possibly most patients with pancreatic sufficiency.

Characteristics of the screening system Various concerns were expressed by the Task Force, but these can now be answered: The IRT value does decrease with time, but as we have shown, levels at 4-6 weeks are still elevated in CF babies. Sweat-testing can be carried out with no delay at this time in N.S.W. There has been no increase in the number of sweat-tests performed, but rather a decrease of 50-60%. Physicians have been most supportive and responsive to our programme, and increasingly so as the reliability and effect of screening have become better known. There has been no evidence of increased delay in diagnosis of the few children who had a false-negative screening test. Also, there has been almost 100% immediate referral to CF clinics, so that concerns about management of patients by physicians untrained in the area appear to be unfounded.

There could possibly be serious medico-legal consequences arising from a false-negative screening result. There could also be serious medico-legal consequences arising from late diagnosis of a child with symptoms. Optimal medical management ought not really to be influenced by medico-legal concerns.

Finally, we have demonstrated a clear-cut cost-benefit in N.S.W. of the screening programme, related to the early years. Overall, the cost-benefit may be hard to determine if very early diagnosis were to lead to improved survival. However, if improved survival were to result, screening could surely be supported on other grounds than cost.

THE ADVANTAGES IN NEWBORN SCREENING FOR CYSTIC FIBROSIS

In summary, we have found that newborn screening leads to reliable detection of cystic fibrosis, with a low false negative rate, thus avoiding substantial diagnostic delay and inappropriate management of babies with symptoms. It leads to earlier attendance at clinics, improved health in early childhood, and earlier genetic counselling. It is cost-effective, at least in New South Wales. For children with CF and their families there appear to be great advantages and no obvious disadvantages. The only drawback we see is

the low positive predictive value of the test performed at five days, and much energy should be put in to improving this aspect of a very valuable new tool in the management of cystic fibrosis.

REFERENCES

Ad Hoc Committee Task Force on Neonatal Screening. Cystic Fibrosis Foundation : (1983) Neonatal screening for cystic fibrosis: Position paper. Pediatrics 72: 741-745

Allen JL, Phelan PD (1985). Incidence of cystic fibrosis in ethnic Italians and Greeks and in Australians of predominantly British origin. Acta Paed Scand 14: 286-289

Bodegard G, Fyro K, Larsson A, (1983). Psychological reactions in 102 families with a newborn who has a falsely positive screening test for hypothyroidism. Acta Paed Scand suppl. 304 1-21

Boland CL (1986) Newborn screening for cystic fibrosis: Some effects on the mother-child relationship. Honours thesis; Macquarie University, Sydney, Australia

Crossley JR, Elliott RB, Smith PA (1979). Dried-blood spot screening for cystic fibrosis in the newborn. Lancet i: 472-74

Crossley JR, Smith PA, Edgar BW, Gluckman PD, Elliott RB (1981). Neonatal screening for cystic fibrosis by immuno-reactive trypsin assay in dried blood spots. Clin Chim Acta 113; 111-121

Dankert-Roelse J, Knolk K, ten Kate LP (1982). Second report on follow-up of cystic fibrosis: screening versus non-screening. Proc 11th Annual meeting EWG/CF Brussels, Belgium

Dodge JA, Ryley HC (1982). Screening for cystic fibrosis. Arch dis child 57; 774-780

Ganeshananthan M, Dauphinais R, Jezyk P (1985a). Diagnosis of cystic fibrosis: results of screening vs no screening in a three year programme. CF Club Abstracts, 26: 93

Ganeshananthan M, Dauphinais R, Muangman P (1985b). Clinical course of cystic fibrosis in patients diagnosed by screening, meconium ileus, and non-specific symptoms. CF Club Abstracts 26: 91

Gitzelman R (1981). Why we should not screen our newborns for cystic fibrosis. Helv paediat Acta 36: 493-494

Hammond K (1985). Early detection and follow-up of cystic fibrosis in newborns. The Colorado experience. Birth Defects Institute Symposium XVI Albany New York

Hassemer DJ, Laessig RH, Hoffman G: Neonatal screening for cystic fibrosis. Fourth National Symposium on Newborn Screening, Columbus Ohio (In press)
Holtzman N (1984). Routine screening of newborns for cystic fibrosis: Not yet. Pediat 73: 98-99
Kuzemko J (1985). Screening, early neonatal and prenatal diagnosis. J R Soc Med Supp 12 79: 2-5
Mastella G, Barlocco G, Braggion C, et al (1983). A long-term follow-up study of CF patients diagnosed by neonatal screening, or meconium ileus or symptoms In: Adam G (ed) Proc 12th Annual Meeting EWG/CF, Athens
Naylor EW (1985). Recent developments in neonatal screening. Seminars in Perinatology 9: 232-249
Phelan PD (1984). Screening for cystic fibrosis. In: Meadow R (ed) "Recent Advances in Paediatrics". Churchill Livingstone London Edinburgh pp103-120
Pritchard DJ, Hickman GR, Nelson R (1983). Sex ratio and heterozygote advantage in cystic fibrosis families. Arch Dis Child 58: 290-293
Reardon MC, Hammond KB, Accurso FJ et al (1984). Nutritional deficits exist before 2 months of age in some infants with cystic fibrosis identified by screening test. J Ped 105: 271-274
Rosenstein B, Langbaum T (1984) Diagnosis. In: LM Taussig (ed) "Cystic Fibrosis" Thieme-Stratton, New York pp85-114
Schiotz PO, Hoiby N, Flensborg EW (1981). Cystic fibrosis in Denmark. In: Warwick WJ (ed) "1000 years of cystic fibrosis" Univ of Minnesota, Minneapolis pp141-145
Sorenson JR, Levy HL, Mangione TW, Sepe SJ (1984). Parental response to repeat testing of infants with false positive results in a newborn screening program. Pediat 73: 183-187
Sturgess JM, Czegledy-Nagy E, Corey M, Thompson MW (1985). Cystic fibrosis in Ontario. Am J Med Genet 22: 383-393
Travert G (1986). Cystic fibrosis screening in the newborn. Medical Laboratory 14: 17-22
Wilcken B, Brown ARD, Urwin R, Brown DA (1983). Cystic fibrosis screening by dried blood-spot trypsin assay: results in 75,000 newborn infants. J Ped 102: 383-387
Wilcken B, Towns SJ, Mellis CM (1983). Diagnostic delay in cystic fibrosis: lessons from newborn screening. Arch Dis Child 58: 863-866
Wilcken B, Chalmers G (1985). Reduced morbidity in patients with cystic fibrosis detected by neonatal screening. Lancet ii: 1319-1321

CURRENT TRENDS IN INFANT SCREENING

Proceedings of the 7th International Screening Symposium, Sao Paulo, Brazil, 6-9 November 1988

Editors:

Benjamin J. Schmidt

Department of Pediatrics
Escola Paulista de Medicina
and
Head of the Laboratory
Association of Parents and Friends of the Mentally Handicapped, APAE
Sao Paulo, Brazil

Aron J. Diament

Head of the Service of Child Neurology
Department of Neurology
Faculty of Medicine
University of Sao Paulo
Sao Paulo, Brazil

Nicole S. Loghin-Grosso

Head of the Genetics Laboratory
Association of Parents and Friends of the Mentally Handicapped, APAE
Sao Paulo, Brazil

1989

Excerpta Medica, Amsterdam – New York – Oxford

Current trends in infant screening. B.J. Schmidt et al, eds.

AN ANALYSIS OF FALSE NEGATIVE SCREENING TESTS FOR CYSTIC FIBROSIS

BRIDGET WILCKEN AND ANDREW BROWN.
Oliver Latham Laboratory, New South Wales Health Department, Sydney, Australia.

Possible causes of missed cases

The discovery by Crossley, Elliott and associates [1] that immunoreactive trypsin is elevated in newborns with cystic fibrosis led to the speculation that this finding was due to the obstruction of pancreatic ducts and ductules by inspissated secretions, with a consequent back-leakage of enzymes from the still functioning acini. This view has never been seriously challenged, or, until recently, investigated. Thus we could speculate that the very mildly affected patients with little pancreatic duct blockage in the newborn period would be the cases of CF most likely to be missed by screening. On this reasoning, we would not predict that the more severely affected patients would be missed, since, histologically, although there is some fibrosis pancreatic acinar cells appear only mildly affected in CF patients dying in the newborn period, and are presumably still functioning.

Detection of missed cases

The criticisms of screening for cystic fibrosis voiced in 1983 by the Cystic Fibrosis Foundation's task-force on neonatal screening [2] suggested that testing by IRT assay would have a 10% obligatory false negative rate (because of the approximately 10% of patients who do not have obvious pancreatic involvement) but the reality seems lower [3]. It is also clear that the screening test does detect patients who have adequate pancreatic function [4].

In analysing who are the missed cases and why they were missed two issues are important: firstly, whether or not screening program directors would hear about all missed cases upon final diagnosis. In the case of our program in New South Wales, Australia, the answer is certainly yes. All patients in whom CF is suspected or diagnosed are nowadays referred to one of the CF clinics, and new diagnoses are notified to us both by the clinics and by the laboratories performing sweat tests. This may not be so in other regions and other countries. Such a situation could lead to an apparent false negative rate which is misleadingly low. Secondly, it is important to discover whether an appropriate cut-off point has been selected in the protocol of the test programme. This is a potential problem where a very rigid and unchangeable protocol has been adopted for purposes of a trial.

Overall, the known false negative rate in various programs ranges from 0% to 30%, a huge variation. However, if one discards data from the one screening program with apparent difficulties which are untypical, the mean rate for the

remaining 14 programs is 6.1% of all CF cases presenting without meconium ileus [3]. The test certainly is problematic in patients presenting with meconium ileus and this will be discussed later.

Expected age at diagnosis of missed cases

As screening programs have been in place for only a few years, it is important to consider what percentage of missed CF cases would be expected to be diagnosed at any given age. We cannot gain a good idea of this until we know the special characteristics, if any, of missed cases. However, in Australia we have comprehensive figures about the age at diagnosis (because of symptoms) of patients first seen in the years 1980-1986 from the neighbouring State of Victoria where, up until now, screening has not been carried out. Considering only patients in whom CF was not suspected at birth - that is, excluding babies with meconium ileus or with an affected sibling, 5% of the total were older than 3 years at diagnosis. The figures from Sydney for a three-year period before screening began are similar (Table 1).

TABLE 1

Age at diagnosis of CF patients because of symptoms.
(Patients with a CF sibling and those presenting with meconium ileus are excluded.)

Age (months)	Victoria, Australia 1980 - 1986 n = 115	New South Wales, Australia 1978 - 1981 n = 53
	%	%
0 - 11	73	58
12 - 23	13	25
24 - 35	8	13
36+	5	4

In the New South Wales screening program, among the cohort of children who were screened and who are already over the age of 3 years, there have been 122 CF cases, 98 of whom presented without meconium ileus. Of these 98 children, 4 (4%) were missed by the screening tests. On the basis of the foregoing data, we would not expect many more children from this cohort to be diagnosed as having CF. With a total of 216 patients now known, later results have not been so good and the overall figures now indicate a false negative rate of 6.5% in New South Wales.

Test characteristics of the missed cases

The immediate reasons for a false negative test have been analysed for eight

different centres - New South Wales, New Zealand, Peterborough, Caen, Connecticut, Colorado, Verona and Milan (Table 2).

TABLE 2. False negative results in eight programs (excluding meconium ileus patients)

1st IRT test not elevated		21
Very low (>50% below cut off)	5	
Low (49-10% below cut off)	12	
Borderline (<10% below cut off)	4	
2nd IRT test not elevated (1st elevated)		6
Very low	1	
Low	1	
Borderline	4	
Laboratory error		7
Technical	5	
Clerical	2	

Amongst these eight laboratories the known false-negative rate varied between 1.4% and 10% of CF cases (excluding meconium ileus). The resample rate for these laboratories varied between 0.3% and 0.7%, but there was no significant correlation between this and the false negative rate.

As can be seen from Table 2, 21 of the 34 false negative results occured with the initial blood sample, and most of these had values at least 10% below the cut-off value. Only our own program had false-negative cases with borderline values. Thus, by and large, an inappropriate cut-off point is not likely to be associated with these false negative cases. Although there is no data on what would be the resample rate if a more generous cut-off point had been adopted, so as to take in these CF cases, it is likely in most programs to be too large for practical purposes.

Clinical Characteristics

Of the eleven New South Wales cases who were missed, the negative result was due to a technical error in one (a repeat assay of the original blood spot carried out when the diagnosis was made after 13 months showed an unequivocally elevated IRT level). Clinical information relating to the time of the negative test was available for 9 of the remaining 10 cases. Six were acutely unwell when the test was performed. Four of the six had acute gastroenteritis probably due to rotavirus and one of those, together with two further patients, had an acute respiratory illness, also probably viral. Two patients developed chest infections within a week of the negative test. One patient was well. He has

very mild manifestations of CF, and was missed on the 2nd IRT test in which the result was borderline (less than 10% below the cut-off value). Soon after diagnosis at 16 months of age his pancreatic function test revealed pancreatic sufficiency, with lipase and colipase secretion approximately 25% of normal.

This finding of six of nine missed cases having intercurrent infections is very different from findings in CF patients identified by the screening test. In the first sixty patients identified in our programme only seven were said to have significant symptoms at the time the screening sample was collected, and of these, four had borderline test results which were above but within 10% of the cut-off point.

Among the 23 missed cases from other centres there was information of neonatal problems in eleven. Two had diarrhoea, two a chest infection and in one there were problems in establishing feeding. One patient had intrapartum distress and the remaining five had no problems. Thus, in all 20 cases where the neonatal status was known, 60% were either sick or known to be feeding poorly at the time the negative test sample was taken.

The later outcome of patients missed by the screening test was very variable. Certainly they cannot be classified, overall, as either very mild or very severe. Amongst our own ten patients with negative tests one is in the severely affected category, two have evidence of pancreatic sufficiency and can be categorised as mild and the remaining seven fall between these extremes. Of patients from the other seven groups, two are very mild, with pancreatic sufficiency, six are said to fall into mild category, seven are average, and the remainder were not classified (they were either too young, or their status was not known). However, four patients from the Verona program, who had initialy high IRT levels, proceeded straight to sweat-testing because the meconium lactase level was also elevated. At sweat-testing their repeat IRT levels were very low, and they can be classified as potentially missed cases. All had severe pancreatic insufficiency during the first month of life.

Age at diagnosis of missed cases

It is very important to discover whether the diagnosis of CF in a patient missed by the screening test is in any way delayed because of the false reassurance of a normal test result. There was anecdotal evidence of delay in diagnosis because of this in two of our patients. However, when the age at diagnosis of missed cases is examined, there is very little suggestion of delay, when compared with unscreened patients. Amongst the 33 missed cases from seven centres for whom a date of diagnosis is known, 20 (60%) were diagnosed during the first year and a further seven (21%) during the second years. Whilst there is a trend for the diagnosis to be slightly later than in the contemporaneus unscreened group from Australia already discussed (Table 1) this difference is

not significant. (Less than 1 yr vs 1 year and over : $X^2 = 1.6, p>0.1$.)

The problem of meconium ileus

Meconium ileus patients frequently have negative IRT screening tests, but results using different reagents have varied. In our experience using the Auckland University reagents 39 out of 41 MI patients had a positive test, whereas using Agen Biomedical reagents, results more closely resembled those obtained with other well-established reagents, and we have found only 3 of 7 MI patients to have a positive test.

The fact that the IRT screening test cannot with confidence be used to diagnose CF in a baby with meconium ileus (MI) seems at first to be entirely unimportant. These babies with meconium ileus have already declared themselves as almost certainly having cystic fibrosis and a screening test seems superfluous. However, our experience in New South Wales leads us to feel that a positive test in MI babies is often useful.

Amongst the 216 babies with cystic fibrosis identified during the screening programme, there were six very early deaths. All these six babies had presented with meconium ileus, and four died in the neonatal period, three following a volvulus and one a premature baby with severe bronchopulmonary dysplasia. Cystic fibrosis was not being actively considered in two of these patients until the screening result was made known. It is possible that parents of such babies who die early might never be in touch with CF clinics and thus could be denied the full benefits of genetic counselling and prenatal diagnosis in subsequent pregnancies.

Other cases with suboptimal screening outcome

There are other ways in which a screening program may fail. The screening test may detect a case, but the baby may become lost to follow-up before confirmatory testing can occur, or there may be undue delay in referral for confirmatory testing and treatment. In addition to these possibilities, a baby may never be tested, although a screening program is in place in the area.

Amongst 167 New South Wales babies with CF not presenting with meconium ileus four fell into these categories. One was detected, could not be found for follow-up, and presented at 10 months with failure to thrive. Two had delayed follow-up by the obstetric service and became seriously ill before this was accomplished. And one baby did not have a blood sample taken for screening. Thus although our known false-negative rate is 6.5%, our total 'suboptimal screening outcome' rate is 9% (15 cases). It is salutary to note that among these 15 cases the suboptimal result was clearly avoidable in 5.

DISCUSSION AND CONCLUSIONS

How do the actual findings accord with the theoretical considerations?

Amonst the missed cases reported, while there is a tendency for these to be at the milder end of the CF spectrum, there is no overwhelming preponderance of especially mild cases. We do not have enough information to know whether we are detecting those few CF patients who would normally not be diagnosed until teenage or later. It seems certain that if there were no pancreatic involvement at all in infancy, such cases would be missed. No such cases have come to light so far. However, patients with mild pancreatic disease and overall pancreatic sufficiency certainly are detected by the test [4].

There was also no preponderance of patients with a severe later outcome among the missed cases. On known neonatal histology, indicating minimal acinar cell damage at this time, we would not expect to miss these.

The finding that 60% of missed cases about whom there was appropriate information had an intercurrent illness or poor feeding at the time of blood sampling is important. It is difficult to see why intercurrent illness should be associated with lower than expected IRT values, as there is evidence that the serum IRT level is not related to food intake, and no diurnal variation is observed, in healthy adults [5]. The same appears to apply in newborns. We examined the relationship of blood-spot IRT levels to feeding in 1,666 newborn babies. There was no significant difference between the mean IRT levels or the distribution of values in 1,399 babies sampled immediately before feeding (mean IRT 77.8 AU/L) and 767 sampled immediately after feeding (mean IRT 82.6AU/L, standard error of the mean 4.139). Individual babies were sampled once only. None of these babies suffered from CF. For babies with CF the situation could well be different because of the partial pancreatic ductal blockage, and this merits study. In the meantime it appears prudent to suggest that the results of an IRT screening test performed on a blood sample taken when a baby is sick should be interpreted with care.

Among the false negative cases we analysed, this adverse result was clearly avoidable in 7 (20%) and may well have been avoidable in a further 8 (those with borderline results). However, there remains a small percentage of patients who have CF but whose IRT level in the neonatal period is probably too low for them to be detected by a screening test and who therefore could be considered as obligatory false negatives with the current protocols in use. Our evidence leads us to conclude that these are a mixed bag, including some exceptionally severe and some exceptionally mild cases and some who have an intercurrent illness. This should not deter us from feeling happy with the overall performance of CF screening by blood-spot IRT testing, but merely make us doubly aware of its potential fallability in a small percentage of cases.

REFERENCES

1. Crossley JR, Elliott RB, Smith PA. Dried blood spot screening for cystic fibrosis in the newborn. Lancet 1979; i: 472-4
2. Ad Hoc Committee Task Force on Neonatal Screening, Cystic Fibrosis Foundation: Neonatal Screening for Cystic Fibrosis; position paper. Pediatrics 1983; 72: 741-45
3. Travert G. Analyse de l'experience mondiale de depistage neonatal de la mucoviscidose. Proceedings of Conference Internationale Mucoviscidose: Depistage neonatal et prise en charge precoce . Caen 1988
4. Gaskin K, Waters D, Gruca M, Wilcken B, Dorney S. Pancreatic function in infants with cystic fibrosis identified by a neonatal screening program. Proceedings of Conference Internationale Mucoviscidose: Depistage neonatal et prise en charge precoce. Caen 1988
5. Moller-Petersen J, Pedersen S. Cathodic trypsin-like immunoreactivity in serum: influence of sex, age, renal function, food and diurnal variation. Clin Chim Acta 1982; 124: 31-37

PANCREATIC FUNCTION IN INFANTS IDENTIFIED AS HAVING CYSTIC FIBROSIS IN A NEONATAL SCREENING PROGRAM

DONNA L. WATERS, R.N., STUART F.A. DORNEY, M.B., B.S., KEVIN J. GASKIN, M.B., CH.B., MARGIE A. GRUCA, M.SC., MARY O'HALLORAN, B.SC., AND BRIDGET WILCKEN, M.B., CH.B.(EDIN.)

Abstract The use of the dried-blood immunoreactive-trypsin assay for the detection of cystic fibrosis in newborns has been questioned on the grounds that it may fail to identify patients with enough pancreatic function to have normal fat absorption. To investigate this possibility, we assessed pancreatic function in 78 patients identified in a neonatal screening program as having cystic fibrosis. The diagnosis of cystic fibrosis was confirmed by abnormal results on a sweat chloride test.

The results of measurements of fecal fat excretion, pancreatic-stimulation tests, and estimations of the serum level of pancreatic isoamylase indicated that 29 of the 78 children (37 percent) had substantial preservation of pancreatic function. These children (median age, four years) had growth that was close to normal and comparable to growth in children with severe pancreatic insufficiency who received oral enzyme therapy. Pancreatic insufficiency subsequently developed in 6 of the 29 patients, at 3 to 36 months of age.

We conclude that the serum immunoreactive-trypsin assay used in neonatal screening programs identifies patients with cystic fibrosis who have sufficient pancreatic function to have normal fat absorption and that a substantial proportion of infants identified as having cystic fibrosis are in this category. (N Engl J Med 1990; 322: 303-8.)

NEONATAL screening for cystic fibrosis with use of the immunoreactive-trypsin assay on dried blood samples has been adopted by many medical centers worldwide.[1-5] The rationale for screening is that early diagnosis permits early therapy, which in turn may decrease morbidity and mortality, and makes possible more effective genetic counseling. Screening is not universally accepted, however, because improvements in outcome with screening have yet to be demonstrated. Moreover, a 1983 task force considered data indicating that the trypsin screening technique might have an inherent false negative rate of 10 percent, since it had not been validated for the estimated 10 percent of patients with cystic fibrosis who have "normal" pancreatic function.[6]

This subgroup of patients has become increasingly important in the investigation of gene mutation,[7] pathophysiology,[8] and pancreatic dysfunction[9] in this disease. Although they have normal fat absorption, these patients do not, in fact, have normal pancreatic function, as is evident from their range of values for pancreatic colipase and lipase secretion, which may vary from as low as 1 and 2 percent of normal values, respectively, to values within the normal range.[10] They are therefore considered to have sufficient pancreatic function to prevent malabsorption ("pancreatic sufficiency"). In contrast to those with malabsorption, these patients have milder lung disease,[11] lower mortality,[12] and less severe deficits in fluid and electrolyte secretion,[13] and they rarely have liver disease, malnutrition, or intestinal obstruction.[14-16] Failure to identify such patients in screening programs could not only lead to longer delays in diagnosis but also interfere with the longitudinal assessment of the patients and the screening program.

This study was undertaken to investigate whether screening with the immunoreactive-trypsin assay could detect patients with pancreatic sufficiency and, if so, to determine the proportion of such infants in the population with cystic fibrosis and to identify their clinical characteristics in early life.

From the Departments of Gastroenterology (D.L.W., S.F.A.D., K.J.G., M.A.G.) and Biochemistry (M.O'H.), Royal Alexandra Hospital for Children, Camperdown, Sydney, Australia, and the New South Wales Department of Health, Oliver Latham Laboratory, Sydney, Australia (B.W.). Address reprint requests to Dr. Gaskin at the James Fairfax Institute of Paediatric Nutrition, Royal Alexandra Hospital for Children, Pyrmont Bridge Rd., Camperdown, NSW 2050, Australia.

Supported by a grant (87/0501) from the National Health and Medical Research Council of Australia.

Presented in part at the 10th International Cystic Fibrosis Conference, Sydney, Australia, March 5–10, 1988.

METHODS

Diagnosis of Cystic Fibrosis

In the New South Wales state screening program, immunoreactive trypsin is measured in dried blood samples, and a second blood sample is analyzed for babies whose initial values are elevated.[2] If the trypsin levels are elevated on both occasions, the patients are referred for sweat electrolyte testing by the standard Gibson–Cooke method.[17] A sweat chloride level of 60 mmol per liter or more is regarded as diagnostic of cystic fibrosis. For patients with borderline results (40 to 60 mmol per liter), the diagnosis is established by the demonstration of exocrine pancreatic acinar dysfunction, ductal dysfunction, or both on direct pancreatic-stimulation testing and by repetition of the sweat test at a later date.

Patients

Ninety-eight infants identified as having cystic fibrosis by immunoreactive-trypsin screening were referred to the cystic fibrosis clinic at the Children's Hospital, Camperdown, between July 1981 and July 1988. This clinic specializes in the care of patients in New South Wales who have cystic fibrosis, but referrals are made according to geographic area by a network of physicians and are unrelated to clinical considerations.

By January 1984, 41 patients had been referred to the clinic by the screening program. Thirty-four of these patients remained at the clinic and subsequently underwent routine assessments of their pancreatic function. A total of 57 infants were referred after January 1984, and the parents of 44 consented to a pancreatic-testing procedure near the time of diagnosis. Thus, a total of 78 patients (80 percent of those referred by the screening program) formed the study group. Informed consent was obtained from the parents of all infants who underwent pancreatic-assessment procedures.

Assessment of Pancreatic Function

Fecal-Fat Estimation

The parents of infants who were fed a standard cow's-milk formula, solid food, or both completed records of weighed food intake for five days; on days 3 through 5, the children's feces were collected. Fat intake was estimated by reference to computerized Australian

Reprinted from the *New England Journal of Medicine*
322:303-308 (February 1), 1990

food-composition tables.[18] Fecal fat was estimated by the standard technique,[19] and fat malabsorption was defined as a loss of 8 percent or more of fat intake in the feces, as defined previously.[10] In exclusively breast-fed infants, the fecal fat loss was considered abnormal if the excretion of fat exceeded 1.5 g per day during a three-day period, an amount derived from selected studies compiled by Fomon.[20] In patients who were receiving oral pancreatic-enzyme supplements, this medication was discontinued 24 to 48 hours before the beginning of the fat-balance study.

Pancreatic-Stimulation Tests

Pancreatic-stimulation tests were performed to determine the degree of pancreatic function in infants with borderline values for fecal fat excretion, to confirm the results of measurements of pancreatic isoamylase, or as the primary test of pancreatic function. After a four-hour fast (or an overnight fast in older children), the infants were sedated with oral secobarbital (10 mg per kilogram of body weight) and were simultaneously given oral metoclopramide (0.25 mg per kilogram). The test used a marker-perfusion system (10 μg of gentamicin per milliliter in 5 percent mannitol), a modification of the technique described previously.[21] After a 20-minute equilibration period, an intravenous infusion of secretin (KabiVitrum; 0.0625 clinical unit per kilogram per minute) and cholecystokinin (Kinevac, Squibb; 0.01 μg per kilogram per minute) was administered over a period of one hour. Colipase and lipase concentrations were measured in the duodenal aspirate, and the rates of secretion of these substances were calculated as described previously.[10,21] A rate of colipase secretion greater than 120 units per kilogram per hour (i.e., more than 1 percent of average normal values) was considered evidence of pancreatic sufficiency.[10]

Measurements of Serum Pancreatic Isoamylase

Estimations of the level of pancreatic isoamylase in serum, with use of the Phadebas Isoamylase test (Pharmacia Diagnostics), were made for 27 children over two years of age. By this method, the mean value for the pancreatic isoamylase level in normal children 2 to 4.9 years old is 51 U per liter (reference range, 13 to 89 U per liter [mean ±2 SD]).[22]

Overall, five patients had all three tests of pancreatic function, eight had pancreatic-stimulation tests and fecal-fat estimations, eight had fecal-fat estimations and measurements of serum pancreatic isoamylase, and two had pancreatic-stimulation tests and isoamylase estimations. Thirty-nine had only fecal-fat estimations, 4 had only pancreatic-stimulation tests, and 12 had only measurements of serum pancreatic isoamylase.

Growth Assessment

The height and weight of the 69 children who were at the clinic within six months of the end of the study period (July 1988) were standardized for age and sex against international reference values assembled by the U.S. National Center for Health Statistics and published by the World Health Organization.[23] Because weight distributions are not symmetrical in this reference population, separate standard deviations have been calculated for the upper and lower halves of each distribution by the U.S. Centers for Disease Control[24] to ensure approximation to normal values. Using this correction, we calculated standard-deviation scores (z-scores) by the equation $z = (y - x)/SD$ of x, where y is the child's height or weight, x is the median height or weight of a normal child of the same age and sex, and SD of x is one SD above or below the median reference value.

Statistical Analysis

Where applicable, the results have been presented as means ±SD. Comparisons of means have been made with the two-tailed Student's t-test, with the rejection level set at 0.05.

Results

Pancreatic function was assessed in all 78 patients in the study group. The mean age at testing in the 34 patients who received the diagnosis of cystic fibrosis before 1984 was 2.3 years (range, 1 month to 4 years), whereas the mean age at pancreatic assessment in the 44 patients who received the diagnosis from 1984 to 1988 was 2 months (range, 1 to 4 months). Of the 78 patients, 74 had the diagnosis of cystic fibrosis confirmed by sweat testing. Four patients had borderline results on the sweat test. Two of these patients later had repeat sweat tests that were diagnostic. The remaining two patients have not had repeat sweat tests, but had results on pancreatic-stimulation tests that were consistent with the diagnosis of cystic fibrosis. Twenty-nine infants (37 percent) were found to have evidence of pancreatic sufficiency.

Fecal Fat Excretion

The estimates of fecal fat excretion in 60 infants are shown in Figure 1. For 26 children who were bottle-fed or eating a normal diet for their age, estimates of fecal fat excretion, expressed as a percentage of oral fat intake, ranged from 3 to 65 percent. Of the eight patients in whom fat excretion was normal or borderline (3 to 7.8 percent), six had their level of pancreatic function confirmed by pancreatic-stimulation test (open squares in Fig. 1). In 33 exclusively breast-fed infants and one other 17-month-old child (with a fat loss of 20 g per day), the results are expressed as grams of fat excreted in the feces per day. The fat excretion of the breast-fed infants ranged from 0.1 to 15 g per day, and the results were normal or borderline in 16. Pancreatic-stimulation tests were performed in four of the breast-fed group (open squares) to determine their levels of pancreatic function.

Pancreatic-Stimulation Tests

Pancreatic-stimulation tests were performed in 19 children (Fig. 2). Colipase secretion ranged from less

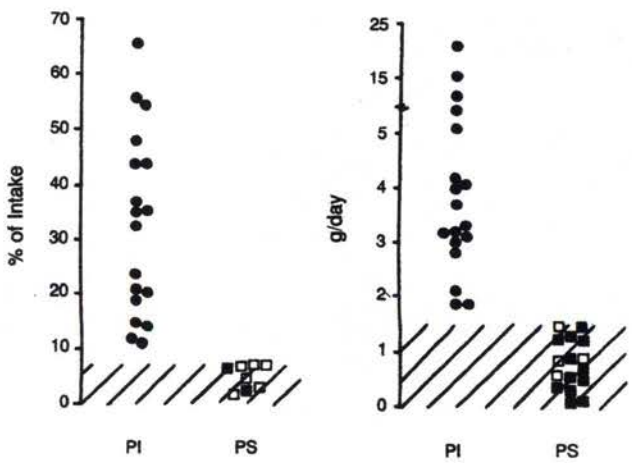


Figure 1. Fecal Fat Excretion in 60 Infants.

On the left, fecal fat loss is expressed as a percentage of fat intake for bottle-fed infants and those fed solids; on the right, it is expressed in terms of grams of fat per day for breast-fed infants. Patients with pancreatic insufficiency (PI) are represented by solid circles. For patients with pancreatic sufficiency (PS), solid squares indicate that pancreatic sufficiency was confirmed by fecal-fat analysis, and open squares that it was confirmed by both fecal-fat analysis and pancreatic-stimulation testing. The hatched area indicates the normal range.

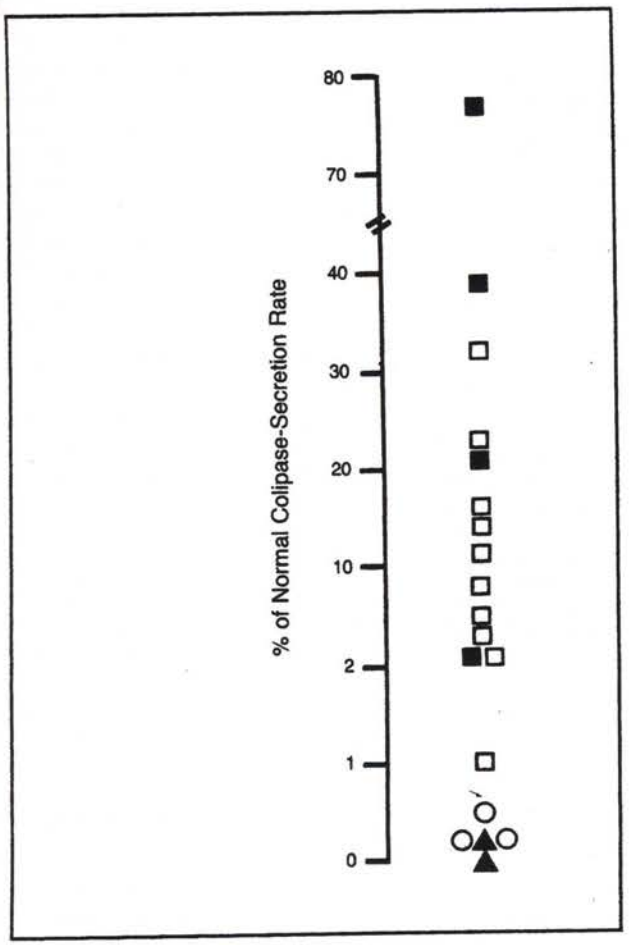


Figure 2. Results of Pancreatic-Stimulation Tests, Expressed as a Percentage of Normal Colipase-Secretion Rates.

Solid squares indicate patients with pancreatic sufficiency who had only a pancreatic-stimulation test, open squares those with pancreatic sufficiency who had both a fecal-fat analysis and a pancreatic-stimulation test, and triangles patients with pancreatic insufficiency. For three patients who initially had pancreatic sufficiency, the results of tests performed at the time pancreatic insufficiency developed (Table 1) are indicated by open circles.

than 1 percent to 77 percent of the previously defined average normal values,[10] with actual secretion rates ranging from less than 1 unit to 9800 units per kilogram per hour.

Pancreatic sufficiency was confirmed in 14 patients. Ten of these (represented by open squares in Fig. 1) had undergone fat-balance studies and had normal fat excretion. In one other child, the pancreatic-stimulation test was used to confirm the results of the isoamylase assay; the pancreatic-stimulation test was the only test of pancreatic function in the remaining three children (these four patients are represented by solid squares). The actual colipase-secretion rates in the group with pancreatic sufficiency varied from 127 to 9800 units per kilogram per hour (>1 percent to 77 percent of normal). Four of the patients with pancreatic sufficiency had initially borderline results on sweat chloride tests (49 to 60 mmol per liter), and the pancreatic-stimulation test was used to confirm their diagnoses. Repeat sweat tests subsequently verified the diagnosis of cystic fibrosis in two of these patients.

Pancreatic insufficiency (<1 percent of normal colipase secretion) was confirmed in five patients — in two at the initial assessment (triangles) and in the remaining three at a follow-up assessment two to three months after the documentation of normal fecal fat excretion (open circles).

Measurements of Serum Pancreatic Isoamylase

We estimated the serum levels of pancreatic isoamylase in 27 children more than two years of age. Figure 3 shows that, in general (with one obvious exception), the patients can be divided according to the results into two groups: one with undetectable or low-normal isoamylase levels (0 to 32 U per liter) — the patients with steatorrhea; and one with normal or high

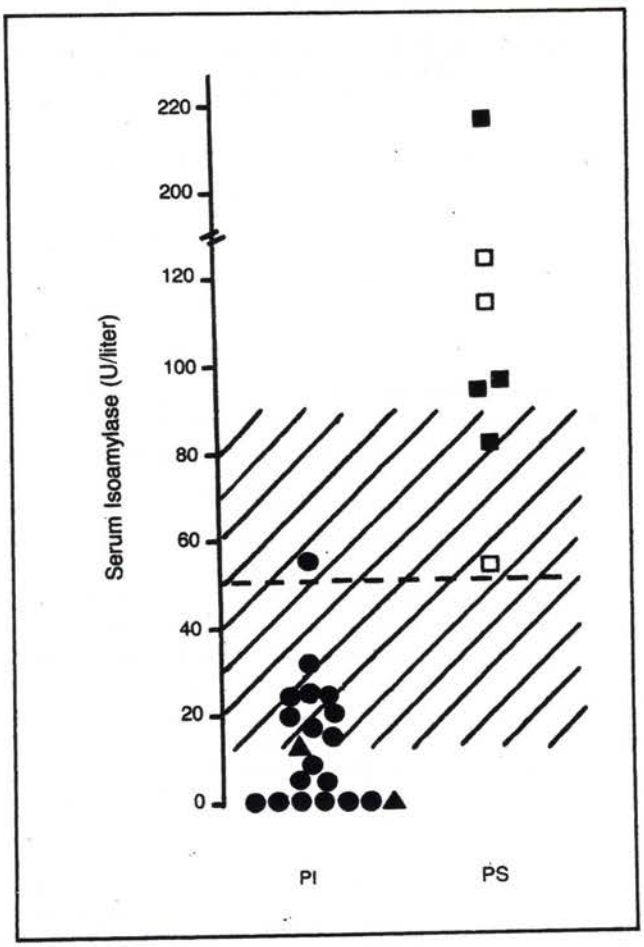


Figure 3. Serum Levels of Pancreatic Isoamylase in 27 Infants More than Two Years of Age.

Solid circles represent patients with pancreatic insufficiency, solid triangles two additional patients whose pancreatic insufficiency was confirmed by pancreatic-stimulation testing, open squares patients with pancreatic sufficiency who had both a fecal-fat analysis and a pancreatic-stimulation test, and solid squares three patients with pancreatic sufficiency whose pancreatic function was confirmed by fecal-fat analysis and one whose pancreatic function has not yet been confirmed. The dashed line indicates the age-specific mean for normal children, and the hatched area the mean ±2 SD for normal children.

levels (81 to 216 U per liter) — the patients with pancreatic sufficiency.

Repeat Tests

By the end of the study (July 1988), the median duration of follow-up for those who maintained sufficient pancreatic function to prevent malabsorption of fat was 2.8 years. In six infants who had pancreatic sufficiency at the initial assessment, pancreatic insufficiency subsequently developed (Table 1). Pancreatic function was retained only for short periods in three patients in whom the deterioration of exocrine pancreatic activity occurred between the demonstration of a normal fat balance and the time the pancreatic-stimulation test was performed (open circles in Fig. 2). These children were between three and five months of age, were not receiving enzyme supplements, and had had no symptoms of malabsorption. The other three patients (age, 16 months to 3 years) were reexamined when symptoms of malabsorption were reported. Previously, these patients had been thriving without pancreatic-enzyme–replacement therapy.

Immunoreactive-Trypsin Screening

The diagnosis of cystic fibrosis was delayed beyond the neonatal period in 4 of the 78 patients referred by the screening program. The delay was due to a negative result on the immunoreactive-trypsin assay in the cases of two infants, was attributable to postal delays in the case of another, and was due to a borderline result on the immunoreactive-trypsin assay and repeatedly borderline results on sweat tests in the case of the fourth. The last child has retained pancreatic sufficiency, but the others were found to have fat malabsorption when tested at the time of the diagnosis.

From 1981 until early 1987, the test used to screen newborns in New South Wales was a radioimmunoassay. During this period there was no significant difference between the mean immunoreactive-trypsin value in samples from the 38 patients with pancreatic sufficiency (418±175 arbitrary units per liter) and the mean value in samples from the 23 patients with pancreatic insufficiency (406±131).

Table 1. Development of Pancreatic Insufficiency in Six Infants.*

Patient No.	First Assessment			Second Assessment			Age at Loss of Function
	Date	CS	FF	Date	CS	FF	
1	6/84	2.2%	7.8%	11/84	—	22.9%	3 yr
2	7/84	—	0.28 g/day	8/85	—	28.0%	18 mo
3	8/84	—	0.95 g/day	10/84	0.2%	—	3 mo
4	9/84	—	1.2 g/day	11/84	0.2%	—	4 mo
5	3/87	4%	0.82 g/day	1/88	—	17.3%	16 mo
6	1/88	—	0.6 g/day	4/88	0.5%	—	5 mo

*CS denotes the colipase-secretion rate (expressed as a percentage of the normal value) detected on pancreatic-stimulation testing, and FF the fecal-fat loss as either the percentage of fat intake or grams per day.

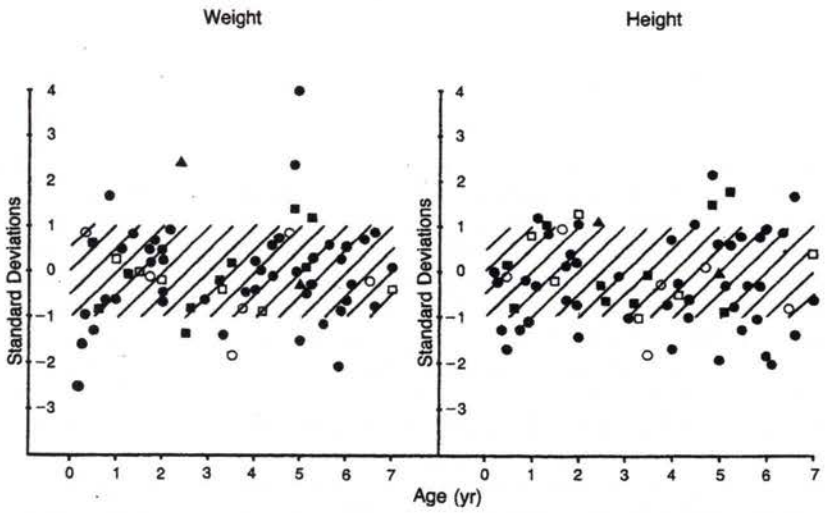


Figure 4. Cross-Sectional Growth Characteristics of the Screened Infants, Standardized for Age and Sex.

Solid circles represent patients who had pancreatic insufficiency (PI) from the time of the diagnosis of cystic fibrosis, open circles patients who initially had pancreatic sufficiency (PS) but in whom pancreatic insufficiency subsequently developed, triangles patients with pancreatic insufficiency whose level of functioning was confirmed by a pancreatic-stimulation test and measurement of serum amylase, solid squares patients whose preservation of pancreatic function was confirmed by a single test, and open squares those in whom pancreatic sufficiency was confirmed by both fecal-fat analysis and pancreatic-stimulation testing. The hatched areas indicate the mean ±SD. Only the 69 children who were at the clinic within six months of the end of the study period (July 1988) are included.

Growth

Figure 4 illustrates the cross-sectional growth characteristics of the infants with cystic fibrosis up to seven years after the screening program began. Patients with pancreatic sufficiency (squares) clearly conform to normal growth standards. With only one exception, the values for this group at a median age of four years were consistently within 1 SD of the median reference values, or higher. The patients who initially had pancreatic sufficiency but now require pancreatic-enzyme replacement for adequate food absorption are designated by open circles.

When adjusted for age and sex, the height and weight measurements of the screened population as a whole were not substantially different from those of the normalized reference population. The mean (±SD) standardized weight was −0.03±1.04, and the height was −0.1±1.09; neither was significantly different from the standardized normal value of zero. Only two children (with pancreatic insufficiency) had weights more than 2 SD below the reference value.

Discussion

The results of the present study demonstrate that the neonatal screening program for cystic fibrosis, in which the dried-blood immunoreactive-trypsin assay

is used, detects infants with pancreatic sufficiency, who make up a relatively large proportion (37 percent) of the infants referred to our clinic. Furthermore, these patients have a wide range of exocrine pancreatic-function levels, as is evident from their colipase-secretion rates (from 1 to 77 percent of average normal values, consistent with data on older patients not referred by a screening program[10]). These results thus confirm those of other investigators,[1,2] who suspected that patients from their screening programs had sufficient pancreatic function on the basis of normal trypsin activity in the stool. Our findings should allay concerns that a cystic fibrosis screening program may not be able to identify such patients.

Currently, we cannot precisely assess the rate of false negative results for patients with pancreatic sufficiency in screening programs, since the diagnosis may be delayed beyond childhood in such patients. However, on the basis of available data, we suspect that the rate is low, particularly because, in the present study, the proportion of patients with pancreatic sufficiency was high and their average immunoreactive-trypsin level was not lower than that of the group with pancreatic insufficiency. Furthermore, in New South Wales, the observed incidence of cystic fibrosis has previously been reported as 1 in 2600 screened infants.[25] This rate is close to that expected for our population and is similar to the 1 in 2500 reported from a neighboring state without a screening program.[26] Moreover, we have yet to see any additional patients with pancreatic sufficiency who were not identified by screening over the seven-year study period. Thus, considering all the evidence, it is unlikely that the screening program is failing to identify large numbers of such patients.

The proportion of infants with pancreatic sufficiency in our study group far exceeds both the figure anticipated by the task force and the 10 to 15 percent reported from cross-sectional studies of older, nonscreened patients.[11] However, this high proportion is in accord with Andersen's original findings at autopsy[27] that half of patients with cystic fibrosis who died in early life had less than 90 percent involvement of the exocrine pancreas on histologic examination. Moreover, the discrepancy between the young and older populations could be explained by a decline in pancreatic function with increasing age, as was documented in a preliminary report on an older group of patients[28] and as is evident in the results of our study. Thus, we anticipate that pancreatic insufficiency would develop in an even higher proportion of our patients over a longer follow-up period. Such findings would account for the variable older age at which some patients in nonscreened populations present with symptoms of malabsorption.

We observed distinct variations in levels of pancreatic function in patients with cystic fibrosis. Three groups were apparent: patients who had pancreatic insufficiency at or soon after birth; those in whom pancreatic function was initially sufficient but in whom insufficiency developed at a variable older age; and those who continued to have pancreatic sufficiency. A recent study suggests that in patients with pancreatic insufficiency a single mutational event occurs at the locus for the cystic fibrosis gene, whereas in patients with pancreatic sufficiency there are multiple mutations at this locus.[7] How these differences influence the development of pancreatic disease remains unknown. Studies in older patients have demonstrated that independent of the degree of acinar dysfunction, there is impaired secretion of bicarbonate and chloride and an associated, dependent fluid-secretion deficit that may predispose the patient to protein precipitation, ductal obstruction, inflammation, and eventual acinar destruction.[9,13,21] Furthermore, quantitative histologic studies of preterm neonates with cystic fibrosis have found dilated pancreatic ducts but morphologically normal acinar tissue.[29] It is possible that the severity of the secretion defect and thus the degree of obstruction of the pancreatic ducts is critical in the evolution of pancreatic disease.[30] This study emphasizes the need for further investigation of the relation between genetic differences and secretory anomalies in this disease.

The findings of the present study have important implications for the clinical management of cystic fibrosis. Our results indicate that pancreatic function should be assessed in all patients, since it is incorrect to assume that all patients with cystic fibrosis have malabsorption and require enzyme therapy. Not only is this an unnecessary burden for the patient with pancreatic sufficiency, but it is also an extra health care expense. The data on cross-sectional growth provide clear evidence that patients with pancreatic sufficiency can thrive without oral enzyme supplementation, a fact that attests to the long-term adequacy of their own endogenous enzyme production. Obviously, such patients need careful follow-up: poor weight gain or the onset of symptoms of malabsorption would necessitate the reevaluation of the patient's absorptive status. Finally, our results contrast with those of Sproul and Huang,[31] who found that the median heights and weights of nonscreened patients with pancreatic insufficiency were consistently below the 10th percentile for the normal population. In our study, the patients with malabsorption were clearly able to achieve normal growth with the early institution of enzyme therapy and without specialized or expensive diets.

We are indebted to the physicians and staff of the cystic fibrosis clinic for their help and support; to the parents and children for their continued cooperation; and to Miss Cheryl Frazer for assistance in the preparation of the manuscript.

References

1. Crossley JR, Smith PA, Edgar BW, Gluckman PD, Elliott RB. Neonatal screening for cystic fibrosis, using immunoreactive trypsin assay in dried blood spots. Clin Chim Acta 1981; 113:111-21.
2. Wilcken B, Brown AR, Urwin R, Brown DA. Cystic fibrosis screening by dried blood spot trypsin assay: results in 75,000 newborn infants. J Pediatr 1983; 102:383-7.
3. Mastella G, Barlocco EG, Antonacci B, et al. Is neonatal screening for cystic fibrosis advantageous? The answer of a wide 15 years follow-up study. Proceedings of the International Conference on Neonatal Screening for Cystic Fibrosis, Caen, France, October 5–6, 1988. Caen, France: University of Caen, 1988:127-43.
4. Reardon MC, Hammond KB, Accurso FJ. Nutritional deficits exist before two months of age in some infants with cystic fibrosis identified by screening test. J Pediatr 1984; 105:271-4.

5. Ryley HC, Deam SM, Goodchild MC, Weller PH, Bradley D, Carter RA. Screening for cystic fibrosis in Wales and the West Midlands. In: Mellis C, Thompson S, eds. Proceedings of the 10th International Cystic Fibrosis Congress, Sydney, Australia, March 5–10, 1988. Hong Kong: Excerpta Medica Asia, 1988:20. abstract.
6. Ad Hoc Committee Task Force on Neonatal Screening, Cystic Fibrosis Foundation. Neonatal screening for cystic fibrosis: position paper. Pediatrics 1983; 72:741-5.
7. Kerem BS, Buchanan JA, Durie P, et al. DNA marker haplotype association with pancreatic sufficiency in cystic fibrosis. Am J Hum Genet 1989; 44:827-34.
8. Durie PR, Forstner GG. Pathophysiology of the exocrine pancreas in cystic fibrosis. J R Soc Med 1989; 82:Suppl 16:2-10.
9. Kopelman H, Durie P, Gaskin K, Weizman Z, Forstner G. Pancreatic fluid secretion and protein hyperconcentration in cystic fibrosis. N Engl J Med 1985; 312:329-34.
10. Gaskin KJ, Durie PR, Lee L, Hill R, Forstner GG. Colipase and lipase secretion in childhood-onset pancreatic insufficiency: delineation of patients with steatorrhea secondary to relative colipase deficiency. Gastroenterology 1984; 86:1-7.
11. Gaskin K, Gurwitz D, Durie P, Corey M, Levison H, Forstner G. Improved respiratory prognosis in patients with cystic fibrosis with normal fat absorption. J Pediatr 1982; 100:857-62.
12. Corey M, Gaskin K, Durie P, Levison H, Forstner G. Improved prognosis in CF patients with normal fat absorption. J Pediatr Gastroenterol Nutr 1984; 3:Suppl 1:S99-S105.
13. Kopelman H, Corey M, Gaskin K, Durie P, Weizman Z, Forstner G. Impaired chloride secretion, as well as bicarbonate secretion, underlies the fluid secretory defect in the cystic fibrosis pancreas. Gastroenterology 1988; 95:349-55.
14. Roy CC, Weber AM, Morin CL, et al. Hepatobiliary disease in cystic fibrosis: a survey of current issues and concepts. J Pediatr Gastroenterol Nutr 1982; 1:469-78.
15. di Sant'Agnese PA, Hubbard VS, Lowe ME. Recent developments in clinical and basic research in cystic fibrosis. Monogr Paediatr 1981; 14: 1-25.
16. Zentler-Munro PL. Cystic fibrosis — a gastroenterological cornucopia. Gut 1987; 28:1531-47.
17. Gibson LE, Cooke RE. A test for concentration of electrolytes in sweat in cystic fibrosis of the pancreas utilizing pilocarpine by iontophoresis. Pediatrics 1959; 23:545-9.
18. Thomas S, Corden M. Metric tables of composition of Australian foods. Canberra, Australia: Australian Government Publishing Service, 1977:1-60.
19. van der Kamer JH, ten Bokkel Huinick H, Weyers HA. Rapid method for the determination of fat in faeces. J Biol Chem 1949; 177:347-55.
20. Fomon SJ, ed. Infant nutrition. 2nd ed. Philadelphia: W.B. Saunders, 1974:152-81.
21. Gaskin KJ, Durie PR, Corey M, Wei P, Forstner GG. Evidence for a primary defect of pancreatic HCO_3-secretion in cystic fibrosis. Pediatr Res 1982; 16:554-7.
22. O'Donnell MD, Miller NJ. Plasma pancreatic and salivary-type amylase and immunoreactive trypsin concentrations: variations with age and reference ranges for children. Clin Chim Acta 1980; 104:265-73.
23. Annex 3: reference data for the weight and height of children. In: Measuring change in nutritional status: guidelines for assessing the nutritional impact of supplementary feeding programmes for vulnerable groups. Geneva: World Health Organization, 1983:61-101.
24. Dibley MJ, Goldsby JB, Staehling NW, Trowbridge FL. Development of normalized curves for the international growth reference: historical and technical considerations. Am J Clin Nutr 1987; 46:736-48.
25. Wilcken B. An evaluation of screening for cystic fibrosis. In: Buchwald M, O'Riordan J, eds. Genetic and epithelial cell dysfunction in cystic fibrosis. New York: Alan R. Liss, 1987:201-15.
26. Allan JL, Robbie M, Phelan PD, Danks DM. The incidence and presentation of cystic fibrosis in Victoria 1955-1978. Aust Paediatr J 1980; 16:270-3.
27. Andersen DH. Cystic fibrosis of the pancreas and its relation to celiac disease: a clinical and pathologic study. Am J Dis Child 1938; 56:344-99.
28. Durie P, Corey M, Forstner G, Laufer D, Moore D, Cleghorn G. Longitudinal changes in exocrine pancreatic function in cystic fibrosis patients with pancreatic sufficiency: evidence for deterioration of pancreatic function. In: Mellis C, Thompson S, eds. Proceedings of the 10th International Cystic Fibrosis Congress, Sydney, Australia, March 5–10, 1988. Hong Kong: Excerpta Medica Asia, 1988:161. abstract.
29. Imrie JR, Fagan DG, Sturgess JM. Quantitative evaluation of the development of the exocrine pancreas in cystic fibrosis and control infants. Am J Pathol 1974; 95:697-707.
30. Zuelzer WW, Newton WA Jr. The pathogenesis of fibrocystic disease of the pancreas: a study of 36 cases with special reference to the pulmonary lesions. Pediatrics 1949; 4:53-69.
31. Sproul A, Huang N. Growth patterns in children with cystic fibrosis. J Pediatr 1964; 65:664-76.

Pediatric Pulmonology Supplement 7:69–71 (1991)

Assessment of Pancreatic Function in Screened Infants With Cystic Fibrosis

K. Gaskin, MB, ChB,[1] D. Waters, RN,[1] S. Dorney, MB, BS,[1] M. Gruca, MSc,[1] M. O'Halloran, BSc,[1] and B. Wilcken, MB, ChB[2]

Summary. Previously we have reported that 37% of infants with cystic fibrosis diagnosed by neonatal screening with the dried blood spot immunoreactive trypsin assay have pancreatic sufficiency. However, 34 of the 78 infants had pancreatic function tests an average 2.3 years after diagnosis, thus it was possible that the percentage with neonatal pancreatic sufficiency was underestimated, due to the loss of pancreatic function with time in some infants.

To assess this hypothesis we have assessed pancreatic function at the time of diagnosis in a further 20 infants since the completion of the previous study. Results of fecal fat determinations and/or pancreatic stimulation tests indicate that 10 (50%) of these infants have pancreatic sufficiency. Combining these results with those of the previous study, 31 of 64 patients (48%) have pancreatic sufficiency at this early age.

We have also monitored the progression of pancreatic disease in the 39 patients with pancreatic sufficiency recognized to date. Eleven have developed pancreatic insufficiency and require enzyme replacement therapy. Five others have shown further improvement of colipase secretion with age.

We conclude that the dried blood immunoreactive trypsin screening program for cystic fibrosis does recognize patients with pancreatic sufficiency, and at diagnosis nearly half the patients are in this category. To date, 28% of patients with pancreatic sufficiency have demonstrated a variable decline in pancreatic function with age. **Pediatr Pulmonol. 1991; Supp 7: 69–71.**

Key words: Cystic fibrosis, pancreatic sufficiency, colipase screening program

INTRODUCTION

Newborn screening for cystic fibrosis is now operative in a number of centers throughout the Western world, using the dried blood immunoreactive trypsin (IRT) assay.[1–5] However, concern has been raised regarding this technique, specifically as to whether it can detect infants with pancreatic sufficiency.[6]

Pancreatic sufficient patients are those with preservation of exocrine pancreatic function sufficient to produce normal fat absorption, and do not require oral enzyme replacement therapy.[7] Quantitative pancreatic stimulation tests in these patients and subsequent analyses of pancreatic fluid with very sensitive assays for lipase and its cofactor, colipase, have demonstrated the precise degree of preservation of pancreatic function.[8] The latter, in terms of colipase secretion, most often is not normal, but is greater than 1% of average normal values.[8] Patients with pancreatic sufficiency are interesting, in that they have a milder illness than patients with pancreatic insufficiency. Characteristically, they have less severe pulmonary disease, a better prognosis, milder defects in fluid and electrolyte secretion, uncommonly develop liver disease or intestinal obstruction syndrome, and rarely become malnourished.[7]

Previously, we have reported that 29 of 78 (37%) patients referred from the screening program have pancreatic sufficiency.[9] However, as pancreatic assessments were not performed in the first 2½ years of screening, 34 patients had pancreatic function tests on average 2.3 years after diagnosis, and only 8 (24%) had pancreatic sufficiency. In contrast, of the remaining 44 tested at the time of diagnosis, 21 (48%) had pancreatic sufficiency. With the continuation of the screening program we wished to assess if there was a higher proportion of pancreatic sufficient patients at the time of diagnosis.

METHODS

Patients

Subsequent to the previous study which was completed in July 1988, a further 24 patients have been referred to our hospital from the New South Wales state screening program. All patients had their diagnosis confirmed by

From the [1]Children's Hospital, Camperdown, NSW, Australia and [2]Oliver Latham Laboratory, Department of Health, Sydney, Australia.

Supported by Grants 87/0501 and 90/0206 from the National Health and Medical Research Council of Australia.

Address correspondence and reprint requests to Dr. K. Gaskin, James Fairfax Institute of Paediatric Nutrition, The Children's Hospital, Pyrmont Bridge Road, Camperdown, NSW 2050, Australia.

sweat testing by the Gibson–Cooke method with a sweat chloride of 60 mmol/L or above accepted as diagnostic of cystic fibrosis.[10]

Pancreatic Function Procedures

Fecal fat estimation

Infants on cow's milk formula underwent a 5-day fat balance study where a record was kept of all formula consumed. Stools were collected over three days and fecal fat estimated as described previously,[11] and expressed as a percent of fat intake. Fat malabsorption (pancreatic insufficiency) was defined as a fecal fat of 8% or more of fat intake.[8] In breast-fed infants only the 3-day fecal fat was estimated and was considered abnormal if fat excretion was more than 1.5 g/day.[9] All patients remained off oral enzyme therapy during the fat balance study.

Pancreatic stimulation tests

Pancreatic stimulation tests were performed by a quantitative marker perfusion technique, utilizing a double lumen duodenal tube as described previously.[9] Pancreatic sufficiency was confirmed by a colipase secretion value of greater than 120 U/kg/hr during stimulation of the pancreas using intravenous cholecystokinin and secretin.

Twenty of the 24 patients participated in this study and informed consent was obtained from all parents who underwent these procedures.

RESULTS

Of the 20 new patients tested, 9 had evidence of pancreatic sufficiency based on the fecal fat data, and in 6 of the 9 this was confirmed by pancreatic stimulation tests. One other patient had only a pancreatic stimulation test to confirm pancreatic sufficiency.

These pancreatic stimulation test data are summarized in Table 1.

During the last 2 years we have also continued to monitor pancreatic function in patients with pancreatic sufficiency from the current group, and also from the previous study. Over the total 9 years of screening, 8 children have developed pancreatic insufficiency as evidenced by a decline of colipase secretion to less than 1% of average normal or the development of fat malabsorption. A further 3 children have probably developed pancreatic insufficiency as shown by the development of oily stools, but have yet to be formally tested. All 11 patients are now receiving enzyme replacement therapy. Of the 6 children who had pancreatic stimulation tests to confirm pancreatic sufficiency and who subsequently have become insufficient, colipase secretion at the time of diagnosis was always less than 10% (range 1.2–6%) of average normal values.

TABLE 1—Colipase Secretion (% of Average Normal) in Patients With Pancreatic Sufficiency

Patient #	Colipase secretion (%)
1	2.2
2	5.8
3	9.0
4	4.6
5	5.4
6	55.0
7[a]	60.0

[a]Fecal fat not performed.

TABLE 2—Improvement in Colipase Secretion (% of Average Normal) in 5 Patients With Pancreatic Sufficiency

	Colipase secretion (%)	
Patient	First test	Second test
a	29	64
b	11	32
c	23	29.5
d	14	19
e	70	100

A further 5 patients with pancreatic sufficiency have shown an improvement in their colipase secretion with age, as shown in Table 2.

DISCUSSION

The current results confirm the findings of the previous study,[9] namely (1) that patients with pancreatic sufficiency are detected by the dried blood spot immunoreactive neonatal screening program, and (2) they form a relatively large proportion of the total population identified. The findings emphasize that the concern regarding a 10% false-negative rate, due to the screening program failing to detect pancreatic sufficient patients, is unwarranted. To date we have found only one patient who was missed by screening, and thus there is unlikely to be a large false-negative rate, although the screening program has only been in operation for 9 years.

The present results also demonstrate that nearly 50% of the infants undergoing pancreatic assessment at the time of diagnosis had pancreatic sufficiency. This proportion is considerably higher than that reported in studies of older patients from other clinics,[7] and is also higher than our previous results.[9] However, as alluded to in the Introduction, in the early phase of the previous study 34

patients had pancreatic studies some 2 years after the neonatal diagnosis of their cystic fibrosis. After that period of the study 44 had pancreatic function tests at the time of diagnosis and 21 (48%) were pancreatic sufficient. When these latter results are combined with the current results, 31 of 64 patients had pancreatic sufficiency at the time of diagnosis, a significantly ($P < 0.01$) higher proportion than those studied at an older age. It is thus likely that in the first 2 months of life nearly half of the patients are sufficient, but over a period of time varying up to 3 years, some will become insufficient, as shown by the loss of pancreatic function in 11 out of 39 pancreatic sufficient patients to date.

In older CF populations, most patients homozygous for ΔF_{508} are pancreatic insufficient, and most pancreatic sufficient patients are either heterozygous for ΔF_{508}, or have one or two of the other mutations.[12] It will be interesting to determine the genotypes of our younger group of patients to determine whether there is any correlation between genotype and pancreatic function, particularly in those who lose function over time. We suspect that those who are in this latter category manifest a more severe secretion deficit which predisposes to ductal protein precipitation, duct obstruction, and, finally acinar degeneration. As these patients all had quite low function at the initial testing, it is highly probable that even at that age they had already experienced considerable deterioration in acinar function.

The preservation of pancreatic function in these infants at least has important clinical implications. All patients at diagnosis should have a fat balance study to determine if malabsorption is present prior to instituting oral enzyme replacement therapy. Moreover, pancreatic function testing in those with normal absorption may provide useful information in pinpointing the group of patients most likely to develop pancreatic insufficiency at a later time.

REFERENCES

1. Crossley JR, Smith PA, Edgar BW, et al. Neonatal screening for cystic fibrosis, using immunoreactive trypsin assay in dried blood spots. Clin Chim Acta 1981; 113:111–121.
2. Wilcken B, Brown AR, Urwin R, et al. Cystic fibrosis screening by dried blood spot assay: results in 75,000 newborn infants. J Pediatr 1983; 102:383–387.
3. Mastella G, Barlocco EG, Antonacci B, et al. Is neonatal screening for cystic fibrosis advantageous? Proceedings of the International Conference on Neonatal Screening for Cystic Fibrosis, Caen, France, October 5–6, 1988. Caen, France: University of Caen, 1988: 127–143.
4. Reardon MC, Hammond KB, Accurso FJ. Nutritional deficits exist before two months of age in some infants with cystic fibrosis identified by screening test. J Pediatr 1984; 105:271–274.
5. Ryley HC, Deam SM, Goodchild MC, et al. Screening for cystic fibrosis in Wales and the West Midlands. In: Mellis C, Thompson S, eds. Proceedings of the 10th International Cystic Fibrosis Congress, Sydney, Australia, March 5–10, 1988. Hong Kong: Excerpta Medica Asia, 1988: 20 (abstract).
6. Ad Hoc Committee Task Force on Neonatal Screening, Cystic Fibrosis Foundation. Neonatal screening for cystic fibrosis: Position paper. Pediatrics 1983; 72:741–745.
7. Gaskin K, Gurwitz D, Durie P, et al. Improved respiratory prognosis in patients with cystic fibrosis with normal fat absorption. J Pediatr 1982; 100:857–862.
8. Gaskin KJ, Durie PR, Lee L, et al. Colipase and lipase secretion in childhood-onset pancreatic insufficiency: delineation of patients with steatorrhoea secondary to relative colipase deficiency. Gastroenterology 1984; 86:1–7.
9. Waters DL, Dorney SFA, Gaskin KJ, et al. Pancreatic function in infants identified as having cystic fibrosis in a neonatal screening program. N Engl J Med 1990; 322:303–308.
10. Gibson LE, Cooke RE. A test for concentration of electrolytes in sweat in cystic fibrosis of the pancreas utilizing pilocarpine by iontophoresis. Pediatrics 1959; 23:545–549.
11. van der Kamer JH, ten Bokkel Huinick H, Weyers HA. Rapid method for the determination of fat in faeces. J Biol Chem 1949; 177:347–355.
12. Kerem E, Corey M, Kerem B, et al. The relation between genotype and phenotype in Cystic Fibrosis–Analysis of the most common mutation (ΔF_{508}). N Engl J Med 1990; 323:1517–1522.

Screening, 2 (1993) 43–62

SCREEN 00038

Review

Newborn screening for cystic fibrosis: Its evolution and a review of the current situation

Bridget Wilcken

New South Wales Health Department, Oliver Latham Laboratory, and The Children's Hospital, Camperdown, Sydney, Australia

(Accepted 5 December 1992)

Almost 30 years after neonatal screening for cystic fibrosis (CF) was first proposed we at last have a suitable screening procedure, but are still not sure whether we should use it. The immunoreactive trypsin (IRT) test in neonatal dried blood samples has been greatly improved by the introduction of a two-tier IRT/DNA strategy, and there is now little doubt that the sensitivity and specificity is acceptable. However, the benefits of early diagnosis have not been fully established. There is evidence of reduced morbidity in the first year or two of life, as well as the opportunity to correct nutritional and biochemical deficiencies which have been documented in CF neonates. Long-term benefits have been difficult to demonstrate, and may only be evident in areas where conventional diagnosis is particularly likely to be delayed. Answers may yet be forthcoming in randomised controlled trials which are underway. The likelihood of any significant drawback to screening is receding. Early diagnosis does not appear to be associated with psychological harm but rather the reverse, and the numbers of false-positive tests have been greatly reduced by the use of DNA testing. The costs of screening are modest, and a positive short-term cost benefit is possible, but not yet substantiated.

Key words: Cystic fibrosis; Neonatal screening

The scope of this review

There has been an interest in screening for cystic fibrosis for almost 30 years. Two methods of biochemical detection – the measurement of electrolytes in sweat and of albumin in meconium – were developed in the late 1950s [38,41] and in 1964 the measurement of meconium albumin was mentioned as a possible test for screening of neonates [98]. It was not until 1979 that a test using a dried blood spot ushered in the 'modern era' of CF screening [21].

Correspondence to: Dr. B. Wilcken, Oliver Latham Laboratory, P.O. Box 53, North Ryde (Sydney), NSW 2113, Australia.

This review will consider only very briefly the use of meconium tests and sweat tests, but will cover in greater depth biochemical tests used for newborn screening in dried blood samples, including DNA tests. The performance of these screening tests worldwide will be discussed, but the main purpose of the review will be to examine the rationale for neonatal screening, and the currently perceived benefits and drawbacks of this form of early diagnosis. In addition, there will be a review of new insights neonatal screening programmes have given to the natural history of the disease. The use of sweat and other tests to confirm the diagnosis of CF will not be covered in depth. Screening for cystic fibrosis heterozygotes, and the prenatal diagnosis for cystic fibrosis are also not covered in this review.

A description of the disorder

Cystic fibrosis is a recessively inherited disease now known to involve the chloride channels in the apical membrane of epithelial cells. It is characterized principally by pancreatic and respiratory dysfunction. As a disease entity it has only been clearly delineated in the last 50 years. The term cystic fibrosis was first used by Anderson in 1938 in describing 49 patients with "cystic fibrosis of the pancreas" [8]. Clinically cystic fibrosis is characterized by the triad of pancreatic insufficiency, recurrent and eventually chronic lung disease, and increased sweat electrolytes. The pancreatic disease causes meconium ileus in about 20% of CF babies. It is also responsible for steatorrhoea, failure to thrive, and various specific nutritional deficiencies and, later in the course of the disease, pancreatic endocrine insufficiency, causing a condition resembling diabetes mellitus in about 10% of older patients. However, some 10 to 15% of CF patients show only partial pancreatic insufficiency, and retain enough function to prevent the syndrome of malabsorption. The respiratory disease of CF is associated with thick and sticky mucus, causing a suppurative lung disease, with *Staphylococcus aureus* and *Pseudomonas aeruginosa* being the significant organisms isolated, which may eventually colonize the lung chronically. Elevated levels of electrolytes in the sweat were first described in 1953 [27], and this feature even now forms the basis of the definitive test for cystic fibrosis, although this is now to some extent being supplemented by mutation analysis. Perhaps in the future a more definitive test is likely to be related to the gene product. Sweat testing for cystic fibrosis will be mentioned briefly a little later in this review. Amongst the many other clinical features of the multi-system disease are biliary cirrhosis, which occurs not uncommonly in older patients, and an absence or abnormality of the vas deferens.

Survival in cystic fibrosis

In the late 1930s and into the 1940s the mean age of death from CF was said to be well under 12 months, but the disorder, not then well known, was still regarded as one of the "new diseases in paediatrics" as recently as 1955 [51]. Nowadays, with treatment in expert centres, the median survival is approx. 30 years. Patients with CF can now in many cases grow up, obtain a job, and females may reproduce

(although sterility is the norm for males because of the absence or abnormality of the vas deferens). Nevertheless, children with CF still experience very considerable morbidity, and all will die prematurely. The burden of this disorder remains very great.

Genetics

The inheritance of CF is autosomal recessive, and the incidence varies among different populations. It is relatively common in Caucasians, with an estimated incidence of around 1:2000 newborns in the United Kingdom and parts of Northern Europe. It is less common in Southern Europe and in the Middle East, although the precise incidence amongst the newborn population is not known in many of these areas. In Northern Italy the incidence appears to be 1:3100 (Mastella, personal communication) and the situation could be similar in Egypt, in much of the Middle East, and quite possibly India, but at the moment this is unknown. Cystic fibrosis is rare in blacks and virtually unknown in Chinese, Japanese, and probably in most oriental groups.

The precise localization of the CF gene on the long arm of chromosome 7, region q31, facilitated various molecular strategies to isolate the gene, and in 1989 the gene responsible for CF was identified [53,71,76]. The most common mutation in the CF gene which causes the disorder of cystic fibrosis is a three base pair deletion, resulting in the loss of the phenylalanine residue at position 508. This mutation, the ΔF508 mutation, accounts for up to 80% of the CF mutant chromosomes in some populations. Where CF is commonest, this mutation is also common, and in several geographical populations 92–96% of CF sufferers have at least one copy of this mutation. There has been rapid progress since 1989 and over 200 other DNA mutations causing CF have now been characterized. The frequency of the common mutation has been established for a wide range of populations [23] but the incidence of each of the other individual CF mutations is rare in most populations so far studied.

The CF gene product, and the pathogenesis of CF

The predicted CF gene product has been called the cystic fibrosis transmembrane conductance regulator (CFTR). Several studies have indicated that the basic defect in cystic fibrosis is associated with the cAMP-mediated regulation of chloride ion transport. In normal epithelial cells chloride channels are opened by an increase in the intracellular level of cAMP. In CF cells chloride channels are present but cannot be opened by this mechanism. Recent studies make it seem possible that CFTR is itself acting as a chloride channel [9,52]. The pathophysiology of the disordered function in CF is not fully worked out at present, but this major defect results in impaired secretion of water, electrolytes and bicarbonate in the pancreas, with resultant protein hyperconcentration and blockage of acinar ducts. The mucous secretions in the lungs are viscid, and indeed the water content is probably low in all mucus, and this contributes to other intestinal symptoms including meconium ileus.

The proposed rationale for screening

The possible advantages that might come from pre-symptomatic diagnosis were considered in the early 1980s by a number of authors [28,37,66]. Prior to this there had been little evidence presented [81], mainly because of the difficulty of conducting controlled trials of early treatment. In 1977, Orenstein and colleagues [62] examined the outcome in 16 sibling pairs and found that the younger, early diagnosed siblings had, on the whole, a much better outcome at the age of 7 years, and they felt this was significant as they could identify no therapeutic advances which took place between the diagnoses of the older and younger siblings.

Phelan [66] discussed several theoretical advantages: (i) screening could alter favourably the natural history of the disease; (ii) parents of a CF child could take advantage of genetic counselling to plan further additions to the family. Without screening CF children would be diagnosed too late for their families to take advantage of this before the birth of the next child; (iii) the period of pre-diagnosis anxiety would be eliminated; (iv) there would be scientific advantages in discovering the true incidence, and in forming groups for the critical study of new forms of treatment. The possible disadvantages included: (i) interference with developing family relationships; (ii) psychological damage following false-positive tests; (iii) 'labelling' of patients with mild disease. A number of authors expressed strong views against the introduction of any screening programs [40,50], and in 1983, the Cystic Fibrosis Foundation (USA) convened a task force to consider this question [4]. Their paper cautioned that considerable research was needed to answer questions on possible usefulness and harm in early diagnosis before any mass screening should be undertaken.

The tests used in neonatal screening for cystic fibrosis

Over the last 30 or so years, a number of tests have been proposed for the early diagnosis of cystic fibrosis. These can be divided into (i) tests of electrolyte content of tissues; (ii) tests of pancreatic function in meconium and stool; (iii) tests of immunoreactive trypsin in blood; and (iv) DNA tests. Only the latter two will be fully discussed.

1. The sweat test and other tests of tissue electrolytes

The sweat test still remains the major test for diagnosing or excluding CF in a child with suggestive symptoms. This test requires the collection of sweat stimulated by iontophoresis and the measurement of sodium and chloride in the sweat. Elevated levels of electrolytes, especially chloride, are considered diagnostic of CF. Attempts to adapt this test for use in mass testing of infants have not been successful, largely due to cost and logistic problems. Similarly, measurement of electrolytes in nail clippings and saliva were also too complex for screening purposes. The methods of sweat testing, and other tests of electrolytes, have been reviewed by Heeley and Watson [46] and will not be further considered here, but it is important to note that

the availability of accurate sweat testing (and this is a notoriously difficult test to perform accurately) is an essential facet of CF screening.

2. Meconium albumin and lactase, and faecal trypsin

Elevated levels of albumin were noted in babies with meconium ileus in 1958 [41], and by 1975 Stephan et al [84] was able to promote a chromatographic strip test, the BM-test (Boehringer Mannheim), which soon came into widespread use. There have been many studies of this test and, although it has been quite successfully used in some programs [60], there is undeniably a high rate of both false-positive and false-negative results [49,67]. Modifications of the test, and an alternative test using meconium lactase, have not given better results [43], although the lactase test has been used effectively as an adjunct of a blood screening test [64] (see below). Again, one major problem of the meconium test is a logistic one. For good performance, it is not only necessary to collect the first or second meconium – an opportunity not afforded for long in the life of the neonate – but to transport the sample to a central laboratory at a temperature not exceeding 4°C. The test is not now considered useful in newborn screening programmes, except in some restricted circumstances. An improved test of faecal trypsin content was devised by Crossley et al. [20], and had some advantages over meconium testing in that samples did not need refrigeration and could be re-collected, if necessary. However, this test also had high false positive and false negative rates [34]. These meconium and faecal tests were also reviewed by Heeley and Watson [46].

3. The immunoreactive trypsin assay

In 1973, Crossley et al. [21] published an exciting report showing that, while in older patients the level of serum immunoreactive trypsin (IRT) was lower than normal, the level was raised early in the course of CF. They also reported the measurement of IRT in dried blood spots by radioimmunoassay. This important study renewed interest in screening for CF.

The basis for the assay. The species of trypsin actually measured as circulating IRT appears to be largely the zymogen trypsinogen [10]. Trypsin complexed with α_2-macroglobulin is not immunoreactive [83], but trypsin bound to α_1-antitrypsin shows considerable cross-reactivity with antibodies raised against human trypsin. Crossley and colleagues used trypsin prepared from human pancreas, and the final solution contained no chymotrypsin immunoreactivity [22]. Commercially available kits used various different methods of preparation. The pancreas secretes both cationic and anionic trypsinogens, and these are immunologically distinct. The specificity of the several different IRT assays available varies, and it is sometimes not clear which species of trypsinogen are being measured. Because of this, and because there is no international reference material for this assay, the results obtained with the different assay reagents are not comparable. This aspect of CF screening has been excellently reviewed by Heeley and Bangert [44a].

Since it was shown that blood IRT is raised in neonates with CF, it has been assumed that this is secondary to thickened secretions blocking pancreatic ductules, sufficient to impede the flow of all or much of the pancreatic juice into the duodenum.

The decline in IRT levels with increasing age presumably reflects secondary atrophy and degeneration of the pancreatic acini, and their replacement by fibrous tissue.

Confirmation of the early findings. The findings of Crossley and co-workers were quickly confirmed, and the assay method modified to make it more sensitive and more suited to newborn screening [54,10,29,55]. The assay could be performed on a dried blood sample as small as 3 mm in diameter [55], and, in retrospective studies, all investigators found a clear differentiation between blood from CF neonates and controls born at the same time. Kirby and colleagues found elevated levels of IRT in a small number of CF patients up to the age of 1 year [55].

The assay of Crossley was performed using an in-house radioimmunoassay (RIA). Subsequently, two other RIA kits (Behring-Hoechst and Sorin) were developed and found to be suitable for mass screening. A number of these investigators also assessed the effects of different variables on the IRT level on normal neonates. Dominici and colleagues [29] found no differences with gender, age between 3–30 days, weight from 2700–4650 g, and age of the sample to about 8 weeks. Kirby et al. [55], however, demonstrated a decline in recoverable IRT in samples stored for 5–6 weeks compared to fresher samples 1–2 weeks old. Neither of these workers specified storage conditions. Later work demonstrated that, while there was no change in IRT concentration after 32 weeks gestational age or during the postnatal days 1 to 5, levels were significantly lower in eutrophic babies 31 weeks gestation or less [17].

Prospective studies. In the early 1980s several prospective studies were reported [5,11,18,42,45,63,88,95]. All of these used RIA methods. In addition, all used a protocol similar to that reported by Crossley et al. in a small prospective study [22]. This protocol included IRT assay in the routine newborn screening sample, taken at approx. 4 to 5 days of age, re-assay and a second sample sought when the IRT was elevated, and sweat test performed if the result was elevated in this second sample. This will be referred to as a two-step IRT protocol. Heeley et al. [45] used such a protocol, with reagents from CIS (UK) Ltd, when they tested 39,000 infants of whom 0.33% required a second test, and diagnosed CF in 16. Two other babies without CF had persistently elevated IRT levels. We found similar results in 75,000 Australian babies, but we required a second test in 0.58%. There were only two false-positive results after the second test, and one false-negative result, and 35 infants were found to have cystic fibrosis (incidence 1:2143) [95].

The need for a two-step IRT protocol is mandated by the large proportion of infants who have an elevated blood spot IRT at 3–5 days, but who do not have CF. In CF, levels typically remain elevated for several weeks, while rapidly declining in non-CF infants. However, care needs to be taken with the timing of the second sample (see below).

Modification to the IRT test. Certain modifications to the two-stage IRT test were proposed. An extended protocol using a meconium lactase test in babies who had a positive IRT was successful for one group (where meconium was being routinely collected), and this increased the specificity [64]. Blood immunoreactive pancreatic lipase (IRPL) measurements were also explored [43,6] and a combined test with IRT and IRPL was proposed [43] but not used in the longer term [42].

In 1987, Bowling and colleagues developed a monoclonal antibody-based trypsino-

gen enzyme immunoassay [15] and this compared well with the previous IRT method [14]. We found some differences, however, particularly in a lower positive predictive value after the second test, and an increase in the apparent rate of decline in IRT with age, making the timing of the repeat test critical (unpublished data). A monoclonal antibody-based time resolved fluorescence immunoassay (DELFIA) has been described [65] which has not yet had widespread field trials, but which in a limited trial showed an increased discrimination between CF and non-CF babies [89]. A DELFIA IRT test has also been developed as part of a quadruple test, used in one assay system for simultaneous measurement of three other analytes – thyroid stimulating hormone, creatine kinase and 17-hydroxyprogesterone [99]. This development will be watched with interest.

The most important new modification to CF screening, however, is the combined use of an IRT test with DNA analysis for the common CF mutation, delta F508. This is likely to be the form of CF screening generally used from now on (see below).

The age-related decline in IRT values. It has always been appreciated that the high levels of IRT found in neonates with cystic fibrosis decline with age. Rock and colleagues [74] studied this well-known decline of IRT values by examining initial and follow-up samples from 36 infants with cystic fibrosis. In nine, the follow-up sample had an IRT value which had declined to below the cut-off value within the first 2.5 months. Hammond et al. [42] had similar results, although only five of 53 infant values fell to below their original cut-off value (which they later lowered) within 20 to 80 days. Our own findings with our initial assay, using the reagents of Crossley, were still less divergent with only two of 45 infants having levels below the cut-off point by age 3 months [92]. Using a monoclonal antibody-based ELISA assay, however, we have noted a more rapid decline in a higher proportion of infants.

The IRT test and meconium ileus. The IRT test result is often below the cut-off level (i.e., negative) in CF babies with meconium ileus (MI). To some extent this may seem unimportant, in that they have already 'declared' themselves. However, a negative test result can be misunderstood by primary care physicians. In our experience, the proportion of babies with MI who have negative IRT tests varies according to the reagents used: using a polyclonal RIA assay 39 of 41 MI patients had a positive result; using a monoclonal ELISA assay only about half were positive, a finding more akin to that of other groups [93a].

False-positive and false-negative results. Curiously little has been written about the causes of false results in IRT screening. A careful study done in Wisconsin [75] showed that perinatal stress factors accounted for a proportion, perhaps 25%, of babies with elevated IRT levels but without CF, and we and others have observed elevated IRT levels associated with renal failure, bowel atresias, congenital infections and with trisomies 13 and 18 [39,44b,93]. There is also a familial occurrence, and recently, with the advent of DNA technology (see below), it has been suggested that in babies with elevated IRT levels, without CF, heterozygosity for CF may be more common than expected [56]. Clearly more work needs to be done in this area. Despite all these associations, most babies with 'falsely' elevated IRT levels appear healthy.

False negative results are harder to categorise. We studied 11 CF babies who had

had negative tests. In one there had been a laboratory error, and in six there had been an acute illness, either respiratory or gastrointestinal, at the time of the falsely negative test (most of these were second tests). This contrasted with the occurrence of acute illness in CF babies with positive tests – only seven of 65 had any adverse symptoms at the time of the test. Others had similar but less marked experiences [93a].

Case finding with a two-stage IRT assay

A number of follow-up reports of prospective studies have been published, and overall they give a similar picture [47,19,33,42,87,92]. The situation in 1988 was reviewed by Travert [87] who collated the results of screening over 2 million babies. The analytical systems that were extensively used included three radioimmunoassays and, more recently, one ELISA kit. Three screening strategies were considered: the two-stage IRT assay (12 of 15 laboratories from which results were analysed), a sweat test performed after a single positive screening test result (two laboratories) and a complementary meconium lactase assay with the screening test (one laboratory). Among 2.2 million babies screened by the two-stage IRT test, 730 CF cases were identified, of whom 559 were detected by screening, 124 had meconium ileus and 47 (6.4% of the total) were missed by a screening test. There were no striking differences in performance among the three radioimmunoassays, but there were very striking differences in the performance of individual laboratories, with false negative rates varying from 1.5% to 24%. The positive predictive value (PPV) of the first test (excluding one laboratory where there appeared to be particular problems) varied between 10.3% and 2.7% with an overall figure of 5.1%. The PPV of the second test was much higher with a mean of 52%. Still there were many variations. In our program during this period the PPV of the second test was 79%. Travert has reported a similar figure of 76% [86] and Heeley et al. 80% [47]. However, Hammond et al. recently reported a PPV of only 24% for the second test [42], although earlier they have achieved 80% [43]. Similarly, the performance of the assay in our program in recent years gives a PPV of under 50% for the second test (unpublished data), whereas initially it was 92% [95]. In our case this appears to be partly due to the use of a different antibody.

It must be remembered that the choice of a cut-off point is a major feature affecting sensitivity and specificity, as well as PPV. Such choices are usually made empirically and different laboratories have opted for different approaches, some using fixed and some using floating cut-off points. In addition, there has been no agreement about the use of a lower cut-off point for the second test, probably because of the variations which have been observed in the rate of decline of IRT with age (see above). There are considerable problems in evaluating these collected data, particularly in evaluating the sensitivity. For example, we noted six missed cases in the cohort born between July 1981 and June 1986, giving a false negative rate of 3.8% [92]. In the subsequent 6 years, however, four missed cases have emerged from this cohort giving a revised false negative rate of 6%. Since 34% of the total cystic fibrosis cases from that cohort had either meconium ileus, or were siblings of known cases, the false negative rate amongst totally unsuspected cases was 9%. This modification of figures would almost certainly apply to all reported

series, as some screening programs may not have adequate surveillance methods in place. Our figures agree well with the figure of 11.5% from Hammond et al. of unsuspected cases being missed [42]. Nevertheless, at least one program, with good surveillance practices, has maintained an enviably low false negative rate: Heeley and colleagues report that for the first 9 years of their program (1980–89) they have detected 79 patients with CF (excluding surgical meconium ileus) and missed only one case, a sensitivity of 98.75% [47] By and large, however, reported false negative rates must be accepted with caution.

The current situation with a two-stage IRT protocol. We can conclude from the above that the case finding potential of a two-stage IRT test may be excellent in some hands, but that the screening protocol is a demanding one, and the sensitivity may not be ideal. Even so, about 95% of all CF babies achieve a diagnosis by 2–3 months of age. The problem in this strategy is the very low positive predictive value after the first test, which ranged from 3–10% in the collected data.

The current situation has been well reviewed recently in reports emanating from the randomised controlled trial of CF screening at present being conducted in the state of Wisconsin in the USA [31–33]. These reports and other recent reappraisals [42] show that while the IRT test in the two-step form is reasonably efficacious, there are inherent uncertainties in its performance.

The two-tiered IRT/DNA approach

The characterization of the CF gene and the discovery of the common mutation immediately made a single-time two-tier screening protocol, using both IRT and a mutational analysis, seem an attractive possibility for CF screening, and this approach has been adopted by most screening programs. Reports from Australia, Wisconsin (USA) and France have been very promising [33,56,68].

The common DNA mutation, ΔF508, is present overall in about 70% of all CF chromosomes, but there is marked variation in this proportion in different populations. Broadly, this mutation is common in populations derived from Northern Europe (about 70–80% of CF chromosomes) and less common in Southern and Eastern European populations, where typical prevalence levels are 40–60% [23]. Each of the other 200 or so CF mutations now discovered is rare. Even so, in many populations up to 94% of all CF babies will have at least one copy of the common mutation (e.g., if the prevalence of this is 75% of all CF genes) and this is the reason why a DNA approach included in screening can be successful.

The protocols successfully adopted have involved an IRT test on a filter paper blood spot obtained as usual on day 3 to 5, followed by a mutation analysis (on the same sample) carried out in those samples with elevated IRT levels. As a relatively low cut-off level can then be employed, the sensitivity of this method depends largely on the likelihood of the occurrence of the common mutation in at least one chromosome. In South Australia, in samples from 12,000 babies, those samples with an IRT in the top 1% of values had an analysis for two mutations, ΔF508 (prevalence 72% in South Australia) and ΔI506 (1%) [68]. Seventeen of 148 samples so tested had at least one of these mutations. Four of these babies were homozygotes, or compound heterozygotes, and thus clearly had CF, while among the remaining 13

two had positive sweat tests and the other eleven were carriers of CF. The families of the carrier infants were offered genetic counselling and further testing, and all accepted this. Similar results have been reported by others [32,56].

The precise approach and the results will, of course, depend upon the genetic makeup of the community being screened, but the advantages of the approach are clear, at least in communities derived mainly from Northern Europe where one could expect to identify around 95% of the CF population. Such a test protocol is likely to have somewhat increased sensitivity and greatly increased specificity with a much greater PPV. In the South Australian experience quoted above, the PPV was 35% and the false positive rate was only 0.09%. The only drawback is that the 'false-positive' infants are all carriers of CF and identifying them with the consequent genetic counselling and testing of their families creates new issues. Furthermore, although usually only one parent will be a carrier, the other parent, with no identified mutation, cannot be absolutely excluded as a carrier. However, Robertson et al. found no apparent adverse effects from this screening protocol in South Australia [73]. DNA testing can probably be incorporated into most existing routine screening programs: the author is aware of this occurring successfully in a number, both in Australia and elsewhere. The IRT/DNA approach appears likely to be generally acceptable in communities. It answers most of the problems of specificity. However, there will be a small percentage of unavoidably 'missed cases' – no more in number than with the current two-step IRT protocols (in fact probably fewer) but because the protocol causes these to be *unavoidably* missed, careful explanation and education will be needed to practising physicians.

Evaluating the benefits and drawbacks of screening

The results of some earlier prospective studies in regard to apparent benefits and drawbacks of early diagnosis of cystic fibrosis have been evaluated, but since these studies [13,25,26,59,96] did not have completely comparable controls, there are problems in achieving a satisfactory evaluation. Two controlled randomised (or virtually randomised) trials have now been progressing for some time [19,31–33,79] and data from these will have more weight. Nevertheless, important data are available from the early studies.

Clinical outcome

Early studies. In our study several years ago we used historical controls for analysis of early morbidity before and during neonatal screening [96]. We compared hospital days during the first 2 years of life among 48 patients without meconium ileus born in the 3 years before screening began, and diagnosed clinically, with 34 born in the first 2 years of screening. The difference was substantial with a mean of 27.7 hospital days for the clinically diagnosed patients and only 4.8 hospital days for the screened patients. Only one of 34 screened patients but 20 of 48 non-screened patients were hospitalised for over 21 days. Even after the diagnosis of CF, clinically

diagnosed patients had a mean of 16.3 days of hospital in their first 2 years [94]. We considered the possibility that changes in management might have occurred during the period of screening to explain marked reduction in hospital days. However, two lines of argument suggest that our findings were valid. Firstly, there was no trend with time to an improving rate of hospitalisation. The change was abrupt, and coincided with the onset of screening. Secondly, patients presenting with meconium ileus were considered separately, and although the numbers were small, the hospital days in the first 2 years (excluding initial hospitalisation related to surgery) were similar for patients born before and after the onset of screening. Thus, these patients constituted a form of internal control. A similar study from Queensland, Australia had substantially the same findings [13].

Dankert-Roelse and colleagues [25] compared the clinical outcome in patients born in the same region, and in the same time period, 16 of whom were screened by meconium analysis, and 20 of whom were not. All screened patients were included, whether or not they were detected by the test. Treatment was not the same for all patients initially, but it was after 1980. In 1987, screened patients had better survival, and better clinical and chest radiograph scores than the non-screened. There were no differences in height, weight, liver function tests or fat absorption, but IgG was significantly lower and serum vitamin A significantly higher in the screened patients. Other studies have shown similar intimations of possible benefit [26,16].

Perhaps the most interesting non-randomised study, and certainly the one with the longest follow-up period, has been that of Mastella and colleagues in North Eastern Italy [59]. They have studied CF patients diagnosed in a defined region from 1973 to 1988, and all managed by three CF centres which had a homogenous management policy. The decision to test for CF or not in neonates was a local area decision, and never based on individual babies. The test was initially a meconium test, and from 1981 was a modified two-stage IRT test. Since 1978, about 80% of all babies were screened. CF patients were evaluated in groups – screened, not screened, missed by screening and meconium ileus. The median age at diagnosis for screened patients was 1.8 months ($n=144$), while for unscreened it was 13.5 months ($n=75$) and for false negative cases 10.5 months ($n=37$). Fifty-three patients with meconium ileus were diagnosed at a median age of 0.8 months. Comparing screened and unscreened, significant differences were seen in survival, clinical score (Scwachman and Kulsychi), chest X-ray score, and weight score. Screened patients did better in each comparison, but differences were not apparent until the ages of 7 to 12 years, after which changes become very marked until evaluation ceased at around age 15 years, as the numbers in any group fell to below 20. The same differences were seen when patients detected by screening were compared with those missed by the test (34 meconium and three IRT tests were falsely negative). What was not different in any group was the risk of colonisation by *Pseudomonas* sp.

The very striking findings indicating an advantage to screened children in an older age group seem incontrovertable. What is not clear is whether their findings apply only to the cohort born in the early part of their study, or whether they are truly age-related. It could be that there is indeed a better clinical outcome in screened patients, which only becomes apparent in middle to later childhood. Alternatively,

it may be that the finding would only apply where clinical diagnosis was typically late, as may have occurred in the 1970s in Italy, or before current methods of treatment were available to all.

Randomised controlled trials. Two controlled trials are at present underway. In Wisconsin a randomised controlled trial began in 1985 [44]. With parental consent, babies have been randomised to a screened and a comparison group. Babies from both groups have an IRT test performed as part of the routine neonatal screening program. Those in the screened group have the IRT test completed at once, while for the comparison group the test is performed, but the calculations to obtain a result not completed, and the data is stored in a computer until the child reaches the age of 4 years. The screened babies with a positive test, i.e., an IRT level at or above the 99.8th percentile, have sweat tests at 6 weeks of age. This test protocol is thus a direct strategy, and not a two-step IRT test. More recently this group has adopted an IRT/DNA approach. A great deal of information has already come out of this study with regard to false-positive tests (Apgar scores were lower in comparison to all live born babies in Wisconsin) [75], decline of IRT with time (see above) [74] and psychosocial impact [48,85]. It appears that clinical outcome results are still inconclusive; but the screened group has not shown any consistent, statistically significant differences in nutritional and pulmonary status compared to the standard diagnosis (control) group [32,33].

This trial has been impeccably designed, but even so there could be problems in evaluating the data. One problem may be the 'halo' effect. This effect, recently referred to by Rosenstein [78], is seen when the introduction of a screening program in an area raises the general awareness of the disorder. In Wisconsin it appears that the median age at the time of clinical diagnosis decreased significantly in the non-screened group coincident with the prospective screening trial. This, of course, may be of great benefit overall, but may blunt or eradicate what might otherwise have been a net benefit to the population detected by neonatal screening. An additional difficulty may have been the very high cut-off selected for the IRT test. This was selected to minimise the number of sweat tests required by their protocol, but it will inevitably have increased the false negative rate.

In the UK a differently designed trial started also in 1985, with babies in two separate geographical areas, the West Midlands and Wales, being screened only on alternate weeks. Whilst this is not strictly randomised, no obvious biases have been introduced [79]. During a 5-year period 227,000 babies were screened, and an estimated 247,000 were not screened [19]. This trial design has the drawback that CF patients in the non-screened group may remain undiagnosed for a considerable length of time, and thus not assessed. Some could die undiagnosed. The Wales/West Midland group have recently compared clinical outcomes of 58 screened and 44 clinically diagnosed cases [19]. Nine patients who were screened but had false-negative test results were included with the clinically diagnosed cases. The infants have been assessed yearly for 1–4 years. Only 12 screened and seven unscreened have so far been included in the 4-year-old assessment. There were no significant differences between the groups for height and weight, clinical or chest radiograph scores, or in any laboratory test. However, early morbidity was significantly different,

with increased numbers of hospital admissions and total hospital days in the first year among the clinically diagnosed patients. In addition, three of the clinically diagnosed patients had died. In assessing these data, it is relevant to remember the probability of not-yet-diagnosed cases, particularly among the unscreened group. The considerable inbalance in numbers between the groups indicates a substantial shortfall in clinically diagnosed cases. As well, there is the possible effect of including screened, but missed, cases among the unscreened. Another factor in this trial which could influence results is the likely non-uniformity of treatment, much of which may not have been carried out at experienced centres.

Even with these problems of interpreting the data, the interim outcome of these two trials does not indicate a striking clinical benefit from early diagnosis. However, there is a suggestion of reductions in the early morbidity, hospitalisation and early mortality in the screened population.

Genetic counselling and reproductive decisions

A possible benefit of screening is the avoidance of a further CF child by early diagnosis of the first affected child in a family. There have been three studies of reproductive decision-making in CF related to neonatal screening. Dankert-Roelse and colleagues [24] studied 44 families, but this was done when prenatal diagnosis was not yet available. They concluded that genetic counselling may well have led to decisions against further high-risk pregnancies. The likelihood was greater in families of screened babies than in the non-screened, but the numbers were small, and this did not reach significance. Another study [7] examining attitudes to prenatal diagnosis and screening among the families of 18 screened and 11 clinically diagnosed babies found that only 15 of the 20 couples (52%) would undertake antenatal diagnosis, and not all of these would consider termination. The question of a benefit from early genetic counselling remains unresolved. It seems certain that a proportion of families would benefit, but that proportion might be very small indeed.

Pre-diagnosis anxiety

From several studies of parental attitudes to neonatal screening [7,48,93] it is evident that parents overwhelmingly favoured screening, and there was no difference in attitude between those who had and had not experienced screening for their child. One major reason for favouring screening given by those whose child had not been screened was the avoidance of stress and anxiety associated with undiagnosed illness in the child during the period before diagnosis. This is graphically described in a study of the effects of screening on maternal behaviour [12]. Delay in CF diagnosis was well documented [77,97] some 10 years ago, and is not likely to be different now in most regions where screening does not occur.

Contributions to scientific research as a benefit from screening

With near-universal diagnosis by age 2 months or so in regions with CF screening, the early natural history of CF has become more completely described. It became evident that a high percentage of CF babies detected by screening already had symptoms at diagnosis – in our study 92 of 114 babies had symptoms at a mean

age of 5 to 7 weeks (predominantly gastrointestinal symptoms) [93]. Many nutritional deficits have been documented in screening-diagnosed neonates, including low levels of serum albumin, total protein and rapid-turnover protein, cholesterol, linoleic acids and fat-soluble vitamins [1,57,61,69,82]. Birth weight in CF newborns is significantly lower than in controls [57] as is total body potassium, an index of body cell mass [80]. Interestingly, at diagnosis almost 50% of screened babies have pancreatic sufficiency, although there is a rapid fall-off to insufficiency during the first year [16,36,90]. Studies of the early respiratory course have recently been reviewed [3]. Abman and colleagues [2] followed longitudinally 42 infants diagnosed by screening and found that the age of first positive *Pseudomonas* culture was 21 months, and that 19% had positive serial cultures by a mean of 27 months. These authors have also suggested that hypoalbuminaemia at diagnosis is associated with a severe respiratory course [1], although others have not substantiated this [70].

The negative effects of screening

The Cystic Fibrosis Foundation's position paper of 1983 (USA) considered the potential harmful effects of screening on two groups: infants with false-positive tests, and asymptomatic infants with CF [4]. There might also be the problem of 'labelling' CF patients with very mild disease not destined to cause symptoms for many years [66]. The task force also canvassed several other potential drawbacks, as did Farrell [30]. These included an increased demand for sweat testing, an obligatory false-negative rate of 10% (because of the supposed inability of the IRT test to detect babies with pancreatic sufficiency), possible litigation related to false-negative test results, and delay in the clinical diagnosis of 'missed' cases because of false reassurance. Some of these suppositions have been shown to be unfounded, or not important. There has not been any evidence so far of the labelling of very mild cases, and these would have to constitute a small number, as the majority of diagnosed patients do in fact have early symptoms [19,72,93]. Nor has there been any documentation of an increased demand for sweat tests. In our state, sweat-test numbers actually decreased [93], but this could have been due to physicians gaining reassurance from the negative test finding, and thus may not have been an entirely desirable outcome. There is now good evidence that babies with pancreatic sufficiency are detected by the IRT test [36,90], but a widespread adoption of an IRT/DNA approach would indeed see an obligatory false negative rate, which would vary according to the region and the protocol adopted but which is likely to be around 5–6%.

The important potential negative effects of screening undoubtedly related to the psychosocial aspects. Boland and Thompson [12] studied maternal behaviour and the mother-child relationship in mothers who had one CF child (without meconium ileus) who had been diagnosed either by clinical symptoms or by neonatal screening. All 71 mothers who fulfilled the criteria were approached. Five declined to be interviewed, and eight others were excluded because of language or geographical reasons, or they were unable to be contacted. There were finally 29 in each group. The results indicated that newborn screening had not increased maternal over-protectiveness or maternal denial. The authors concluded that neonatal screening, "far from having detrimental effects on the mother-child relationship, has in some

cases been beneficial". They related this finding in part to the self-blame and self-doubt which had arisen during the period between onset of symptoms and diagnosis in the non-screened children. Helton and colleagues [48] studied a variety of parental attitudes to newborn screening, and found that parents of screened children perceived that the diagnosis of CF had caused them to feel emotionally closer to the child, and some felt that the diagnosis had actually enhanced family relationships.

There has rightly been great concern about the number of false-positive tests generated by CF screening. One hundred and four parents in the Wisconsin program were asked about their reactions to the disclosure of a positive screening result and subsequent negative sweat test [85]. Almost all parents showed anxiety, and many also felt depression, shock, disbelief, confusion and anger. An important finding was that there was a significant difference in understanding of negative sweat test results when these were communicated in person rather than by telephone. Parents also had many misconceptions when information about newborn screening was given by brochure. It must be remembered that because of the experimental protocol, all infants with positive tests needed sweat tests. It could be that the anxiety provoked by simply repeating the blood test is less, but this has not been investigated. With the availability of an IRT/DNA approach, the actual numbers of 'false-positive' tests will fall dramatically, and these will be confined to babies who are CF carriers, introducing a whole new set of problems, albeit more easily focused and much smaller in number (see above).

Costs

Using a two-stage IRT assay, the cost of screening in Colorado was estimated as US $1.71 per infant in direct cost of testing. The incremental cost, including indirect costs, was estimated as US $2.52, and the cost of identifying a single infant through testing was $8.877 [42]. There have been few other estimates of cost. Dauphinais [26a], with a higher incidence of CF, estimated a cost of approx. US $6200 to identify a case, but this is reduced to US $5000 if costs associated with billing are excluded. In our program, the cost of identifying each otherwise unsuspected case in 1986 was A $2300 (approx. US $1900) [93]. Our cost per test within the laboratory is currently A $1.25 (US $0.94) and the cost per identified patient A $3250. This does not include the costs of sweat testing, patient retrieval for retesting, and so forth. The costs of the IRT/DNA strategy have not been documented fully, and are likely to be changing with rapidly improving technology. Ranieri and colleagues [68] reported that the DNA test materials cost 3-times as much as those for the IRT test, but that this was offset by savings in administration and clinical costs of recalling fewer infants (only 0.11% of infants required recall vs. 0.8% with the two-step IRT strategy).

Most recently, Farrell and Mischler have published costs of screening in Wisconsin [33]. With their IRT protocol, which included sweat tests on all babies with positive tests, the cost of diagnosing one case (estimated incidence 1:4000) was US $7244, while the predicted costs of using an IRT/DNA strategy was similar at US $7042. They also examined the medical costs for patients diagnosed by screening, and compared medical costs of two patients aged 21 months, one diagnosed by screening,

and one 'missed' by the test. Their estimates serve to underline the difficulties of analysing costs and benefits. We identified a likely cost-saving in the short term, considering on the one hand costs of testing and treatment for 6 months, and on the other, averted costs of hospitalisation [93]. A similar finding has been reported from Connecticut [35].

Conclusions

Almost 30 years after neonatal screening for cystic fibrosis was first proposed it seems that, although we now have a suitable screening procedure, we are still not sure whether we should use it. This review has explored the current position.

The test of immunoreactive trypsin (IRT) in neonatal dried blood samples has proved adequate for case-finding in many hands, even when used as a two-stage IRT test. When combined with a DNA test, however, there is no doubt of its good performance, with a sensitivity of about 95% and a very high specificity. This is an evolving position and different regions will find differing protocols to maximise the case-finding efficiency of this test.

The benefits of early diagnosis have not been fully established, however. There is a lot of evidence of short-term benefit, with reduced morbidity in the first year or two of life, and there is also the opportunity to correct nutritional and biochemical deficiencies that have been documented in babies even in the early weeks. Benefit in the longer term has been difficult to demonstrate, and it is possible that this may only occur when conventional diagnosis is particularly likely to be delayed. Further results from on-going randomised controlled trials are awaited.

The likelihood of any significant drawback to early diagnosis by screening seems to be receding. There is now evidence that there is no psychological harm amongst families of neonates so identified, while, conversely, the adverse psychological impact of delayed diagnosis has been well documented. The problem of false-positive testing and the attendant anxiety can be largely overcome with the adoption of an IRT/DNA strategy. Screening seems to be widely and strongly approved of by parents.

The costs of screening have not been studied in detail, and although it has been suggested that screening is cost beneficial in the short term, it is not yet clear that this is so. All-in-all, the scales appear to be tipping in favour of neonatal screening for CF, but the final verdict is not yet in.

References

1 Abman SH, Reardon MC, Accurso, FJ, Hammond KB, Sokol RJ. Hypoalbuminaemia at diagnosis as a marker for severe respiratory course in infants with cystic fibrosis identified by newborn screening. J Pediatr 1985;107:933–935.

2 Abman SH, Ogle JW, Harbeck RJ, Butler-Simon N, Hammond KB, Accurso FJ. Early bacteriologic, immunologic, and clinical course of young infants with cystic fibrosis identified by neonatal screening. J Pediatr 1991;119:211–217.

3 Accurso F, Sokol RJ, Hammond KB, Abman SH. Early respiratory course in infants with cystic fibrosis. Pediatr Pulmonol 1991;Suppl 7:42–45.
4 Ad Hoc Committee Task Force on Neonatal Screening, Cystic Fibrosis Foundation. Neonatal screening for cystic fibrosis: Position paper. Pediatr 1983;72:41–45.
5 Adriaenssens K, Janssens H, van Soom H. Two-tier screen for cystic fibrosis. Lancet 1981;ii:823.
6 Adriaenssens K, Van Riel L. Serum pancreatic lipase as a screening test for cystic fibrosis. Arch Dis Childh 1982;57:553–555.
7 Al-Jader LN, Goodchild MC, Ryley HC, Harper PS. Attitudes of parents of cystic fibrosis children towards neonatal screening and antenatal diagnosis. Clin Genet 1990;38:460–465.
8 Anderson DH. Cystic fibrosis of the panceas and its relation to coeliac disease. Am J Dis Child 1938;56:344.
9 Anderson MP, Rich DP, Gregory RJ. et al. Generation of cAMP-activated chloride currents by expression of CFTR. Science 1991;251:679–682.
10 Applegarth DA, Davidson AGF, Kirby LT. Dried blood spot screening for cystic fibrosis. Lancet 1979;ii:1936–1937.
11 Barbier Y, Gilly R, Chazalette JP. Neonatal screening for cystic fibrosis. Monogr Paediatr 1981;14:177–181.
12 Boland C, Thompson NL. Effects of newborn screening of cystic fibrosis on reported maternal behaviour. Arch Dis Child. 1990;65:1240–1244.
13 Bowling F, Cleghorn G, Chester A, et al. Neonatal screening for cystic fibrosis. Arch Dis Child 1988;63:196–198.
14 Bowling FG, Brown ARD. Newborn screening for cystic fibrosis using an enzyme linked immunosorbent assay (ELISA) technique. Clin Chim Acta 1988;171:257–261.
15 Bowling FG, Walson AR, Rylatt DB, Elliott JE, Bunch RS, Bunderson PG. A monoclonal antibody-based trypsinogen EIA for screening of neonates for cystic fibrosis. Lancet 1987;i:826–827.
16 Bronstein FJ, Sokol RJ, Abman SH, et al. Pancreatic insufficiency, growth, and nutrition in infants identified by newborn screening as having cystic fibrosis. J. Pediatr. 1992;120:533–540.
17 Bourguignon JP, Deby-Dupont G, Reuter A, Santerre J, Gerard A, Franchimont P. Variations in dried blood spot immunoreactive trypsin in relation to gestational age and during the first week of life. Eur J Pediatr 1986;144:547–549.
18 Cassio A, Bernardi F, Piazzi S, et al. Neonatal screening for cystic fibrosis by dried blood spot trypsin assay: results in 47127 newborn infants from a homogenous population. Acta Paediatr Scand 1984;73:554–558.
19 Chatfield S, Owen G, Ryley HC, et al. Neonatal screening for cystic fibrosis in Wales and the West Midlands: clinical assessment after five years of screening. Arch Dis Child 1991;66:29–33.
20 Crossley JR, Berryman CC, Elliott RB. Cystic fibrosis screening in the newborn. Lancet 1977;ii:1093–1095.
21 Crossley JR, Elliott RB, Smith PA. Dried blood spot screening for cystic fibrosis in the newborn. Lancet 1979;i:472–474.
22 Crossley JR, Smith PA, Edgar BW, Gluckman PD, Elliott RB. Neonatal screening for cystic fibrosis, using immunoreactive trypsin assay in dried blood spots. Clin Chim Acta 1981;113:111–121.
23 The Cystic Fibrosis Genetic Analysis Consortium. Worldwide survey of the F508 mutation. Am J Hum. Genet 1990;47:354–359.
24 Dankert-Roelse JE, te Meerman GJ, Knol K, ten Kate LP. Effect of screening for cystic fibrosis on genetic counselling. Clin Genet 1987;32:271–275.
25 Dankert-Roelse JE, te Meerman GJ, Martijn A, ten Kate LP, Knol K. Survival and clinical outcome in patients with cystic fibrosis, with or without neonatal screening. J Pediatr 1989;114:362–367.
26 Dauphinais RM, Ganeshananthan M, Jezyk P. Cystic fibrosis: early detection and clinical course. In: Carter TB and Willey AM, eds. Genetic disease: screening and management. New York: Alan R Liss, 1986;65–79.
26a Dauphinais RM. A cost analysis of blood-spot screening newborns for cystic fibrosis. J Clin Immunoassay 1992;15:121–125.
27 di Sant'Agnese PA, Darling RC, Perera GA, et al. Abnormal electrolyte composition of sweat in

cystic fibrosis of the pancreas. Clinical significance and relationship to the disease. Pediatrics 1953;12:549.
28 Dodge JA, Ryley HC. Screening for cystic fibrosis. Arch Dis Child 1982;57:774–780.
29 Dominici R, Monaco F, Morano S, Antonozzi I. Immunoreactive trypsin on dried blood spots as a possible neonatal test for cystic fibrosis. La Ricerca Clin Lab 1980;10:551–559.
30 Farrell PM. Early diagnosis of cystic fibrosis: to screen or not to screen – an important question. Pediatrics 1984;73:115–117.
31 Farrell PM, Mischler EH, Fost NC, et al. Current issues in neonatal screening for cystic fibrosis. Pediatr Pulmonol 1991;Suppl 7:11–18.
32 Farrell PM, Mischler EH, Wilfond BS, et al. Neonatal screening for cystic fibrosis: clinical features, evaluation of testing procedures, and randomised investigation of benefits and risks. In: Wilcken B, Webster D, eds. Neonatal screening in the nineties. Sydney: Kelvin Press, 1992;156–165.
33 Farrell PM, Mischler EH. Neonatal screening for cystic fibrosis. Adv Pediatr 1992;39:35–70.
34 Forrest DC, Wilcken B, Turner G. Screening for cystic fibrosis by a stool trypsin method. Arch Dis Child 1981;56:151–153.
35 Ganashananthan M, Dauphinais R, Jezyk P. Diagnosis of cystic fibrosis: results of screening vs no screening in a three year programme. CF Club Abstracts 1985;26:93.
36 Gaskin K, Waters D, Dorney S, Gruca M, O'Halloran M, Wilcken B. Assessment of pancreatic function in screened infants with cystic fibrosis. Pediatr Pulmonol 1991;Suppl 7:69–71.
37 Gibson LE. Screening of newborns for cystic fibrosis. Am J Dis Child 1980;134:925–926.
38 Gibson LE, Cooke RE. A test for concentration of electrolytes in sweat in cystic fibrosis of the pancreas utilizing pilocarpine by iontophoresis. Pediatrics 1959;23:545–9.
39 Giglio L, Faraguna D, Pederzini F. Neonatal screening for cystic fibrosis: IRT false positive cases. In: Mucoviscidose; Depistage Neonatal et Prise en Charge Precoce Caen; CHRU de Caen 1988;235.
40 Gitzelman R. Why we should not screen our newborns for cystic fibrosis. Helv Pediatr Acta 1981;36:493–494.
41 Green MN, Clark JT, Shwachman H. Studies in cystic fibrosis of the pancreas: protein pattern in meconium ileus. Pediatr 1958;21:635–641.
42 Hammond KB, Abman SH, Sokol RJ, Accurso FJ. Efficacy of statewide screening for cystic fibrosis by assay of trypsinogen concentrations. N Engl J Med 1991;325:769–774.
43 Hammond KB, Ask CG, Watts DC. Immunoreactive pancreatic lipase test for neonatal detection of cystic fibrosis. Lancet 1984;i:42.
44 Hassemer DJ. Neonatal screening for cystic fibrosis in Wisconsin. In: Therrell BL, ed. Advances in neonatal screening. Amsterdam: Excerpta Medica, 1987;393–394.
44a Heeley AF, Bangert SK. The neonatal detection of cystic fibrosis by measurement of immunoreactive trypsin in blood. Ann Clin Biochem 1992;29:361–376.
44b Heeley AF, Fagan DG. Trisomy 18, cystic fibrosis, and blood immunoreactive trypsin. Lancet 1984;ii:169–170.
45 Heeley AF, Heeley ME, King DN, Kuzemko JA, Walsh MP. Screening for cystic fibrosis by dried blood spot trypsin assay. Arch Dis Child 1982;57:18–21.
46 Heeley AF, Watson D. Cystic fibrosis – its biochemical detection. Clin Chem 1983;29:2011–2018.
47 Heeley ME, Keenan CA, Heeley AF. An update of cystic fibrosis screening in East Anglia. Pediatr Pulmonol 1991;Suppl 7:90.
48 Helton J, Harmon RJ, Robinson N, Accurso F. Parental attitudes towards newborn screening for cystic fibrosis. Pediatr Pulmonol 1991;Suppl 7:23–28.
49 Holsclaw DS, Keith HH, Palmer J. Meconium screening for cystic fibrosis. Pediatr Ann 1978;7:1, 15–20.
50 Holtzman NA. Routine screening of newborns for cystic fibrosis: not yet. Pediatrics 1984;73:98–99.
51 Hutchinson JH. Some new diseases in paediatrics. Br Med J 1955;ii:339–349.
52 Kartner N, Hanrahan JW, Jensen TJ. Expression of the cystic fibrosis gene in non-epithelial invertebrate cells produces a regulated anion conductance cel 1991;64:681–691.
53 Kerem BS, Rommens JM, Buchanan JA, et al. Identification of the cystic fibrosis gene: genetic analysis. Science 1989;245:1073–1080.

54 King DN, Heeley AF, Walsh MP, Kuzemko JA. Sensitive trypsin assay for dried blood specimens as a screening procedure for early detection of cystic fibrosis. Lancet 1979;ii:1217–1219.
55 Kirby LT. Applegarth DA, Davidson AGF, Wong LTK, Hardwick DF. Use of a dried blood spot in immunoreactive-trypsin assay for detection of cystic fibrosis in infants. Clin Chem 1981;27:678–680.
56 Laroche D, Travert G. The application of PCR amplification and the polymorphic marker KM.19 to dried blood spots: comparison with deletion 508 for the confirmation of the neonatal screening test for cystic fibrosis. Pediatr Pulmonol 1991;Suppl 7:19–22.
57 Marcus MS, Sondel SA, Farrell PM, et al. Nutritional status of infants with cystic fibrosis associated with early diagnosis and intervention. Am J Clin Nutr 1991;54:578–585.
58 Reference deleted.
59 Mastella G, Barlocco EG, Antonacci B, et al. Is neonatal screening for cystic fibrosis advantageous? The answer of a wide 15 years follow-up study. In: Mucoviscidose: Depistage Neonatal et Prise en Charge Precoce Caen; CHRU de Caen, 1988;127–143.
60 McHill G. Experience with the nation-wide early detection of cystic fibrosis In: Naruse H, Irie M, eds. Neonatal screening. Amsterdam; Exerpta Medica, 1983;350–351.
61 Mischler EH, Marcus MS, Sondel SA, et al. Nutritional assessment of infants with cystic fibrosis diagnosed through screening. Pediatr Pulmonol 1991;Suppl 7:56–63.
62 Orenstein DM, Boat TF, Stern RC, et al. The effect of early diagnosis and treatment in cystic fibrosis. Am J Dis Child 1977;131:973–975.
63 Pederzini F, Armani P, Barbato A, et al. Newborn screening for cystic fibrosis: two methods compared in 229,626 newborns. Ital J Pediatr 1983;9:445–447.
64 Pederzini F, Faraguna D, Giglio L, Pedrotti D, Perobelli L, Mastella G. Development of a screening system for cystic fibrosis: Meconium or blood spot trypsin assay or both? Acta Paediatr Scand 1990;79:935–942.
65 Petterson K, Soderholm J, Suonpaa M. Organisation of an immunofluorometric assay based on the DELFIA technique for the measurement of immunoreactive trypsin in dry blood spots. Pediatr Pulmonol 1991;57:90.
66 Phelan PD. Screening for cystic fibrosis In: Meadow, R, ed. Recent advances in paediatrics. Edinburgh: Churchill Livingstone, 1984;103–120.
67 Prosser R, Owens H, Bull F, et al. Screening for cystic fibrosis by examination of meconium. Arch Dis Child 1974:49;597–601.
68 Ranieri E, Ryall RG, Morris CP et al. Neonatal screening strategy for cystic fibrosis using immunoreactive trypsinogen and direct gene analysis. Br Med J 1991;302:1237–1240.
69 Reardon MD, Hammond KB, Accurso FJ, et al. Nutritional deficits exist before 2 months of age in some infants with cystic fibrosis identified by newborn screening. J Pediatr 1984;105:271–274.
70 Reisman J, Petrou C, Corey M, Stringer D, Durie P, Levison H. Hypoalbuminaemia at initial examination in patients with cystic fibrosis. J Pediatr 1989;115:755–758.
71 Riordan JR, Rommens JM, Kerem BS, et al. Identification of the cystic fibrosis gene: cloning and characterisation of complementary DNA. Science 1989;245:1066–1073.
72 Roberts G, Stanfield M, Black A, Redmond A Screening for cystic fibrosis: a four-year experience. Arch Dis Child 1988;63:1438–1443.
73 Robertson EF, Anderson D, Harper G, et al. DNA strategies in cystic fibrosis screening. In: Wilcken B, Webster D, eds: Neonatal Screening in the nineties. Sydney: Kelvin Press, 1992;166–170.
74 Rock MJ, Mischler EH, Farrell PM, et al. Newborn screening for cystic fibrosis is complicated by age-related decline in immunoreactive trypsinogen levels. Pediatrics 1990;85:1001–1007.
75 Rock MJ, Mischler EH, Farrel PM, Bruns WT, Hassemer DJ, Laessig RH. Immunoreactive trypsinogen screening for cystic fibrosis: characterisation of infants with a false-positive screening test. Pediatr Pulmonol 1989;6:42–48.
76 Rommens JM, Iannuzzi MC, Kerem BS, et al. Identification of the cystic fibrosis gene: chromosome walking and jumping. Science 1989;245:1059–1065.
77 Rosenstein BJ. Diagnostic delay in CF. J Am Med Assoc 1982;247:3313.
78 Rosenstein B. IVth International Conference on Newborn Screening for Cystic Fibrosis. Pediatr Pulmonol 1991:Suppl 7:84–86.

79 Ryley HC, Deam SM, Williams J. Neonatal screening for cystic fibrosis in Wales and the West Midlands: I. Evaluation of the immunoreactive trypsin test. J Clin Pathol 1988;41:726–729.
80 Shepherd RW, Holt TL, Greer R, Cleghorn GJ, Thomas BJ. Total body potassium in cystic fibrosis. J Pediatr Gastroenterol Nutr 1989;9:200–205.
81 Shwachman H, Redmond A, Khaw K-T. Studies in cystic fibrosis. Report of 130 patients diagnosed under 3 months of age over a 20 year period. Pediatrics 1970;46:335–343.
82 Sokol RJ, Reardon MC, Accurso FJ, et al. Fat-soluble vitamin status during the first year of life in infants with cystic fibrosis identified by screening of newborns. Am J Clin Nutr 1989;50:1064–1071.
83 Starkey PM, Barrett AJ. Proteinases in mammalian cells and tissues In: Barrett AJ, ed. Research monographs in cell and tissue physiology. Amsterdam: North Holland Publishing Co, 1977;663.
84 Stephan U, Busch EW, Kollberg H, Hellsing K. Cystic fibrosis detection by means of a test-strip. Pediatrics 1975;55:35–38.
85 Tluczek A, Mischler EH, Farrell PM, et al. Parents knowledge of neonatal screening and response to false-positive cystic fibrosis testing. Dev Behav Pediatr 1992;13:181–186.
86 Travert G. Cystic fibrosis screening in the newborn. Med Lab 1986;14:17–22.
87 Travert G. Analyse de l'experience mondiale de depistage neonatal de la mucoviscidose par dosage de la trypsine immunoreactive sanguine. In Mucoviscidose; Depistage Neonatal et Prise en Charge Precoce Caen; CHRU de Caen 1988;1–23.
88 Travert G, Duhamel JF. Depistage neonatal systematique de la mucoviscidose par dosage de la trypsine immunoreactive sanguine: bilan de 80,000 tests. Arc Fr Pediatr 1983;40:95–98.
89 Urwin R, Sim LE, Robinson J, Wiley V, Wilcken B. Screening for cystic fibrosis using dried blood samples: comparison of a time-resolved fluoroimmunoassay with an enzyme immunoassay for immunoreactive trypsin. In: Wilcken B, Webster D eds. Neonatal screening in the nineties. Sydney: Kelvin Press, 1992;355.
90 Waters DL, Dorney SFA, Gaskin KJ, Gruca MA, O'Halloran M, Wilcken B. Pancreatic function in infants identified as having cystic fibrosis in a neonatal screening program. N Engl J Med 1990;322:303–308.
91 Wesley AW, Smith PA, Elliott RB. Experience with neonatal screening for cystic fibrosis in New Zealand using measurement of immunoreactive trypsinogen. Aust Paediatr J 1989;25:151–155.
92 Wilcken B. Screening for cystic fibrosis in New South Wales: Evaluation of the results of screening 400,000 babies. In: Therrell BL, ed. Advances in neonatal screening. Amsterdam: Excerpta Medica, 1987;385–390.
93 Wilcken B. An evaluation of screening for cystic fibrosis. In: Buchwald M, Riordan J, eds. Genetics and epithelial cell dysfunction in cystic fibrosis. New York: Alan R Liss, 1987;201–215.
93a Wilcken B, Brown ARD. An analysis of false negative screening tests for cystic fibrosis In: Schmidt BJ, Diament AJ, Loghin-Grosso NS. eds. Current trends in infant screening. Amsterdam: Excerpta Medica, 1989;19–25.
94 Wilcken B, Chalmers G. Reduced morbidity in cystic fibrosis. Lancet 1986:i:439.
95 Wilcken B, Brown ARD, Urwin R, Brown DA. Cystic fibrosis screening by dried blood spot trypsin assay: results in 75000 newborn infants. J Pediatr 1983;102:383–387.
96 Wilcken, B, Chalmers G. Reduced morbidity in patients with cystic fibrosis detected by neonatal screening. Lancet 1985;ii:1319–1321.
97 Wilcken B, Towns SJ, Mellis CM. Diagnostic delay in cystic fibrosis: Lessons from newborn screening. Arch Dis Child 1983;58:863–866.
98 Wiser WC, Beier FR. Albumin in the meconium of infants with cystic fibrosis. A preliminary report Pediatrics 1964;33:115.
99 Xu YY, Petterson K, Blomberg K, Hemmila I, Lovgren T. Triple-label time-resolved fluoroimmunoassay of thyrotropin, 17-hydroxyprogesterone, and immunoreactive trypsin from dried blood spots. In: Wilcken B, Webster D, eds. Neonatal screening in the nineties. Sydney: Kelvin Press, 1992;223–225.

Neonatal screening for cystic fibrosis: A comparison of two strategies for case detection in 1.2 million babies

Bridget Wilcken, FRACP, Veronica Wiley, PhD, Gina Sherry, BSc, and Ursula Bayliss, BHA

From the New South Wales Newborn Screening Program, Royal Alexandra Hospital for Children, Sydney, Australia

Objectives: **To review the overall performance of a neonatal screening program for cystic fibrosis (CF) from 1981 to 1994, and to compare two strategies of case detection.**

Program design: **Initially, immunoreactive trypsin (IRT) was measured in dried blood spots, and because of the low sensitivity of this test at days 3 to 5, a second sample was needed from babies with positive test results. Since 1993 a positive IRT assay result has been followed by direct gene analysis for the common CF mutation, ΔF508, with the use of the same sample. Cases with false-negative results were actively sought throughout the period.**

Results: **With IRT alone, 1,015,000 babies were tested. Of 389 babies with CF, 30 had a clinical diagnosis of CF made after a negative screening test result or an administrative error. Early diagnosis was achieved in 92%. With the IRT/DNA protocol, 59 of 62 infants had a positive screening test result (44 were homozygous for ΔF508) among 189,000 babies tested. Three babies with CF had no copy of this mutation, but two were identified early because of meconium ileus. The false-positive rate was much greater for IRT alone than for the IRT/DNA test (0.69% vs 0.054%). All false-positive cases in the IRT/DNA protocol were, of necessity, CF carriers.**

Conclusion: **The percentage of babies with CF who had an early diagnosis was similar with the two protocols, but we concluded that the advantages of the IRT/DNA test for screening, particularly in the avoidance of the need for second IRT samples, outweighed the drawback of unwanted carrier detection. (J PEDIATR 1995;127:965-70)**

Effective neonatal screening for cystic fibrosis by determination of dried blood-spot immunoreactive trypsin levels was first described in 1979.[1] Although measurement of IRT proved to be a sensitive and specific test for CF in babies aged 3 to 5 weeks, the specificity was much lower during the first week of life, when neonatal screening test samples are collected. Therefore the original protocols involved obtaining a second sample from babies whose first IRT levels were above a predetermined cutoff point. Any baby whose IRT level on the second test remained elevated then had a sweat test. The sensitivity of such a two-stage protocol proved satisfactory for the diagnosis of cystic fibrosis, and was adopted by a number of state or regional screening programs.[2-4] The major problem with this otherwise satisfactory method of case detection was that for every baby in whom CF was de-

CF	Cystic fibrosis
IRT	Immunoreactive trypsin

Submitted for publication May 5, 1995; accepted July 7, 1995.

Reprint requests: Bridget Wilcken, FRACP, Royal Alexandra Hospital for Children, Post Office Box 3515, Parramatta (Sydney) 2124, Australia.

0022-3476/95/$5.00 + 0 **9/23/67543**

tected, some 7 to 20 babies needed to have a second test, and this caused anxiety to many families.[5]

The cloning of the CF gene and the discovery of a common mutation, present in well over 90% of patients with CF in some ethnic groups, made possible a different approach to neonatal screening, in which screeners used a combination of IRT determination and direct gene analysis with the same dried blood sample. This approach precludes the necessity of obtaining second samples from babies, and has been used with success by a number of groups.[6-9]

We and others have demonstrated benefits of early diagnosis by neonatal screening in the short term,[10-13] and benefits in the longer term have been shown in at least one program.[14] This article describes the overall results of screening more than 1.2 million babies for a period of 13½ years in New South Wales, Australia, and compares the case-finding efficiency of screening by IRT alone with that of a protocol using the IRT/DNA approach.

METHODS

Population tested. All babies born in New South Wales and the Australian Capital Territory are offered a screening test for several genetic metabolic disorders. The compliance has been more than 99%. Since mid-July 1981 in New South Wales, and January 1986 in the Australian Capital Territory, a test for CF was performed as part of this screening program.

Samples. Blood samples were collected by heel prick onto No. 903 paper (Schleicher & Schuell Inc., Keene, N.H.) when babies were 48 hours of age or more, and usually on days 4 to 6.

Testing protocols

First protocol (IRT alone). From 1981 to 1992, a total of 1,015,000 babies were tested with a dried blood-spot assay for IRT, as described previously.[3] Samples were tested in daily batches. Those with IRT levels in the top 2% of values were retested in duplicate, and if the mean value was within the top 0.6% (first 2 years) or top 0.7% of the week's values, this was considered a positive result, and a second sample was sought by one, or if necessary two, written requests. Second samples were usually taken when the babies were 3 to 6 weeks of age. If the second test result remained elevated, a sweat test was performed.

Second protocol (IRT/DNA). From January 1993 to December 1994, a total of 189,000 babies were tested with an IRT/DNA protocol. The IRT assay was performed on all samples, as for protocol 1, and those samples with final IRT values in the top 1% were further tested by direct gene analysis for the common CF mutation, ΔF508. Direct gene analysis was also carried out if babies had meconium ileus or a sibling with CF. Babies homozygous for the mutation were referred directly to a CF clinic for further evaluation and management. Babies with one copy of the mutation could either have CF (carrying in addition another, rarer mutation) or be a healthy carrier. These babies were referred for sweat testing, preferably at one of four major hospitals. When the result of the sweat test was negative, the parents were offered genetic counseling and testing for CF mutations.

Information for professionals and parents. When protocol 2 was introduced, information was sent to all maternity units, pediatricians, and clinical genetics units in the state. The same information accompanied each notification of a positive result. A pamphlet describing the screening tests is given to mothers during antenatal visits or immediately before the sample is taken. More comprehensive written information is available for parents of babies with a positive CF DNA result.

Immunoreactive trypsin assay. Immunoreactive trypsin was measured by several different methods during the 13 years. Initially, a radioimmunoassay was used with a polyclonal antibody prepared in the Auckland University Department of Paediatrics.[1,3] From October 1987 a two-site sandwich immunoassay with two mouse monoclonal antibodies was employed (Agen Biomedical Ltd., Acacia Ridge, Queensland, Australia),[15] and after July 1992 the assay was a fluoroimmunoassay with different monoclonal antibodies (Wallac, Oy, Turku, Finland).[16]

DNA testing. DNA was extracted from a 3 mm circle punched from the dried blood sample, amplified by a standard polymerase chain reaction technique, and the product was separated by means of polyacrylamide gel electrophoresis.[17] Any sample homozygous or heterozygous for the ΔF508 allele was reanalyzed, with the use of a further punched dried blood circle, to confirm the finding and ensure that no contamination had occurred.

Ascertainment of patients with false-negative screening results. Cases possibly missed by the screening program were systematically sought by contact with all the CF clinics in New South Wales and the Australian Capital Territory, as well as with all the major laboratories conducting sweat tests and CF clinics in the other states. Virtually 100% of patients in New South Wales with diagnosed CF attend a CF clinic for all or part of their care.

Prenatal diagnosis. Prenatal diagnosis for CF has been available for pregnancies at risk in Australia since 1985, with an assay of microvillar enzymes (Dr. W. F. Carey, Adelaide Children's Hospital), and since 1989 with DNA techniques (Prof. R. J. Trent, Royal Prince Alfred Hospital, Sydney). We obtained information about all cases in which the fetus was predicted to be affected on prenatal testing, including the pregnancy outcome (born affected, or termination of pregnancy), and the predicted date of delivery of a term baby in the case of a termination, so as to determine to which cohort that fetus would have belonged. The three babies predicted to be affected and born affected had positive screening results and are included in the high-risk category.

RESULTS

First protocol (IRT alone). We tested 1,015,000 babies using the first protocol, of whom 7362 (0.7%) had a positive result. At least 42 of these babies are known to have died in the neonatal period, including at least 12 with trisomy 13 or 18. Among the remaining 7320 babies, second samples were received from all except 149, a compliance rate of 98%. In an unknown number of the 149, the baby's local physician had arranged a sweat test in lieu of sending the second sample, and we were notified of 30 such infants, two of whom had CF. In total, 389 babies from the cohort have been identified with CF (Table I); 117 (30%) of these were at high risk of having CF and should have been (and usually were) easily identified because of meconium ileus (80; 21%) or having a sibling known to have CF (47; 12%). Some had both risk factors. In the remaining 272 babies, CF was unsuspected, and all but 30 of these were identified by the screening protocol. During the period when this protocol operated, prenatal diagnosis was undertaken in a number of at-risk pregnancies, and 22 fetuses were predicted to be affected. Nineteen terminations occurred of fetuses who would otherwise have been part of this cohort.

Missed cases. Of the 30 "missed" patients, 26 had a false-negative screening result. In 11 the result of the first test was negative, and no further test was performed. However, 5 of the 11 patients had IRT levels that, although not above the cutoff level then operating, would have fallen into the top 1% of values; thus the patients would have received DNA testing under the current protocol, but 6 would not have received such further testing. In these 11 patients, CF was diagnosed clinically at ages 4 to 72 months. Thirteen patients had a falsely negative result on the second test, with clinical diagnosis at 1 to 96 months, and in two patients there was an error in the laboratory. A further four patients had problems that delayed or prevented follow-up (no name on card, no second test returned).

One feature associated with a false-negative result on the second test was late performance of the test. From 1981 to 1991, a total of 191 babies with CF had second tests performed by 6 weeks of age, and in 6 of these the test result was negative. A further 48 babies with CF had their second blood sample taken after 6 weeks, and 8 of these had a negative result. This difference is significant (chi-square value = 12.7; $p < 0.0005$).

False-positive results. Six hundred thirty-five babies had a positive second IRT test result (8.9% of the total who provided a second sample), and 300 of these were found to have CF. Some babies at high risk of having CF did not have a second sample taken but proceeded directly to a sweat test. Thus the total number of babies referred for a sweat test who did not have CF was 335, 0.033% of the screened population. An unknown but apparently small number of babies was referred for sweat testing by their physicians after the notification of the positive first test result (see above).

Table I. Results of neonatal screening for CF, for both protocols

Protocol	IRT/IRT	IRT/DNA
Babies screened	1,015,000	189.000
Total CF cases known	389	62
Total high risk	117	9
Meconium ileus	80*	8
Sibling of known patients	47*	1
Unexpected cases	272	53
Detected by test	242	52
Not detected	30	1†
Babies receiving early diagnosis (%)	92	98
False-positive results		
First sample	6,984 (0.69%)	102‡ (0.054%)
Second sample	335‡ (0.033%)	NS

NS, Not sought.
*Ten babies were in both high-risk categories.
†Three babies had a neagative DNA test result: two of these were at high risk, with meconium ileus.
‡Babies who required sweat tests but did not have CF.

Second protocol (IRT/DNA). From January 1993 to December 1994 we screened 189,000 babies. DNA testing was carried out on 1968 samples (1.04%). Of these, 1807 had no copy of the ΔF508 mutation, and thus had a negative screening result. Forty-four infants were homozygous for the mutation and were referred directly to a CF clinic. The remaining 117 samples had one ΔF508 and one normal allele, and the babies so identified were recalled for a sweat test. Fifteen babies had a positive result on the sweat test. Three babies later identified as having CF had an elevated IRT level, falling into the top 1% of values, but no copy of the ΔF508 mutation. Two of these had meconium ileus, and CF was diagnosed in the first week of life; the third case was recognized when the patient was 3 months of age. No other false-negative results are known from this cohort.

Carrier detection. We identified 102 babies who were heterozygous for the ΔF508 mutation but had normal results on sweat tests and who were thus considered CF carriers. The total prevalence of carriers in the population, derived from the incidence of homozygotes, is expected to be 1:25, and the prevalence of ΔF508 carriers 1:33. On a population basis the expected number would have been 60, and this difference is significant (chi-square value = 11.36; $p < 0.001$). In our neonates the relative risk that a heterozygote would have an IRT value in the top 1% of the population was thus 1.7 (95% confidence limits, 1.4 to 2.1). One patient from this cohort was originally classified as a carrier of ΔF508 on the basis of an equivocal sweat test, and was later found to have CF,

Table II. Sensitivity, specificity, and positive predictive value

	IRT/IRT	IRT/DNA
Sensitivity (%)		
Total cases*	92	98†
Unexpected cases‡	89	98%
Specificity (%)		
First sample	99.31	99.95%
False-positive rate (%)		
First sample	0.69	0.054
PPV (%)		
First sample	5.1	37.4 (overall)§
Second sample	47	NA

PPV, Positive predictive value; *NA*, not applicable.
*Percentage of babies with CF identified early, either by screening or because of high risk (meconium ileus or an affected sibling).
†Only 59 (95%) of 62 babies had a positive DNA test result (see text).
‡Percentage of CF babies, excluding high-risk babies, identified by the screening program.
§Where the DNA result showed homozygosity for the mutation, the positive predictive value was 100%. If the baby had one copy only of the mutation, the positive predictive value was 13%.

carrying one copy of the common mutation and one mild mutation, R117H.[18]

Incidence of CF. The apparent incidence of CF in the neonatal population, based on figures obtained during the IRT protocol, was 1:2609 (95% confidence interval, 1:2291 to 1:2977). Taking into account 19 terminations of pregnancy after prenatal diagnosis of an affected fetus, a "corrected" incidence of 1:2488 (1:2183 to 1:2827) could also be estimated. During the period from 1981 through 1984, no prenatal diagnosis was available, and children in this cohort are now aged at least 10 years, so this might be the most truly representative cohort to consider. The incidence in this cohort was 1:2531 (95% confidence interval, 1:2109 to 1:3064). During the IRT/DNA screening protocol, we found an incidence of 1:2821 (1:2221 to 1:3640) when termination of five affected fetuses was taken into account. These figures are not significantly different. It seems likely that there would be some additional patients with CF, not identified by the screening program and yet to be recognized clinically, especially from this latter cohort.

Sensitivity, specificity, and positive predictive value. The sensitivity in unexpected (not high-risk) cases was 89% for the IRT protocol and 98% for the IRT/DNA test. The specificity of the IRT/DNA protocol was somewhat higher, and the overall positive predictive value of this protocol was 37.4%, considerably greater than that of the IRT test, which was only 5.1% after the first IRT test (Table II).

Cost. In 1994 the incremental laboratory cost of screening for CF—that is, the cost of adding CF screening to an already-existing neonatal screening program—was $A1.42 per baby screened (roughly equivalent to $U.S.1.00). This included all disposable items and the costs of the scientific and clerical staff involved in CF testing, but did not include administrative or general laboratory costs. The component attributable to the DNA testing was 10 cents per baby, or $A10 for each baby who had a DNA test. The cost per unexpected CF case detected was $A5160. The within-laboratory cost of protocol 1 in 1994 terms would have been approximately $A1.44, similar to that of protocol 2. The cost of collecting a second sample (protocol 1) was estimated at $A13.25. Second samples from 75% of babies were collected in maternity or baby health units, with the only incremental cost being postage. The average cost of genetic counseling and of further DNA testing for each family of a heterozygote baby (protocol 2) was approximately $A100. These latter two figures must be regarded as estimates only, but they suggest a similar cost of around $A1.50 per baby screened for each protocol.

DISCUSSION

The IRT testing alone has provided an early diagnosis for 92% of babies with CF, but among patients who were not already at high risk (siblings or babies with meconium ileus), the test failed to identify 11% of affected babies. This group of 11% did not primarily comprise patients with mild CF; the test readily identifies babies with pancreatic sufficiency.[19] The IRT protocol also had the drawback that the positive predictive value of the first test was only 5.1%, and a large number of babies required a second blood test. From the earliest days it was recognized that the IRT level in babies with CF declined rapidly into the normal range.[1] Late performance of the second test (after 6 weeks, despite a timely request) appeared to account for one fourth of the missed cases.

Sensitivity of the IRT test. Babies with meconium ileus often have an IRT level below the normal cutoff value.[20] This was so in approximately 50% of babies with meconium ileus in our study when monoclonal antibodies were used in the testing method, but in only 5% when a polyclonal antibody from Auckland University was used. When babies with meconium ileus were excluded, the sensitivity of the first IRT test was 96.4%. However, with a cutoff value being the top 1% as in protocol 2, only 1.9% of patients would have had a negative result on the first IRT test. This figure is the appropriate one to use when one is considering the theoretical sensitivity of the IRT/DNA protocol.

Our results with an IRT strategy are similar to those reported worldwide.[21, 22] We concluded that this original strategy was a moderately successful one, with a high proportion of babies achieving a diagnosis by 2 to 3 months, but that the protocol was a demanding one, with poor specificity after the first test.

DNA testing. Our results approximate those predicted

theoretically on the basis of the genotyping of patients with a previous diagnosis of CF in New South Wales.[23] We opted to test only for the ΔF508 mutation, which accounted for 76% of mutant alleles, on the grounds of cost-effectiveness. Using ΔF508 alone, we predicted that our detection rate would be 92%, with 2% negative results on the IRT test and 6% negative results on DNA testing, but that 2% to 3% among these would be high-risk cases and so detected despite a negative test result. Adding a test for the next most common mutation might result in the detection of one extra case every 3 years. During the 4½ years of IRT/DNA testing in South Australia, with testing for five different mutations, all 29 babies whose CF has been detected so far would have been identified only with ΔF508 mutation testing.[7]

The prediction of the likely outcome with DNA testing has been calculated with an assumption of random mating, whereas random mating is unlikely to occur in some ethnic groups in our population. Nevertheless, because the incidence of CF itself in the neonatal population is lower in those ethnic groups in which the percentage occurrence of the ΔF508 mutation is also lower, this assumption of random mating introduces very little error with the present population mix in New South Wales.

Babies with elevated IRT levels but normal DNA test results were considered to have an overall negative test result. The positive IRT test result was not reported separately. The risk of such a baby having CF, because of carrying two rare mutations, is less than 1:400, which is relatively low but considerably greater than the risk for a baby with a negative IRT test result. We considered that the only way to deal with this dilemma was to ensure that all health professionals and parents were aware that CF could be missed by this screening test.

Advantages and disadvantages of the two protocols. There are advantages in the IRT/DNA protocol, both for case detection and because of the reduced occurrence of false-positive results. These advantages have also been noted in the Wisconsin research program in CF screening.[13] We have enough experience with the IRT test, and enough information about genotypes in the CF community, to be reasonably sure that protocol 2 will lead to the identification of around 94% of those babies with CF who otherwise would have been recognized clinically during infancy or childhood. We cannot be so sure about patients who have very mild CF. The IRT/DNA protocol generated a slightly higher sweat-testing rate than the IRT protocol, but in the IRT protocol the false-positive rate of the first IRT test was considerable and this constituted the main disadvantage.

The disadvantages of the IRT/DNA test are two: (1) the patients with the CF cases missed by the strategy tend to come from certain predictable ethnic groups, even though overall a higher percentage of cases may be detected, and (2) there is the problem of the unwanted identification of a small number of babies who are healthy CF carriers. This requires careful organization of appropriate genetic counseling and the development of policies for further testing of family members.

Which cases can be detected by neonatal screening tests? The ascertainment of cases missed by screening tests must be pursued rigorously if it is to become clear which cases can, and which cannot, be detected by current screening tests. There was no evidence of extra delay in diagnosis in those patients missed by our screening tests. The ages at which the diagnosis was made had a distribution similar to that of current data from the United States.[13] Sweat testing has been used as the definitive test for confirmation of the diagnosis of CF, but we have found equivocal results not uncommonly in neonates subsequently shown to have CF[24]; in addition, patients have recently been described who have CF but normal sweat test results.[25] Of particular interest is the finding, by almost all programs with an IRT/DNA strategy of detection, of an excessive number of heterozygote babies with normal sweat test results, apparently healthy carriers of CF.[7, 8, 26, 27] In our study the detection of apparent carriers among neonates with elevated IRT levels was approximately 1.7 times that expected. The reason is not clear, but one possible explanation is that some of these patients have a mild form of CF.

Our early experience with the IRT/DNA protocol for neonatal screening for CF in 189,000 infants justifies this approach because of the more efficient and streamlined process it affords, with equal or better early detection rates, greatly improved positive predictive value, and minimal additional cost. The potential for adverse effects from the detection of a small number of heterozygotes seems likely to be more than balanced by reduction in the number of families with anxiety because of false-positive test results in the protocol employing only IRT testing.

REFERENCES

1. Crossley JR, Elliott RB, Smith PA. Dried blood spot screening for cystic fibrosis in the newborn. Lancet 1979;2:472-4.
2. Heeley AF, Heeley ME, King DN, Kuzemko JA, Walsh MP. Screening for cystic fibrosis by dried blood spot trypsin assay. Arch Dis Child 1982;57:18-21.
3. Wilcken B, Brown ARD, Urwin R, Brown DA. Cystic fibrosis screening by dried blood spot trypsin assay: results in 75,000 newborn infants. J PEDIATR 1983;102:383-7.
4. Hammond KB, Abman SH, Sokol RJ, Accurso FJ. Efficacy of statewide screening for cystic fibrosis by assay of trypsinogen concentrations. N Engl J Med 1991;325:769-74.
5. Tluczek A, Mischler EH, Farrell PM, et al. Parents' knowledge of neonatal screening and response to false-positive cystic fibrosis testing. Dev Behav Pediatr 1992;12:181-6.
6. Ranieri E, Ryall RG, Morris CP, et al. Neonatal screening strategy for cystic fibrosis using immunoreactive trypsinogen and direct gene analysis. BMJ 1991;302:1237-40.

7. Ranieri E, Lewis BD, Gerace BL, et al. Neonatal screening for cystic fibrosis using immunoreactive trypsinogen and direct gene analysis : four years' experience. BMJ 1994;308:1469-72.
8. Gregg RG, Wilford BS, Farrell PM, Laxova A, Hassemer D, Mischler E. Application of DNA analysis in a population screening program for neonatal diagnosis of cystic fibrosis: comparison of screening protocols. Am J Hum Genet 1993; 52:616-26.
9. Larsen J, Campbell S, Faragher EB, et al. Cystic fibrosis screening in neonates: measurement of immunoreactive trypsin and direct gene analysis for ΔF508. Eur J Pediatr 1994;153:569-73.
10. Wilcken B, Chalmers G. Reduced morbidity in cystic fibrosis patients detected by neonatal screening. Lancet 1985;2:1319-21.
11. Danckert-Roelse JE, te Meerman GJT, Martijn A, et al. Survival and clinical outcome in patients with cystic fibrosis, with or without neonatal screening. J PEDIATR 989;114:363-7.
12. Bowling F, Cleghorn G, Chester A, et al. Neonatal screening for cystic fibrosis. Arch Dis Child 1988;63:196-8.
13. Farrell PM, Aronson RA, Hoffman G, Laessig RH. Newborn screening for cystic fibrosis in Wisconsin: first application of population-based molecular genetics testing. Wis Med J 1994; 93:415-21.
14. Mastella G, Barlocco E, Antonacci B, et al. Is neonatal screening for cystic fibrosis advantageous? The answer of a wide 15-year follow-up study. In: Travert G, ed. Mucoviscidose: Depistage Neonatal et Prise en Charge Precoce Caen; CHRU de Caen, 1988:127-43.
15. Bowling FG, Walson AR, Rylatt DB, Elliott JE, Bunch RS, Bunderson PG. A monoclonal antibody–based trypsinogen EIA for screening of neonates for cystic fibrosis. Lancet 1987; 1:826-7.
16. Petterson K, Soderholm J, Suonpaa M. Organisation of an immunofluorimetric assay based on the DELFIA technique for the measurement of immunoreactive trypsinogen in dry blood spots [Abstract]. Pediatr Pulmonol 1991;57:90.
17. Nelson PV, Carey WF, Morris CP. Gene amplification directly from blood spots. Lancet 1990;336:1451-2.
18. Fitzgerald D, Van Asperen P, Henry R, et al. Delayed diagnosis of cystic fibrosis in children with a rare genotype (ΔF508/R117H). J Pediatr Child Health 1995;31:168-71.
19. Waters DL, Dorney SFA, Gaskin KJ, Gruca MA, O'Halloran M, Wilcken B. Pancreatic function in infants identified as having cystic fibrosis in a neonatal screening program. N Engl J Med 1990;322:303-8.
20. Wilcken B. Newborn screening for cystic fibrosis: its evolution, and a review of the current situation. Screening 1993;2: 43-62.
21. Travert G. Analyse de L'experience mondiale de depistage neonatale de la mucoviscidose par dosage de la trypsine immunoreactive sanguine. In: Travert G, ed. Mucoviscidose; Depistage Neonatal et Prise en Charge Precoce Caen; CHRU de Caen, 1988:1-23.
22. Hammond KB, Abman SH, Sokol RJ, Accurso FJ. Efficacy of statewide screening for cystic fibrosis by trypsinogen concentrations. N Engl J Med 1991;325:769-74.
23. Cystic Fibrosis Genetic Analysis Consortium. Population variation of common cystic fibrosis mutations. Human Mutation 1994;4:167-77.
24. O'Halloran M, Brown J, Wilcken B. Diagnosis of cystic fibrosis [Letter]. Lancet 1987;2:1156.
25. Highsmith WE, Burch LH, Zhou Z, et al. A novel mutation in the cystic fibrosis gene in patients with pulmonary disease but normal sweat chloride concentrations. N Engl J Med 1994;331: 974-80.
26. Larouche D, Travert G. Abnormal frequency of ΔF508 mutation in neonatal transitory hypertrypsinaemia [Letter]. Lancet 1991;339:55.
27. Spence WC, Paulus-Thomas J, Orenstein DM, Naylor EW. Neonatal screening for cystic fibrosis: addition of molecular diagnostics to increase specificity. Biochem Med Metab Biol 1993;49:200-11.

Cystic Fibrosis—Current Topics

Volume 3

Edited by **J. A. Dodge**
The Queen's University of Belfast, UK

D. J. H. Brock
The University of Edinburgh, UK

J. H. Widdicombe
University of California, San Francisco, USA

JOHN WILEY & SONS
Chichester · New York · Brisbane · Toronto · Singapore

Chapter Nine

Neonatal Screening using Combined Biochemical and DNA-based Techniques

Enzo Ranieri[a], Barry D. Lewis[a], C. Phillip Morris[a] and Bridget Wilcken[b]

[a]*Department of Chemical Pathology, Women's and Children's Hospital, Adelaide, SA and* [b]*Biochemical Genetics and Newborn Screening, Royal Alexandra Hospital for Children, Sydney, NSW, Australia*

INTRODUCTION

Cystic fibrosis (CF) is not always readily identified in infants and children and considerable delays in diagnosis after onset of symptoms have been documented[1-3]. Because of this, there has long been an interest in achieving very early diagnosis by neonatal screening. In 1964, there was a preliminary report of increased albumin in the meconium of infants with CF[4] and this became the basis of tests used in infant screening programmes[5-8]. However, it was only when Crossley et al.[9] first described increased immunoreactive trypsin (IRT) levels in dried blood-spots from infants with CF in 1979 that neonatal screening for CF on a large scale seemed likely to be successful. With the possibility of a dried blood-spot test came much more wide ranging consideration of the advisability of neonatal screening for CF.

Cystic Fibrosis—Current Topics, Volume 3. Edited by J. A. Dodge, D. J. H. Brock and J. H. Widdicombe

THE RATIONALE FOR NEONATAL SCREENING FOR CF

The likely benefits of screening were considered in the early 1980s by several authors[10–13]. Possible advantages included: a favourable alteration in the natural history of the disease; availability of genetic counselling to avoid the birth of a second child with CF in a family where the first CF child remained undiagnosed; eliminating the period of pre-diagnosis anxiety; and the prospect of being able to conduct critical studies of current or new forms of treatment. Theoretical disadvantages of pre-symptomatic diagnosis which were mentioned included: adverse effects on developing family relationships with the new baby; the trauma of false positive tests; and the stigmatisation of patients who might only have mild disease.

Various criteria have been proposed for inclusion of a disease in neonatal screening programmes (World Health Organization)[14–16]. These 10 or more 'commandments' can readily be reduced to two broad considerations: does neonatal diagnosis provide clear-cut benefits and are these benefits reasonably balanced by the costs, of all kinds? In 1983, the Cystic Fibrosis Foundation (USA) task force concluded that the case for screening was not clear-cut and considerable research was needed on the questions of benefit and harm before mass screening should be undertaken[17]. In the meantime, however, several centres had already started screening.

IMPLEMENTATION AND APPARENT PROBLEMS WITH IRT

The papers by Crossley et al. in 1979 and 1981[9,18] showed clearly that increased IRT was a hallmark of neonatal CF (although older patients have greatly reduced levels). This is thought likely to be due to blockage of ductule secretions in the pancreas, as the acinar cells appear to produce normal amounts of zymogen[19]. The decline in IRT levels with age in CF presumably occurs secondary to atrophy of the acini and their replacement with fibrotic tissue. The findings of Crossley et al.[9,18] were confirmed and a number of established neonatal screening programmes added an IRT test for CF to their routine screening procedures[20–28]. In all these prospective but uncontrolled studies, the protocol adopted was similar to that first reported by Crossley et al.[18]. This protocol involved an IRT assay performed on the routine neonatal screening blood-spot collected at age 3–5 days. The specificity of the test in the first week of life was quite low and a second sample was obtained, usually at about 3–5 weeks, from any infant whose first result was higher than a designated cut-off value. At this time the test had much higher specificity and the infant was referred for a sweat test if the IRT

remained increased. Although there were some modifications, generally this two-stage IRT protocol remained in use throughout the 1980s and beyond.

The case-finding performance of this protocol was reasonably satisfactory with around 95% of infants with CF achieving early diagnosis, but there were several problems. The false positive rate after the first test was high, between 0.3% and 1%[29–31] of screened babies needed a second IRT test, with resulting parental anxiety[32,33]. A second problem was the age-related decline of IRT levels in infants with CF. The rate of decline varied with the method used[24,34,35] and was responsible for false negative results in tests performed on repeat samples[36]. The IRT was also only variably increased in infants with meconium ileus[37,38], although this was not a major problem since such patients had already declared themselves. Nevertheless, it was open to misinterpretation.

THE IMPACT OF DNA ANALYSIS ON NEONATAL SCREENING FOR CF

The first neonatal screening programme to incorporate DNA analysis into CF screening was reported by Laroche et al.[39] in early 1990. They used haplotype analysis for the polymorphic marker KM19, in order to eliminate infants with a low risk of CF after the initial IRT measurement[40]. However, the discovery of a common disease mutation in the CFTR gene[41,42] then made possible a single-sample, two-tier screening protocol using both IRT assay and DNA mutation analysis. Such an approach, outlined in theory by Bowling et al.[43], immediately did away with the need for second samples and problems of age-related decline in IRT. Similar IRT/DNA protocols have now been adopted by most existing screening programmes and early reports have been favourable[44–50].

The protocol adopted has been to measure IRT in routinely collected dried blood-spots, followed by mutation analysis in those in which the IRT is above a generous cut-off level, usually the top 1% of values, using the same sample. As an example, the South Australian (SA) protocol is outlined in Figure 1. Infants found to carry two CF mutations are immediately identified as suffering from CF. Infants with only one identified mutation could either have CF, with an additional unidentified mutation, or they could be carriers. These infants are referred for a sweat test to define their CF status.

In populations derived from northern Europe, the frequency of the common ΔF508 mutation varies between 70–90% of all CF alleles[51], thus predicting that about 90–99% of infants with CF will carry at least one copy of this mutation. Screening programmes therefore either test for this mutation alone or, in a few cases, a combination of ΔF508 and other mutations. The frequency of the ΔF508 mutation is crucial in determining whether the IRT/DNA approach is appropriate for a given population.

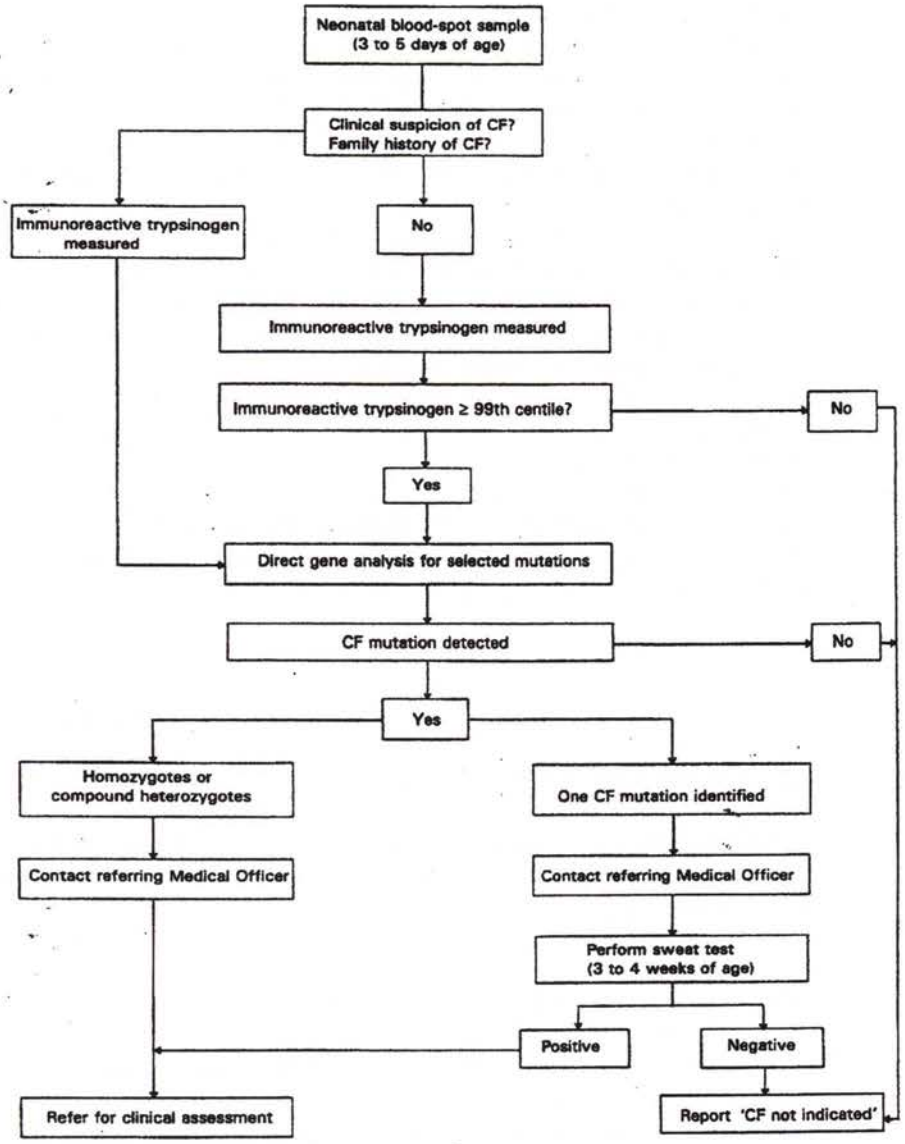


Figure 1 The South Australian IRT/DNA screening protocol

NEONATAL SCREENING STRATEGIES CURRENTLY IN USE

Most screening programmes that now use the combined IRT/DNA approach analyse only the ΔF508 mutation but there are other strategies in use. The SA programme includes four additional mutations[45] which increases the frequency of detectable mutations from 72% to 80% and consequently reduces the false negative rate on DNA analysis from 8% to 4%. In northern France, where the prevalence of the ΔF508 mutation is only 60%[51], DNA haplotyping for the KM19 polymorphism as a second test has proved effective[40]. Infants who are homozygous for the '2' allele are sweat tested, those who are heterozygous have a second IRT measured, and those who are homozygous for the '1' allele are not tested further. This protocol has recently been modified to incorporate mutation analysis and a complex risk estimate for CF is calculated based on ΔF508 genotype, KM19 haplotype, and the extent to which the IRT is increased (G. Travert, personal communication). The complexity of this protocol highlights the limitations of the IRT/DNA approach in populations with a low frequency of ΔF508. In northern Italy, where the prevalence of the common mutation is even lower (about 54%), the DNA approach is not feasible at all and CF screening consists of a combination of IRT and meconium lactase testing. This has better specificity than either test alone[52]. Other regions may adopt different protocols to suit the particular needs of the population to be tested.

THE BENEFITS OF EARLY DETECTION

Short-term benefit has been shown for infants with CF who are diagnosed early by neonatal screening but, as yet, no clear-cut proof of an altered long-term outcome has been established. Early studies of outcome were uncontrolled or used historical controls[53–55]. Nevertheless, these, and other studies, produced compelling evidence of a significant reduction in morbidity in the first years of life[53,54]. Two other studies with long-term follow-up of screened and unscreened patients, followed in parallel, in the same regions and with similar treatment protocols, also showed better survival and better clinical scores in the screened patients[56,57]. However, it is not certain how the data would relate to the present day since the patients in these studies were born largely in the 1970s when clinical diagnosis may not have been so timely nor treatment as effective.

Two randomized controlled trials of screening have been conducted, one in East Anglia[58,59] and one in Wisconsin[2], which were both started in 1985. The follow-up period is thus quite short and, although early morbidity was significantly less in the screened patients[59], so far no striking differences have emerged beyond this time. There have been some problems in assessing these

data[30]. The trial in Wisconsin has now come to an end and CF screening has been incorporated into the routine screening protocol for that state[47]. This is one of a very few screening centres in the USA to offer the test. We must conclude that, so far, only short-term benefits can be shown, but these are enough in the opinion of some to make neonatal screening for CF worth while and it has proved to be popular wherever it is carried out.

BIOCHEMICAL AND DNA-BASED TECHNIQUES

INHERENT PROBLEMS WITH IRT

The use of IRT in neonatal screening for CF has recently been reviewed by Heeley and Bangert[60]. Regardless of the protocol followed, IRT, along with its inherent problems, still remains the cornerstone of neonatal screening for CF from dried blood-spots. However, variability associated with the measurement of IRT can have considerable impact on the sensitivity and specificity of this screening marker. This is due to the considerable sample, individual and population variability in IRT as well as persisting assay problems[60,61].

The problems associated with sample collection are well recognised. In our experience, poor and variable collection technique, improperly dried samples, and sample contamination with faecal trypsin are common problems. There is also a decline in recoverable IRT with age of sample and this decline is variable in infants with CF[60-62]. We have also observed significant differences in IRT in populations from different geographical areas (Figure 2). This is likely to be secondary to variation in timing of the IRT assay after sample collection although there may also be true population differences. Patient factors influencing the performance of IRT assays include age at initial testing (complicated by a shift towards earlier discharge) and the clinical and nutritional status of the infant at time of sample collection[60]. It is also well established that in the normal neonatal period blood IRT concentration is highly positively skewed, with the highest values occurring in the first days of life. The phenomenon of transient (non-CF) neonatal hypertrypsinaemia, which occurs predominantly in the first days of life, is the primary cause of false positive results[30,31]. IRT is also unreliable in screening infants with meconium ileus[37,38].

There is considerable variation in the IRT results obtained by different assays. Some of the commercial assays for the measurement of blood-spot IRT are listed in the legend of Figure 3. The circulating trypsin recognized by these immunoassays appears largely to be trypsinogen plus some trypsin bound to α_1-antitrypsin[60]. The pancreas secretes two forms of trypsinogen (cationic and anionic) which are immunologically distinct and also differ in their stability and half-lives in serum[63,64] The specificity of the different IRT assays

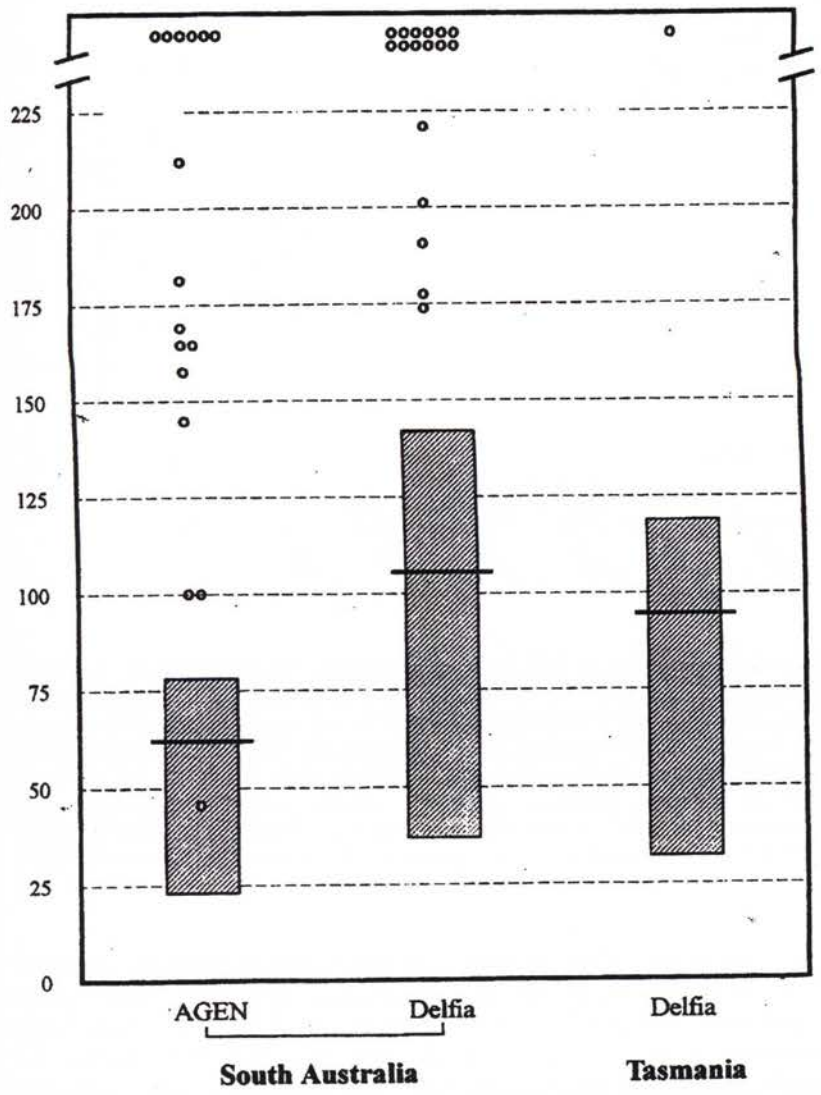


Figure 2 Comparison of neonatal blood-spot IRT distributions using two different commercial immunoassays; NeoScreen EIA, AGEN Biomedical, Australia (AGEN) and DELFIA®, Wallac, Finland (Delfia). Data generated by the South Australian State Screening Laboratory using a sample of routinely collected neonatal blood-spots from South Australia ($n = 10\,000$ for each method) and Tasmania ($n = 3860$ for the Delfia method only). The bar graph represents the 50th centile (bottom), 99th centile (top), and the 97th centile (——) for each distribution. All infants with cystic fibrosis identified by these methods since screening commenced (O). One infant with meconium ileus fell below the 99th centile. There is a significant difference in the IRT distributions between methods and between populations ($p < 0.001$, Mann–Whitney U test)

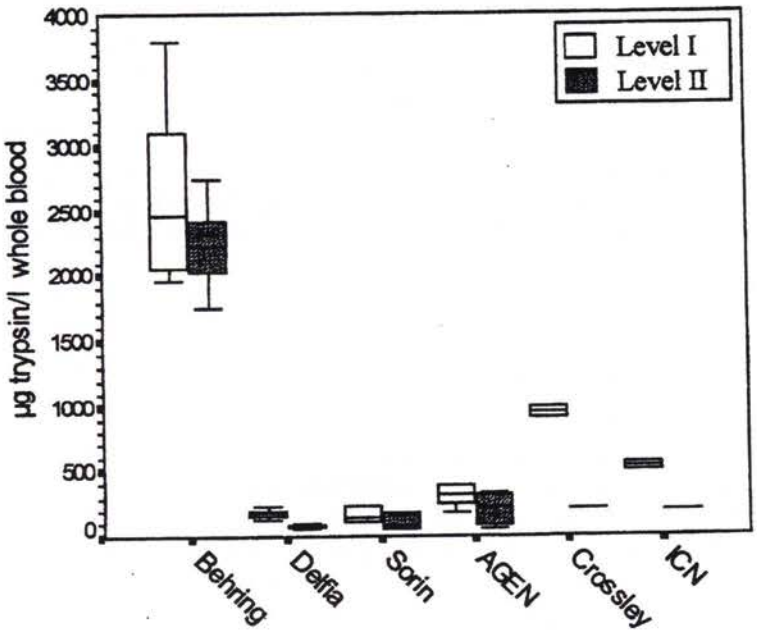


Figure 3 Comparison of neonatal blood-spot IRT assays. Two duplicate blood-spot samples with different increased levels of IRT (level I and level II) were distributed to laboratories participating in the IRTIQAS programme[69]. The data from 26 laboratories is displayed as a box plot, the range of values is defined by the tails and the bar graph identifies the 25th and 75th centiles. The number of participants and methods used were: RIA-gnost Trypsin ($n = 6$), Behring, Germany; NeoIRT ($n = 12$), Delfia®, Wallac, Finland; Trypsik® ($n = 3$), Sorin Biomedica, Italy; NeoScreen EIA ($n = 3$), AGEN Biomedical, Australia; the Crossley ($n = 1$) in-house method[18]; ICN Biomedical ($n = 1$), USA

for these components varies and it is not clear which species of trypsinogen is being measured[60]. The introduction of monoclonal antibodies[65,66] did not significantly improve inter-assay variability[67–69] and the continuing lack of an international IRT reference material complicates assay comparisons. The SA laboratory has compared two IRT monoclonal antibodies from different manufacturers during analysis of several thousand neonatal blood-spots from the SA population (Method 1 detects primarily cationic while Method 2 recognizes cationic and anionic trypsinogen). Using the same in-house calibrators and method of signal detection[70] we obtained estimates for the 50th centile of 23 and 37 µg trypsin/l whole blood, and for the 99th centile of 78 and 142 µg trypsin/l whole blood respectively (Figure 2).

There are also significant differences in the sources of trypsin, and in the methods of purification, used to prepare calibrators for the various IRT assays.

The calibrators also differ in the type of protease inhibitor used, if any, to prevent autolysis and serum antiprotease binding on reconstitution. The calibrators can either be supplied in lyophilized or blood-spot form. Despite the potential problems with stability of assay calibration, especially in the long term, most screening laboratories choose a fixed whole blood IRT concentration when defining their cut-off[60,71]. Others, however, prefer a floating cut-off derived statistically during each analytical run of the assay, while some centres use a combination of both. The assignment of a particular IRT cut-off, above which further action is indicated, is very much dependent on the type of assay used and the screening protocol adopted.

APPLYING DNA TECHNIQUES TO NEONATAL SCREENING FOR CF

DNA analysis of all neonatal samples is currently impractical, being too expensive, and would be complicated by identifying many CF carriers in addition to those affected by CF. The IRT/DNA screening protocols, as they currently exist, do not require mutation analysis on the same scale as some other mass screening projects, such as pilot carrier screening programmes[72]. However, polymerase chain reaction (PCR) amplification of DNA directly from neonatal blood-spots without the need for extraction[73,74] does permit many samples to be processed at the one time. One problem associated with the direct use of blood-spots for PCR amplification is the low and variable amplicon yield, possibly due to contaminants within the blood-spot. Makowski et al.[75] have recently published a simple *in situ* amplification method to detect CF mutations which may overcome this problem. In the meantime, if amplification is carried out directly from blood-spots, a highly robust and sensitive mutation detection technique is desirable.

The methods employed by 19 laboratories participating in a pilot international quality assurance programme for the detection of the ΔF508 mutation are given in Table 1. Heteroduplex analysis is a popular method, perhaps because it is the simplest and cheapest of the methods listed, but it is also most prone to errors and artefacts[76]. The SA laboratory has chosen to continue using allele specific oligonucleotide analysis, one of the most expensive and time consuming methods, because of the highly robust and reliable nature of the assay and because it can be applied to the detection of any mutation. However, the most common method used to detect ΔF508 is the resolution of mutant and wild-type PCR products based on the 3 base pair size difference. This is the method used in New South Wales (NSW). Although there are differences in the methods used by various screening laboratories, it is up to each laboratory to assess the method most appropriate for their individual requirements and level of expertise. The widespread use of mutation analysis in neonatal screening must await the development of a truly automated mutation analysis system.

Table 1 Methods used for blood-spot ΔF508 mutation analysis

Method of ΔF508 determination	No. of laboratories[a]
Heteroduplex	6
Endonuclease digestion	3
Sizing on polyacrylamide gels	9
Allele specific oligonucleotide analysis	1

[a] Results from a survey conducted in 1993 of those laboratories undertaking ΔF508 mutation analysis from dried blood-spots.

QUALITY ASSURANCE PROGRAMMES

There are two external quality assurance programmes currently offered to neonatal screening laboratories to monitor IRT assay performance. The Immunoreactive Trypsin International Quality Assurance Scheme (IRTIQAS) is run by Dr G. Travert from the Laboratoire de Biophysique Médicale, Centre Hospitalier Universitaire de Caen, in conjunction with the Regional Genetic Screening Unit, Peterborough, East Anglia[69]. IRTIQAS distributes dried blood-spots, prepared with blood from patients with pancreatitis or renal failure, to some 33 laboratories in quarterly cycles and compares performance within IRT assay groups as well as assessing outcome of any actions taken about decision levels. Dr D. Webster, The National Testing Centre, Auckland, New Zealand also offers a programme which includes ΔF508 mutation analysis in addition to IRT[77]. These programmes perform reasonably well considering the continuing lack of an international IRT reference material and the instability of IRT in dried blood-spot preparations.

To highlight the variability associated with the measurement of IRT, we circulated two duplicate IRT quality control blood-spots in conjunction with the Human Genetics Society of Australasia Newborn Screening Quality Assurance Program (HGSANSQAP). The blood-spots were prepared at different concentrations of increased IRT using highly purified trypsin from human pancreas (two sources of trypsin; Scripps Laboratories, San Diego, and National Testing Centre, Auckland) and we used a number of proteolytic enzyme inhibitors to maintain IRT stability. The blood-spots also included ΔF508 PCR product to determine if DNA analysis could be incorporated into the programme. The quality control samples were then circulated to laboratories participating in IRTIQAS. The results (Figure 3) confirmed a large variability in estimation of IRT amongst participating laboratories and reflect not only differences in the performance of IRT measurement but also differences in calibration. The successful inclusion of ΔF508 PCR product in the dried blood-spots has now led to its incorporation in the HGSANSQAP and is soon to be added to the IRTIQAS programme.

WORLD IRT/DNA SCREENING EXPERIENCE

To date there have been four prospective studies incorporating DNA analysis (Table 2)[40,44–47] and several other studies that have retrospectively used DNA analysis as an addition to their existing IRT programmes[48–50]. Three prospective studies (Table 2) have reported their experience with an IRT/DNA protocol using a single blood-spot sample analysed for the presence of specific CF mutations[44–47]. These three studies represent close to 400 000 neonates screened to detect 112 out of a total of 113 infants with CF. Each study used essentially the same protocol as described in Figure 1, except that the NSW and Wisconsin studies only screened for the presence of ΔF508, rather than five mutations in the SA study. The sensitivity of all studies is high with only one infant so far known to have been missed in the largest cohort (NSW). The positive predictive value is also high compared to IRT-alone protocols, although a lower positive predictive value was obtained in the Wisconsin study due to a lower IRT cut-off (98%)[47]. This lower cut-off has the added disadvantage of increasing the number of unwanted carriers detected. Interestingly, these workers now recommend an even lower IRT cut-off of 96%[47]. The Normandy study is a prospective study that maintained their existing IRT protocol (Table 2). This study initially incorporated DNA haplotyping and a second blood-spot sample, and has now been modified to include ΔF508 mutation analysis (see Neonatal Screening Strategies Currently in Use). There have been other retrospective studies[48–50] that are generally continuations of IRT-alone protocols that have incorporated an additional DNA analysis in their programme, i.e. they involve a second blood-spot sample. The existence of several different programmes highlights the difficulties faced when introducing CF screening in different populations and the need to tailor any programme to the specific population tested.

It is difficult to combine data from most of these studies in a meaningful way due to procedural variations, or differences in the incidence of CF and frequency of CF mutations in these geographically distinct populations. However, the data from SA and NSW can be combined as these populations and their programmes are very similar (Table 3). Despite their relatively uniform ethnic composition, there seem to be some differences between the two populations that are probably related to founder effects. The most notable difference is the incidence of ΔF508 in the two populations, 72% and 76% in SA and NSW, respectively. The SA programme has incorporated a further four mutations to bring the proportion of mutations detected up to about 80%. The combined studies represent the most extensive body of data for an IRT/DNA protocol with almost 100 infants with CF detected in about 300 000 screened. Interestingly, a higher proportion of CF patients were defined by the presence of two CF mutations than expected based on the frequency of mutations amongst CF patients (80% [$\chi^2 = 3.27$; $0.1 > p > 0.05$] and 71% [$\chi^2 = 5.31$; $p < 0.001$] in SA and NSW, respectively).

Table 2 World experience of prospective IRT/DNA neonatal screening for cystic fibrosis

	Single blood-spot			Dual blood-spot
	SA[a]	NSW[b]	Wisconsin[c]	Normandy[d]
Screening time (years)	5	2	2.5	5[e]
Number of infants screened	108 871	189 000	94 259	251 790
DNA analysis	1 220 (1.12%)	1 968 (1.04%)	1 746 (1.85%)	796 (0.31%)
One identifiable mutation	89	117	106	195
Two identifiable mutations	26	44	NA	21
Infants with cystic fibrosis	33	62	18	54
detected	33	61	18	53
missed	–	1	–	1
Sensitivity (%)	100	98.4	100	98.1
Positive predictive value (%)	28.7	37.8	17.0	24.5

[a] Ranieri et al.[44], Ranieri et al.[45], and unpublished data; [b] Wilcken et al.[46]; [c] Farrell et al.[47]; [d] Laroche and Travert[40], and G. Travert, personal communication; [e] Includes some retrospective analysis for the ΔF508 mutation. NA = not available.

Table 3 Combined South Australian (SA) and New South Wales (NSW) IRT/DNA screening data

	SA	NSW	Combined
Number of infants screened	108 871	189 000	297 871
IRT <99th centile	107 651	187 032	294 683
DNA mutation analysis	1 220 (1.12%)	1 968 (1.04%)	3 188 (1.07%)
No identifiable mutation	1 094	1 798	2 892
One identifiable mutation	89	117	206
sweat test positive	7	15	22
sweat test negative	82	102	184
Two identifiable mutations	26	44	70
Positive predictive value	28.7	37.8	34.1
Total number of infants with cystic fibrosis	33	62	95
High risk neonates with MI	7	9	16
One or two identifiable mutations	33	59	92
No identifiable mutations:			
detected (MI)	–	2	2
missed	–	1	1

MI, meconium ileus; IRT, immunoreactive trypsinogen.

The genotype of 74 out of 95 of these infants with CF was determined by further analysis of a range of other mutations (Table 4). In theory, these programmes expect to miss 4–6% of infants with CF due to the absence of detectable mutations but, unexpectedly, two of three such infants in the NSW population were detected due to meconium ileus. Thus, the known total false negative rate for these Australian studies to date is only about 1%. The single infant missed presented with clinical signs of CF at age three months. Of course, the real number may be higher as patients missed by the initial IRT cut-off or missed due to absence of screened mutations may take many years to present clinically and the DNA screening programmes have only been in operation for five years in SA and two years in NSW. It may be possible to get some idea of the potential miss rate of the initial IRT measurement by retrospective analysis of neonates screened by an IRT-alone protocol. From a population of 1 015 000 neonates screened in NSW between 1981 and 1992, only 6 out of a total of 389 known CF cases (1.5%) would have been missed by the first screen if an IRT cut-off of the top 1% had been operating[46].

To gain an accurate idea of the incidence of CF from figures derived from screening programmes[78], the length of follow-up must be sufficient and the ascertainment of missed cases must be rigorous, and there is now the confounding factor of prenatal diagnosis and the termination of some affected fetuses. In addition, the apparent incidence may be increasing due to early recognition by neonatal screening programmes of mildly affected patients, such as those with borderline sweat tests, who would not previously have been

Table 4 Genotype of screened infants with cystic fibrosis

	SA[a]	NSW[b]
Total	33	62
ΔF508/ΔF508	20	44
ΔF508/G551D	3	
ΔF508/G542X	2	
ΔF508/R117H	1	2
ΔF508/W1282X	1	
ΔF508/ΔI506/7	1	
ΔF508/X	5	13
X/X		3

Additional mutations analysed: [a] ΔF508, ΔI506/7, G551D, G542X, R553X, R1162X, R117H 621 + 1G→T, 1717-1G→A, N1303K, W1282X, R560T, S549I, S549N and S549R; [b] ΔF508 and R117H, ;X no mutation detected.
SA, South Australia; NSW, New South Wales.

recognized as carrying CF mutations. We specifically sought missed cases and considered only a cohort aged more than 10 years, and born before the availability of prenatal diagnosis, in reaching a CF incidence amongst neonates of 1:2500 (95% CI 1:2100–1:3000) for the NSW region[46].

COMPARISON OF IRT VERSUS IRT/DNA SCREENING STRATEGIES

The sensitivity of the IRT/DNA approach depends on two factors: the sensitivity of the IRT test during the first week of life using a generous cut-off point, and the prevalence of the mutation(s) tested. In our studies, we found that over 98% of CF infants without meconium ileus had an IRT in the top 1% of values during the first week of life[46], so that less than 2% would be missed by the test strategy at that point. In theory, we expect a further 4–6% to have a negative DNA test but a proportion of these would be high risk babies. The sensitivity in our population is therefore likely to be 94–95%. This is similar to worldwide figures quoted for the IRT-alone approach[29,30] although many of the programmes included would not have had full ascertainment of missed cases. While using the IRT-alone approach we found that, overall, 92% of babies were diagnosed by two to three months of age but, amongst the unexpected cases (i.e. not high risk), 11% had a negative screening result[46]. This percentage varied over the years with different IRT methods, and some programmes have reported much better results[27] possibly for the same reason. For the IRT/DNA protocol it would be possible to select an even more generous IRT cut-off and in Wisconsin the top 4% has been recommended[47]. However, the specificity and positive predictive value would consequently be reduced and the additional cost of DNA analysis would not be inconsiderable.

The specificity of the two approaches is quite different, however, being much greater with the IRT/DNA strategy. The positive predictive value, i.e. the percentage of babies with a positive screening test who actually have CF, is 100% in those cases where DNA analysis reveals homozygosity, or compound heterozygosity, for a CF mutation, but only about 11% where only one copy of a mutation is found. Nevertheless, the combined positive predictive value was 34% in our study, compared with 5% after the first test in the IRT-alone protocol. The great advantage of the IRT/DNA approach is that the number of infants receiving a false positive screening result is dramatically reduced. The drawback, however, is those infants identified by the screening test as CF carriers and this unwanted carrier detection raises a new set of problems. The overall numbers are small though with about 0.06% or 1:1600 families affected by this. These families need genetic counselling.

EXCESS CARRIER DETECTION

One of the consistent observations from our IRT/DNA screening programme and others[44–48,79] is the higher than expected number of carriers amongst infants with increased IRT values (these are infants with an IRT in the top 1% of values, who carry one identifiable mutation, but are sweat-test negative). Our combined data indicates a carrier frequency of 1 in 17 compared to the expected frequency in our population of 1 in 32 ($\chi^2 = 72.8$; $p < 0.001$). Several reasons have been proposed to account for this finding.

A small number of the apparent carriers may actually have mild CF with a rare mutation in addition to the one detected. In support of this, normal or equivocal sweat electrolytes have been reported in ΔF508/R117H and ΔF508/3849 + 10kbC→T compound heterozygotes[80–82] and also in G551S homozygotes[83].

Alternatively, and perhaps more likely, CF heterozygotes may have altered pancreatic secretions causing some ductule blockage with consequent elevation in IRT. Nørgaard-Pedersen et al.[48] have reported slightly but significantly higher IRT in ΔF508 heterozygotes than in control newborns (30% increase in median IRT values), and Figarella and Carrère[19] have shown that this difference is also present in adults and persists during secretin stimulation testing. Another recent report has also confirmed that the distribution of ΔF508 heterozygotes amongst the IRT population on newborn screening is shifted to higher IRT values (G. Travert, personal communication). This evidence strongly supports an association between ΔF508 heterozygosity and altered pancreatic physiology and the basis of this change deserves further investigation.

THE SWEAT TEST IS NOT THE GOLD STANDARD

Since the introduction of neonatal screening for CF in SA, the total number of sweat tests performed by the SA laboratory has declined by approximately 40% (Figure 4) and there was a similar finding in NSW[34]. It must be remembered though that the number of sweat tests generated as a result of the screening programme still remains small in comparison with those requested on clinical grounds.

It is general practice to report a sweat chloride or sodium greater than 60 mmol/l[84] as being consistent with cystic fibrosis and any borderline results between 40 and 60 mmol/l as indeterminate. These borderline patients have usually required further clinical evaluation to confirm or exclude a diagnosis of CF. The sweat electrolyte results from 205 infants recalled as part of our IRT/DNA protocol are shown in Figure 5 along with their genotypes. All sweat tests were performed at three to six weeks of age. The majority of infants with CF have sweat electrolyte concentrations well above 60 mmol/l and the majority

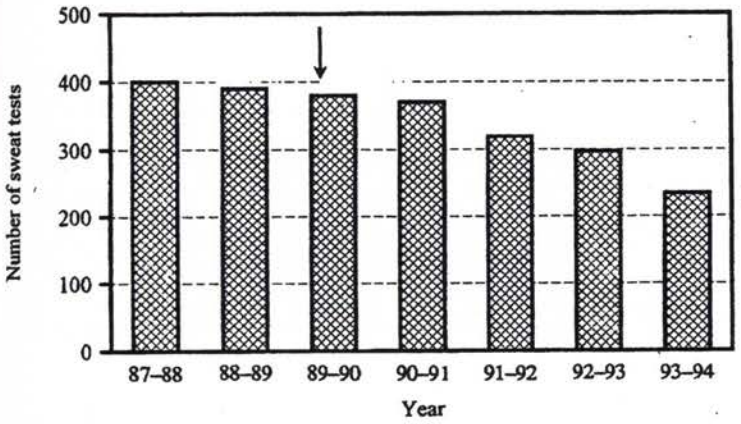


Figure 4 The number of sweat tests performed yearly by the Women's and Children's Hospital, Adelaide, South Australia. The arrow (↓) indicates when neonatal screening for cystic fibrosis was introduced

of apparent heterozygotes are grouped below 35 mmol/l. We have, however, identified several infants who repeatedly have sweat electrolytes in the borderline range and in whom the sweat chloride to sodium ratio is reversed and greater than one. In this group, three infants were ultimately found to be ΔF508/R117H compound heterozygotes with pancreatic sufficient CF and are currently on treatment[80]. Two other unrelated infants are ΔF508 heterozygotes and have recurrent and persistent chest infections without fat malabsorption. In this group, three infants were ultimately found to be ΔF508/R117H, and one infant as ΔF508/R117C, compound heterozygotes, with pancreatic sufficient CF and are currently on treatment. One ΔF508 heterozygote infant with recurrent and persistent chest infections had an abnormal pancreatic stimulation test. In two of these families, older siblings were incidentally found with similar genotypes and borderline sweat electrolytes.

Sweat testing has always identified a group of children with borderline results independent of age-related changes in sweat electrolytes. In the IRT/DNA screening protocol, however, those infants with borderline sweat electrolytes also have an increased IRT and carry one identifiable CF mutation. It is important therefore that sweat testing is performed at designated centres using an appropriate technique, e.g. pilocarpine iontophoresis[85], which is capable of separating those affected from simple carriers. We now routinely repeat the sweat test in any infant when the sweat chloride is between 35 and 60 mmol/l. If the result remains persistently abnormal, and this is invariably accompanied by a reversal of the sweat chloride/sodium ratio, this infant is referred for clinical assessment and investigation, and additional mutation testing is performed. Until the outcome for these children with borderline sweat electrolytes is determined it is important to continue to monitor their progress.

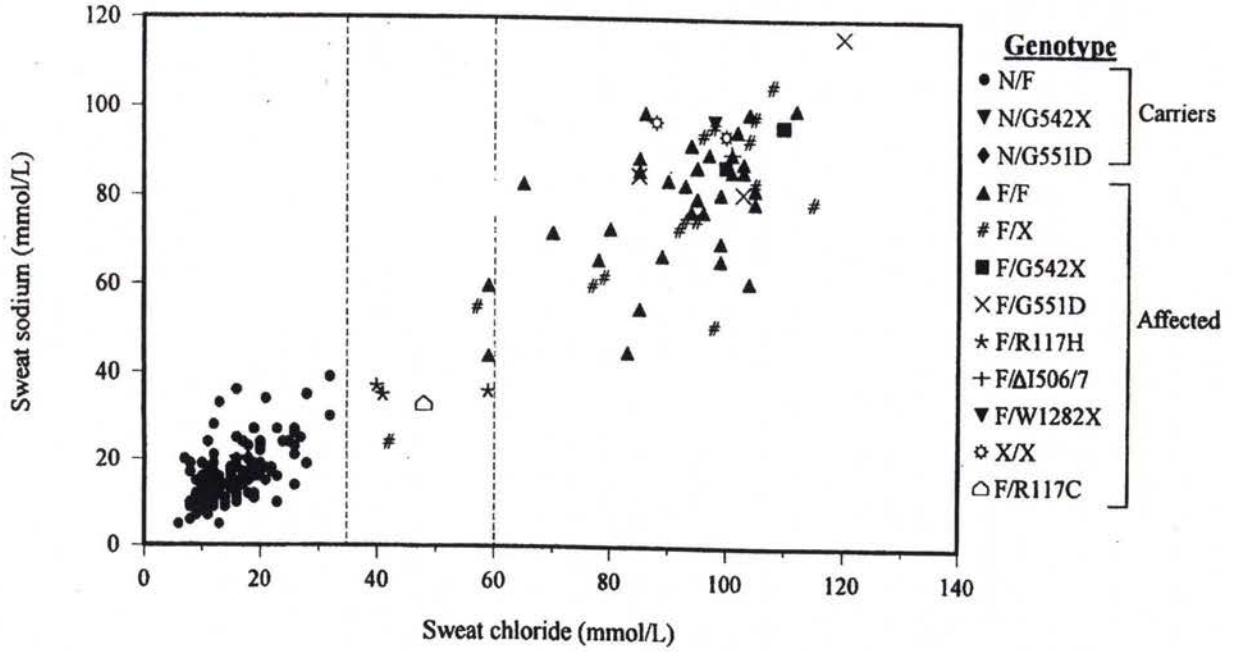


Figure 5 Combined South Australian (SA) and New South Wales (NSW) sweat electrolytes from infants recalled by the IRT/DNA screening protocol. The majority of the SA sweat tests were performed in one centre and for NSW, in four designated centres. Sweat testing was by pilocarpine iontophoresis[85] at three to six weeks of age. The figure contains sweat test results from 60 infants affected by cystic fibrosis and 145 infants who are carriers of a cystic fibrosis mutation. F = ΔF508; X = unidentified mutation; N = presumed normal allele

THE NEED FOR GENETIC COUNSELLING

The false-positive cases in the IRT/DNA screening protocol are carriers and this carrier detection is a disadvantage. Although it is mentioned in the screening information leaflets given to parents, most families are unaware of the possibility of carrier detection. The greatest challenge facing a screening laboratory when introducing an IRT/DNA screening protocol is the need to provide genetic counselling for the families identified with an infant who is a CF carrier. This requires development of appropriate counselling services and a decision as to how best to provide for carrier testing of additional family members.

At some stage during the recall process, either at the time of the sweat test, or soon after, the parents of carrier infants are counselled about the implications of being a CF carrier and what this means for their family. It is often only possible to provide counselling at the time of the sweat test if the family has had to travel some distance to the sweat test centre. Following the initial contact, the family is then provided with a written record of their genetic results as well as a leaflet reiterating the information given at counselling. We encourage the parents to distribute and discuss the information leaflets with other family members and offer to provide additional counselling and CF carrier testing for these people. From our own experience, we have found approximately 63% of parents request carrier testing for themselves, and there is on average an additional 2.7 requests for carrier testing per family. The testing has so far been reactive in responding to family requests but does aim to provide carrier testing of first-degree relatives. This could, however, act as an entry point for other programmes wishing to provide cascade carrier testing[72].

The proposed benefits of detecting CF carriers include alerting the parents to their risks for subsequent pregnancies, providing carrier testing for other family members of reproductive age and providing the same options for the carrier child and its siblings in future years. However, because of possible adverse effects of carrier detection, the counselling should be performed by appropriately qualified personnel at designated counselling centres. There must be a commitment to providing uniform counselling and written information in an attempt to minimize any possible harm as a result of the intervention.

Gregg et al.[86] have implemented a one-year follow-up questionnaire which they send to parents to assess their understanding of the information given during counselling and to determine the incidence of any problems. The information gained from this follow-up is still forthcoming and should indicate if there are significant adverse psychosocial consequences of unwanted carrier detection. We also plan to assess the impact this information has had on our own carrier families and recognize the need to continue this prospectively until it is clear that any adverse effects can be minimized and the counselling process modified appropriately.

Although all screening programmes that incorporate mutation analysis have the added problem of unwanted carrier detection, it may be possible to overcome this and dramatically reduce the number of sweat tests by only sweat testing apparent carriers who still have elevated IRT values after a second blood-spot sample. The NSW programme is attempting to obtain data to resolve this question. This does, however, raise the question of whether it is ethical to withhold any information on carrier status from families if that information is available.

COST

Analysis of costing data reported from several screening centres indicates that the relative costs of an IRT/DNA approach are not prohibitive and that economic considerations are not likely to drive the decision to screen for CF. Gregg et al.[86] reported that the cost of detecting each infant with CF (US$10 187 versus US$10 744) was similar regardless of whether an IRT-alone or IRT/DNA protocol was used. Their analysis did include the additional costs of genetic counselling and sweat tests as well as the laboratory reagent costs. In the NSW programme in 1994, the total within-laboratory cost of adding CF screening to an existing programme was US$1.00 per infant, with the component attributed to DNA analysis being US$0.07 per infant or US$7.20 for each infant that had mutation analysis. The cost for each infant with CF detected by this programme was US$3715[46]. This is similar to SA, which found that to include CF screening into their existing screening programme cost US$1.30 per infant, when assessing reagent costs only[44,45]. This equates to US$4288 for each infant with CF detected. A significant component is the additional time required for genetic counselling and in SA this averages 1.5–2 hours for each family. However, the cost of genetic counselling may be offset by a reduction in the total number of sweat tests performed (Figure 4). It must also be remembered that for each screening programme, the cost of detecting each infant with CF will depend greatly on the actual IRT cut-off used.

Any cost benefit associated with early identification of infants with CF has not yet been fully estimated[87]. The NSW group did identify a likely cost-saving in the short term, when they compared the costs of testing and treatment over six months versus the averted costs of hospitalization[33].

CONCLUSIONS

The experience to date indicates the IRT/DNA approach to neonatal screening for CF is at least as effective as existing IRT-alone screening programmes and

does not add significantly to the cost of such programmes. In populations where the ΔF508 frequency is high, the IRT/DNA approach has proven attractive because of its superior positive predictive value, high sensitivity, specificity and significantly reduced recall rate. However, the advantages of the IRT/DNA protocol need to be balanced against the problem of unwanted detection of CF carriers and the consequent need for genetic counselling. The difficulties associated with the genetic information generated by this protocol can be overcome, though, if a standardized programme of counselling and monitoring recalled families is followed. The genetic information gained is also helping to redefine the spectrum of clinical disease in those with CF and may prove important in the future when developing suitable treatment regimens.

If neonatal screening for CF is widely adopted, especially by those centres with populations of predominantly northern European origin, then the IRT/DNA protocol will prove most useful. The future development of simplified or automated amplification and mutation detection systems will also speed the introduction of DNA testing into existing CF screening programmes, especially those that currently do not have the expertise or resources to undertake DNA analysis. In addition, such developments will allow the analysis of very large numbers of blood-spot samples and raises the prospect of developing DNA-based neonatal screening tests for other common genetic disorders. The experience gained in screening for CF by DNA analysis will have direct application to this area.

ACKNOWLEDGEMENTS

We thank G. Travert, Centre Hospitalier Universitaire de Caen, for providing unpublished data and E. Edkins, Princess Margaret Hospital for Children, Perth, for analysis of the less common CF mutations. We are also grateful to R. L. Gerace, R. G. Ryall, P. V. Nelson, W. F. Carey and E. F. Robertson, the Department of Chemical Pathology, Women's and Children's Hospital, SA, and to V. Wiley, G. Sherry and U. Bayliss from the NSW Newborn Screening Programme, for their scientific, clinical and administrative help.

REFERENCES

1 Rosenstein BJ. Diagnostic delay in cystic fibrosis. *JAMA* 1982; **247**: 3313.
2 Farrell PM, Mischler EH. Neonatal screening for cystic fibrosis. *Adv Pediatr* 1992; **39**: 35–70.
3 Wilcken B, Towns SJ, Mellis CM. Diagnostic delay in cystic fibrosis: lessons from newborn screening. *Arch Dis Child* 1983; **58**: 863–6.
4 Wiser WC, Beier FR. Albumin in the meconium of infants with cystic fibrosis. A preliminary report. *Pediatrics* 1964; **33**: 115.
5 Stephen U, Busch EW, Kollberg H, Hellsing K. Cystic fibrosis detection by means of a test strip. *Pediatrics* 1975; **55**: 35–8.

6 Holsclaw DS, Keith HH, Palmer J. Meconium screening for cystic fibrosis. *Pediatr Ann* 1978; **7**: 15–20.
7 Shwachman H, Autonowicz I, Mahmoodian A, Ishida S. Studies in meconium: an approach to screening tests to detect cystic fibrosis. *Am J Dis Child* 1978; **132**: 1112–14.
8 McHill G. Experience with the nation-wide early detection of cystic fibrosis. In: *Neonatal Screening* (eds H Naurse, M Irie), Excerpta Medica, Amsterdam, 1983, pp 350–1.
9 Crossley JR, Elliott RB, Smith PA. Dried blood spot screening for cystic fibrosis in the newborn. *Lancet* 1979; **i**: 472–4.
10 Gibson LE. Screening of newborns for cystic fibrosis. *Am J Dis Child* 1980; **134**: 925–6.
11 Dodge JA, Ryley HC. Screening for cystic fibrosis. *Arch Dis Child* 1982; **57**: 774–80.
12 Farrell PM. Early diagnosis of cystic fibrosis: to screen or not to screen—an important question. *Pediatrics* 1984; **73**: 115–17.
13 Phelan PD. Screening for cystic fibrosis. In: *Recent Advances in Paediatrics* (ed R Meadow), Churchill Livingstone, Edinburgh, 1984, pp 103–20.
14 World Health Organization. Technical report services no. 497. Genetic disorders: preventions, treatment and rehabilitation. Geneva, 1972.
15 Cadman D, Chambers L, Feldman W, et al. Assessing the effectiveness of community screening programs. *JAMA* 1984; **251**: 1580–5.
16 Therrell BL, Panny SR, Davidson A, et al. US Newborn screening system guidelines: statement of the council of regional networks for genetic services. *Screening* 1992; **1**: 135–47.
17 Ad Hoc Committee Task Force on Neonatal Screening, Cystic Fibrosis Foundation Neonatal screening for cystic fibrosis: Position paper. *Pediatrics* 1983; **72**: 41–5.
18 Crossley JR, Smith PA, Edgar BW, et al. Neonatal screening for cystic fibrosis using immunoreactive trypsin assay in dried blood spots. *Clin Chim Acta* 1981; **113**: 111–21.
19 Figarella C, Carrère J. The evolution of pancreatic disease in cystic fibrosis. In: *Cystic Fibrosis—Current Topics*, Vol 2 (eds JA Dodge, DJH Brock, JH Widdicombe), John Wiley, Chichester, 1994, pp 255–75.
20 Adriaenssens K, Janssens H, van Soom H. Two-tier screen for cystic fibrosis. *Lancet* 1981; **ii**: 883.
21 Barbier Y, Gilly R, Chazalette JP. Neonatal screening for cystic fibrosis. *Monogr Paediatr* 1981; **14**: 177–81.
22 Wilcken B, Brown ARD, Urwin R, Brown DA. Cystic fibrosis screening by dried blood spot trypsin assay: results in 75 000 newborn infants. *J Pediatr* 1983; **102**: 383–7.
23 Travert G, Duhamel JF. Despistage neonatal systematique de la mucoviscidose par dosage de la trypsine immunoreactive sanguine: bilan de 80,000 tests. *Arch Fr Pediatr* 1983; **40**: 95–8.
24 Hammond KB, Abman SH, Sokol RJ, Accurso FJ. Efficacy of state wide screening for cystic fibrosis by assay of trypsinogen concentrations. *N Engl J Med* 1991; **325**: 769–74.
25 Pederzini F, Armani P. Barbato A, et al. Newborn screening for cystic fibrosis: two methods compared in 229,626 newborns. *Ital J Pediatr* 1983; **9**: 445–7.
26 Cassio A, Bernardi F, Piazzi S, et al. Neonatal screening for cystic fibrosis by dried blood spot trypsin assay: results in 47 127 newborn infants from a homogenous population. *Acta Paediatr Scand* 1984; **73**: 554–8.

27 Heeley AF, Heeley ME, King DN, et al. Screening for cystic fibrosis by dried blood spot trypsin assay. *Arch Dis Child* 1982; **57**: 18–21.
28 Wesley AW, Smith PA, Elliott RB. Experience with neonatal screening for cystic fibrosis in New Zealand using measurement of immunoreactive trypsinogen. *Aust Paediatr J* 1989; **25**: 151–5.
29 Travert G. Analyse de l'experience mondiale se despistage neonatal de la mucovisidose par dosage de la trypsine immunoreactive sanguine. In: *Mucovišidose; Despistage Neonatal et Prise en Charge Precose Caen*; CHRU de Caen 1988; 1–23.
30 Wilcken B. Newborn screening for cystic fibrosis: Its evolution and a review of the current situation. *Screening* 1993; **2**: 43–62.
31 Rock MJ, Mischler EH, Farrell PM, et al. Immunoreactive trypsinogen screening for cystic fibrosis: Characterisation of infants with false-positive screening test. *Pediatr Pulmonol* 1989; **6**: 42–8.
32 Tlucek A, Mischler EH, Farrell PM, et al. Parent's knowledge of neonatal screening and response to false-positive cystic fibrosis testing. *Dev Behav Pediatr* 1992; **13**: 181–6.
33 Al-Jader LN, Goodchild MC, Ryley HC, Harper PS. Attitudes of parents of cystic fibrosis children towards neonatal screening and antenatal diagnosis. *Clin Genet* 1990; **38**: 460–5.
34 Wilcken B. An evaluation of screening for cystic fibrosis. In: *Genetics and Epithelial Cell Dysfunction in Cystic Fibrosis* (eds M Buchwald, J Riodan), Alan P Liss, New York, 1987, pp 201–15.
35 Rock MJ, Mischler EH, Farrell PM, et al. Newborn screening for cystic fibrosis is complicated by age-related decline in immunoreactive trypsinogen levels. *Pediatrics* 1990; **85**: 1001–7.
36 Wilcken B, Brown A. An analysis of false negative screening tests for cystic fibrosis. In: *Current trends in infant screening*. (eds BJ Schmidt, AJ Diament, NS Login-Grossa) Proceedings of the 7th International Screening Symposium, Sao Paulo, Brazil, 6–9 November 1988: Amsterdam: Excerpta Medica, 1989; p 19 (abstract).
37 Duhamel JF, Travers G, Delmas P, et al. Special features of the time related evolution of IRT blood levels in six out of seven newborns with cystic fibrosis complicated by meconium ileus. In: *Cystic Fibrosis: Horizons* (ed D Lawson), John Wiley, Chichester, 1984, p 209 (abstract).
38 Rusakow LS, Abman SH, Sokol RJ, et al. Immunoreactive trypsinogen levels in infants with cystic fibrosis complicated by meconium ileus. *Screening* 1993; **2**: 13–18.
39 Laroche D, Peres O, Baird ML, et al. A new strategy of neonatal screening for cystic fibrosis. The association of immunoreactive trypsin and molecular biology. *Arch Fr Pediatr* 1990; **47**: 251–3.
40 Laroche D, Travert G. The application of PCR amplification and the polymorphic marker KM.19 to dried blood spots: comparison with deletion 508 for confirmation of the neonatal screening test for cystic fibrosis. *Pediatr Pulmonol* 1991; Suppl 7: 19–22.
41 Riodan JR, Rommens JM, Kerem BS, et al. Identification of the cystic fibrosis gene: cloning and characterization of complementary DNA. *Science* 1989; **245**: 1066–73.
42 Rommens JM, Iannuzzi MC, Kerem BS, et al. Identification of the cystic fibrosis transmembrane conductance regulator (CFTR) gene: chromosome walking and jumping. *Science* 1989; **245**: 1059–65.

43 Bowling FG, McGill JJ, Shepherd RW, Danks DM. Screening for cystic fibrosis: use of ΔF508 mutation. *Lancet* 1990; **335**: 925–6.
44 Ranieri E, Ryall RG, Morris CP, et al. Neonatal screening strategy for cystic fibrosis using immunoreactive trypsinogen and direct gene analysis. *Br Med J* 1991; **302**: 1237–40.
45 Ranieri E, Lewis BD, Gerace RL, et al. Neonatal screening for cystic fibrosis using immunoreactive trypsinogen and direct gene analysis: four years' experience. *Br Med J* 1994; **308**: 1469–72.
46 Wilcken B, Wiley V, Sherry G, Bayliss U. Neonatal screening for cystic fibrosis: a comparison of two strategies for case detection in 1.2 million babies. *J Pediatr* 1995; **127**: 965–70.
47 Farrell PM, Aronson RA, Hoffman MS, Laessig RH. Newborn screening for cystic fibrosis in Wisconsin: First application of population-based molecular genetic testing. *Wis Med J* 1994; **93**: 415–21.
48 Nørgaard-Pedersen B, Høgdall E, Iitia A, et al. Immunoreactive trypsin and a comparison of two ΔF508 mutation analyses in newborn screening for cystic fibrosis: An anonymous pilot study in Denmark. *Screening* 1993; **2**: 1–11.
49 Spence WC, Paulus-Thomas J, Orenstein DM, Naylor EW.Neonatal screening for cystic fibrosis: addition of molecular diagnosis to increase specificity. *Biochem Med Metab Biol* 1993; **49**: 200–11.
50 Larsen J, Campbell S, Faragher EB, et al. Cystic fibrosis screening in neonates—measurement of immunoreactive trypsin and direct genotype analysis for ΔF508 mutation. *Eur J Pediatr* 1994; **153**: 569–73.
51 The Cystic Fibrosis Genetic Analysis Consortium. Population variation of common cystic fibrosis mutations. *Hum Mutat* 1994; **4**: 167–77.
52 Pederzini F, Faraguna D, Giglio L, et al. Development of a screening system for cystic fibrosis: Meconium or blood spot trypsin assay or both? *Acta Paediatr Scand* 1990; **79**: 935–42.
53 Wilcken B, Chalmers G. Reduced morbidity in patients with cystic fibrosis detected by neonatal screening. *Lancet* 1985; **ii**: 1319–21.
54 Bowling F, Cleghorn G, Chester A, et al. Neonatal screening for cystic fibrosis. *Arch Dis Child* 1988; **63**: 196–8.
55 Dankert-Roelse JE, te Meerman GJ, Martijn A, et al. Screening for cystic fibrosis: A comparative study. *Acta Paediatr Scand* 1987; **76**: 209–14.
56 Dankert-Roelse JE, te Meerman GJ, Martijn A, et al. Survival and clinical outcome in patients with cystic fibrosis, with or without neonatal screening. *J Pediatr* 1989; **114**: 362–7.
57 Mastella G, Barlocco EG, Antonacci B, et al. Is neonatal screening for cystic fibrosis advantageous? The answer of a wide 15 year follow-up study. In: *Mucoviscidose Despistage Neonatal et Prise en Charge Prescoce Caen*, CHRU de Caen, 1988, pp 127–43.
58 Ryley HC, Deam SM, Williams J, et al. Neonatal screening for cystic fibrosis in Wales and the West Midlands: 1. Evaluation of immunoreactive trypsin test. *J Clin Pathol* 1988; **41**: 726–9.
59 Chatfield S, Owen G, Ryley HC, et al. Neonatal screening for cystic fibrosis in Wales and the West Midlands: clinical assessment after five years of screening. *Arch Dis Child* 1991; **66**: 29–33.
60 Heeley AF, Bangert SK. The neonatal detection of cystic fibrosis by measurement of immunoreactive trypsin in blood. *Ann Clin Biochem* 1992; **29**: 361–76.
61 Dhondt JL, Farriaux JP. What do immunoreactive trypsin assays measure? *Screening* 1994; **3**: 33–8.

62 Kirby LT, Applegarth DA, Davidson AGF, et al. Use of dried blood spot in immunoreactive trypsin assay for detection of cystic fibrosis in infants. *Clin Chem* 1981; **27**: 678–80.
63 Mallory PA, Travis J. Human pancreatic enzymes. *Biochemistry* 1973; **12**: 2847–51.
64 Itkonen O, Koivunen E, Hurme M, et al. Time-resolved immunofluorometric assays for trypsinogen-1 and 2 in serum reveal preferential elevation of trypsinogen-2 in pancreatitis. *J Lab Clin Med* 1990; **115**: 712–17.
65 Bowling FG, Watson ARA, Rylatt DB, et al. Monoclonal antibody-based enzyme immunoassay for trypsinogen in neonatal screening for cystic fibrosis. *Lancet* 1987; i: 826–7.
66 Petterson K, Soderholm J, Suonpaa M. Optimisation of an immunofluorometric assay based on the DELFIA® technique for the measurement of immunoreactive trypsin in dry blood spots. In: Proceedings of the IVth International Conference on Newborn screening for cystic fibrosis. *Pediatr Pulmonol* 1991; **7**: 90 (abstract).
67 Cabrini G, Pederzini F, Perobelli L, Mastella G. An evaluation of an enzyme-immunoassay method for immunoreactive trypsin in dried blood spots. *Clin Biochem* 1990; **23**: 1–7.
68 Dhondt JL, Farriaux JP, Paux E. Comparison of three kits for neonatal screening of cystic fibrosis. In: Proceedings of the IVth International Conference on Newborn screening for cystic fibrosis. *Pediatr Pulmonol* 1991; Suppl 7: 89 (abstract).
69 Heeley ME, Travert G, Ferre C, Lemonnier F. The international quality assurance program (IRTIQAS) for the assay of immunoreactive trypsin in dried blood spots. *Pediatr Pulmonol* 1991; Suppl 7: 72–5.
70 Ryall RG, Gjerde EM, Gerace RL, Ranieri E. Modifying an enzyme immunoassay of immunoreactive trypsinogen to use time resolved fluorescence. *Clin Chem* 1993; **39**: 224–8.
71 Hassemer DJ, Laessig RH, Hoffman GL, Farrell PM. Laboratory QC issues related to screening newborns for cystic fibrosis using immunoreactive trypsin. *Pediatr Pulmonol* 1991; Suppl 7: 76–83.
72 Super M, Schwarz MJ, Malone G, et al. Active cascade testing for carriers of cystic fibrosis gene. *Br Med J* 1994; **308**: 1462–8.
73 Nelson PV, Carey WF, Morris CP. Gene amplification directly from Guthrie blood spots. *Lancet* 1990; **336**: 1451–2.
74 Schwartz EI, Khalchitsky SE, Eisensmith RC, Woo SLC. Polymerase chain reaction amplification from dried blood spots on Guthrie cards. *Lancet* 1990; **336**: 639–40.
75 Makowski GS, Aslanzadeh J, Hopfer SM. *In situ* PCR amplification of Guthrie card DNA to detect cystic fibrosis mutations. *Clin Chem* 1995; **41**: 477–9.
76 Hatcher SLS, Lambert QT, Teplitz RL, Carlson JR. Heteroduplex formation: a potential source of genotyping error from PCR products. *Prenatal Diag* 1993; **13**: 171–7.
77 Webster D, Dhondt J-L, Hannon W, Torresani T. Description of quality assurance programs in newborn screening. *Screening* 1993; **2**: 219–28.
78 Green MR, Weaver LT. Early and late outcome of cystic fibrosis screening. *J R Soc Med* 1994; Suppl 21, **87**: 5–10.
79 Laroche D, Travert G. Abnormal frequency of ΔF508 mutation in neonatal transitory hypertrypsinaemia. *Lancet* 1991; **337**: 55.
80 Fitzgerald D, Van Asperen P, Henry R, et al. Delayed diagnosis of cystic fibrosis in children with a rare genotype (ΔF508/R117H). *J Paediatr Child Health* 1995; **31**: 165–71.

81 Stewart B, Zabner J, Shuber AP, et al. Normal sweat chloride values do not exclude the diagnosis of cystic fibrosis. *Am J Respir Crit Care Med* 1995; **151**: 899–903.
82 Highsmith WE, Burch LH, Zhou Z, et al. A novel mutation in the cystic fibrosis gene in a patient with pulmonary disease but normal sweat chloride concentrations. *N Engl J Med* 1994; **331**: 974–80.
83 Strong TV, Smith LS, Turpin S, et al. Cystic fibrosis gene mutation in two sisters with mild disease and normal sweat electrolyte levels. *N Engl J Med* 1991; **325**: 1630–4.
84 Behrman RE, Vaughan VC Nelson (eds). *Nelson Textbook of Pediatrics*, edn 13, WB Saunders, 1987, p 929.
85 Gibson LE, Cooke RE. A test for concentration of electrolytes in sweat in cystic fibrosis of the pancreas utilising pilocarpine by iontophoresis. *Pediatrics* 1959; **23**: 545–9.
86 Gregg RG, Wilfond BS, Farrell PM, et al. Application of DNA analysis in a population-screening program for neonatal diagnosis of cystic fibrosis (CF): comparison of screening protocols. *Am J Hum Genet* 1993; **52**: 616–26.
87 Dauphinais R. Cost analysis for screening newborns for cystic fibrosis. *J Clin Immunoassay* 1992; **15**: 121–5.

Pancreatic function and extended mutation analysis in ΔF508 heterozygous infants with an elevated immunoreactive trypsinogen but normal sweat electrolyte levels

R. John Massie, PhD, FRACP, Bridget Wilcken, MBChB, FRACP, Peter Van Asperen, MD, FRACP, Stuart Dorney, MBBS, FRACP, Margie Gruca, MSc, Veronica Wiley, PhD, and Kevin Gaskin, MD, FRACP

Background: Newborn screening for cystic fibrosis (CF) with immunoreactive trypsinogen (IRT) and ΔF508 analysis followed by sweat testing misses some infants with CF and detects more ΔF508 carriers than expected. Some of the apparent ΔF508 carriers may be ΔF508 compound heterozygotes with normal sweat electrolyte levels.

Methods: Infants identified by newborn screening with an elevated IRT level, one ΔF508 allele, and a sweat chloride level <60 mmol/L underwent CF mutation analysis, pancreatic stimulation testing, and repeat IRT analysis followed by clinical review and repeat sweat test at 12 months.

Results: Over a 24-month period we identified 122 ΔF508 heterozygotes and recruited 57; 4 had borderline sweat chloride levels (40 to 60 mmol/L), 5 (8.8%, 95% CI 1.4, 16.2) had a second CF mutation (R117H), and 11 (20%, 95% CI 10, 30) had the intron 8 5T allele. Three had clinical CF at 12 months (initial sweat chloride levels: 53, 51, and 32 mmol/L). Pancreatic electrolyte secretion in the subjects with a borderline sweat chloride level was similar to that in patients with known CF.

Conclusion: The excess of ΔF508 heterozygotes detected by IRT/DNA screening is associated with the presence of a second mutation or the 5T allele in some infants. Screened infants with borderline sweat chloride levels almost certainly have CF, but long-term follow-up of the infants with the genotype ΔF508/R117H and ΔF508/5T is required to determine their outcome. In the meantime, newborn screening should be confined to severe mutations associated with classic CF. (J Pediatr 2000;137:214-20)

From the Department of Respiratory Medicine, the New South Wales Newborn Screening Program, the Department of Gastroenterology, and the James Fairfax Institute of Paediatric Nutrition, Royal Alexandra Hospital for Children, Sydney, Australia.

Supported by project grants from The Australian Cystic Fibrosis Foundation, The Royal Alexandra Hospital for Children Research Fund, and The Ronald Geoffrey Arnott Foundation. Dr Massie was supported by a National Health and Medical Research Council of Australia postgraduate medical research scholarship.

Submitted for publication May 6, 1999; revisions received Sept 16, 1999, and Feb 17, 2000; accepted Mar 6, 2000.

Reprint requests: R. John Massie, PhD, FRACP, Department of Respiratory Medicine, Royal Children's Hospital, Flemington Rd, Parkville, VIC 3052, Australia.

0022-3476/2000/$12.00 + 0 **9/21/107162**

doi:10.1067/mpd.2000.107162

Most neonatal screening programs for cystic fibrosis involve measurement of neonatal blood-spot immunoreactive trypsinogen,[1] followed by gene mutation analysis for ΔF508 in those with an elevated IRT level.[2,3] Infants with an elevated IRT level and 2 ΔF508 alleles have CF. Those with one ΔF508 allele require a sweat test to distinguish true carriers from infants with CF who have a second but as yet unrecognized mutation (compound heterozygotes).

CAVD	Congenital absence of the vas deferens
CF	Cystic fibrosis
CFTR	Cystic fibrosis transmembrane conductance regulator protein
Cl	Chloride
IRT	Immunoreactive trypsinogen
NSW	New South Wales
PS	Pancreatic sufficient

We have previously reported the detection of a greater number of ΔF508 heterozygotes with an elevated neonatal IRT level, but normal sweat electrolyte levels, than would be expected from the population frequency of this gene mutation in New South Wales.[4] Other centers that screen for CF have reported similar findings.[5-7] The excess of ΔF508 carriers with an elevated neonatal IRT level may include some infants with a second mutation that less severely affects sweat duct function. Sweat test-negative CF has been described,[8-10] and specific mutations associated with normal sweat electrolyte levels include R117H,[11,12] A455E,[13] 3849+10kbC→T,[8,14] and the

intron 8 allele 5T.[15] Separating sweat test-negative patients with CF from true carriers is important for the early recognition of complications in those with the disease and for reassurance of carriers and their parents.

To find infants missed by screening and to understand why the detection of carriers is increased, we studied infants who were identified by a statewide neonatal screening program as ΔF508 heterozygotes but who were not identified as having CF because of normal sweat electrolyte levels.

METHODS

Newborn Screening for Cystic Fibrosis: 1995 to 1996

From January 1995 to December 1996, 186,949 infants were screened for CF in NSW. There were 44 ΔF508 homozygotes (3 were identified before birth) and 142 ΔF508 heterozygotes. Eighteen ΔF508 heterozygotes had CF with sweat chloride levels >60 mmol/L, 122 had sweat Cl levels <60 mmol/L, and 2 had parents who declined sweat testing. CF was diagnosed in 3 additional infants because of meconium ileus and sweat Cl levels >60 mmol/L, but they did not have a ΔF508 mutation. In total, 65 patients in the years 1995 to 1996 had CF diagnosed by ΔF508 genotyping and sweat electrolyte analysis, and 122 were considered carriers.

Subjects

All 122 ΔF508 heterozygote infants with sweat Cl levels <60 mmol/L were invited to join the study. Notification of the screening result was made at 4 to 6 weeks of age. The study was approved by the Ethics Committee of the Royal Alexandra Hospital for Children, Camperdown, NSW. Written informed consent was obtained before the study was begun. A complete history and physical examination were performed (J.M.), and if the subjects had not had a sweat test performed at our center, the sweat test was repeated. Sweat tests at our center were performed by Gibson and Cooke pilocarpine iontophoresis, with >100 mg of sweat collected in all cases. Measurement of Cl levels was done by colorimetry, and measurement of sodium levels was done by flame cytometry.[16]

Gene Mutation Analysis

Blood was taken and DNA extracted[17] for an extended cystic fibrosis transmembrane conductance regulator protein gene mutation analysis as described previously.[18] The following mutations were included: ΔF508, ΔI507, R117H, G551D, A455E, G542X, N1303K, W1282X, 1717-1G→A, R560T, R347P, R334W, R553X, R1162X, S549N, 3849+10C→T, and 621+1G→T. Newborn blood spots were used for the intron 8 polythymidine sequences, with DNA extracted by Chelex 100 (Bio-Rad)[19] and subjected to semi-nested polymerase chain reaction of CFTR intron 8 and exon 9.[20] The polythymidine sequence was determined by allele-specific oligonucleotide hybridization[21] (or direct dot blot) with enhanced chemiluminescence detection (Amersham Pharmacia Biotech).

Pancreatic Stimulation Testing

Pancreatic stimulation testing was performed by quantitative marker perfusion technique,[22,23] with measurement of bicarbonate and Cl secretion, water flow, and colipase activity.[24] Control and CF range values for pancreatic sufficient and pancreatic insufficient patients with CF were taken from studies that used the same technique of quantitative marker perfusion testing performed at our center from 1983 to 1994 (published and unpublished data).[25] Patients with established CF (sweat Cl levels >60 mmol/L) were studied to document pancreatic exocrine function. Children in the control group had pancreatic stimulation tests for gastrointestinal disease thought to be caused by pancreatic insufficiency, although all members of the control group had sweat Cl levels <40 mmol/L. From these previous studies, 39 (93%) of 42 patients with sweat test–confirmed CF had values of summed pancreatic HCO_3^- + Cl^- <0.7 mmol/kg/h, whereas 41 (98%) of 42 members of the control group had summed pancreatic HCO_3^- + Cl^- >0.7 mmol/kg/h. We have used this figure as an arbitrary cutoff for CF range pancreatic electrolyte secretion, and it is approximately 2 SDs below the mean of control values. The results of these studies were comparable to those of published studies that used the same methods.[22,23]

Repeat IRT and Follow-up

Repeat IRT measurement[26] was made at initial assessment (generally between 6 and 8 weeks of age). Parental genotyping (ΔF508 only, unless other mutations were identified in the infant) and genetic counseling were offered to all parents. Clinical follow-up and repeat sweat testing were performed at 12 months of age.

Statistical Analysis

Subjects were divided into groups based on genotype and sweat test results. The ΔI507 mutation was considered equivalent to ΔF508, and subjects with non-5T polythymidine alleles and no additional exonic mutation were grouped together. The frequency of mutations identified was compared by χ^2 test with the reported frequency of the mutation.[27] The subjects in the study, divided into genotype and sweat test groups, were compared with members of the control group and patients with known CF with the use of nonparametric analysis (Mann-Whitney). Results of repeat sweat tests were compared with results of initial tests by paired Student *t* test. All *P* values were 2-tailed, and significance was taken as $P < .05$.

RESULTS

ΔF508 Heterozygote Subject Characteristics

The 122 ΔF508/– infants with sweat Cl levels <60 mmol/L were eligible for

Table I. Results of the extended CF gene mutation analysis

Exonic mutations	No.	Male, female	Intron 8 polythymidine sequences
ΔF508/R117H	5	2, 3	9T/7T
ΔI507/–	2	1, 1	7T/7T
ΔI507/–	1	0, 1	7T/5T
ΔF508/–	10	4, 6	9T/5T
ΔF508/–	37	18, 19	9T/7T
ΔF508/–	2	1, 1	9T/9T
Total	57		

Table II. Pancreatic bicarbonate and chloride secretion, water flow, and colipase activity by subject group and with control group and known patients with CF

Genotype group	$HCO_3^- + Cl^-$ (mmol/kg/h)	H_2O (mL/kg/h)	Colipase activity (%)
Control group (n = 43)	1.24 (0.6-1.88)	9.02 (4.97-14.5)	94 (20-200)
ΔF508/R117H (n = 4)	0.71* (0.62-0.89)	7.0 (5.8-9.4)	53* (48-59)
ΔF508/5T (n = 10)	0.81* (0.67-0.97)	7.9 (6.5-9.0)	32* (24-49)
ΔF508/– (n = 34)	0.84* (0.68-0.96)	7.5 (5.8-8.9)	30* (21-42)
Borderline (n = 3)	0.44† (0.37-0.52)	6 (5.9-6.2)	33* (23-50)
Cystic fibrosis (PS, n = 22)	0.32 (0.22-0.45)	2.9 (1.9-5.2)	35 (8-68)
Cystic fibrosis (PI, n = 20)	0.18 (0.10-0.25)	2.2 (1.0-2.9)	0 (0-0)

Results are presented as median values with interquartile ranges presented in parentheses.
PS, Pancreatic sufficient; *PI*, pancreatic insufficient.
*$P < .01$ for comparison with both control group and pancreatic insufficient patients with CF.
†$P \geq .3$ for comparison with pancreatic sufficient patients with CF.

the study. Four had sweat Cl levels of 40 to 60 mmol/L (borderline), and 118 had sweat Cl levels <40 mmol/L. Of the 122, 3 had a distant family history of CF but unknown gene mutations. None had a family history of recurrent chest infections, pancreatitis, nasal polyps, chronic liver disease, or male infertility. Sweat Cl levels for the 122 eligible infants ranged from 9 to 53 mmol/L, with a mean of 16.8 ± 9.1 (SD) mmol/L (95% CI 15.1,18.5 mmol/L).

Fifty-seven subjects were enrolled in the study including all 4 with borderline sweat Cl levels; 51 underwent the complete protocol (extended mutation analysis, repeat IRT analysis, and pancreatic stimulation testing), and 6 had genotyping and repeat IRT analysis only. At initial assessment (6 to 8 weeks of age), none had delayed passage of meconium, steatorrhea, rectal prolapse, or prolonged neonatal jaundice. One infant with a borderline sweat Cl level had *Staphylococcus aureus* pneumonia at 6 weeks of age. The principal reason for 65 infants not being enrolled in the study was distance from our center (which was often >500 km). The IRT and sweat Cl levels of the 65 infants not enrolled were not different from those of the enrolled group, and none of the infants has subsequently presented to any of the CF clinics in NSW.

Subject Genotypes

Details of the gene mutation analysis and polythymidine tract alleles are presented in Table I. Five (8.8%) infants (95% CI 1.4, 16.2) had an additional exonic mutation, R117H. This is significantly higher than the reported frequency of R117H in the Australian CF population (1%, $P < .01$)[27] and the reported frequency of R117H in a community-based screen of a predominantly white population (0.6%, $P < .01$).[28] The intron 8 polythymidine sequences of the ΔF508/R117H infants were 9T/7T. Three other infants thought to be ΔF508 heterozygotes were discovered to be ΔI507 heterozygotes instead because of the similar physical properties of the 2 mutations in the polyacrylamide gel electrophoresis used for screening. Eleven (20%) infants (95% CI 10, 30) without an identified second exonic mutation had the intron 8 5T allele (10 ΔF508/–, *9T/5T* and 1 ΔI507/–, *7T/5T*), which is in contrast to 8% of the population in NSW ($P < .01$).[29] None of the infants with borderline sweat Cl levels had additional exonic mutations or the 5T allele.

We divided the subjects into groups on the basis of genotype and initial sweat Cl values: initial borderline (40 to 60 mmol/L) sweat Cl levels (borderline, n = 4), subjects with a second mutation (ΔF508/R117H, n = 5), subjects with the 5T allele (ΔF508/5T, n = 11), and subjects with no other mutation and initial sweat Cl level <40 mmol/L (ΔF508/–, n = 37).

Pancreatic Stimulation Tests

Fifty-one infants underwent pancreatic stimulation testing including 3 of 4 infants with an initial borderline sweat Cl level, 4 of 5 infants with the R117H mutation, 10 of 11 infants with the 5T allele, and 34 with no additional mutation and sweat Cl levels <40 mmol/L (Table II).

Subjects with initial borderline sweat Cl levels had pancreatic $HCO_3^- + Cl^-$ secretion within the CF range (all ≤0.7 mmol/kg/h), and no statistical difference was found between the median summed $HCO_3^- + Cl^-$ secretion between subjects with an initial borderline sweat Cl level and previously studied

PS patients with CF (P = .36). In those subjects with an initial sweat Cl level <40 mmol/L, including those with R117H or the 5T allele, the median pancreatic electrolyte secretion, water flow, and colipase activity were midway between values for members of the control group and pancreatic insufficient patients with CF (Figure). Three of the 4 R117H subjects and 4 of the 11 5T subjects who had pancreatic stimulation testing had pancreatic electrolyte secretion <0.7 mmol/kg/h (CF range). Eleven (35%) of the 31 (95% CI 18, 52) subjects with no additional gene mutation or the 5T allele and sweat Cl levels <40 mmol/L had HCO_3^- + Cl^- secretion <0.7 mmol/kg/h.

Repeat IRT Measurement

Repeat IRT measurements were obtained when the subjects were between 14 and 117 days of age. The median IRT level was 23 μg/L, and the range was 8 to 72 μg/L. All levels were <80 μg/L, which was the cutoff used at 6 to 8 weeks of age, on the 2-tiered IRT screen (used before the *CFTR* protein gene was discovered).[30]

Clinical Follow-up and Repeat Sweat Tests

The subjects were monitored for 12 months. Two of the patients with initial borderline sweat test results had clinical evidence of CF (chronic productive cough, deep pharyngeal suction aspirate positive for *Staphylococcus aureus*, and chest radiographic changes of hyperinflation and thickened bronchi), and both had sweat Cl levels >60 mmol/L. None of the R117H infants had symptoms, nor did those with the 5T allele. One infant, who had no additional mutation, had clinical evidence of CF (chronic productive cough and chest radiographic changes of hyperinflation and thickened bronchi), with the sweat Cl level increasing from an initial value of 32 mmol/L to 54 mmol/L at 12 months. He also had CF range pancreatic duct function on testing at 6 weeks. Three other infants had an initial sweat Cl level between 30 and 40 mmol/L, and repeat sweat Cl level was in the same range (n = 2) or lower (n = 1) at 12 months; all were well on clinical evaluation. The infants whose initial sweat Cl level was <30 mmol/L, irrespective of genotype, were also well on clinical evaluation, and the results of repeat sweat tests remained unchanged (mean initial sweat Cl level, 13.8 ± 4.9 mmol/L; mean follow-up sweat Cl level, 14.0 ± 5.8 mmol/L; paired Student t test, P = .38).

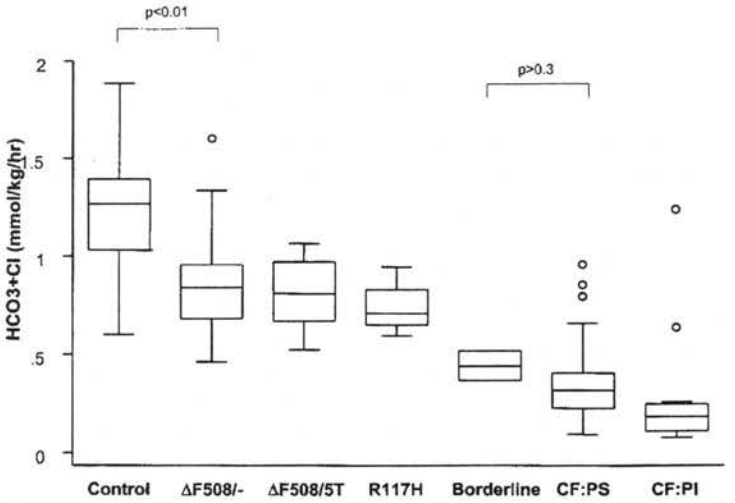


Figure. Boxplot comparison of pancreatic electrolyte secretion (summed HCO_3^- + Cl^-) for control group, infants with no additional mutation and initial sweat Cl^- levels <40 mmol/L (ΔF508/–), infants with 5T allele (ΔF508/5T), infants with second mutation (ΔF508/R117H), infants with borderline initial sweat test results (*Borderline*), patients with PS CF (*CF:PS*), and pancreatic insufficient patients with CF (*CF:PI*). Median of pancreatic summed HCO_3^- + Cl^- for each group is indicated by *horizontal line*, with values within interquartile range contained in box, and 95% CI whiskers. Outlying values are plotted as *open circles*.

DISCUSSION

We identified 1.6 (95% CI 1.2, 2.2) times the expected number of ΔF508 heterozygotes from the newborn screening program and found 8.8% of those studied to have a second mutation and 20% to have the 5T allele. Four other infants had initial borderline sweat Cl levels, of whom 2 had clinical CF, and a third, CF range pancreatic electrolyte secretion. The median pancreatic electrolyte secretion and exocrine function of the infants with no additional mutation was half that of members of the control group who did not have CF, and 35% had CF range pancreatic electrolytes, with one of these having CF confirmed at 12 months.

The results of the pancreatic stimulation test and the clinical follow-up suggest that ΔF508 heterozygote infants with borderline sweat electrolyte levels have CF. A sweat Cl level of 40 mmol/L is 6 SDs above that of the infant control group,[31] and consideration should be given to lowering the diagnostic sweat Cl cutoff value to 40 mmol/L in screened infants. This is supported by other groups who have directly measured transepithelial Cl transport from rectal suction biopsy specimens and have demonstrated reduced Cl transport through CFTR in patients with sweat Cl levels of 30 to 60 mmol/L.[32,33]

We found more ΔF508/R117H compound heterozygotes than expected. For patients with R117H, the length of the intron 8 polythymidine sequences determined the proportion of CFTR messenger RNA transcripts containing exon 9, which is critical in determining

the clinical outcome.[34] The ΔF508/R117H genotype may cause PS CF, usually when R117H is in *cis* with the 5T allele,[21] and congenital absence of the vas deferens when in *cis* with the 7T allele.[20,35] The ΔF508/R117H subjects in this study had the 9T/7T alleles and could be expected to have ~5% CFTR activity, an amount sufficient to prevent serious respiratory and pancreatic disease.[36] However, variable splicing is common,[15] and patients with R117H in *cis* with 7T have been described with PS CF,[21] recurrent pancreatitis,[37] and allergic bronchopulmonary aspergillosis.[38] Therefore it is possible that the R117H subjects from this study may have variants of CF and represent infants missed by screening.

We found more infants with the 5T allele than expected, suggesting an influence of 5T on neonatal pancreatic function. Other groups have reported similar findings in infants with an elevated neonatal IRT level.[39,40] ΔF508, in *trans* with 5T, has been reported in male subjects with congenital absence of the vas deferens and is associated with approximately 5% to 10% CFTR activity.[36] The 5T allele has also been found in individuals with CF, ranging from an atypical to typical phenotype; and presentations with nasal polyps, asthma, chronic sinusitis, and chronic bronchitis have been reported.[15,41] The predicted level of CFTR activity in subjects with the genotype ΔF508/5T is at the threshold of disease-producing levels, but no prospective information is available to define the risk of disease in these infants with no symptoms.

The subjects with R117H or 5T meet some of the criteria for a diagnosis of CF according to the criteria of the Cystic Fibrosis Foundation,[42] namely a positive newborn screening result (elevated IRT level), mutations in each *CFTR* gene known to cause CF, and for some R117H subjects (n = 3) and 5T subjects (n = 4), evidence of in vivo abnormalities of ion transport. These laboratory findings must be seen in the context of the clinical condition of the subjects, and to date, no associated clinical manifestations have been found in either group. Measurement of nasal potential differences is useful in cases of diagnostic uncertainty,[42] but we found the study too difficult to perform at this age. We are unable to assess the state of the vas deferens either on clinical evaluation or ultrasonography in the male infants with R117H (n = 2) or the 5T allele (n = 5) to provide evidence of disease expression. Extended follow-up will be required to determine whether clinical manifestations are associated with the gene mutations.

The infants with no additional mutations are likely to be true carriers; however, we found 35% to have low pancreatic electrolyte secretion. One of these has clinical CF and must have a second *CFTR* mutation. Whether CF will develop in the other infants with low pancreatic electrolyte levels is unknown. The 16-mutation analysis covers 97.6% of mutations in the Australian population and includes the mutations associated with normal sweat electrolyte levels. We have shown that the ΔF508 heterozygotes with neonatal hypertrypsinogenemia have reduced pancreatic function. Reduced pancreatic function in some carriers of *CFTR* mutations may be the explanation for the association with chronic panceatitis,[37,43,44] and there are reports of other disease manifestations in apparently true carriers, such as disseminated bronchiectasis[45] and allergic bronchopulmonary aspergillosis[38] and perhaps asthma.[46]

We conclude that the excess of apparent carriers from the IRT/ΔF508 screening program is partly explained by the presence of a second mutation (R117H) and the 5T allele in some infants and that pancreatic duct function is reduced in ΔF508 heterozygotes with neonatal hypertrypsinogenemia. Some of the infants missed by screening are ΔF508 heterozygotes with borderline sweat Cl levels, and changing the sweat test cutoff values for infants would improve the diagnosis of CF from newborn screening. Long-term follow-up is required to determine whether the infants with 2 mutations will have CF and represent infants missed by screening. In the meantime, newborn screening should be confined to severe mutations associated with classic CF.

We thank Prof Craig Mellis and Dr Gillian Nixon for reviewing the manuscript; Ursula Bayliss and Gina Sherry from the NSW Newborn Screening Laboratory for help with identification of subjects; Donna Waters for assistance with the pancreatic stimulation tests; Viive Howell from the NSW Newborn Screening Laboratory for performing the polythymidine tract sequences (1996); Dr Bill Carey from Royal Women and Children's Hospital, Adelaide, for performing the polythymidine tract sequences (1995); and Dr Ted Edkins, Dr Roger Martin, and Kathy Gray from the Molecular Genetics Department of Princess Margaret Hospital for Children, Perth, for performing the extended cystic fibrosis gene mutation analysis.

REFERENCES

1. Crossley JR, Elliot RB, Smith PA. Dried blood spot screening for cystic fibrosis in the newborn. Lancet 1979;1: 472-4.
2. Wilcken B. Newborn screening for cystic fibrosis: its evolution and a review of the current situation. Screening 1993;2:43-62.
3. Ranieri E, Lewis BD, Morris P, Wilcken B. Neonatal screening using combined biochemical and DNA-based techniques. Cystic fibrosis: current topics. Vol 3. In: Dodge JA, Brock DJM, Widdicombe JM, editors. New York: John Wiley & Sons Ltd; 1996. p. 182-206.
4. Wilcken B, Wiley V, Sherry G, Bayliss U. Neonatal screening for cystic fibrosis: a comparison of two strategies for case detection in 1.2 million babies. J Pediatr 1995;127:965-70.
5. Laroche D, Travert G. Abnormal frequency of DF508 mutation in neonatal transitory hypertrypsineamia. Lancet 1991;337:55.
6. Farrell PM, Aronson RA, Hoffman MS, Leassig RH. Newborn screening for cystic fibrosis in Wisconsin: first application of population based molecular genetic testing. Wis Med J 1994;93:415-21.
7. Spence WC, Paulus-Thomas J, Orenstein DM, Naylor EW. Neonatal

screening for cystic fibrosis: addition of molecular diagnostics to increase specificity. Biochem Med Metab Biol 1993;49:200-11.
8. Stewart B, Zabner J, Shuber AP, Welsh MJ, McCray PB. Normal sweat chloride values do not exclude the diagnosis of cystic fibrosis. Am J Respir Crit Care Med 1995;151:899-903.
9. Huff DS, Huang NN, Arey JB. Atypical cystic fibrosis of the pancreas with normal levels of sweat chloride and minimal pancreatic lesions. J Pediatr 1979;94:237-9.
10. Sarsfield JK, Davies JM. Negative sweat tests and cystic fibrosis. Arch Dis Child 1975;50:463-6.
11. Fitzgerald D, Van Asperen P, Henry R, Waters D, Freelander M, Wilson M, et al. Delayed diagnosis of cystic fibrosis in children with a rare genotype (DF508/R117H). J Pediatr Child Health 1995;31:168-71.
12. The Cystic Fibrosis Genotype-Phenotype Consortium. Correlation between genotype and phenotype in patients with cystic fibrosis. N Eng J Med 1993;329:1308-13.
13. Gan K-H, Veeze H, van den Ouweland AMW, Halley DJJ, Scheffer H, van der Hout A, et al. A cystic fibrosis mutation associated with mild lung disease. N Engl J Med 1995;333:95-9.
14. Highsmith WE, Burch LH, Zhou Z, Olsen JC, Boat T, Spock A, et al. A novel mutation in the cystic fibrosis gene in patients with pulmonary disease but normal sweat chloride concentrations. N Engl J Med 1994;331: 974-80.
15. Kerem E, Harel-Rave N, Augarten A, Madgar I, Nissim-Rafina M, Yahav Y, et al. A cystic fibrosis transmembrane conductance regulator splice variant with partial penetrance associated with variable cystic fibrosis presentations. Am J Respir Crit Care Med 1997;155: 1914-20.
16. Gibson LE, Cooke RE. A test for concentration of electrolytes in sweat in cystic fibrosis of the pancreas utilising pilocarpine by iontophoresis. Pediatrics 1959;23:545-9.
17. Wang L, Hirayasu K, Ishizawa M, Kobayashi Y. Purification of genomic DNA from human whole blood by isopropanol-fractionation with concentrated NaI and SDS. Nucleic Acids Research 1994;22:1774-5.
18. Goldblatt J, Creegan R, Edkins T, Landau L, Ryan G, Walpole I. Mutation analysis of Western Australian families affected by cystic fibrosis. Med J Aust 1995;162:12-4.
19. Walsh PS, Metser DA, Higuchi R. Chelex 100 as a medium for simple extraction of DNA for PCR-based typing from forensic material. Biotechniques 1991;10:506-11.
20. Dork T, Fislage R, Neumann T, Wulf B, Tummler. Exon 9 of the CFTR gene: splice site haplotypes and cystic fibrosis mutations. Hum Genet 1994; 93:67-73.
21. Kiesewetter S, Macek MJ, Davis C, Curristin S, Chu C-S, Graham C, et al. A mutation in CFTR produces different phenotypes depending on chromosomal background. Nature Genet 1993;5:274-8.
22. Gaskin KJ, Durie PR, Corey M, Wei P, Forstner GG. Evidence for a primary defect of pancreatic HCO_3 secretion in cystic fibrosis. Pediatric Res 1982; 16:554-7.
23. Kopelman H, Corey M, Gaskin K, Durie P, Weizman Z, Forstner G. Impaired chloride secretion, as well as bicarbonate secretion, underlies the fluid secretory defect in the cystic fibrosis pancreas. Gastroenterology 1988;95: 349-55.
24. Gaskin KJ, Durie PR, Hill RE, Lee LM, Forstner GG. Colipase and maximally activated pancreatic lipase in normal subjects and patients with steatorrhoea. J Clin Invest 1982;69: 427-34.
25. Waters DL, Dorney SFA, Gaskin KJ, Gruca M, O'Halloran M, Wilcken B. Pancreatic function in infants identified as having cystic fibrosis in a neonatal program. N Eng J Med 1990; 332:303-8.
26. Petterson K, Solderholm J, Suonpaa M. Organisation of an immunofluorometric assay based on the DEFLFIA technique for the measurement of immunoreactive trypsinogen in dry blood spots [abstract]. Pediatr Pulmonol 1991;(Suppl 7)90.
27. The Cystic Fibrosis Genetic Analysis Consortium. Human Mutation 1994;4: 167-77.
28. Witt DR, Schaefer C, Hallam P, Wi S, Blumberg B, Fishbach A, et al. Cystic fibrosis heterozygote screening in 5,161 pregnant women. Am J Hum Genet 1996;58:823-35.
29. Wiley V, Sherry G, Howell V, Wilcken B. The incidence of 5T mutation in the CFTR gene in the newborns in NSW. Proceedings of the Australian Society for Inborn Errors of Metabolism: Perth; 1997. p. 71.
30. Wilcken B, Brown ARD, Urwin R, Brown DA. Cystic fibrosis screening by dried blood spot trypsin assay: analysis of 75,000 newborn infants. J Pediatr 1983;102:383-7.
31. Farrell PM, Koscik RE. Sweat chloride concentrations in infants homozygous or heterozygous for F_{508} cystic fibrosis. Pediatrics 1996;97:524-8.
32. Veeze HJ, Sinaasappel M, Bijman J, Bouquet J, De Jonge HR. Ion transport abnormalities in rectal suction biopsies from children with cystic fibrosis. Gastroenterology 1991;101: 398-402.
33. Veeze H, Halley DJ, Bijman, de Jongst JC, Sinaasappel M. Determinants of mild clinical symptoms in cystic fibrosis patients. J Clin Invest 1994; 93:461-6.
34. Chu C-S, Trapnell BC, Curristan S. Genetic basis of variable exon 9 skipping in cystic fibrosis transmembrane conductance regulator mRNA. Nat Genet 1993;3:151-6.
35. Chillon M, Casals T, Mercia B, Bassas L, Lissens W, Silber S, et al. Mutations in the cystic fibrosis gene in patients with congenital absence of the vas deferens. N Engl J Med 1995;332: 1475-80.
36. Davis PB, Drumm M, Konstan M. State of the art: cystic fibrosis. Am J Respir Crit Care Med 1996;154: 1229-56.
37. Cohn JA, Freidman KJ, Noone PG, Knowles MR, Silverman LM, Jowell PS. Relation between mutations of the cystic fibrosis gene and idiopathic pancreatitis. N Engl J Med 1998;339: 653-8.
38. Miller PW, Hamosh A, Macek M, Greenberger PA, MacLean J, Walden SM, et al. Cystic fibrosis transmembrane conductance regulator (CFTR) gene in allergic bronchopulmonary aspergillosis. Am J Hum Genet 1996;59: 45-51.
39. Castellani C, Bonizzato A, Mastella G. CFTR mutations and IVS8-5T variant in newborns with hypertrypsinaemia and normal sweat test. J Med Genet 1997;34:297-301.
40. Chin S, Ranieri E, Gerace RL, Nelson PV, Carey WF. Frequency of intron 8 CFTR polythymidine sequence variant in neonatal blood specimens. Lancet 1997;350:1368-9.
41. Friedman KJ, Heim RA, Knowles MR, Silverman LM. Rapid characterisation of the variable length polythymidine tract in the cystic fibrosis (CFTR) gene: association of the 5T allele with

selected CFTR mutations and its incidence in atypical sinopulmonary disease. Hum Mutation 1997;10:108-15.
42. Rosenstein BJ, Cutting GR for the Cystic Fibrosis Foundation Consensus Panel. The diagnosis of cystic fibrosis: a consensus statement. J Pediatr 1998; 132:589-95.
43. Sharer N, Schwarz, Malone G, Howarth A, Painter J, Super M, et al. Mutations of the cystic fibrosis gene in patients with chronic pancreatitis. N Engl J Med 1998;339:645-52.
44. Castellani C, Bonizzato A, Cipolli M, Sgarbi D, Cavallini GC, Bassi C, Pederzoli P, Mastella G. CFTR mutations and IVS8-5T prevelence in idiopathic chronic and recurrent pancreatitis. Pediatr Pulmonol 1997;S14:248.
45. Pignatti PF, Bombieri C, Marigo C, Benetazzo M, Luisetti M. Increased incidence of cystic fibrosis gene mutations in adults with disseminated bronchiectasis. Hum Mol Genet 1995;4:635-9.
46. Dahl M, Tybjaeg-Hanson A, Lange P, Nordestgaard BG. ΔF508 heterozygosity in cystic fibrosis and susceptibility to asthma. Lancet 1998;351:1911-3.

50 Years Ago in The Journal of Pediatrics

DIAGNOSTIC VALUE OF THREE DRAWING TESTS FOR CHILDREN

Silver AA. J Pediatr 1950;37:129-43

This is a classic review article of 3 pencil and paper drawing tests that were available in 1950 (and are still available) to obtain information about the developmental or emotional status of children. Specifically, the author describes the use of: (1) the Gesell Drawing Tests (to estimate visual-motor maturational levels of young children), (2) the Bender-Gestalt Test (to measure visual-motor function and integration), and (3) the Goodenough Draw A Man Test (to approximate mental age or intelligence). Using his clinical experience from the Children's Psychiatric Unit of Bellevue Hospital, the author then supplies case vignettes and drawings to illustrate the use and interpretation of results of one or more of these tests in the clinical care of children, particularly in a psychiatric setting.

In the last 50 years, the importance of children's drawings has clearly been recognized not only for assessment of developmental maturity but also as a projective technique for understanding the child's feelings and perceptions about himself, his family, and, in some cases, his illness. In fact, in some pediatric offices, children's drawings are used routinely.[1] This classic review article from 1950 reminds us that helping children to express themselves to us as caregivers is an important task and there are many effective methods.

Marianne E. Felice, MD
Department of Pediatrics
University of Massachusetts
Worcester, MA 01655

REFERENCE

1. Welsh JB, Instone SL. Use of drawings by children in the pediatric office. In: Dixon S, Stein M, editors. Encounters with children. 3rd ed. St Louis: Mosby Press; 2000. p. 571-89.

PAEDIATRIC RESPIRATORY REVIEWS (2003) 4, 272–277
doi: 10.1016/S1526-0542(03)00084-8
Available online at www.sciencedirect.com
SCIENCE DIRECT®
MINI-SYMPOSIUM: SCREENING IN CYSTIC FIBROSIS
Paediatric Respiratory Reviews

Newborn screening methods for cystic fibrosis

Bridget Wilcken[1,2*] and Veronica Wiley[1]

[1]New South Wales Newborn Screening Programme, The Children's Hospital at Westmead, Sydney, Australia; [2]Discipline of Paediatrics, University of Sydney, Sydney, Australia

KEYWORDS
cystic fibrosis;
neonatal screening;
immunoreactive trypsin;
DNA; sweat test

Summary An efficient newborn screening test for detecting cystic fibrosis has been available for over 20 years but is only now coming into widespread use. Blood immunoreactive trypsin is elevated in babies with cystic fibrosis and its measurement in dried blood spots is the primary screening tool. Poor discrimination in the first week requires a re-sampling step. The identification of the cystic fibrosis transmembrane conductance regulator gene and the discovery of a common mutation has allowed a combination of the primary screening test with a DNA test using the same sample. Differing genetic backgrounds have led to the development of population-specific protocols. A false-negative rate of around 5% is usual. Specificity is high. In all protocols involving a DNA test, confirmation of the diagnosis by sweat test is necessary when only one mutation is identified, identification of some carriers therefore being unavoidable. Careful counselling is needed for the families of these carriers.

SCOPE OF THE REVIEW AND SEARCH STRATEGY

This review does not consider the rationale for newborn screening for cystic fibrosis (CF), the clinical outcomes or, in detail, the methods for diagnosing CF in symptomatic or high-risk individuals. We concentrate here on the methods and protocols for diagnosis by the population (mass) screening of newborns. Early methods, using meconium as the primary screening vehicle, are briefly reviewed.

We searched Medline from 1966 to 2003, week 16, and Embase from 1980 to 2003, week 16, using the search terms "newborn screening", "cystic fibrosis", "immunoreactive trypsin" and "DNA" to find the most relevant references for this short review.

SCREENING FOR CYSTIC FIBROSIS: A BRIEF HISTORICAL OVERVIEW

Newborn screening blood tests have been routinely carried out in most of the developed world for many years in order to detect phenylketonuria, hypothyroidism and some other disorders. Most commonly, babies have a heel-prick blood sample taken some time between 24 h and 5 days of age on to filter paper manufactured to a uniform thickness and absorbency so that analytes can be eluted and quantitated. DNA can also be obtained from such samples. Samples are ideally analysed in large central laboratories.

Although saltiness had been appreciated from mediaeval times (the child who tasted salty when kissed being, according to Swiss and Austrian folklore, said to be bewitched and soon to die) Paul di Sant'Agnese, in 1953, was the first to show excess salt in the sweat of CF patients.[1] Based on the accumulating knowledge of CF, two methods of biochemical detection were developed in the 1950s: the measurement of sweat electrolytes, pioneered by Gibson and Cooke[2] (a seminal development for the secure diagnosis of CF), and the measurement of albumin in meconium. In the 1960s and 70s the measurement of meconium albumin was being actively explored as a test for screening neonates for CF.[3] Only in 1979 did a test using a dried blood spot in the measurement of immunoreactive trypsin signal the real possibility of widespread neonatal population screening.[4] The discovery in 1989 of the gene mutated in CF and of a common mutation, ΔF508, provided further possibilities for refining newborn screening. The implementation of newborn

*Correspondence to: B. Wilcken. Tel.: +61 2 9845 3654; Fax: +61 2 9845 3121; E-mail: bridgetw@chw.edu.au

1526-0542/$ – see front matter

screening was, however, controversial at first, as the benefits of early diagnosis were thought not to be clear[5] but the results of two randomised-controlled trials indicated nutritional and possibly other benefits[6,7] and newborn screening for CF is gradually becoming widespread.

EARLY SCREENING TESTS: MECONIUM ALBUMIN AND LACTASE, AND FAECAL TRYPSIN

Elevated meconium albumin levels in babies with meconium ileus were first noted in 1958 and Stephan *et al.* in 1975 [3] reported on the use of a chromatographic test strip, the Boehringer Mannheim (BM) test strip for meconium testing in 69 000 babies. Sixty of those with a positive test result were shown to have CF and four with CF had a negative test result. Following this, there were many reports of meconium testing but sensitivity and specificity were generally unacceptable.[8] Apart from the undeniably high-false positive and false-negative rates, there were other drawbacks to any meconium-based test, including problems of collecting meconium and, for an accurate result, storing it at 4°C until it was analysed. In addition, babies could not be resampled.

In some circumstances however, it may still be useful. The World Health Organization discussed the likely underdiagnosis of CF in some less-developed parts of the world and recommended screening using a meconium test as being practical and affordable.[9] Nazer had employed the BM test in Jordan, where CF was scarcely recognised, and diagnosed three cases among 7682 newborns, observing that, despite the drawbacks of the test, this screening had increased awareness of the disorder in the region.[10] A faecal trypsin test has also been described but has not proved useful.[11]

BLOOD SPOT IMMUNOREACTIVE TRYPSIN

A report in the *Lancet* in 1979, from Crossley *et al.* in New Zealand, showed that serum immunoreactive trypsin (IRT), which was known to be reduced in older CF patients, was significantly elevated in the newborn period and could be measured in dried blood spots.[4] This indicated for the first time that newborn screening for CF could perhaps be achieved using the routine dried blood sample already being collected by newborn screening programmes all over the developed world for the detection of phenylketonuria, hypothyroidism and other disorders.

Crossley used a radioimmunoassay with a polyclonal antibody to analyse IRT in the dried blood spots. Several other groups[12,13] confirmed the results of Crossley *et al.* and then modified the method to improve sensitivity. However, although the radioimmunoassay method had a high screening sensitivity, it was time-consuming. This led to the development of an alternative immunoassay approach using a monoclonal antibody-based enzyme immunoassay in a microtitre plate ELISA technique.[14]

The ELISA assay used a solid phase, two-site sandwich enzyme immunoassay of two mouse monoclonal antibodies with specificity for human trypsin and trypsinogen. Dried blood spot samples were assayed against a lyophilised standard of human pancreatic trypsin. The labour component for ELISA was half that of the RIA method and provided further advantages. The assay was easier to handle, with samples in a 96-well plate format rather than individual tubes, thus reducing the possibility of misidentification. The assay steps were simpler and the reagents more stable than those used for radioimmunoassay, with the added benefit of reducing occupational health and safety concerns as isotopic labels were not required.[15]

The major disadvantage of the ELISA assay remained the level of interaction required by the analyst. One of the critical stages was the development of the colour reaction. These problems were addressed with the development of a dissociated enhanced lanthanide fluoroimmunoassay (DELFIA).[16,17] IRT from the sample was reacted simultaneously with immobilised monoclonal antibody and europium-labelled monoclonal antibody in assay buffer. The microtitre plates were washed and enhancement solution added which dissociated the europium ions, providing highly fluorescent chelates, which were stable for hours. This assay has subsequently been fully automated as an AutoDELFIA assay, significantly reducing the analyst time required.

Different assays may give very different results. The major form of trypsin circulating in blood is trypsinogen. The acinar cells of the pancreas secrete two major isoenzymes of enzymatically inactive trypsinogen, one cationic, the other anionic. Once activated, trypsin is bound to α_1-antitrypsin, α_2-macroglobulin and inter-α-trypsin inhibitor. Commercially available antibodies have variable binding to each subspecies so results from different kits thus may not be comparable.[18]

No matter which assay is used for IRT measurement, there is a decline in the level of IRT with age[4,19,20] and samples collected soon after birth may demonstrate a falsely elevated level of IRT. Crossley *et al.* showed that children with CF had an abnormally high level of IRT in the first year of life and those with residual pancreatic function had an elevated level 3–5 years after birth, IRT was, however, below that of age-matched controls for CF children older than 2 without residual pancreatic activity.[4] There is also a decline in IRT with storage of the samples, the level of decline depending on the temperature, time stored, individual variation and assay antibody used.[18,21]

There is conflict over the need to establish specific action levels for IRT concentration in babies with meconium ileus as most babies with meconium ileus have CF and should therefore undergo additional testing regimens to exclude the disorder. It has been shown that although the IRT level

of meconium ileus patients as a group is elevated above that of normal infants, up to half the infants with meconium ileus have an IRT level below the cut-off.[22] Conversely many babies with prematurity have a falsely elevated level of IRT.[23]

DNA TESTS ON DRIED BLOOD SPOTS

Direct genotyping can readily be performed on dried blood spots and has been used for this purpose in newborn screening for some time. McCabe and colleagues demonstrated the feasibility of this approach for CF testing by performing direct genotyping on DNA from dried and liquid blood specimens and intact 4 mm discs punched from filter paper dried blood spots. They found complete agreement on the presence or absence of the mutation ΔF508, the common mutation found in cystic fibrosis.[24] DNA was amplified by a polymerase chain reaction (PCR) and analysed by polyacrylamide gel electrophoresis. Many other methods are possible. To give one example, Gasparini and colleagues have reported using an oligonucleotide ligation assay and sequence-coded separation following PCR, for a semi-automated approach.[25]

SCREENING PROTOCOLS

Several different protocols have been described using the IRT assay as the primary screening tool. These include a two-stage IRT assay involving a second test for babies with a high result on the first test (IRT/IRT), a single IRT test followed by a sweat test, an IRT test followed by DNA analysis on the same sample (IRT/DNA), a three-stage assay (IRT/DNA/IRT) and tests involving the use of IRT and meconium lactase, with or without DNA. Nowadays, almost all screening programmes use DNA mutation analysis as part of their testing protocol. All protocols have their different advantages and drawbacks, the choice depending on many factors, including cost and, where DNA tests are concerned, the specific genetic background of the population to be screened. Almost universally, the screening test will not detect all CF cases and a few cases could still present clinically after a negative screening test. All babies with a positive screening test result need a definitive diagnostic test before CF can be confirmed. This will usually be a sweat test (see below) but in some babies two disease-causing CF mutations will have been identified and confirmed and a sweat test may not be necessary.

IRT/IRT

The IRT level in babies at 2–5 days of age is poorly predictive of CF. Early studies suggested a positive predictive value (PPV) of 5–10% for an initially elevated IRT and this has been borne out over time.[19,26] The initial protocol adopted by most screening programmes was the two-stage IRT test. An elevated IRT level led to a request for a second sample for IRT testing and only when the second sample showed an elevated IRT level was a sweat test requested for definitive diagnosis. This early protocol was followed in Australia, Austria, Belgium, France, Germany, Italy, New Zealand, the UK, the USA and others. It relies on the greater differentiation between IRT levels in CF and non-CF babies at about 4–6 weeks old than occurs at 2–5 days. Case-finding with IRT/IRT was reviewed by Travert in 1988.[27] He collated the results of screening over 2 million babies by 15 laboratories, using four different analytical systems, three being radioimmunoassays. Overall, 6.4% of the 730 babies identified to have CF were missed by the screening tests. The PPV of the first test varied between 3% and 10% and that for the second test had a mean of 52%. There was a huge variation between laboratories. Some of this variation could be due to the differing performance of different antibodies.[5]

The PPV and specificity of any test depend upon the selection of an appropriate cut-off point. For some disorders, for example phenylketonuria, there is a strong need to detect every single case and the cut-off point is selected accordingly. For IRT/IRT CF testing, the re-testing is part of the whole test and cut-off points have been selected by programmes on different bases, some using fixed and some floating cut-offs (see above), most adopting a cut-off point to produce a re-sample rate acceptable to the community. The re-sample rate in these early programmes usually varied between 0.3% and 1%.[27] As discussed above, the IRT level declines over time,[19,20] and some programmes adopted a lower cut-off point for the second test. Another important point discussed earlier is the possibility of a negative test result in babies with meconium ileus.[5] Because these babies have "declared themselves" to be at high risk of CF, this is not a major issue but nevertheless one that must be understood by primary care physicians.

With an IRT/IRT protocol, about one baby in every 200 will need a second newborn screening sample, causing anxiety for the parents. The Wisconsin programme found that perinatal stress factors accounted for some 25% of false-positive results.[20] Other factors include renal failure, bowel atresia and congenital infection as well as trisomies 13 and 18.[5] The performance of the IRT/IRT protocol has been reviewed in several studies, some of which compare this with other protocols.[19,26–29] Overall, about 95% of babies can now be readily detected by this testing protocol but there is a high false-positive rate.

Single IRT test

A single IRT test followed by a sweat test was the approach adopted in Wisconsin during the first 4 years of a randomised, controlled trial of the efficacy and outcome of neonatal screening for CF.[20] It resulted in a sensitivity of only 85% but whereas lowering the cut-off point led to improved sensitivity, the corresponding increase in the number of false-positive results was deemed undesirable.

IRT/DNA protocol

The identification of the CF transmembrane conductance regulator (CFTR) gene in 1989 and the discovery of a common mutation, ΔF508, made possible efficient single-sample screening for CF. IRT could be assayed on the routine blood sample and a DNA mutation analysis performed for the common mutation on those filter-paper blood samples with an elevated IRT level. As a second blood sample was not needed, a relatively generous cut-off point could be adopted, usually of the top 1% of values, the sensitivity then depending largely on the prevalence of the mutation among the total number of mutations in the *CFTR* gene. Babies with two copies of the mutation clearly had CF and were referred for treatment. Babies with one copy might have CF or might simply be a carrier, a sweat test being needed to determine this. Babies with no mutation were recorded as having a negative screening test. In the Australian population in New South Wales, using an IRT/DNA protocol and seeking only for ΔF508, it was predicted that 6% of CF babies would have a negative DNA test but that some of these would be high risk because of family history or meconium ileus. The detection rate would be around 95%.[26]

A further characterisation of the CFTR mutations present in different populations has improved the potential for screening and most screening programmes now screen for more mutations.[25,30–33] The ΔF508 mutation is present in some 70–87% of CFTR genes in populations from Northern Europe but is less common in Southern and Eastern Europe (around 50–70%), populations in the Middle East, South America, Africa and elsewhere showing a much lower occurrence.[34] There is a significant positive correlation between the percentage occurrence of the ΔF508 mutation and the population prevalence of CF. This is reviewed with reference to newborn screening by Bobadilla and colleagues.[34] They found that, in most populations with sufficient data, there were fewer than 20 mutations that each accounted for more than 0.5% of all CFTR mutations. They recommended an ideal screening panel for the USA of 50 mutations. This sort of approach or an even more extensive one will soon become cost-effective and therefore more widely used.[25] An illustration of the importance of knowing the genetic make-up of populations in order to introduce CF screening can be taken from the Middle East, where there is now interest in screening. In the United Arab Emirates, two mutations would identify over 95% of patients and account for 88% of CFTR alleles, with S549R occurring in the Bedouin population and ΔF508 in those of Baluch lineage.[35] In neighbouring Saudi Arabia, a large number of mutations occur, with different sets of mutations in different regions.[34]

One problem with any protocol involving the analysis of a large number of mutations is the potential for detecting cases of CF not destined to become significantly symptomatic until late in life, if at all. If there is any doubt about the usefulness of screening for classical, severe CF, there must be much more doubt about detecting the very mild cases. Screening programmes that include mild mutations in their screening test will need to think very carefully about their approach to this problem. This is especially apparent with the R117H mutation, with which the clinical consequences vary according to the haplotype background – including either IVS8-5T (clinical CF usually eventuating) or IVS8-7T (often without any clinical problem).[36]

With IRT/DNA testing, babies with false-positive results (those with one detected mutation and a truly negative sweat test; see below) will all be carriers of the CFTR gene(s) being tested for. Babies carrying ΔF508 have, as a group, slightly higher IRT levels and they are up to twice as likely to be identified by the screening programme than would be predicted by population frequency. All screening programmes have found this.[26,37] Although the number of carriers detected by this screening approach is quite small, counselling them has not proved easy.[38] Depending on the precise DNA approach, babies with a falsely negative result will tend to belong to specific ethnic groups.

IRT/DNA/IRT protocol

The major advantages of the IRT/DNA approach are clear. There is usually an increased sensitivity, with greatly increased PPV and specificity. The cost is not substantially different [26,39] and, importantly, the anxiety and misunderstanding engendered by repeat testing are avoided. There are, however, drawbacks, the most obvious being the unwanted but obligate detection of a few carriers, with uncertainty persisting while a sweat test is arranged.[37] To address this partially, Pollitt and colleagues explored an IRT/DNA/IRT approach.[29] They combined a retrospective DNA analysis of samples from babies who had been screened by an IRT/IRT protocol with a prospective study of a three-stage protocol in which babies with one copy of the ΔF508 mutation on IRT/DNA proceeded to a repeat IRT test. They concluded that IRT/DNA/IRT was the superior approach, with a 92% reduction in the number of second tests compared with the IRT/IRT protocol and an 80% reduction in negative sweat tests, i.e. the false-positive cases identified as carriers. This approach is now used by a number of screening programmes.[40]

A comparison of these different screening protocols is given in given in Table 1.

Use of meconium lactase

Italian workers investigated the use of a meconium lactase test as an adjunct to screening with IRT and/or DNA. They retrospectively investigated which cluster of tests would be most effective among various combinations of IRT, with or without a repeat test, mutation analysis and meconium lactase. They concluded that a combination of IRT testing at birth, together with meconium lactase testing and an

Table 1 Comparison of advantages and drawbacks of the different screening protocols

Screening protocol	Probable advantages*	Drawbacks*
Meconium-based alone	Low cost and easy availability in developing regions	Relatively poor specificity and sensitivity
IRT/IRT	Good PPV after second test	Re-test rate high, engendering anxiety
IRT/DNA	No re-testing; good specificity. Sensitivity good but dependent on the population and DNA mutations tested	All "false positives" are carriers of cystic fibrosis and families need genetic counselling
IRT/DNA/IRT	Better specificity and better PPV after full testing than IRT/DNA. Better sensitivity, depending on DNA tests, as above	Some re-testing needed, engendering anxiety. False positives are carriers of cystic fibrosis and families need genetic counselling

IRT, immunoreactive trypsin; PPV, positive predictive value. IRT/IRT, IRT/DNA, IRT/DNA/IRT: cystic fibrosis screening protocols. See text for explanation.
* Actual values for specificity and sensitivity are not given as these can vary greatly from laboratory to laboratory.

analysis of three mutations, gave a higher sensitivity and considerably increased the specificity.[41] This approach of course requires two different samples and has not been taken up elsewhere.

Other approaches

PAP/IRT

Pancreatitis-associated protein (PAP) is increased in the blood of babies with CF. Sarles *et al.* investigated the use of PAP combined with IRT in newborn screening. In a limited study, they found very high specificity and reasonable sensitivity. They suggested that the avoidance of a genetic test would be especially useful when an explicit informed consent was required by law.[42]

Cord blood screening

IRT is elevated in the cord blood of babies with CF and is according to one report reasonably discriminatory.[43] As other newborn screening tests require postnatal blood, this is unlikely to be widely useful.

SWEAT TESTING FOR CONFIRMATION

For a confirmation of CF detected by newborn screening, a sweat test is necessary unless two disease-causing CFTR mutations have been found. The many problems associated with conducting the sweat test will not be reviewed here. Suffice it to say that the volume of sweat must be accurately measurable and the mass of sweat, representing the rate of sweating, is important. For an interpretable result, both chloride and sodium should ideally be measured and conductivity is not enough to give a secure result. In healthy neonates, a sweat chloride level should be less than 30 mmol/l, and any level above this must be regarded with suspicion. Infants with a sweat chloride of 30–39 mmol/l, one identified CF mutation and a positive IRT test result need further review. Those with a sweat chloride level of 40–59 mmol/l almost certainly have CF.[44]

CONCLUSION

Newborn screening for CF has been feasible for many years, with efficient case-finding. Table 1 summarises the possible approaches currently used and their advantages and disadvantages. Neonatal screening programmes are understandably very concerned to choose the right approach for case detection and aim for the highest specificity and sensitivity achievable. Although efficacy of case-finding is important, we must not forget that the primary reason for newborn screening is to provide measurable clinical benefit to the babies screened.

REFERENCES

1. di Sant'Agnese PA, Darling RC, Perera GA, Shea E. Abnormal electrolyte composition of sweat in cystic fibrosis of the pancreas. *Pediatrics* 1956; **12**: 549–563.
2. Gibson LE, Cooke RE. A test for the concentration of electrolytes in sweat in cystic fibrosis of the pancreas utilizing pilocarpine iontophoresis. *Pediatrics* 1959; **23**: 545–549.
3. Stephan U, Busch EW, Kollberg H, Hellsing K. Cystic fibrosis detection by means of a test strip. *Pediatrics* 1975; **55**: 35–38.
4. Crossley JR, Elliott RB, Smith PA. Dried blood spot screening for cystic fibrosis in the newborn. *Lancet* 1979; **i**: 472–474.
5. Wilcken B. Newborn screening for cystic fibrosis: its evolution and a review of the current situation. *Screening* 1993; **2**: 43–62.
6. Farrell PM, Kosorock MR, Rock MJ *et al.* Early diagnosis of cystic fibrosis through neonatal screening prevents severe malnutrition and improves long-term growth. *Pediatrics* 2001; **107**: 1–13.
7. Chatfield S, Owen G, Ryley HC *et al.* Neonatal screening for cystic fibrosis in Wales and the West Midlands: clinical assessment after five years of screening. *Arch Dis Child* 1991; **66**: 29–33.
8. Ryley HC, Neal LM, Brogan TD, Bray PT. Screening for cystic fibrosis in the newborn by meconium analysis. *Arch Dis Child* 1979; **54**: 92–97.
9. World Health Organization/International Cystic Fibrosis (Mucoviscidosis) Association Feasibility study of community control programmes for cystic fibrosis. *Bull WHO* 1990; 68: 709–713.
10. Nazer H. Cystic fibrosis in Jordan – incidence and prevalence. *Trans Roy Soc Trop Med* 1986; **4**: 1–4.

11. Forrest DC, Wilcken B, Turner G. Screening for cystic fibrosis by a stool trypsin method. *Arch Dis Child* 1981; **56**: 151–153.
12. Applegarth DA, Davidson AGF, Kirby LT. Dried blood spot screening for cystic fibrosis. *Lancet* 1979; **ii**: 1936–1937.
13. King DN, Heeley AF, Walsh MP, Kuzemko JA. Sensitive trypsin assay for dried blood specimens as a screening procedure for early detection of cystic fibrosis. *Lancet* 1979; **ii**: 1217–1219.
14. Bowling FG, Watson ARA, Rylatt DB, Elliot JE, Bunch RJ, Bundesen PG. Monoclonal antibody-based trypsinogen enzyme immunoassay for screening of neonates for CF. *Lancet* 1987; **i**: 826–827.
15. Bowling FG, Brown ARD. Newborn screening for cystic fibrosis using an enzyme linked immunoabsorbent assay (ELISA) technique. *Clin Chim Acta* 1988; **171**: 257–262.
16. Soini E, Kojola H. Time-resolved fluorometer for lanthanide chelates — a new generation of nonisotopic immunoassays. *Clin Chem* 1983; **29**: 65–68.
17. Hemmila I, Dakubu S, Mukkala VM, Slitari H, Lovgren T. Europium as a label in time-resolved immunofluorometric assays. *Anal Biochem* 1984; **137**: 335–343.
18. Heeley AF, Bangert SK. The neonatal detection of cystic fibrosis by measurement of immunoreactive trypsin in blood. *Ann Clin Biochem* 1992; **29**: 361–376.
19. Hammond KB, Abman SH, Sokol RJ, Accurso FJ. Efficacy of statewide screening for cystic fibrosis by assay of trypsinogen concentrations. *N Engl J Med* 1991; **325**: 769–774.
20. Rock MJ, Mischler EH, Farrell PM *et al.* Newborn screening for cystic fibrosis is complicated by age-related decline in immunoreactive trypsinogen levels. *Pediatrics* 1990; **85**: 1001–1007.
21. Dhondt JL, Farriaux JP. What does immunoreactive trypsin assays measure? *Screening* 1994; **3**: 33–38.
22. Rusakow LS, Abman SH, Sokol RJ, Seltzer W, Hammond KB, Accurso FJ. Immunoreactive trypsinogen levels in infants with cystic fibrosis complicated by meconium ileus. *Screening* 1993; **2**: 13–17.
23. Bourguignon JP, Deby-Dupont G, Reuter A, Santerre J, Gerard A. Variations in dried blood spot immunoreactive trypsin in relation to gestational age and during the first week of life. *Eur J Pediatr* 1986; **144**: 547–549.
24. Seltzer WK, Accurso F, Fall MZ *et al.* Screening for cystic fibrosis: feasibility of molecular genetic analysis of dried blood specimens. *Biochem Med Metab Biol* 1991; **46**: 105–109.
25. Gasparini J, Arbustini E, Restagno G *et al.* Analysis of 31 CFTR mutations by polymerase chain reaction/oligonucleotide ligation assay in a pilot screening of 4,476 newborns for cystic fibrosis. *J Med Screen* 1999; **6**: 67–69.
26. Wilcken B, Wiley V, Sherry G, Bayliss U. Neonatal screening for cystic fibrosis: a comparison of two strategies for case detection in 1.2 million babies. *J Pediatr* 1995; **127**: 965–970.
27. Travert G. Analysis of worldwide experience of neonatal screening for cystic fibrosis by measurement of blood immunoreactive trypsin. In: Travert G (ed.) *Mucoviscidose: Depistage Neonatal et Prise en Charge precoce*. Caen: CHRU de Caen, 1988; pp. 1–23.
28. Narzi L, Lucarelli M, Lelli A *et al.* Comparison of two different protocols of neonatal screening for cystic fibrosis. *Clin Genet* 2002; **62**: 245–249.
29. Pollitt RJ, Dalton A, Evans S, Hughes HN, Curtis D. Neonatal screening for cystic fibrosis in the Trent region: two-stage immunoreactive trypsin screening compared with a three-stage protocol with DNA analysis as an intermediate step. *J Med Screen* 1997; **4**: 23–28.
30. Scotet V, de Braekeleer M, Roussey M *et al.* Neonatal screening for cystic fibrosis in Brittany, France: assessment of 10 years' experience and impact on prenatal diagnosis. *Lancet* 2000; **356**: 789–794.
31. Gregg RG, Simantel A, Farrell PM *et al.* Newborn screening for cystic fibrosis in Wisconsin: comparison of biochemical and molecular methods. *Pediatrics* 1997; **99**: 819–824.
32. Spence WC, Paulus-Thomas J, Orenstein DM, Naylor EW. Neonatal screening for cystic fibrosis: addition of molecular diagnostics to increase specificity. *Biochem Med Metab Biol* 1993; **49**: 200–211.
33. Ranieri E, Lewis BD, Gerace RL *et al.* Neonatal screening for cystic fibrosis using immunoreactive trypsinogen and direct gene analysis: four years' experience. *BMJ* 1994; **308**: 1469–1472.
34. Bobadilla JL, Macek M, Fine JP, Farrell PM. Cystic fibrosis: a worldwide analysis of CFTR mutations – correlation with incidence data and application to screening. *Hum Mutat* 2002; **19**: 575–606.
35. Frossard PM, Girodon E, Dawson KP *et al.* Identification of cystic fibrosis mutation in the United Arab Emirates. *Hum Mutat* 1997; **11**: 412–413.
36. Massie RJ, Poplawski N, Wilcken B, Goldblatt J, Byrnes C, Robertson C. Intron-8 polythymidine sequence in Australasian infants with CF mutations R117H and R117C. *Eur Respir J* 2001; **17**: 1195–1200.
37. Castellani C, Bonizzato A, Mastella G. CFTR mutations and IVS8-5T variant in newborns with hypertrypsinaemia and normal sweat test. *J Med Genet* 1997; **34**: 297–301.
38. Ciske DJ, Haavisto A, Laxova A, Rock LZ, Farrell PM. Genetic counselling and neonatal screening for cystic fibrosis: an assessment of the communication process. *Pediatrics* 2001; **107**: 699–705.
39. Farrell PM, Mischler EH. Neonatal screening for cystic fibrosis. *Adv Pediatr* 1992; **39**: 35–70.
40. Corbetta C, Seia M, Bassotti A, Ambrosioni A, Giunta A, Padaon R. Screening for cystic fibrosis in newborn infants: results of a pilot programme based on a two tier protocol (IRT/DNA/IRT) in the Italian population. *J Med Screen* 2002; **9**: 60–63.
41. Castellani C, Bonizzato A, Cabrini G, Mastella G. Newborn screening strategy for cystic fibrosis: a field study in an area with high allelic heterogeneity. *Acta Paediatr* 1997; **86**: 497–502.
42. Sarles J, Barthellemy S, Ferec C *et al.* Blood concentrations of pancreatitis associated protein in neonates: relevance to neonatal screening in cystic fibrosis. *Arch Dis Child Fetal Neonatal Ed* 1999; **80**: F118–F122.
43. Ploier R, Emhofer J, Licka B, Rezanka E, Turnher H. Regional mucoviscidosis screening using immunoreactive trypsin in umbilical cord blood. *Pädiatr Pädol* 1991; **26**: 263–266.
44. Massie RJ, Gaskin K, Van Asperen P, Wilcken B. Sweat testing following newborn screening for cystic fibrosis. *Pediatr Pulmonol* 2000; **29**: 452–456.

J Inherit Metab Dis (2007) 30:537–543
DOI 10.1007/s10545-007-0584-0

NEWBORN SCREENING

Newborn screening for cystic fibrosis: Techniques and strategies

Bridget Wilcken

Received: 13 February 2007 / Submitted in revised form: 29 March 2007 / Accepted: 4 April 2007 / Published online: 12 May 2007

Summary Newborn screening for cystic fibrosis has been carried out for over 25 years, and clinical and cost benefits have been documented. There is still much variation in the methods and strategies adopted. All current screening programmes use a measurement of immunoreactive trypsin as a primary screening test, and in most, a second tier test involves analysing DNA mutations. The choice of DNA mutations depends on the genetic background in the region, and considerations of cost. Using DNA analysis as part of a screening procedure has introduced unwanted carrier detection, and protocols have now been devised in an attempt to avoid this. There are at least seven distinct protocols in use, all of which have different advantages and disadvantages, and no method or strategy will suit every region. Further careful study of performance and costs of various strategies is needed.

Abbreviations

CF	cystic fibrosis
CFTR	cystic fibrosis transmembrane conductance regulator
DGGE	denaturing gradient gel electrophoresis
DNA	deoxyribonucleic acid
IRT	immunoreactive trypsin
PAP	pancreatic-associated protein

Communicating editor: Rodney Pollitt

Competing interests: None declared

B. Wilcken (✉)
Biochemical Genetics and Newborn Screening,
The Children's Hospital at Westmead, Hawkesbury Road,
Westmead, NSW 2145, Australia
e-mail: bridgetw@chw.edu.au

B. Wilcken
The University of Sydney, Sydney, NSW, Australia

Introduction

Newborn screening for cystic fibrosis has been carried out for over a quarter of a century, and some benefits were quickly established. Only recently has it become less controversial as clinical and cost benefits have become clearer, the range of benefits seen has extended, harms have been found to be minimal, and benefits are seen to outweigh any harms. Now, as Farrell has so cogently written, the key question is no longer 'should we screen' but 'how should we screen?' (Farrell 2004). Regions where screening has long been established, such as north-western France, East Anglia (UK), Australasia, areas of Italy, and Colorado and Wisconsin (USA), still use a variety of strategies, and as screening is becoming more widespread, and is being planned in many regions, information about the advantages and drawbacks of different screening strategies is very important.

Screening for cystic fibrosis (CF) is currently carried out universally in Australia and New Zealand, in 27 of 51 states in the USA, with partial screening or planned screening in a further 7 (as of February 2007, http://genes-r-us.uthscsa.edu/nbsdisorders.pdf), and in 26 regional or state programmes in Europe (Southern et al 2007). There is interest in the Middle East (Nazer 1992). It is planned for screening to become universal in England this year (http://www.newbornscreening-bloodspot.org.uk/, accessed February 2007) (it is already carried out in Scotland, Wales and Northern Ireland); two Canadian provinces will begin screening

this year; and it is likely to be taken up by many more states in the USA. We reviewed CF screening methods in 2003, (Wilcken and Wiley 2003) and a synopsis of that review is presented here. There was a similar review in the same year dealing with strategies (Southern and Littlewood 2003). Since that time there have been proposals of new strategies, as well as studies comparing different strategies, and these are discussed.

Dried blood spot immunoreactive trypsin as the primary screening test

Newborn screening for CF was being discussed in the 1960s and 1970s, when the only method applicable to mass screening was the measurement of meconium albumin, a test with many drawbacks including poor sensitivity and specificity. A test-strip for screening was developed in 1975 (Stephan et al 1975). However, the discovery in 1979 of elevated blood immunoreactive trypsin (IRT) in babies with CF (a surprise initially, as low levels had been predicted) immediately led to the possibility of effective mass newborn screening (Crossley et al 1979).

All CF screening strategies now use dried blood spot IRT as the primary screening test. Crossley's original test used a polyclonal antibody and radioimmunoassay. This was a successful approach from the point of view of case finding, but it was time-consuming. Initial modifications included the use of a monoclonal antibody-based enzyme immunoassay using a 96-well microtitre plate ELISA technique (Bowling et al 1987). Further improvements led to the widespread, and now near-universal, use of an automated dissociated enhanced lanthanide fluoroimmunoassay (autoDELFIA, Perkin Elmer Life and Analytical Sciences, Wallac Oy, Turku, Finland) (Soini and Kojola 1983).

The problem with a primary IRT assay is one of poor specificity in the first few days of life. The positive predictive value at 2–5 days of age is about 3–10% (Travert 1988). The initial strategy using IRT was a two-stage assay, IRT/IRT: an initial elevation of IRT led to a repeat test at around 2–4 weeks, when the positive predictive value was approximately 50%, a much more manageable proportion. The Wisconsin group reported that perinatal stress factors accounted for some 25% of false-positive results in the first IRT test, with other causes including renal failure, congenital infection, bowel atresia and some aneuploidies (trisomies 13 and 18) (Rock et al 1990). A sweat-test was required for babies with two positive IRT test results, to confirm or exclude cystic fibrosis. Although the specificity of an initial positive IRT test was not good, the sensitivity proved to be high—we found a sensitivity of 98.1% for the first IRT test after screening over 1 million babies, if a cut-off of 1% was used (Wilcken et al 1995).

Adding DNA testing

Mutation testing can readily be performed on dried blood spots. The identification of the cystic fibrosis transmembrane conductance regulator (CFTR) gene in 1989 and the discovery of a common mutation, ΔF508, immediately made it possible to screen babies for CF using a single sample, carrying out mutation testing on samples with an elevated IRT level, without the immediate need for requesting a second sample (Seltzer et al 1991). In areas where the common mutation was particularly frequent—among populations derived mainly from Northern Europe—it was possible to screen using only that mutation. Babies with two copies had cystic fibrosis and were referred directly to a CF clinic, while babies with only one copy needed a sweat test to differentiate affected babies from carriers. In regions where the common mutation was less frequent, more mutations could be tested for in the initial blood spot, depending on the CF mutation background in the population (e.g. Scotet et al 2000; Spence et al 1993). Cost was one major consideration. Initially, adding new mutations was very costly, and this is to some extent still the case. This type of strategy is referred to as IRT/DNA.

Problems associated with DNA testing

The advantages of using a DNA test as second-tier testing are clear. Not only was there good sensitivity with little increase in cost but there was of course no need to recall babies for a second test, abolishing parental anxiety for many. The main disadvantage was the detection of carriers of CF—an outcome not included in the aim of screening. The population of newborn babies with elevated IRT levels is enriched with carriers (Castellani et al 2005; Parsons et al 2003). In several studies, about one carrier is detected for every CF patient identified. This still represents only a small proportion of all carriers—about 1%. While some people see this as a possible advantage, it has also been seen as undesirable, causing anxiety and possible harm to the developing mother–baby relationship. (There is in fact no evidence to support the latter assertion; e.g. Parsons et al 2003).

Several strategies have been suggested to avoid some of the problems of carrier detection, most

notably the IRT/DNA/IRT strategy (Pollitt et al 1997). Babies with an elevated level of IRT and one copy of the common mutation proceeded to a repeat IRT test. Only if this remained elevated was a sweat test requested. There was a 92% reduction in the number of second tests requested compared to an IRT/IRT strategy with no DNA component, and an 80% reduction in sweat tests. Thus very few carriers were formally detected; parents of babies with one copy of a CFTR mutation were told that cystic fibrosis was very unlikely, and were offered referral for genetic counselling. This approach has proved popular (e.g. Corbetta et al 2002), and other variations have melded IRT/IRT and IRT/DNA/IRT approaches by requesting a second sample from all babies. This latter protocol surrenders the advantages of a reduction in the number of second samples requested (Littlewood et al 1995).

Another problem has been a legal one: in France, bioethics laws mean that it is not permitted to perform any DNA testing without written, informed consent. To fulfil these obligations, the Ethics and Genetics committee of the French Association for Neonatal Screening recommended that informed consent should be obtained for all neonates at birth by having the parents sign directly on the sampling paper. Using this approach, the refusal rate was low, and declined from 0.8% at the start of the program to 0.2% at the end of the first year of screening (Dhondt 2005). Specific consent for screening for CF was introduced in Massachusetts, with a similar fall in refusals over the first year of screening.

Which DNA mutations?

A further difficulty in choosing a screening protocol is the realization that (again as in all screening) mild cases will be detected. This seems particularly so when a panel of DNA mutations is used. A consideration of the different DNA mutations chosen for inclusion in a screening strategy is of course dependent on the local population, and on the funding available, and is beyond the scope of this review. An example of a region where using just the common mutation might seem justified is the Hunter region of New South Wales, Australia, where it was shown that some 98% of babies with CF carried at least one copy of the common mutation, ΔF508, which comprised 80% of the mutations occurring (Henry et al 1996). In New South Wales as a whole, the occurrence of the common mutation is also high, comprising about 75% of all mutations. In some contrast, a study in north-eastern Italy found that it was necessary to use the 16 most common CFTR mutations to cover 86.6% of the mutations occurring in the CF patients—the 'common' mutation being much less common in this region (Bombieri and Pignatti 2001). There has been a move to avoid inclusion of mild mutations, such as the fairly common R117H, from panels even although this mutation can be quite severe, depending upon the haplotype background (O'Sullivan et al 2006).

Alternative methods

Other approaches suggested in areas with high allelic heterogeneity include use of meconium lactase as an extra test to improve an IRT/DNA strategy (Castellani et al 1997), and to avoid the use of DNA altogether by the use of pancreatic-associated protein (PAP) as an initial test, with a follow-up IRT test for samples with an elevated PAP (Sarles et al 2005). Neither of these approaches has been used outside the region where they were developed, although both seem to be effective.

Strategies for CF screening

The main strategies in use have already been mentioned: IRT/IRT, IRT/DNA, either with single or multiple mutations, and an extension of this, IRT/DNA/IRT, in which a second blood spot for IRT is requested when the initial IRT/DNA testing has revealed only one mutation. Further interesting variations on these have also been used. In Massachusetts, a strategy with IRT/DNA (using multiple CFTR mutation testing) also employs a third 'failsafe' step of referring for sweat testing not only those with two mutations or one mutation detected but also those with an extremely elevated IRT (Comeau et al 2004). This differs from the IRT/DNA/IRT strategy proposed by Pollitt, in that the second IRT is requested for patients without any mutation detected but with a very high initial IRT. On a different tack, the randomized controlled trial of screening in Wisconsin first used a strategy in which an elevated IRT with rather a high cut-off led immediately to a sweat test (IRT strategy, Rock et al 1990). Another proposed strategy employs an IRT/multiple DNA test, with the addition of an extended mutation analysis for samples where only one mutation has been found (Merelle et al 2006). The apparent benefits and drawbacks of these different approaches are shown in Table 1.

Table 1 Comparison of different CF screening strategies: actual benefits and drawbacks will depend on cut-off points adopted for the IRT assay, local genetic variation, and many other factors

Strategy	Steps	Likely benefits	Likely drawbacks
IRT/IRT	1. IRT on initial blood spot. If elevated: 2. Resample: IRT on second blood spot. If elevated: 3. Sweat test	Good specificity and sensitivity second test No carriers detected	Poor specificity first test: thus, more families with anxiety
IRT	1. IRT on initial blood spot. If elevated: 2. Sweat test	Good specificity No carriers detected	Either poor sensitivity *or* greatly increased false-positive rate. High sweat test rate
IRT/DNA Single (common) mutation	1. IRT on initial blood spot. If elevated: 2. DNA on same blood spot. Single (common) mutation: 3. If one copy of mutation: sweat test. If no copy of mutation: CF not indicated	Good sensitivity in some communities	Detection of some carriers Will have increased chance of missing CF for certain ethnic groups
IRT/DNA Multiple mutations	1. IRT on initial blood spot. If elevated: 2. DNA on same blood spot. Multiple mutations, chosen for genetic background: 3. If only one copy of any mutation: sweat test. If no copy of any mutation: CF not indicated	Better sensitivity compared with single mutation test	Increased cost Lower specificity: increased number of carriers identified
IRT/DNA/IRT	1. IRT on initial blood spot. If elevated: 2. DNA on same blood spot. Single or multiple mutations: 3. If one mutation only detected: new sample for IRT 4. If IRT elevated: sweat test If IRT not elevated: CF not indicated	Increased specificity compared with IRT/DNA Reduced number of sweat tests No carriers have automatic sweat test	Repeat blood sample for a small number of babies False-negative result possible if no mutations detected, or if second IRT is normal (2 severe mutations, one undetected Some carriers are identified, and offered genetic counselling. Some of these could have another mutation
IRT/DNA/ failsafe step	1. IRT on initial blood spot. If elevated: 2. DNA on same blood spot. Single or multiple mutations: 3. If one mutation detected or if no mutation, but extremely high IRT: sweat test	Increased sensitivity compared with IRT/DNA or IRT/DNA/IRT strategies	Increased number of sweat tests
IRT/DNA/DNA (experimental strategy)	1. IRT on initial blood spot. If elevated: 2. DNA on same blood spot. Single or multiple mutations: 3. If one mutation only detected: extended mutation analysis CF identified if two mutations detected 4. Uncommonly, DNA variants of unknown pathogenicity detected: sweat test	Increased specificity Better sensitivity with multiple mutations in second tier: only patients with two rare mutations missed	Increased cost compared with IRT/IRT A few new DNA variants of unknown pathogenicity detected: need for a sweat test Carriers are identified within the laboratory but are not reported to parents

For references see text.

Table 2 Studies comparing different strategies

Comparison	Region	Finding	Reference
IRT vs IRT/DNA	Wisconsin, USA	Fewer sweat tests and fewer false-positive subjects contacted using IRT/DNA	Gregg et al (1993)
IRT/IRT vs IRT/DNA	New South Wales, Australia	Similar sensitivity. Some carrier detection but no recall samples using IRT/DNA	Wilcken et al (1995)
IRT vs IRT/DNA	Wisconsin, USA	Improved positive predictive value, fewer false-positives, quicker diagnosis, no recall specimens using IRT/DNA	Gregg et al (1997)
IRT/DNA vs IRT/DNA/IRT	Trent region, UK	92% reduction in need for second blood sample, 80% reduction in sweat tests, using IRT/DNA/IRT; similar (very slightly lower) detection rate	Pollitt et al (1997)
Different mutation strategies	Veneto and Trentino Alto Adige regions, Italy	Complete gene screening detected 90% mutations vs 86.6% using 16 most common mutations, in one area	Bombieri and Pignatti (2001)
IRT/IRT vs IRT/DNA/IRT	Lazio region, Italy	Increased sensitivity using IRT/DNA/IRT	Narzi et al (2002)
IRT/multiple DNA/failsafe vs IRT/single DNA/failsafe	Massachusetts, USA	Increased sensitivity; increased carrier detection	Comeau et al (2004)

Comparison of strategies

Several studies have been conducted to compare strategies, and a selection of these are summarized in Table 2. It is clear that there is a balancing act (as always in screening) between, on the one hand, adopting protocols with the highest sensitivity and, on the other, keeping the 'noise' of false-positive results as low as possible. (An ideal approach might be an IRT test at 2–3 weeks, when it has high discrimination, followed by a sweat test at, say, 4 weeks on all babies with a positive IRT test; but alas this would be impracticable and expensive.)

Comparison of costs

Costing health care is very complex. There have been a few small studies examining costs of screening diagnosis versus clinical diagnosis (e.g. Lee et al 2003, more recently extended by Rosenberg and Farrell 2005). An interesting study modelled the costs of four screening strategies and assessed these in relation to health effects (Van den Akker-van Marle et al 2006). This type of economic-modelling exercise requires a number of assumptions to be made, and in this case it was assumed that there would be a gain of 40 life-years for every early non-meconium-ileus CF death prevented. The authors compared an IRT/IRT strategy with IRT/DNA, IRT/DNA/IRT, and IRT/DNA/DGGE strategies. The DGGE step was extended mutation analysis by denaturing gradient gel electrophoresis (see Table 1). They concluded that the most favourable cost–effectiveness ratio was achieved with IRT/IRT and that IRT/DNA/DGGE achieved more health effects at a lower cost than IRT/DNA/IRT. These costs will of course vary greatly from country to country and will depend heavily on the performance of the tests employed (the sensitivity and specificity, which can to some extent be manipulated by cut-off points) and the health-care system in the country, but these authors do conclude that screening is a good economic option.

An alternative to newborn screening

Offering an effective prenatal carrier-screening programme is considered by some as an alternative strategy to newborn screening. A programme in the East Lothian area of Scotland was in place for many years. Different approaches to carrier testing under

these circumstances include screening one member of a couple, normally the pregnant woman; if she is found to carry the common mutation, her partner is offered screening. Another approach is couple testing; both members of the couple are tested, but the result is given only as 'couple at risk' if both are carriers or 'couple at low risk' for other couples. In the United Kingdom and in subjects originating in Northern Europe, in a couple with one member carrying the common mutation and the other member not a carrier of this, the risk of CF in a baby is about 1:400. Currently the programme in E. Lothian has a very low uptake (A. Mehta, personal communication 2007) and a similar proposal in Denmark (Schwartz et al 1993) was not implemented (F. Skovby, personal communication 2007).

Aims of screening, methods and strategies

When considering the way forward for cystic fibrosis screening, it is important to keep in mind the general aims of newborn screening: to provide a benefit to an affected baby from early diagnosis, together with minimal harm to the baby and the community. Documenting outcomes is vital, so that benefits can be measured. Happily, good studies of CF outcome have been undertaken, as documented in the accompanying article (McKay 2007). The possible harms of screening are few, and include the financial costs borne by the community and the psychological and other costs that stem from false-positive results. Both of these are very dependent on the screening method and the strategy adopted. No method or strategy will suit every region. However, it is to be hoped that carefully designed studies will provide guidance, and that both established programmes and developing ones will benefit from the knowledge emerging about the results of different approaches to cystic fibrosis screening.

References

Bombieri C, Pignatti PF (2001) Cystic fibrosis mutation testing in Italy. *Genet Test* **5**: 229–233.

Bowling FG, Watson AR, Rylatt DB, et al (1987) Monoclonal antibody-based enzyme immunoassay for trypsinogen in neonatal screening for cystic fibrosis. *Lancet* **1**(8537): 826–827.

Brock DJ (1990) A consortium approach to molecular genetic services. Scottish Molecular Genetics Consortium. *J Med Genet* **27**: 8–13.

Castellani C, Bonizzato A, Cabrini G, et al (1997) Newborn screening strategy for cystic fibrosis: a field study in an area with high allelic heterogeneity. *Acta Paediatr* **86**: 497–502.

Castellani C, Picci L, Scarpa M, et al (2005) Cystic fibrosis carriers have higher neonatal immunoreactive trypsinogen values than non-carriers. *Am J Med Genet A* **135**(2): 142–144.

Corbetta C, Seia M, Bassotti A, et al (2002) Screening for cystic fibrosis in newborn infants: results of a pilot programme based on a two tier protocol (IRT/DNA/IRT) in the Italian population. *J Med Screen* **9**: 60–63.

Comeau AM, Parad RB, Dorkin HL, et al (2004) Population-based newborn screening for genetic disorders when multiple mutation DNA testing is incorporated: a cystic fibrosis newborn screening model demonstrating increased sensitivity but more carrier detections. *Pediatrics* **113**: 1573–1581.

Crossley JR, Elliott RB, Smith PA (1979) Dried-blood spot screening for cystic fibrosis in the newborn. *Lancet* **1**(8114): 472–474.

Dhondt JL (2005) Implementation of informed consent for a cystic fibrosis newborn screening program in France: low refusal rates for optional testing. *J Pediatr* **147**(Suppl): S106–S108.

Farrell PM (2004) Cystic fibrosis newborn screening: shifting the key question from "should we screen?" to "how should we screen?". *Pediatrics* **113**(6): 1811–1812.

Gregg RG, Wilfond BS, Farrell PM, et al (1993) Application of DNA analysis in a population-screening program for neonatal diagnosis of cystic fibrosis (CF): comparison of screening protocols. *Am J Hum Genet* **52**(3): 616–626.

Gregg RG, Simantel A, Farrell PM, et al (1997) Newborn screening for cystic fibrosis in Wisconsin: comparison of biochemical and molecular methods. *Pediatrics* **99**: 819–824.

Henry RL, Hettiarachchi LC, Colley P, et al (1996) Genotype of the cystic fibrosis population of the Hunter Region of New South Wales. *J Paediatr Child Health* **32**: 416–418.

Lee DS, Rosenberg MA, Peterson A, et al (2003) Analysis of the costs of diagnosing cystic fibrosis with a newborn screening program. *J Pediatr* **142**: 617–623.

Littlewood JM, Littlewood AE, McLaughlin S, Shapiro L, Connolly S (1995) Twenty years' continuous neonatal screening in one hospital: progress of the 37 patients and their families. *Pediatr Pulmonol* **19**(Suppl): 213.

McKay K (2007) Cystic fibrosis: benefits and clinical outcome. *J Inherit Metab Dis* **30** (in press).

Merelle ME, Scheffer H, De Jong D, et al (2006) Extended gene analysis can increase specificity of neonatal screening for cystic fibrosis. *Acta Paediatr* **95**: 1424–1428.

Narzi L, Lucarelli M, Lelli A, et al (2002) Comparison of two different protocols of neonatal screening for cystic fibrosis. *Clin Genet* **62**: 245–249.

Nazer HM (1992) Early diagnosis of cystic fibrosis in Jordanian children. *J Trop Pediatr* **38**(3): 113–115.

O'Sullivan BP, Zwerdling RG, Dorkin HL, et al (2006) Early pulmonary manifestation of cystic fibrosis in children with the DeltaF508/R117H-7T genotype. *Pediatrics* **118**: 1260–1265.

Parsons EP, Clarke AJ, Bradley DM (2003) Implications of carrier identification in newborn screening for cystic fibrosis. *Arch Dis Child Fetal Neonatal Ed* **88**(6): F467–F471.

Pollitt RJ, Dalton A, Evans S, et al (1997) Neonatal screening for cystic fibrosis in the Trent region (UK): two-stage immunoreactive trypsin screening compared with a three-stage protocol with DNA analysis as an intermediate step. *J Med Screen* **4**: 23–28.

Rock MJ, Mischler EH, Farrell PM, et al (1990) Newborn screening for cystic fibrosis is complicated by age-related decline in immunoreactive trypsinogen levels. *Pediatrics* **85**: 1001–1007.

Rosenberg MA, Farrell PM (2005) Assessing the cost of cystic fibrosis diagnosis and treatment. *J Pediatr* **147**(Suppl): S101–S105.

Soini E, Kojola H (1983) Time-resolved fluorometer for lanthanide chelates—a new generation of nonisotopic immunoassays. *Clin Chem* **29**: 65–68.

Sarles J, Berthezene P, Le Louarn C, et al (2005) Combining immunoreactive trypsinogen and pancreatitis-associated protein assays, a method of newborn screening for cystic fibrosis that avoids DNA analysis. *J Pediatr* **147**: 302–305.

Schwartz M, Brandt NJ, Skovby F (1993) Screening for carriers of cystic fibrosis. Results of a pilot study among pregnant women. *Eur J Hum Genet* **1**(3): 239–244.

Scotet V, de Braekeleer M, Roussey M, et al (2000) Neonatal screening for cystic fibrosis in Brittany, France: assessment of 10 years' experience and impact on prenatal diagnosis. *Lancet* **356**(9232): 789–794.

Seltzer WK, Accurso F, Fall MZ, et al (1991) Screening for cystic fibrosis; feasibility of molecular genetic analysis of dried blood specimens. *Biochem Med Metal Biol* **1**: 105–109.

Southern KW, Littlewood JM (2003) Newborn screening programmes for cystic fibrosis. *Paediatr Respir Rev* **4**(4): 299–305.

Southern KW, Munck A, Pollitt R, et al (2007) A survey of newborn screening for cystic fibrosis in Europe. *J Cyst Fibros* **6**(1): 57–65.

Spence WC, Paulus-Thomas J, Orenstein DM, et al (1993) Neonatal screening for cystic fibrosis: addition of molecular diagnostics to increase specificity. *Biochem Med Metab Biol* **49**: 200–211.

Stephan U, Busch EW, Kollberg H, et al (1975) Cystic fibrosis detection by means of a test-strip. *Pediatrics* **55**(1): 35–38.

Travert G (1988) Analysis of worldwide experience of neonatal screening for cystic fibrosis by measurement of blood immunoreactive trypsin. In: Travert G, ed. *Mucoviscidose: depistage neonatal et prise en charge precoce*. Caen: CHRU de Caen, 1–23.

van den Akker-van Marle ME, Dankert HM, Verkerk PH, et al (2006) Cost-effectiveness of 4 neonatal screening strategies for cystic fibrosis. *Pediatrics* **118**: 896–905.

Wilcken B, Wiley V (2003) Newborn screening methods for cystic fibrosis. *Paediatr Respir Rev* **4**(4): 272–277.

Wilcken B, Wiley V, Sherry G, et al (1995) Neonatal screening for cystic fibrosis: a comparison of two strategies for case detection in 1.2 million babies. *J Pediatr* **127**: 965–970.

The Lancet · Saturday 14 December 1985

REDUCED MORBIDITY IN PATIENTS WITH CYSTIC FIBROSIS DETECTED BY NEONATAL SCREENING

BRIDGET WILCKEN GWEN CHALMERS
Royal Alexandra Hospital for Children, Sydney, and New South Wales Health Department, Oliver Latham Laboratory, Sydney, Australia

Summary The effects of neonatal diagnostic screening on cystic-fibrosis (CF) -related morbidity were evaluated by comparing hospital admissions for CF-related illness in the first 2 years of life in 40 patients detected by means of neonatal screening and 56 patients born in the 3 years before screening began. Unscreened patients without meconium ileus had a mean of 27·25 hospital days for CF-related illness, and screened patients a mean of 3·9 days. There was no trend with time towards fewer days spent in hospital: the change was sudden. The difference was significant and could not be attributed to non-comparability of groups, changes in admission policy, or changes in management. In patients with meconium ileus there was no significant difference in hospital admissions between the groups. Neonatal screening significantly reduces CF morbidity in the first 2 years of life.

Introduction

NEONATAL screening for cystic fibrosis (CF) with the dried-blood-spot immunoreactive trypsin assay is very effective for case-finding,[1,2] but the benefits of early diagnosis are disputed.[3,4] The principal reasons for this uncertainty are the lack of data which unequivocally demonstrate medical benefit to patients and the possibility of psychological harm to families by diagnosis of a serious disorder before symptoms have become apparent to the parents.

A randomised controlled trial to answer these questions would raise difficulties. Randomised neonatal screening for CF in any one region would require informed parental consent, which would pose considerable logistic problems for screening laboratories, and the attendance of both screened and unscreened patients at the same clinic would be difficult for both staff and patients. Comparing screened patients from one region with unscreened patients from another remote one might not be valid even with agreed treatment protocols. We have conducted a sequential trial of screening, in which unscreened patients born in the three years before screening began were compared with screened patients, and all patients were treated in the same clinic.

Because a number of the unscreened patients were diagnosed late, some data for this group, such as height and weight, were not complete for the early years. We have looked at data on hospital admissions as an accurate index of morbidity during the first 2 years of life. We now report our findings.

Methods

CF screening was introduced into New South Wales, Australia, in July, 1981, as part of a routine neonatal screening programme. Immunoreactive trypsin was assayed in dried blood spots collected on the 5th day of life, with a method modified from Crossley,[5] described fully elsewhere.[2]

CF patients in the unscreened group included all those born in the three years immediately before screening started (between mid-July, 1978, and mid-July, 1981) in whom CF had been diagnosed by October, 1985, and who were referred at diagnosis to the CF clinic at the Royal Alexandra Hospital for Children, Sydney. Patients in the screened group included all those children born in New South Wales after screening began in whom CF was diagnosed, who were referred to the same clinic at diagnosis, and who are now aged 2 years or more. In this group was 1 patient missed by the screening test and 1 patient accidentally never screened because of neonatal illness.

There were 57 patients in the unscreened group. 1 girl who was immediately transferred to another centre has been omitted from the analysis. The other 56 patients were 59% of the total number of CF children expected for the State of New South Wales, assuming an incidence of 1:2500 births. In the screened group there were 40 children, 58% of the total CF children known to have been born after screening started who were aged 2 years or more. Thus the percentage of all CF children from New South Wales attending our clinic was the same in the two groups. 8 of the unscreened and 6 of the screened patients presented with meconium ileus.

During the period of the study, July, 1978, to July, 1985, two identifiable changes in management took place. From March, 1983, oral pancreatic enzyme supplementation was gradually changed from 'Viokase' (A. H. Robbins Pty Ltd) to 'Pancrease' (Ethnor Pty Ltd) in most patients who required enzymes, and from mid 1983 a gastroenterologist began regular attendances at the clinic. (A dietitian had provided diet management at the clinic before this.) Unscreened patients were all aged over 2 years before these changes occurred. 13 of the screened patients were exposed to the changes for 0–6 months during their first 2 years, and 21 for 7–24 months.

Details of hospital admissions were sought by examining the hospital records and recording the number of admissions, the number of days in hospital (including both day of admission and day of discharge), and the reason for admission. Parents were interviewed at the clinic about admissions to other hospitals, and details of admissions were obtained from those hospitals.

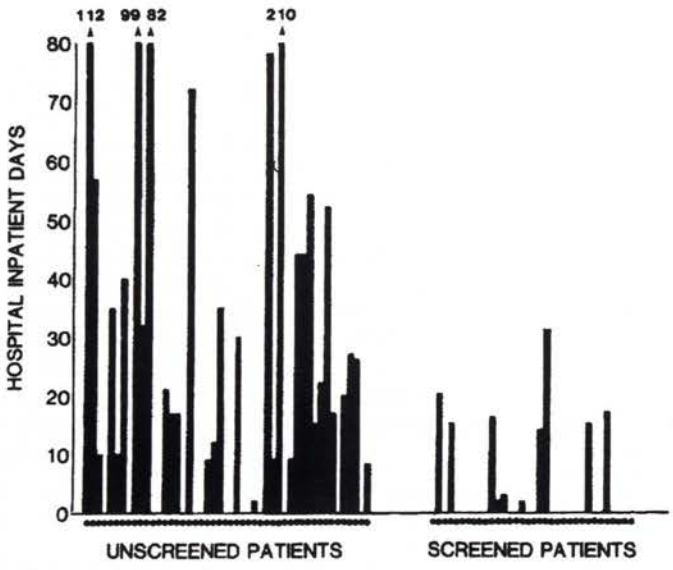
All admissions for chest infections, intravenous antibiotic therapy, malabsorption, diarrhoea, failure to thrive, and meconium ileus equivalent were included in our analysis. Admissions for sweat tests or instruction about cystic fibrosis therapy on diagnosis were recorded separately. Excluded were hospital days immediately following birth, hospital days for treatment of meconium ileus (initial treatment and closure of ileostomy), and admissions for illnesses unrelated to cystic fibrosis.

Results

Days in hospital for CF-related illness in the first 2 years of life for each patient are shown in the figure. Excluding patients who presented with meconium ileus, the 48 unscreened patients had a mean of 27·25 hospital days, whereas 34 screened patients had a mean of 3·9 days. The difference was significant ($p<0·01$ by Student's *t*-test and Wilcoxon rank sum test). Unscreened patients had a mean of 1·96 admissions, compared with a mean of 0·4 admissions for screened patients. By contrast, hospital days for the screened and unscreened patients who presented with meconium ileus were not significantly different ($p>0·5$ by Student's *t*-test and Wilcoxon rank sum test).

The mean of all hospital days in each group, including admissions for sweat tests and instruction, was 27·6 days for unscreened and 4·8 for screened patients (both without meconium ileus), an apparent "saving" of 22·8 days per patient (table I).

The difference between the groups in patients without meconium ileus was not solely due to 1 or 2 patients with long stays in hospital. Even when the 3 patients with most days in hospital are omitted from the unscreened group the difference between the groups is still highly significant. Also, the proportion of patients in each group staying in hospital for less than 1, less than 2, and less than 3 weeks was significantly smaller for the screened group (table II).



Hospital days for CF-related illness in the first 2 years of life in 34 screened and 48 unscreened patients without a meconium ileus presentation.

Individual CF patients are represented by closed circles.

TABLE I—HOSPITAL ADMISSIONS IN THE FIRST 2 YEARS OF LIFE IN SCREENED AND UNSCREENED PATIENTS (MEAN DAYS)

Reason for admission	Unscreened patients	Screened patients	
Excluding meconium ileus patients:	(n=48)	(n=34)	
CF-related illness	27·3	3·9	p<0·01*
Sweat test, instruction	0·4	0·9	
Total	27·7	4·8	p<0·01*
Meconium ileus patients:	(n=8)	(n=6)	
CF-related illness (excluding treatment of meconium ileus)	11·5	15·8	NS*

*Values of mean days for screened patients compared with unscreen patients.

TABLE II—SCREENED AND UNSCREENED PATIENTS ADMITTED TO HOSPITAL IN THE FIRST 2 YEARS

Hospital days	Unscreened patients (n=48)	Screened patients (n=34)	χ^2	p
0	15 *(31%)*	24 *(71%)*	12·3	<0·0005
<7	16 *(33%)*	27 *(79%)*	16·9	<0·0005
<14	23 *(48%)*	27 *(79%)*	8·3	<0·005
<21	28 *(58%)*	33 *(97%)*	15·7	<0·0005

There was no trend with time towards less time in hospita In the unscreened group without meconium ileus the first 2 patients had a mean of 27·5 days, and the next 24 a mean c 27·0 days in hospital. In the screened group, the first 17 ha slightly fewer days (mean 3·4) than the next 17 (mean 4·4

For the screened group this lack of trend with time was als important in considering the identified changes c management. 13 patients for whom the changes operated fc 6 months or less had the same number of hospital days as di 21 patients for whom the changes operated for 7 to 24 month (mean 3·9 for both groups).

In the unscreened group CF was diagnosed by age 1 months in 70% of patients. Among those who did not preser with meconium ileus, the diagnosis was made before an hospital admission in 22 patients, 13 of whom had n admissions at any time during their first 2 years. Diagnosi was made during the first admission in 20 patients, of whom had subsequent admissions during the first 2 years. In patients the diagnosis was made during the second (4) or thir (2) admission. In 67 of 94 admissions only respirator indications were present, and in a further 12 there were boti respiratory and gastrointestinal indications. In the screene patients without meconium ileus the pattern was the same, 1 of the 14 admissions being for respiratory indications only.

Discussion

This study shows that the number of days spent in hospital the number of separate admissions, and the number o patients ever admitted to hospital during the first 2 years o life were considerably smaller for patients in whom CF wa diagnosed by screening than for those in whom the diagnosi was indicated by symptoms.

There are various possible explanations for this finding: th control and screened groups may not have been comparable there may have been a change in practice relating t admission to hospital; or there may have been a significan change in management offered to patients attending th clinic.

The patients in the unscreened group are now aged between 4·25 and 7·25 years. A proportion of patients born during the 3 years of entry into this group will have remained undiagnosed and will probably be mildly affected. In a previous study[6] we found that only 1 patient out of 33

consecutively diagnosed patients at our clinic was aged over 4·25 years at diagnosis, so this bias is likely to be small in our group of 56 patients. In the screened group there may also be undiagnosed patients, but as the incidence of CF during our screening study, 1:2504 births, is similar to that found in a comprehensive long-term Australian study,[7] there will probably be few of these, and they too are likely to be mildly affected. In the two groups the percentages of all New South Wales patients attending our clinic are similar—59% and 58%, respectively. Referral patterns have remained stable over recent years; they are partly geographical and partly due to physician networks but do not depend on severity of illness. Thus it seems very unlikely that the patients in the unscreened group represented a more severely affected phenotype than the screened patients.

There may have been a change of practice relating to hospital admission over the 5-year period, but two pieces of internal evidence argue against this. Firstly, there is no trend with time during either period towards fewer days in hospital; the change occurred suddenly, with the advent of screening. Secondly, the patients presenting with meconium ileus, who may represent a different and perhaps more severe phenotype, are detected soon after birth, and their management is thus uninfluenced by screening. They can, therefore, serve as an "internal control". In these patients there was no difference, between the two periods under consideration, in the number of days in hospital for CF-related reasons other than management of meconium ileus. The meconium ileus patients as a group had more hospital days for illness than the screened patients and fewer than the unscreened.

There were some slight changes in medical management during the latter part of the screening period in this clinic—the introduction of a different pancreatic enzyme supplement and the attendance of a gastroenterologist at the clinic. However, no differences in hospital admissions could be demonstrated between patients for whom these changes operated for 0–6 months and those in whom they operated for 7–24 months, so that it seems unlikely that the reduction in time spent in hospital could be explained by changes in management.

70% of CF cases among unscreened patients were diagnosed by 12 months of age, and this is similar to findings in other States in Australia and in other countries.[6] In the majority of patients (42 of 48 without meconium ileus) CF was diagnosed either before or during the first admission to hospital, and this suggests that the number and length of their admissions during the first 2 years of life was a reflection of the severity of their symptoms and that failure to diagnose on presentation to hospital was not a factor.

Despite the sequential nature of the design of this study, dictated by ethical considerations, the results strongly suggest that the screening programme has reduced morbidity in cystic fibrosis patients during the first 2 years of life. There appeared to be an average saving of 22 days in hospital for each child screened. The reasons for this reduced morbidity are likely to be related to the early initiation of appropriate management regimens.

We thank the Cystic Fibrosis Association of New South Wales, who kindly supported this study; Dr J. Brown, Dr C. Mellis, and the other physicians of the CF clinic who kindly made patient records and families available to us; and Mrs L. Cant for secretarial assistance.

Correspondence should be addressed to B. W., Oliver Latham Laboratory, PO Box 53, North Ryde, NSW, Australia.

References at foot of next column

TREATMENT OF CADAVERIC RENAL TRANSPLANT RECIPIENTS WITH TOTAL LYMPHOID IRRADIATION, ANTITHYMOCYTE GLOBULIN, AND LOW-DOSE PREDNISONE

BARRY LEVIN, GEOFFREY COLLINS, MARK WAER, THEODORE GIRINSKY, RICHARD T. HOPPE, ELIZABETH MILLER, CHARLES BIEBER, SAMUEL STROBER

Departments of Medicine, Radiology, and Cardiovascular Surgery, Stanford University School of Medicine, Stanford, and Renal Transplantation Unit, Pacific Medical Center, San Francisco, California, USA

Summary The ability of preoperative total lymphoid irradiation (TLI) to reduce the need for chronic immunosuppression after cadaveric renal transplantation was examined in 25 recipients who were given a brief course of antithymocyte globulin (ATG) postoperatively with daily low-dose prednisone (0·1–0·2 mg/kg) as the sole maintenance immunosuppressive drug. Patients were selected for the study on the basis of their low levels of cytotoxic antibodies. Grafts were not HLA-matched, and the mean interval between completion of TLI and transplantation was 9 days. During an observation period of up to 25 months, 2 grafts were lost because of rejection. There were two deaths due to disseminated viral infections and two to late cardiovascular complications. At the last observation point, the mean serum creatinine of the 19 patients with functioning grafts was 1·5 mg/dl, and the mean dose of prednisone was 10·2 mg/day. 10 of these patients did not have a rejection episode. Comparison of patients given TLI with a group given cyclosporin at the same institution showed similar graft survival but better graft function in the TLI group.

Introduction

THE possibility of reducing or dispensing with the chronic administration of immunosuppressive drugs after organ transplantation is desirable because of the side-effects that have been associated with their long-term use. In the past two decades a combination of steroids and azathioprine has been the commonest post-renal-transplant regimen.[1] Overall this regimen gave a 1-year cadaveric renal allograft survival of 56% in North America in 1980.[2] Serious side-effects associated with the use of prednisone in transplant recipients have included diabetes, osteoporosis, bacterial and viral infections, aseptic necrosis of bone, and hypertension,[3,4] and those attributed to azathioprine have been bacterial infections, blood-element depressions, hepatic toxicity, and increased risk of malignancy.[3,4] The substitution of cyclosporin for azathioprine has increased 1-year cadaveric renal allograft survival by 20–30%.[5,6] However, long-term

B WILCKEN AND G CHALMERS: REFERENCES

1. Heeley AF, Heeley MF, King DN, Kuzemko JA, Walsh MP. Screening for cystic fibrosis by dried blood spot trypsin assay. *Arch Dis Child* 1982; **57:** 18–21.
2. Wilcken B, Brown ARD, Urwin R, Brown DA. Cystic fibrosis screening by dried blood spot trypsin assay: Results in 75,000 newborn infants. *J Pediatr* 1983; **102:** 383–87.
3. Ad Hoc Committee Task Force on Neonatal Screening, Cystic Fibrosis Foundation. Neonatal screening for cystic fibrosis. Position paper. *Pediatrics* 1983; **72:** 741–45
4. Phelan PD. Screening for cystic fibrosis In. Meadow R, ed. Recent advances in paediatrics. Edinburgh: Churchill Livingstone, 1984. 103–20.
5. Crossley JR, Smith PA, Edgar BW, Gluckman PD, Elliott RB. Neonatal screening for cystic fibrosis using immunoreactive trypsin assay in dried blood spots. *Clin Chim Acta* 1981, **113:** 111–21
6. Wilcken B, Towns SJ, Mellis CM. Diagnostic delay in cystic fibrosis: Lessons from newborn screening. *Arch Dis Child* 1983, **58:** 863–66
7. Allan JL, Robbie M, Phelan PD, Danks DM. The incidence and presentation of cystic fibrosis in Victoria 1955–1978. *Aust Paediatr J* 1980, **16:** 270–73

ORIGINAL ARTICLES

Clinical outcomes of newborn screening for cystic fibrosis

Donna L Waters, Bridget Wilcken, Les Irwig, Peter Van Asperen, Craig Mellis, Judy M Simpson, John Brown, Kevin J Gaskin

Abstract

***Aim*—To determine how early diagnosis of cystic fibrosis, using neonatal screening, affects long term clinical outcome.**

***Methods*—Fifty seven children with cystic fibrosis born before neonatal screening was introduced (1978 to mid 1981) and a further 60 children born during the first three years of the programme (mid 1981 to 1984), were followed up to the age of 10. The cohorts were compared on measures of clinical outcome, including height, weight, lung function tests, chest *x*-ray picture and Shwachman score.**

***Results*—Age and sex adjusted standard deviation scores (SDS) for height and weight were consistently higher in children screened for cystic fibrosis than in those born before screening. At 10 years of age, average differences in SDS between groups were 0.4 (95% CI −0.1, 0.8) for weight and 0.3 (95% CI −0.1, 0.7) for height. This translates to an average difference of about 2.7 cm in height and 1.7 kg in weight. Mean FEV_1 and FVC (as percentage predicted) were significantly higher in the screened cohort at 5 and 10 years of age, with an average difference of 9.4% FEV_1 (95% CI 0.8, 17.9) and 8.4% FVC (95% CI 1.8, 15.0) at 10 years. Chest *x*-ray scores were not different between the groups at any age, but by 10 years screened patients scored an average 5.3 (95% CI 1.2, 9.4) points higher on the Shwachman score.**

***Conclusion*—Although not a randomised trial, this long term observational study indicates that early treatment made possible by neonatal screening may be important in determining subsequent clinical outcomes for children with cystic fibrosis. For countries contemplating the introduction of neonatal screening for cystic fibrosis, its introduction to some areas in a cluster randomised design will permit validation of studies performed to date.**

(*Arch Dis Child Fetal Neonatal Ed* 1999;**80**:F1–F7)

Keywords: cystic fibrosis; neonatal screening; early diagnosis; clinical outcomes

James Fairfax Institute of Paediatric Nutrition Royal Alexandra Hospital for Children, Sydney, PO Box 3515, Parramatta 2124, New South Wales, Australia
D L Waters
K J Gaskin

Department of Biochemical Genetics
B Wilcken

Department of Respiratory Medicine
P Van Asperen
C Mellis
J Brown

Department of Public Health, University of Sydney
L Irwig
J M Simpson

Correspondence to:
Professor K J Gaskin.

Accepted 11 August 1998

Neonatal screening programmes for the common genetic disorder, cystic fibrosis (CF), have been developed and implemented in many centres over the past 25 years on the premise that early diagnosis and pre-symptomatic treatment improve clinical outcome.[1] Diagnosis of CF in the newborn period has permitted detailed investigation into the early natural history of this condition,[2] and programmes for screening[3–6] seem to be comparable with others used for neonatal detection of disease. However, concerns have been expressed about possible harm to early parent–child relationships from newborn screening programmes for CF,[7–9] and strategies for communicating screening test information to parents have been problematic in some regions.[10 11]

Previous observational studies suggest generally favourable short term outcomes from neonatal screening for CF. Prevention of nutritional deficits,[12–14] the early detection and treatment of pancreatic disease,[6] and reduced morbidity and inpatient stay in the first two years of life[15] have been reported. Significantly better outcomes for screened infants have also been reported in two long term observational studies. In one, however, results are complicated by non-standardisation of methods, with increasing proportions of infants joining the programme over the 15 year follow up period.[16] In another,[17] a total of only 13 cases were available for comparison at 10 year follow up. Two randomised controlled trials of neonatal screening for CF have been conducted in the USA and UK, but both have failed to demonstrate conclusively long term benefits. In the UK trial[18] incomplete ascertainment of CF cases within the non-screened population probably occurred, while in the US study,[19–21] only four children in the early diagnosis group and nine controls were available for follow up at 10 years of age.[21] Furthermore, pre-diagnosis anthropometric data were not available for the control group. Therefore, analysis at earlier ages omitted around 35% of infants (possibly those with less severe disease), violating the "intention to treat" principle of randomised controlled trial analysis. Until such methodological issues are addressed, long term benefits attributed to children receiving early treatment for CF can only be speculative. The reporting of accumulated long term observa-

tional or cohort data is therefore important to the ongoing debate.

The State of New South Wales (NSW) in Australia screens over 90 000 infants a year and has tested neonates for CF since July 1981. Although screening was not introduced in a randomised controlled trial, the size of the programme and its longevity mean that a longitudinal 10 year follow up of a cohort of screened infants diagnosed with CF can be undertaken. These children were compared with an historical control group of similarly treated infants born before screening was introduced, to determine whether early diagnosis was associated with any long term benefits for outcomes such as respiratory function and growth.

Methods

An historical cohort study comparing a group of 57 patients with CF born in the three years immediately before the NSW Newborn Screening Programme was introduced (July 1978 to July 1981; "non-screened") with 60 infants born during the first three years of this programme (July 1981 to July 1984; "screened"). Both groups comprise those infants born in NSW who were referred to the CF Clinic at the Royal Alexandra Hospital for Children, Sydney, and who were diagnosed before October 1995. Referral to this clinic is based on physician network and geographic proximity, with attendance by about 60% of all new cases of CF diagnosed in NSW. The statewide percentage for all children with CF referred to the centre was the same in both groups, as referral patterns did not change with the introduction of the screening programme.

Excluded from the study were three children in the screened group who transferred to clinics in other states before 12 months of age. Patients transfer between clinics primarily to attend the one closest to their place of residence. Seventeen of the children (7 screened, 10 non-screened) transferred to a new CF clinic opened in western Sydney in 1982 for this reason. These children remain in the cohort and were included in the analysis. Inpatient rates and length of stay for CF related illness have been reported before for this cohort which in 1985 comprised infants aged 2 years or more.[15]

In the current study, groups were compared on measures of clinical status up to 10 years of age. During the study period, neonatal screening was performed using dried blood immunoreactive trypsin (IRT) assay which was repeated if raised. Infants with persistently raised IRT were referred for sweat testing to confirm the diagnosis of CF.[1] The false positive rate of this test protocol was between five and six per cent.

The study was approved by the ethics committees of the Royal Alexandra Hospital for Children and the University of Sydney.

Changes to patient management throughout the 16 years of the study (apart from the introduction of neonatal screening) were the gradual phasing out of low fat diets, beginning 1981–2, the introduction of new generation pancreatic enzyme replacement treatment (enteric coated microspheres) from 1982–3, and the regular attendance of a paediatric gastroenterologist (KG) at both the Sydney and western Sydney clinic from mid 1983. These changes were not introduced homogenously, as individual differences in physician, parent, and patient acceptance dictated treatment preferences over a period of years. The use of nebulised antibiotics has been a feature of outpatient management since 1988, its use applying equally to both groups. No other major changes to staff, treatment philosophy, antibiotic or chest treatments occurred during each patient's 10 year follow up period.

MEASURES OF CLINICAL OUTCOME

Apart from visits for ill health, clinic patients routinely attend an interval medical check on (or around) each birthday and every six months inbetween. In this study routinely collected data on lung function, nutritional status, and growth recorded at diagnosis and closest to the first, fifth, and tenth birthdays were retrieved from each patient's medical record. Measures of clinical outcome include height, weight, chest *x*-ray score, Shwachman clinical score and (in children able to perform reliable pulmonary function testing) forced expiratory volume in one second (FEV_1), forced expiratory flow rate in the middle half of forced vital capacity (FEF_{25-75}), and forced vital capacity (FVC). Lung function measures are expressed as a percentage of predicted normal values for height and sex.[22] The presence of symptoms at diagnosis was transcribed from the physician's record of the first clinic visit. For the 100 children continuing to receive care at the original referral centre during the 10 years of follow up, height (in centimetres) and weight (in kilograms) were measured on the same stadiometer and scales over the period. Variables not recorded or not measured at the designated times are coded as missing values.

Anthropometric variables were converted to age and sex appropriate standard deviation scores (SDS) for height and weight using the method of Dibley *et al.*[23] Pancreatic function was determined by fecal fat estimation and/or pancreatic stimulation tests, as described before[6] in 40 (67%) of the screened and 15 (26%) of the non-screened children. We assumed pancreatic insufficiency—that is, the presence of malabsorption—in seven patients (3 screened, 4 non-screened) born with meconium ileus. Four of the remaining 55 patients (2 screened, 2 non-screened) were categorised as pancreatic sufficient (PS) on the basis of a normal serum pancreatic isoamylase and/or no reported symptoms of malabsorption and normal growth without pancreatic enzyme replacement treatment for varying periods between diagnosis and five years. A further five screened and 19 non-screened patients were classified as pancreatic insufficient (PI) after 2 years of age on the basis of low pancreatic isoamylase measurements, as defined before.[6] In the absence of definitive testing the remaining 27 patients (10 screened, 17 non-screened) were classified as PI on the basis of clinical symptomatology (oily stools and poor growth),

Table 1 Comparison of screened and non-screened birth cohorts

	Screened (n=60)	*Non-screened (n=57)*	*p Value*
Male:female	34:26	29:28	0.5
Meconium ileus	7	6	0.8
Age of diagnosis (months)			
Median (range)	1.8 (0.1-81.0)	5.7 (0.1-51.5)	<0.0001
Excluding meconium ileus	1.8 (0.1-81.0)	6.2 (1.1-51.5)	<0.0001
Excluding 3 patients missed by screening (n=57)	1.8 (0.1-7.6)		
Pancreatic function			
PS:PI at diagnosis	16:44	6:51	0.03
PS becoming PI	10	3	0.6
Symptoms at diagnosis (including meconium ileus)	30	56	<0.0001
Height SDS at diagnosis			
Combined mean (SD)	−0.2 (1.6)	−1.2 (1.6)	0.006
Pancreatic sufficient [PS]	−0.4 (1.4)	−0.06 (0.9)	
Pancreatic insufficient [PI]	−0.2 (1.7)	−1.4 (1.7)	
Weight SDS at diagnosis			
Combined mean (SD)	−0.1 (1.4)	−1.2 (1.4)	<0.0001
Pancreatic sufficient [PS]	0.3 (0.9)	0.1 (0.9)	
Pancreatic insufficient [PI]	−0.2 (1.6)	−1.3 (1.3)*	
Deaths due to CF	1	4	0.2
Total lost to follow up (includes deaths and transfers)	10	17	0.1

SDS = standard deviation score; SD = standard deviation; Students *t* test: * p<0.01 for difference in SDS between PS and PI.

requiring treatment with pancreatic enzyme replacement.

ANALYSIS OF DATA

Cross sectional results at diagnosis, 1, 5, and 10 years of age are given as mean and standard deviation (SD) or median and range for the total sample of 117 children: 60 screened and 57 non-screened. Three children in the screened group were missed by the screening test but were included in the analysis of this group from the time of their diagnosis (analysis by intention to screen). Their inclusion did not alter the findings. As anthropometric variables are standardised for age (standard deviation score or SDS), all patients were included in the cross sectional analysis at diagnosis regardless of their age at diagnosis. Similarly, 11 patients diagnosed between 12 and 18 months of age, including two cases missed by screening and eight non-screened (median age of diagnosis 1.3 years), are incorporated into the cross sectional analysis at 1 year. Apart from these exceptions, all other data are from the annual interval medical check performed on (or around) the first, fifth, and tenth birthdays.

Comparison between cohorts was done using Student's *t* test for normally distributed variables, Mann-Whitney U test for non-normally distributed variables, or the χ^2 test (Fisher's exact test for expected cell size <5) for categorical variables. Standard regression techniques using MINITAB Statistical Software (Minitab Inc, State College, PA, USA) were used to compare the cohorts while adjusting for individual patient differences, such as gender, presentation with meconium ileus, or pancreatic function. Estimates of the mean difference between groups (positive values indicate magnitude of advantage to screened group) are given with 95% confidence intervals (CI).

Longitudinal comparison of outcome variables for children remaining in the cohort to 10 years of age was by generalised estimating equations[24] (GEE) using SPIDA (Statistical Computing Laboratory, Eastwood, NSW, Australia). The GEE analysis allows all data at the three time periods to be used, even when a measurement is missing for that child at another time. A two dependence correlation structure was assumed, allowing for a common correlation at consecutive time points, but a different (always smaller) correlation between one and 10 year measurements.

Results

The birth cohorts are compared in table 1. The screened group includes 57 infants referred directly after notification of raised IRT plus three infants missed by the screening programme, and aged 1.3, 1.4, and 6.8 years at diagnosis (n=60). The median age of diagnosis was significantly lower in screened patients whether those born with meconium ileus are included or excluded (table 1). In the non-screened group (n=57) 53% of patients were diagnosed before 6 months of age, 70% before 12 months, and 91% by 2 years. Eighteen (32%) of the non-screened group (aged between 4 months and 4 years) were diagnosed and started treatment during the screening phase of the study (after July 1981). The screened and non-screened groups were not significantly different with respect to the proportion of boys, babies born with meconium ileus, number of pancreatic sufficient (PS) patients becoming pancreatic insufficient (PI), or number of children lost to follow up (including transfers out and deaths) over the 10 years of the study. For children transferring to the western Sydney clinic, the average age of transfer was 2.3 and 3.7 years for seven screened and 10 non-screened patients, respectively.

CROSS SECTIONAL OUTCOMES

As expected, symptoms were present at diagnosis in a significantly higher proportion of non-screened infants (the one asymptomatic child was referred after diagnosis of CF in a sibling). For screened infants, gastrointestinal and respiratory symptoms were recorded at diagnosis in 22 (36%) and 8 (13%), respectively. Height and weight SDS show that

Table 2 Cross sectional anthropometry by age and pancreatic function: mean (SD)

	N =	*Height SDS*	*Weight SDS*
One year			
Screened	57	**−0.4(1.0)**	**−0.1(0.9)‡**
PS	11	0.1(1.1)§	−0.1(0.9)
PI	46	−0.5(0.9)	−0.1(0.9)
Non-screened	48	**−0.8(1.2)**	**−0.6(1.1)**
PS	4	0.6(0.9)*	0.5(0.6)*
PI	44	−0.9(1.2)	−0.7(1.1)
Five years			
Screened	51	**−0.1(1.0)†**	**0.1(1.1)**
PS	6	0.5(1.2)	0.1(1.1)
PI	45	−0.2(1.0)	0.1(1.1)
Non-screened	52	**−0.5(1.0)**	**−0.3(1.0)**
PS	4	0.3(0.4)	0.4(0.4)
PI	48	−0.6(1.0)	−0.3(1.0)
10 years			
Screened	51	**−0.1(1.1)†**	**−0.2(0.9)**
PS	6	0.3(1.2)	0.1(1.3)
PI	45	−0.2(1.0)	−0.2(0.8)
Non-screened	41	**−0.6(0.9)**	**−0.5(0.9)**
PS	3	0.1(0.5)	0.2(0.4)
PI	38	−0.6(0.9)	−0.6(0.9)

SDS = standard deviation score; PS = pancreatic sufficient; PI = pancreatic insufficient. **Student's *t* test: † p<0.05 and ‡ p<0.01 for difference in SDS between screened and non-screened group.** § p<0.05 and * p<0.01 for difference in SDS between PS and PI.

Table 3 Mean differences between screened and non-screened group: cross sectional (CS) and longitudinal (L) data adjusted for pancreatic function (positive values indicate advantage for screened)

	One year Mean (95% CI) difference of estimate	*Five years Mean (95% CI) difference of estimate*	*10 years Mean (95% CI) difference of estimate*	*All ages (95% CI) unadjusted*
Height SDS				
CS	0.3 (−0.1, 0.7)	0.4† (0.1, 0.8)	0.4 (−0.1, 0.8)	
L	0.2 (−0.3, 0.7)	0.4 (−0.1, 0.8)	0.4 (−0.1, 0.8)	**0.3 (−0.1, 0.7)**
Weight SDS				
CS	0.5‡ (0.2, 1.0)	0.4 (−0.1, 0.8)	0.3 (−0.1, 0.7)	
L	0.5† (0.1, 0.9)	0.2 (−0.2, 0.7)	0.3 (−0.1, 0.7)	**0.4† (0.1, 0.7)**
FEV_1 (% predicted)				
CS		9.0‡ (2.6, 15.4)	9.4† (0.8, 17.9)	
L		7.3† (0.4, 14.3)	8.3 (−0.1, 16.6)	**8.6‡ (2.1, 15.0)**
FVC (% predicted)				
CS		8.5† (1.9, 15.2)	8.4† (1.8, 15.0)	
L		4.1 (−2.4, 10.7)	7.3† (1.1, 13.5)	**6.5† (1.0, 11.9)**
FEF_{25-75} (% predicted)				
CS		21.2‡ (9.0, 33.4)	9.3 (−5.7, 24.2)	
L		19.3‡ (5.8, 32.9)	8.0 (−7.1, 23.1)	**13.2† (1.6, 24.7)**
CXR score				
CS	1.0 (−0.4, 1.9)	1.2 (−0.1, 2.6)	1.0 (−1.0, 2.5)	
L	1.0 (−0.4, 2.2)	1.1 (−0.4, 2.6)	1.0 (−1.2, 2.4)	**1.0 (−0.1, 1.9)**
Shwachman score				
CS	2.4‡ (0.7, 4.1)	3.8‡ (1.0, 4.1)	5.3† (1.2, 9.4)	
L	2.3† (0.4, 4.1)	2.8 (−0.3, 6.0)	5.0† (0.8, 9.2)	**3.6‡ (1.4, 5.7)**

SDS = standard deviation score. **Student's *t* test: † p<0.05 and ‡ p<0.01 for difference between screened and non-screened group.**

children referred from the screening programme were both longer and heavier at diagnosis than the non-screened children, after adjusting for age and sex. These differences remained significant after adjustment for the presence of symptoms (including meconium ileus) and for pancreatic function. There were significantly more screened than non-screened infants who were PS at diagnosis, but for those diagnosed early by screening, no significant size difference was apparent between the PS and PI. For the non-screened group, however, six PS infants were significantly heavier than those categorised as PI at diagnosis (table 1).

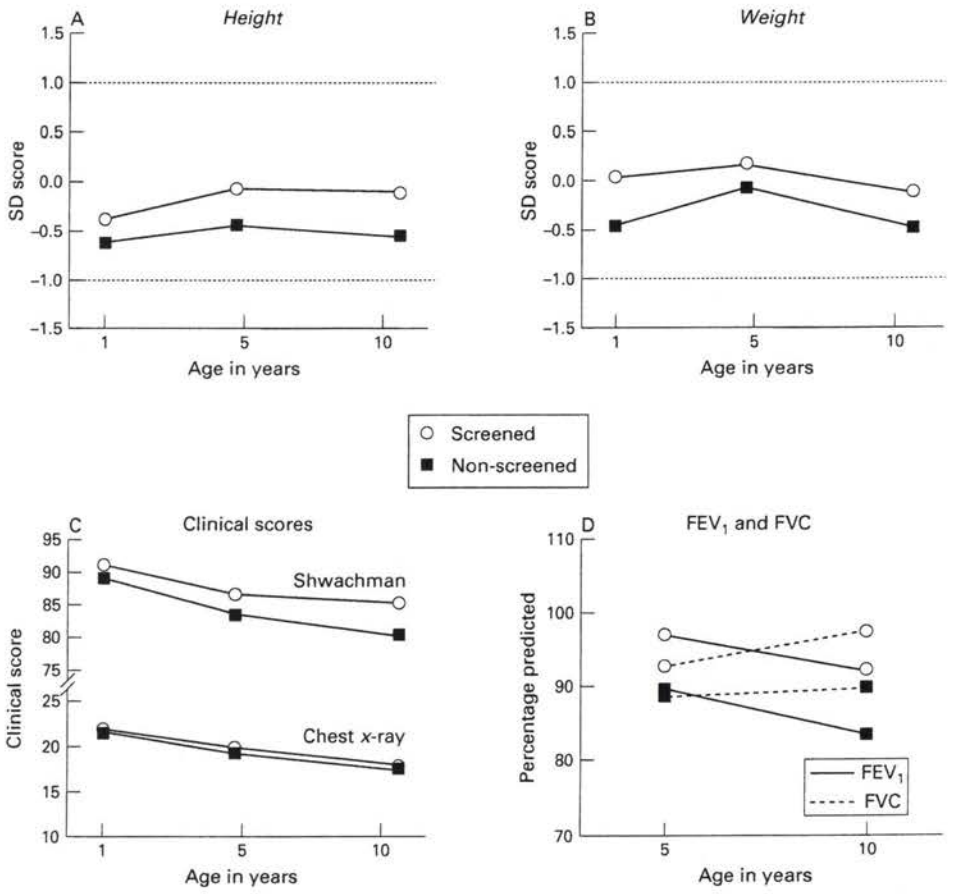


Figure 1 Longitudinal group data for screened and non-screened patients at 1, 5, and 10 years of follow up. Mean SD scores are shown for height (A) and weight (B); clinical scores for Shwachman and chest radiograph (C) and FEV_1 and FVC, as percentage predicted (D).

Table 2 shows cross sectional anthropometric data for the screened and non-screened group at each study age, stratified by pancreatic function (PS or PI) within each group. The mean (SD) age at each occasion of study was 1.1 (0.21) years, 5.0 (0.25) years, and 10.0 (0.29) years. At 1 year, screened patients remained significantly heavier, but were not significantly taller than the non-screened patients (table 2). A height advantage was apparent for PS patients within the screened group, while in the non-screened group at this age, both height and weight SDS were significantly greater in infants with PS. At 5 years, screened patients were significantly taller than the non-screened patients (table 2), but differences in weight SDS were no longer significant. PS remained a significant predictor of height at 5 years, but only six screened and four non-screened children retained pancreatic function, and intragroup anthropometric differences due to PS were not significant. At 10 years of age, the significant height advantage for screened patients represented an average difference of about 2.7 cm in height and an average (but not significant) difference of around 1.7 kg in weight.

The mean differences between scores for the screened and non-screened patients (adjusted for pancreatic function) are shown for the cross sectional and longitudinal group data at each age in table 3. While anthropometric results remained consistent with those shown in table 2, adjusting for PS reduced the mean difference in height SDS at 10 years, which was no longer significant in table 3.

All measured parameters of lung function were significantly higher in the screened cohort at 5 years of age, and by 10 years, there was an average difference of 9.4% in the predicted value of FEV_1 and of 8.4% in the predicted value of FVC, in favour of the screened group. While the average chest radiograph score was similar for screened and non-screened groups at any age (table 3), the combined Shwachman Score was an average 5.3 points higher in the screened group by 10 years of age (mean Shwachman Score 86).

LONGITUDINAL OUTCOMES

The anthropometry of the 51 screened patients remaining in the cohort for 10 years was not significantly different from a normal reference population 50th percentile,[25] with mean height SDS of −0.1 (95% CI −0.4, 0.2) and mean weight SDS of −0.1 (95% CI −0.4, 0.1). In contrast, the 41 non-screened patients were significantly below reference normal values, with a mean height SDS of −0.5 (95% CI −0.9 to −0.3, p=0.0006) and mean weight SDS −0.5 (95% CI −0.8, −0.2, p=0.001). Analysis of all 92 patients in the longitudinal cohort at 1, 5, and 10 years of age confirms significant differences between groups, favouring the screened cohort for weight SDS (p=0.02) and Shwachman Score (p=0.001), but there were no significant differences in height SDS (p=0.1) or chest *x*-ray score (p=0.7). The magnitude of the difference between groups did not change significantly over time, based on tests of interaction between group and time (figs 1A–C). As expected, mean FEV_1 declined in both groups over time, but the magnitude of the differences between groups for both mean FEV_1 and FVC were similar to those found in the cross sectional data (fig 1D). An estimate of the size of the differences between groups over all ages is given in table 3.

Discussion

This study comprises one of the largest systematic longitudinal comparisons of outcomes between screened and non-screened patients receiving treatment at the same CF facility over an extended period. The screened and unscreened groups were comparable in size, as the numbers were not different (57 and 60) for each of the three year periods. In addition, the groups were also comparable in terms of the severity of disease, as at 5 and 10 years the proportion with pancreatic sufficiency was similar (table 2). The results of both the cross-sectional and longitudinal analyses suggest that early treatment following diagnosis of CF by neonatal screening may improve the clinical outcome of patients during the first 10 years of life.

Examination of clinic policy revealed uniform respiratory management over the follow up period. Prophylactic flucloxacillin was used in the first year after diagnosis with oral, inhaled, and intravenous antibiotics for treatment of chest infections. Nutritional management did change in the early phase of this study with the gradual introduction of normal fat diets in 1981–2 and a change over to microspheric forms of enzyme replacement from 1982–3. While it is conceivable that the 80% of non-screened infants born before 1983 may have been disadvantaged by varying periods of low fat diets and changes to pancreatic enzyme replacement, two factors argue against this. Firstly, the non-screened group were older at diagnosis and would have received normal milk based infant and toddler diets before diagnosis. After diagnosis relatively short periods of dietary restriction would have been imposed and treatment with (non-microspheric) enzymes begun. Moreover, 32% of the non-screened group were diagnosed after the screening programme had begun, and in addition to a normal diet before diagnosis, they continued to receive the same (non-restricted) diet and pancreatic enzyme replacement after diagnosis as the screened group. Secondly, while all screened infants continued their normal milk based diet after diagnosis, those born during the first year of screening (36%) were treated with the same (non-microspheric) forms of pancreatic enzymes as the non-screened children. Thus, if the type of enzyme had contributed to the observed differences in growth, then one would also expect to see some detrimental effect on the growth of screened infants in their first year. The results of this study suggest that this did not occur. The interaction of time, age of diagnosis, and individual differences in physician, patient, and parent acceptance during change over periods renders it unlikely that differences in growth

between the screened and non-screened groups could be directly attributed to these changes in clinic policy.

Clearly, the screened group also had a higher proportion of patients with PS at the time of diagnosis. Pancreatic sufficient patients have been described as having milder disease phenotypes[26] and this may contribute to the more favourable outcome of screened patients. Statistical adjustments were therefore made to control for the clinical superiority of PS patients in both groups. Furthermore, we have already shown that, at the time of neonatal screening diagnosis, the proportion of patients with PS was much higher than in older CF populations—that is, patients with PS lose their pancreatic function over time and develop PI.[6] A higher proportion of PS patients in the screened group would therefore be expected due to their younger age of diagnosis. In the current study, loss of pancreatic function over time was such that by 10 years of age, as stated above, the screened and non-screened groups had the same proportion of PS patients.

Other hidden or less quantifiable effects may have contributed to patient outcomes, including physician and parental perceptions of the screened *versus* non-screened infants. Inpatient stays were lower in the screened patients during their first two years of life,[15] and while this could indicate that the non-screened group received more aggressive treatment, it is difficult to conceive how this would be disadvantageous in the long term. Finally, our population of CF patients represented 60% of the cases diagnosed statewide. Referral to clinics in NSW is based mainly on geographic factors. The three CF clinics in NSW are located at three paediatric teaching hospitals and the established referral patterns of local paediatricians depend almost entirely on their attachment to one of these three hospitals. As referral patterns did not change with the introduction of newborn screening, it is unlikely that referral of "milder" cases occurred during the screening period. The comparability of numbers in each patient group and their similarity on indicators of disease severity, such as the proportion with meconium ileus and PS, also support unbiased referral. We believe ascertainment in the cohorts studied is complete. All patients were over 10 years of age and the screening programme has a policy of actively seeking missed cases through regular contact with clinics in NSW and interstate, and with the major laboratories conducting sweat tests.

This study was not a randomised controlled trial and the non-screened comparison group was drawn from historical controls. Despite these limitations, the study is unique in the relatively large numbers of patients available for study at 10 year follow up. The results support an early nutritional and pulmonary function advantage to screened infants and suggest that initial benefits may be maintained over time. These results agree with those of previous controlled trials[18 21] and observational studies[16 17] which despite also experiencing design limitations and minor differences in patient profiles, all arrive at positive conclusions for improved nutrition and reduced inpatient stays among screened infants. We speculate that the early detection, treatment, and perhaps prevention of CF related illness in infants diagnosed early may break the cycle of infection and failure to thrive which previously alerted a medical diagnosis in symptomatic (pre-screened) patients. It may therefore be appropriate for screening programmes to focus on the diagnosis of more severe or classic cases of CF (ΔF508 or common population specific mutations).[4 5 27 28] If neonatal screening programmes for CF are to be more widely adopted, whether screening will provide any benefit to patients with "milder" disease genotypes or phenotypes needs to be considered, as many patients in these categories maintain near normal lung function and remain PS into adulthood.[29 30]

The advantages in nutrition and pulmonary function shown by the screened cohort in this study are yet to be explained. Preliminary results from a randomised controlled study of continuous prophylactic antibiotic treatment started in the neonatal period[31] have suggested that this treatment may benefit screened patients. However, benefits may equally be derived from the early and appropriate introduction of nutritional and pancreatic enzyme replacement in the screened group.

We have not addressed the balance of benefit or harm for the psychological or social aspects of neonatal screening for CF in this report, but are aware that any possible benefits must be weighed against the needs and reactions of parents, costs to the State, and above all, the wellbeing of the infant. Further studies are required to determine how early diagnosis and management influence later clinical, psychological, and social outcomes, and whether clinical benefits can be maintained into future adolescence and adulthood. For countries contemplating the introduction of neonatal screening for CF, its introduction to some areas in a cluster randomised design may help to resolve some of these issues.

We acknowledge Gwen Chalmers, Cheryl Frazer, and Petra Macaskill for their kind assistance in the preparation of this paper.

1 Wilcken B. Newborn screening for cystic fibrosis: its evolution and a review of the current situation. *Screening* 1993;2:43-62.
2 Farrell PM, Mischler EH. Newborn screening for cystic fibrosis: The cystic fibrosisneonatal screening study group. *Adv Pediatr* 1992;39:35-70.
3 Heeley AF, Bangert SK. The neonatal detection of cystic fibrosis by measurement of immunoreactive trypsin in blood. *Ann Clin Biochem* 1992;29:361-76.
4 Gregg RG, Wilfond BS, Farrell PM, Laxova A, Hassemer D, Mischler EH. Application of DNA analysis in a population screening program for neonatal diagnosis of cystic fibrosis (CF): comparison of screening protocols. *Am J Hum Genet* 1993;52:616-26.
5 Wilcken B, Wiley V, Sherry G, Bayliss U. Neonatal screening for cystic fibrosis: a comparison of two strategies for case detection in 1.2 million babies. *J Pediatr* 1995;127:965-70.
6 Waters DL, Dorney SFA, Gaskin KJ, Gruca MA, O'Halloran M, Wilcken B. Pancreatic function in infants identified as having cystic fibrosis in a neonatal screening program. *N Engl J Med* 1990;322:303-8.
7 Ad Hoc Committee Task Force on Neonatal Screening, Cystic Fibrosis Foundation. Neonatal screening for cystic fibrosis: position paper. *Pediatrics* 1983;72:741-5.
8 Al-Jader LN, Goodchild MC, Ryley HC, Harper PS. Attitudes of parents of cystic fibrosis children towards

neonatal screening and antenatal diagnosis. *Clin Genet* 1990;**38**:460-5.
9 Boland C, Thompson NL. Effects of newborn screening of cystic fibrosis on reported maternal behaviour. *Arch Dis Child* 1990;**65**:1240-4.
10 Helton JL, Harmon RJ, Robinson N, Accurso FJ. Parental attitudes toward newborn screening for cystic fibrosis. *Pediatr Pulmonol* 1991;**Suppl** 7:23-8.
11 Perobelli S, Faraguna D, Giglio L, *et al.* False positive screening for cystic fibrosis: reactions in parents and attitudes of professionals. In: Travert G, ed. *Mucoviscidose: Dépistage Néonatal et Prise en Charge Précoce.* Caen: CHRU de Caen, 1988:203-14.
12 Marcus MS, Sondel SA, Farrell PM, *et al.* Nutritional status of infants with cystic fibrosis associated with early diagnosis and intervention. *Am J Clin Nutr* 1991;**54**:578-85.
13 Reardon MC, Hammond KB, Accurso FJ, *et al.* Nutritional deficits exist before 2 months of age in some infants with cystic fibrosis identified by screening test. *J Pediatr* 1984;**105**:271-4.
14 Holliday KE, Allen JR, Waters DL, Gruca MA, Thompson SM, Gaskin KJ. Growth of human milk-fed and formula-fed infants with cystic fibrosis. *J Pediatr* 1991;**118**:77-9.
15 Wilcken B, Chalmers G. Reduced morbidity in patients with cystic fibrosis detected by neonatal screening. *Lancet* 1985;**2**:1319-21.
16 Mastella G, Barlocco EG, Antonacci B, *et al.* Is neonatal screening for cystic fibrosis advantageous? The answer of a wide 15 years follow-up study. In: Travert G, ed. *Mucoviscidose: Dépistage Néonatal et Prise en Charge Précoce.* Caen: CHRU de Caen, 1988:127-43.
17 Dankert-Roelse JE, te Meerman GJ. Long term prognosis of patients with cystic fibrosis in relation to early detection by neonatal screening in a cystic fibrosis centre. *Thorax* 1995;**50**:712-18.
18 Chatfield S, Owen G, Ryley HC, *et al.* Neonatal screening for cystic fibrosis in Wales and the West Midlands: clinical assessment after five years of screening. *Arch Dis Child* 1991;**66**:29-33.
19 Mischler E, Farrell P, Bruns T, *et al.* Progress report: Neonatal screening for cystic fibrosis in Wisconsin. *Wis Med J* 1989;**88**:14-18.
20 Farrell PM, Aronson RA, Hoffman G, Laessig RH. Newborn screening for cystic fibrosis in Wisconsin: first application of population-based molecular genetics testing. *Wis Med J* 1994;**93**:415-21.
21 Farrell PM, Kosorok MR, Laxova A, *et al.* Nutritional benefits of neonatal screening for cystic fibrosis. *N Engl J Med* 1997;**337**:963-9.
22 Corey M, Levison H, Crozier D. Five-to seven-year course of pulmonary function in cystic fibrosis. *Am Rev Respir Dis* 1976;**114**:1085-92.
23 Dibley MJ, Goldsby JB, Staehling NW, Trowbridge FL. Development of normalised curves for the international growth reference: historical and technical considerations. *Am J Clin Nutr* 1987;**46**:736-48.
24 Zeger SL, Liang K-Y, Albert PS. Models for longitudinal data: a generalized estimating equation approach. *Biometrics* 1988;**44**:1049-60.
25 WHO: Annex 3: reference data for the weight and height of children. In: *Measuring change in nutritional status: guidelines for assessing the nutritional impact of supplementary feeding programs for vulnerable groups.* Geneva: World Health Organisation, 1983:61-101.
26 Kristidis P, Bozon D, Corey M, *et al.* Genetic determination of exocrine pancreatic function in cystic fibrosis. *Am J Hum Genet* 1992;**50**:1178-84.
27 Ranieri E, Lewis BD, Gerace RL, *et al.* Neonatal screening for cystic fibrosis using immunoreactive trypsin and direct gene analysis: four years' experience. *BMJ* 1994;**308**:1469-72.
28 Spence WC, Paulus-Thomas J, Orenstein DM, Naylor EW. Neonatal screening for cystic fibrosis: addition of molecular diagnostics to increase specificity. *Biochem Med Metab Biol* 1991;**46**:105-9.
29 Strong TV, Smit LS, Turpin SV, *et al.* Cystic fibrosis gene mutation in two sisters with mild disease and normal sweat electrolyte levels. *N Engl J Med* 1991;**325**:1630-34.
30 The Cystic Fibrosis Genotype-Phenotype Consortium. Correlation between genotype and phenotype in patients with cystic fibrosis. *N Engl J Med* 1993;**329**:1308-13.
31 Weaver LT, Green MR, Nicholson K, *et al.* Prognosis in cystic fibrosis treated with continuous flucloxacillin from the neonatal period. *Arch Dis Child* 1994;**70**:84-9.

More evidence to favour newborn screening for cystic fibrosis

See Articles page 1187

Newborn babies have been screened for cystic fibrosis in some regions for more than 25 years. Benefits have been unequivocally shown, yet screening remains controversial. In today's *Lancet*, Erika Sims and colleagues[1] report more evidence in favour of newborn screening for cystic fibrosis. In a well-designed study, they used data from the UK cystic fibrosis database for 2002 to compare the treatment costs of 184 children aged 1–9 years, who had cystic fibrosis that was identified by newborn screening, with those of 950 children in the same age-group who were identified after clinical presentation of the disease. The findings of this cohort study are clear: savings in the cost of treatment would offset actual costs of the screening programme as it currently exists in Scotland.

Cystic fibrosis remains a life-shortening disorder, but effective treatments are available and the outlook for patients has improved substantially over the past 25 years.[2] Why newborn screening for cystic fibrosis has taken so long to gain popularity is unclear. In the 1980s, those with a role in early screening programmes reported health benefits that were hard to ignore.[3,4] In 1985, two randomised trials started in the USA[5] and UK[6] that assessed outcome in patients with cystic fibrosis who were identified by screening versus those who were identified clinically. Extensive reporting from the US trial showed that screening was associated with benefits in nutrition and growth[7] and in cognitive function.[8] Later, an Australian study[9] showed pulmonary benefits with newborn screening, and another US study showed survival benefits.[10]

If clear clinical benefit does not always persuade governments to implement screening, cost benefits might. The costs of screening are an important part of such decisionmaking. Until now, as Sims and colleagues highlight, cost-effectiveness studies have mainly compared screening with other methods of diagnosis for cystic fibrosis, and have not investigated cost savings that might offset screening costs. Screening for cystic fibrosis has been done throughout Scotland since 2002 and uses an initial measurement of immunoreactive trypsin on a dried blood-spot, followed by a DNA test on samples with the highest levels of immunoreactive trypsin. Sims and colleagues compared the cost of this screening technique with savings in treatment costs for patients with cystic fibrosis who were identified by screening. The costs quoted for the screening test in Scotland seem high (US$4·45 per baby screened). In our screening programme in New South Wales, Australia, incremental costs are about a third of this figure. If screening costs throughout the UK were to be lower than Sims and colleagues estimate, their finding that reduced treatment costs would offset screening costs would be strengthened.

A possible criticism of Sims and colleagues' study is that the children identified by use of newborn screening are not comparable with those identified by clinical presentation because the former probably includes more patients with mild disease, especially in the youngest cohort aged 1–3 years. However, when the researchers analysed only those who were homozygous for the common mutation that is associated with severe classic cystic fibrosis—Phe508del—they found a similar, but slightly less substantial, cost advantage for treatment of screened patients.

Other treatment costs are not addressed by Sims and colleagues' study. For instance, inpatient costs are a substantial proportion of total-care costs,[11] and various other types of cost are not captured by the data on the UK cystic fibrosis database. Nevertheless, the finding that the savings in major treatment would offset the costs of screening is persuasive. There are registries for cystic fibrosis in the USA, France, Australia, Germany, Ireland, and elsewhere. Sims' findings probably apply widely outside the UK, and similar studies in other countries might be useful.

Screening for cystic fibrosis is quickly gaining ground in many countries where it may be relevant: there are at least 26 programmes in Europe,[12] and 27 states in the USA now screen, with a further four likely to begin soon.[13] Newborn screening is done throughout Australia and New Zealand. Some parts of the world—eg, the middle east and possibly parts of India—have a high frequency of cystic fibrosis, and screening might have potential in countries with a stable health-care system.[14] Nevertheless, it remains an enigma that newborn screening for cystic fibrosis has until now been so controversial, when there is so much evidence in favour. By contrast, newborn screening for other disorders, such as toxoplasmosis, has been accepted although evidence of benefit remains unclear.[15]

Bridget Wilcken, Kevin Gaskin
Children's Hospital at Westmead, NSW 2145, and University of Sydney, Sydney, NSW 2006, Australia
bridgetw@chw.edu.au

We declare that we have no conflict of interest.

1 Sims EJ, Mugford M, Clark A, et al. Economic implications of newborn screening for cystic fibrosis: a cost of illness retrospective cohort study. *Lancet* 2007; **369**: 1187–95.
2 Dodge JA, Lewis PA, Stanton M, Wilsher J. Cystic Fibrosis mortality and survival in the United Kingdom, 1947 to 2003. *Eur Respir J* 2006; **29**: 522–26.
3 Abman SH, Reardon MC, Accurso FJ, Hammond KB, Sokol RJ. Hypoalbuminemia at diagnosis as a marker for severe respiratory course in infants with cystic fibrosis identified by newborn screening. *J Pediatr* 1985; **107**: 933–35.
4 Wilcken B, Chalmers G. Reduced morbidity in patients with cystic fibrosis detected by neonatal screeing. *Lancet* 1985; **2**: 1319–21.
5 Rock MJ, Mischler EH, Farrell PM, et al. Newborn screening for cystic fibrosis is complicated by age-related decline in immunoreactive trypsinogen levels. *Pediatrics* 1990; **85**: 1001–07.
6 Chatfield S, Owen G, Ryley HC, et al. Neonatal screening for cystic fibrosis in Wales and the West Midlands: clinical assessment after five years of screening. *Arch Dis Child* 1991; **66**: 29–33.
7 Farrell PM, Kosorok MR, Rock MJ, and the Wisconsin Cystic Fibrosis Neonatal Screening Study Group. Early diagnosis of cystic fibrosis through neonatal screening prevents severe malnutrition and improves long-term growth. *Pediatrics* 2001; **107**: 1–13.
8 Koscik RL, Lai HJ, Laxova A, et al. Preventing early, prolonged vitamin E deficiency: an opportunity for better cognitive outcomes via early diagnosis through neonatal screening. *J Pediatr* 2005; **147** (suppl 3): S51–56.
9 McKay KO, Waters DL, Gaskin KJ. The influence of newborn screening for cystic fibrosis on pulmonary outcomes in New South Wales. *J Pediatr* 2005; **147** (suppl 3): S47–50.
10 Grosse SD, Khoury MJ, Hannon WH, Boyle CA, Farrell PM. Early diagnosis of cystic fibrosis. *Pediatrics* 2001; **107**: 1492.
11 Rosenberg MA, Farrell PM. Assessing the cost of cystic fibrosis diagnosis and treatment. *J Pediatr* 205; **147** (suppl 3): S101–05.
12 Southern KW, Munck A, Pollitt R, on behalf of the ECFS CF Neonatal Screening Working Group. A survey of newborn screening for cystic fibrosis in Europe. *J Cyst Fibros* 2007; **6**: 57–65.
13 US National Newborn Screening and Genetics Resource Center. National newborn screening status report. http://genes-r-us.uthscsa.edu/nbsdisorders.htm (accessed Feb 20, 2007).
14 Nazer HM. Early diagnosis of cystic fibrosis in Jordanian children. *J Trop Pediatr* 1992; **38**: 113–15.
15 Schmidt DR, Hogh B, Andersen O, Fuchs J, Fledelius H, Petersen E. The national neonatal screening programme for congenital toxoplasmosis in Denmark: results from the initial four years, 1999 to 2002. *Arch Dis Child* 2006; **91**: 661–65.

Animal research: the debate continues

In 1875, Charles Dodgson, under his pseudonym Lewis Carroll, wrote a blistering attack on vivisection, which he circulated to the governing body of Oxford University, UK, in an attempt to prevent its establishment of a physiology department. Today, despite the subsequent evolution of one of the most rigorous governmental regulatory systems in the world, little has changed. A report[1] published at the end of 2006—*The use of non-human primates in research*—that was sponsored by the UK Royal Society, Medical Research Council, Wellcome Trust, and Academy of Medical Sciences attempts to establish a sounder basis for the debate on animal research through an in-depth analysis of the scientific basis for research on monkeys. In the UK, no great apes have been used for research since 1986.

Of 3000 monkeys used in animal research every year, 75% are for toxicology studies by the pharmaceutical industry.[1] Although expenditure on biomedical research has almost doubled over the past 10 years, the number of monkeys used for this purpose (about 300) has tended to fall. The report, which discusses mainly the use of monkeys in biomedical research, pays particular attention to the development of vaccines for AIDS, malaria, and tuberculosis, and to the nervous system and its disorders. The report assesses the importance to global health of these issues, together with potential approaches that might avoid the use of animals in research. Other research areas are also discussed, together with ethics, animal welfare, drug discovery, and toxicology.

The report concludes that in some cases there is a valid scientific argument for the use of monkeys in medical research. However, no blanket decisions can be made because of the speed of progress in biomedical science (particularly in molecular and cell biology) and because of the available non-invasive methods for study of the brain. Every case must be considered individually

PAPERS ON NEWBORN SCREENING BY TANDEM MASS SPECTROMETRY

Acta Pædiatr Suppl 432: 48–51. 1999

Newborn screening with tandem mass spectrometry: 12 months' experience in NSW Australia

V Wiley[1], K Carpenter[2] and B Wilcken[2]

NSW Newborn Screening Program[1] and NSW Biochemical Genetics Service[2], The Royal Alexandra Hospital for Children, Sydney, NSW, Australia

Wiley V, Carpenter K, Wilcken B. Newborn screening with tandem mass spectrometry: 12 months' experience in NSW Australia. Acta Pædiatr 1999; 88 Suppl 432: 48–51. Stockholm. ISSN 0803–5326

Since 1998, the NSW Newborn Screening Program has used electrospray tandem mass spectrometry (MS/MS) to analyse samples from all babies born in NSW and the ACT (approximately 95 000 per year) for selected amino acids and acylcarnitines. The software rules editor initially interprets all results where ratio of analyte to internal standard is modified by input from the external standard curves per analyte. The numerical results are then downloaded to the NSW Newborn Screening database, which provides automatic, analyte specific follow-up test cascade. We have analysed samples from 137 120 consecutive newborns received by the program, requested repeat samples from 122 babies, and found abnormal levels in 17 babies with phenylketonuria, 1 tetrahydrobiopterin deficiency, 3 hyperphenylalaninaemia, 1 maple syrup urine disease, 1 tyrosinaemia type II, 1 congenital lactic acidosis, 2 medium-chain acyl CoA dehydrogenase deficiency, 1 short-chain acyl CoA dehydrogenase deficiency, 1 β-ketothiolase deficiency, 2 vitamin B12 deficient babies of vegan mothers and 1 glutaric aciduria type I. Using population data plus that obtained from retrospective samples with proven disorders we have established cut-off levels for each analyte tested. This coupled with the ability of the database to provide ratios of various analytes gives excellent screening specificity and sensitivity for the detection of at least 40 rare inborn errors of metabolism. □ *Mass spectrometry screening, newborn acylcarnitines, newborn amino acids*

V Wiley, The NSW Newborn Screening Program, The Royal Alexandra Hospital for Children, P.O. Box Locked Bag 2012 Wentworthville, NSW, Australia, 2145 (Tel. +61 2 9845 3256, fax. +61 2 9845 3800, e-mail. veronicw@nch.edu.au)

Before introducing any new analyte or disorder to an already existing screening program, standard criteria concerning the associated benefits must be examined. This is no less and perhaps even more important before introducing a technique capable of quantitating levels of many analytes and therefore detecting many disorders on the same sample at the same time.

Tandem mass spectrometry (MS/MS) with two mass spectrometers, on either side of a collision cell, offers an analytical technique capable of measuring many analytes with rapid throughput per sample (approximately 2 min) (1–6). This technique also offers the advantage of simple sample preparation, as the first mass spectrometer is used to separate analytes within the sample, which are then altered in the collision cell and finally detected in the second mass spectrometer. With the recent use of electrospray for sample introduction into the MS/MS, automation became possible, providing a technique now suited to population screening. Further, using a specific operating mode, multiple reaction monitoring (MRM) instead of a scanning mode, it is possible to specify which ions and therefore which analytes will be detected providing the required specificity for a screening tool.

After agreeing on appropriate analytes to be assayed, we tested archived samples from previously diagnosed patients. We then tested samples from a pilot population of approximately 25 000 babies; analysed all data obtained to establish appropriate cut-off levels for follow-up action; and finally introduced the use of MS/MS as a screening tool.

In December 1997 the NSW Newborn Screening Program included a pilot program using electrospray MS/MS to expand the metabolic disorders detected in the routine screening of all babies born in New South Wales and the Australian Capital Territory, Australia (approximately 95 000 per year) from phenylketonuria, congenital hypothyroidism, cystic fibrosis, galactosaemia and homocystinuria by adding other selected amino acids (glycine, alanine, leucine/isoleucine and tyrosine) and acylcarnitines: carnitine, acetyl carnitine (C2), propionyl carnitine (C3), butyryl carnitine (C4), iso-

valeryl carnitine (C5), 3-hydroxy isovaleryl carnitine (C5OH), hexanoyl carnitine (C6), octanoyl carnitine (C8), decanoyl carnitine (C10), decenoyl carnitine (C10:1), myristyl carnitine (C14), tetradecenoyl carnitine (C14:1), and palmitoyl carnitine (C16). Later two additional analytes, citrulline and glutaryl carnitine (C5D), were added.

Methods

Samples

Blood samples were collected by heel prick onto S&S 903 paper (Schleicher and Schuell Inc., Keene, NH) when babies were 48 h of age or more, and usually on day 3 of life. Samples from babies diagnosed with inborn errors of metabolism were retrieved from archival storage in an under building cellar at room temperature where they had been stored for up to 10 y.

Pilot program

From December 1997, 24 965 consecutive samples received by the program were analysed in parallel with the current technique for screening for phenylketonuria and homocystinuria, the bacterial inhibition assay for phenylalanine and methionine respectively (7, 8).

Screening population

Samples from a total of 137 120 babies have been screened using MS/MS. Samples where all results were below the preset cut-offs were reported as no further tests required. Where an analyte was outside the 99.5 percentile the original sample was repeated in-house and, as required, other techniques, including capillary electrophoresis, thin layer chromatography or DNA mutational analysis, were used to determine which babies required further investigation. If the analyte was marginally abnormal, a repeat dried blood spot sample was collected. If this remained abnormal or the original sample was significantly abnormal, further investigation included a repeat dried blood spot sample, urine for a metabolic screen of organic acids and amino acids and plasma for amino acids and/or acylcarnitines.

Sample preparation

The sample preparation was adapted from the method of Millington et al. (1): blood from a 3 mm blood disc was eluted with shaking for 1 h in methanol containing isotopically labelled internal standards in 96 well polypropylene deep-well microtitre plates. The supernatant was decanted to a standard depth polypropylene plate and evaporated to dryness under warm (40 °C) air. Analytes were derivatized using butanolic HCl heated to 60 °C for 15 min before again being evaporated to dryness. The sample was then redissolved in solvent (50% v/v acetonitrile:water) for injection into the MS/MS. Sample preparation required about 2 h person time for the average 400 samples, standards and quality controls per day. Total run time was 2.1 min per sample or less than 14 h overnight per batch.

Acylcarnitine and amino acid standards

Ala, Cit, Gly, Leu/Ile, Met, Phe and Tyr and Carnitine, C2, C3, C4, C5, C5-OH, C5D, C6, C8, C10, C10:1, C14, C14:1, C16 were measured using isotopically labelled standards for Ala, Cit, Leu, Met, Phe and Tyr as well as C2, C8 and C16 (Cambridge Isotopes, Andover, MA). Nine external blood spot standards with varying levels of added amino acids and acylcarnitines were assayed in each batch and contained known spiked concentrations of each amino acid measured as well as carnitine, C2, C3, C5, C6, C8, C10 and C16. These external standards were used to create standard curves per analyte to determine quantitative levels of amino acids and acylcarnitines.

Instrumentation

The mass spectrometer used was a Quattro II (Micromass Ltd, UK), the autosampler a 215 Liquid Sampler (Gilson Instruments, Middleton, WI) and the pump was a series 1100 (Hewlett Packard, Palo Alto, CA).

Computerization

The MS/MS was controlled by a Digital workstation using the Windows NT operating system and Masslynx software (Micromass Ltd, UK). Spectra were initially interpreted by Neolynx software (Micromass Ltd, UK). The results obtained were transferred to the Labmaster database (Wiley Associates Pty Ltd, Sydney, Australia), where each analytical result was checked against predefined reference ranges and caused an analyte-specific tests cascade if the result was outside the range. In addition, various analytical ratios, e.g. phe/tyr or C8/C2, were checked against their predefined reference range and specific cascades were added if the ratio was outside the range.

Results

Pilot program

We have analysed 24 965 consecutive samples received by the program in parallel with the bacterial inhibition assay for phenylalanine and methionine, detecting elevations in all quality control samples, as well as elevations of phenylalanine in samples from the three newly diagnosed patients with phenylketonuria. There were no abnormal levels of methionine in samples from the pilot screening population.

Establishing cut-off levels

The results from the pilot screen were evaluated with initial in-house retesting on the original sample being performed on any sample outside the 99 percentile.

Table 1. Cut-off levels for amino acids and acylcarnitines required for repeat sample collection. Levels were established using both pilot population results and results obtained from testing archived samples from babies with proven disorders.[a]

	Analyte	Cut-off for repeat sample (μmol/l)
Amino acids	Alanine	900
	Citrulline	75
	Glycine	1000
	Leucine/isoleucine	500
	Methionine	80
	Phenylalanine	150
	Tyrosine	500
Acylcarnitines	Carnitine	5–125
	Acetyl carnitine	8–160
	Propionyl carnitine	9
	Butyryl carnitine	1.6
	Isovaleryl carnitine	1.4
	3-hydroxy isovaleryl carnitine	1.0
	Hexanoyl carnitine	0.8
	Octanoyl carnitine	1.0
	Decanoyl carnitine	1.5
	Decenoyl carnitine	0.8
	Myristyl carnitine	1.5
	Tetradecenoyl carnitine	1.5
	Palmitoyl carnitine	8.5

[a] Carnitine and acetylcarnitine have both upper and lower cut-off values; for all other analytes, resamples were requested only if results were higher than the cut-off value.

Using population data, plus data obtained from retrospective samples with proven disorders, we established cut-off levels for each analyte tested (see Table 1). The individual analyte cut-off value was coupled with the ability of the database to provide ratios of various analytes to give further criteria to increase specificity and sensitivity.

Screening

From the 137 120 babies screened we requested a resample from 66 babies for elevated amino acids and 56 babies for acylcarnitines. Thirteen babies required follow-up investigation by the Biochemical Genetics laboratory, including urine metabolic screen and plasma collection. There were 31 babies with a disorder detected. Seventeen babies with phenylketonuria, 1 tetrahydrobiopterin deficiency, 3 hyperphenylalaninaemia, 1 maple syrup urine disease, 1 tyrosinaemia type II, 1 congenital lactic acidosis, 2 medium-chain acyl CoA dehydrogenase deficiency, 1 short-chain acyl CoA dehydrogenase deficiency, 1 β-ketothiolase deficiency, 2 vitamin B12 deficient babies of vegan mothers and 1 glutaric aciduria type I. This last case was detected retrospectively on the original newborn screening sample once screening for glutarylcarnitine was added.

Incremental cost

The initial cost of the instruments required for MS/MS screening remain significant. Within our laboratory the staffing costs associated with MS/MS are: sample preparation and result verification of 15 cents per baby; consumables, including extraction plates, gases, internal standards and all reagents, of 30 cents per baby; and the overheads, including depreciation of the instrument, of 47 cents per baby.

Discussion

Newborn screening is recommended for the detection of a disorder of serious consequences and reasonable frequency with a benefit from early diagnosis. As many disorders can be detected using tandem mass spectrometry, even very rare disorders can be justifiably screened. Using this technique there was one baby with a disorder for every 4500 babies screened. Of the 31 babies detected, 20 would have been detected by our previous protocol of screening for phenylketonuria and homocystinuria, and therefore an additional 11 babies were detected by changing the screening technique.

There needs to be a reliable test with high specificity (healthy persons should have a negative test result) and sensitivity (affected persons should have a positive test result). So far, from the cohort screened there have been two cases of inborn error of metabolism not detected by the original screening test. Apart from the case of glutaric aciduria (glutarylcarnitine was not being screened for) there was one case of cobalamin C defect detected due to neonatal symptoms, for whom the propionylcarnitine was not increased (7.5 μmol/l; cut-off 9 μmol/l) and the C3:C2 ratio was only marginal at 0.3 (population ratio <0.3). There were 122 repeat samples requested using our established cut-off limits, which is less than 0.1%.

Worldwide there is an increasing trend to discharge mother and baby within the first day or two of life. This technique offers the opportunity of using both individual cut-off levels plus ratios, for example, for phenylalanine and tyrosine, so as to discriminate adequately an affected baby from the normal population.

There needs to be a system in place for confirmation of abnormal results, other specific diagnostic tests, counselling, treatment and follow-up. As we are part of an existing program with both Newborn Screening and Biochemical Genetics laboratories, as well as various clinics and medical staff available for treatment of inborn errors, there has been minimal change required by the introduction of MS/MS screening.

The final criterion used for assessing screening is associated with cost justification, including labour, consumables, capital costs and clinical costs. While it is difficult to establish exactly what the clinical cost would be, in our laboratory the cost of adding MS/MS to an already existing program is less than A$1 per baby (92 cents). This includes depreciation of the instruments, the cost of all consumables and the required

labour costs for a technician to prepare the samples and all required verification of the results. Almost one half of this (42 cents) is offset by the cost of the previous methodology for screening for phenylketonuria and homocystinuria. Further it should be noted that we are able to contain the costs of this technique with extensive use of computer algorithms. Labour costs can be significant. Our use of computer software minimizes the required input of experienced and therefore more expensive staff and allows for the use of individually established cut-offs for each of the 21 analytes, plus cut-offs for ratios to other analytes to screen the average 400 samples per day and place all required follow-up sample numbers on the next worklist.

This technique will become even more cost effective if additional analytes can be added. Our laboratory is investigating the use of MS/MS for screening for congenital adrenal hyperplasia and congenital hypothyroidism.

References

1. Millington DS, Kodo N, Norwood DL, Roe CR. Tandem mass spectrometry: a new method for acylcarnitine profiling with potential for neonatal screening for inborn errors of metabolism. J Inher Metab Dis 1990; 13: 321–4
2. Chace DH, Millington DS, Terada N, Kahler SG, Roe CR, Hofman LF. Rapid diagnosis of phenylketonuria by quantitative analysis for phenylalanine and tyrosine in neonatal blood spots by tandem mass spectrometry. Clin Chem 1993; 39: 66–71
3. Chace DH, Hillman SL, Millington DS, Kahler SG, Adam BW, Levy HL. Rapid diagnosis of homocystinuria and other hypermethioninemias from newborns' blood spots by tandem mass spectrometry. Clin Chem 1996; 42: 349–55
4. Chace DH, Hillman SL, Millington DS, Kahler SG, Roe CR, Naylor EW. Rapid diagnosis of maple syrup urine disease in blood spots from newborns by tandem mass spectrometry. Clin Chem 1995; 41: 62–8
5. Rashed MS, Bucknall MP, Little D, Awad A, Jacob M, Alamoudi M, et al. Screening blood spots for inborn errors of metabolism by electrospray tandem mass spectrometry with a microplate batch process and a computer algorithm for automated flagging of abnormal profiles. Clin Chem 1997; 43: 1129–41
6. Rashed MS, Rahbeeni Z, Ozand PT. Application of electrospray tandem mass spectrometry to neonatal screening. Semin Perinatol 1999; 23: 183–93
7. Guthrie R, Susi A. A simple phenylalanine method for detecting phenylketonuria in large populations of newborn infants. Pediatrics 1963; 32: 338–43
8. Guthrie R. Screening for inborn errors of metabolism in the newborn infant—a multiple test program. Birth Defects Original Article Series IV. 1962: 92–8

Carnitine transporter defect diagnosed by newborn screening with electrospray tandem mass spectrometry

Bridget Wilcken, MBChB, FRACP, Veronica Wiley, PhD, Keow Giak Sim, BSc, and Kevin Carpenter, PhD

The carnitine transporter defect is a potentially fatal but treatable disorder. We used electrospray tandem mass spectrometry in the New South Wales (Australia) Newborn Screening Programme to measure free carnitine and acylcarnitine species in the newborn population. Free carnitine levels in dried blood samples from 149,000 neonates did not vary markedly between 2 and 8 days of age. Two of 4 babies subsequently diagnosed clinically with the carnitine transporter defect had a free carnitine level in the neonatal blood sample low enough to be detected by screening. (J Pediatr 2001;138: 581-4)

Carnitine is necessary for the transport of long-chain fatty acids, as acyl coenzyme A esters, across the mitochondrial membrane and is thus an essential component of normal fatty acid oxidation. Carnitine is synthesized from lysine, but the greater part is derived from diet.[1] There is an active transport system across membranes in small intestine, renal tubule, and skeletal muscle, which is also expressed in skin fibroblasts.[2-4] Defective transport of carnitine across cell membranes causes a variety of potentially fatal symptom complexes in infancy and childhood, including neonatal death, hepatic encephalopathy, skeletal myopathy, and more commonly, dilated cardiomyopathy.[5] More than 40 cases have been recorded, variously described as primary carnitine deficiency, carnitine uptake defect, and plasma membrane carnitine transporter defect. The genetic defect is in a sodium ion–dependent carnitine transporter that has been mapped at 5q31.1.[6] Mutations in the gene, OCTN2, have been shown to impair carnitine transport in patients with the carnitine transporter defect.[7]

The carnitine transporter defect is treatable with oral carnitine therapy.[8] Clinical detection is only possible when there are significant symptoms, and affected patients have died of hepatic encephalopathy or cardiac failure before the diagnosis was made.[5] This makes the defect an ideal candidate for newborn screening. The recent introduction of electrospray tandem mass-spectrometry for newborn screening makes possible the diagnosis of a number of inborn errors of metabolism, including fatty acid oxidation defects, but the sensitivity of this technique for the various defects is so far unknown. We introduced this technology to the New South Wales Newborn Screening Programme in 1998.

We recently identified a neonate with carnitine transporter defect by routine newborn screening and retrospectively analyzed newborn screening samples from patients previously identified clinically to investigate the sensitivity of MS/MS for detecting this disorder in neonates.

MS/MS Tandem mass spectrometry

METHODS

Patients

The patient identified by newborn screening (patient 1) was a term baby with a birth weight of 3980 g. The newborn screening sample taken on day 4 had a low level of free carnitine, confirmed in a repeat sample on day 23 (Table I). He was breast fed and was well at all times. A baseline echocardiogram showed no abnormalities at 6 weeks, when carnitine therapy was started.

Five patients from 3 families have been identified clinically in New South

From New South Wales Newborn Screening and Biochemical Genetics Services, The New Children's Hospital, Westmead, Sydney, Australia.

Submitted for publication Mar 17, 2000; revisions received July 6, 2000, and Sept 11, 2000; accepted Sept 20, 2000.

Reprint requests: Bridget Wilcken, MBChB, FRACP, New Children's Hospital, PO Box 3515, Parramatta, New South Wales 2124, Australia.

0022-3476/2001/$35.00 + 0 **9/22/111813**

doi:10.1067/mpd.2001.111813

Table I. Babies with the carnitine transporter defect: Clinical features and free and acetyl carnitine species measured by electrospray MS/MS in routine newborn screening dried blood spots assayed either prospectively or retrospectively

Patient No.	Day of sample	Screened/retrospective	Clinical presentation	Free carnitine (μmol/L)	Acetyl carnitine (μmol/L)
1	4	Prospective	Clinically well, identified by screening	4	5
1	23	Prospective test, repeat sample	Clinically well	1	4
2	3	Pilot study	Cardiac failure at 18 mo	10	45
2	3	Sample re-assayed after storage for 2 y		11	10
3	4	Retrospective	Brother had cardiac failure at 17 mo and died at 18 mo; patient 3 well at 9 mo on diagnosis	7	†
4*	3	Retrospective	Sister had cardiac failure at 6 y; patient 4 diagnosed prenatally*	24*	†

Timing of samples in relation to feeding is unknown.
*Patient 4 was diagnosed prenatally. The mother received carnitine, 1.25 g/d, during pregnancy, and the baby was breast fed.
†Stored card: acetyl carnitine measurement was inaccurate.

Table II. Free carnitine levels (μmol/L whole blood) in the neonatal population, measured by tandem mass spectrometry

Neonate class	Number	Free carnitine median	Free carnitine 5th and 95th centiles	Number (%) <10 μmol/L (0.016 mg/dL)	Number <7 μmol/L	Number <5 μmol/L
Age 2 days	13,311	26.6	14.0-52.0	98 (0.73%)	4	0
Age 3 days	79,422	26.8	14.1–51.8	500 (0.63%)	45	3
Age 4 days	51,248	27.2	14.1–52.3	330 (0.64%)	30	2
Age 5-8 days	5,546	27.5	14.3–49.6	41 (0.74%)	7	1
Birth weight <2000 g	3,264	30.5	14.0–58.6	48 (1.47%)	17	2

Wales since 1990; the diagnosis was suspected after low levels of plasma free carnitine were found and was confirmed by uptake studies on cultured skin fibroblasts (Table I).[9]

Newborn screening blood samples from 3 patients were retrieved and tested retrospectively; one of them (patient 2) had had a prospective MS/MS test of the newborn screening sample during a pilot study.

Protocol

The MS/MS methods were those we routinely use in newborn screening. Briefly, free and acyl carnitines were extracted from blood samples dried on filter paper and converted to butyl esters before analysis by electrospray MS/MS (Micromass Quattro II).[10] Carnitine uptake in cultured skin fibroblasts was measured as described by Stanley et al.[11] Rates of fatty acid oxidation were measured by tritium water release assay as described by Manning et al.[12]

The low cutoff for free carnitine adopted for the newborn population was 10 μmol/L (0.16 mg/dL) for confirmation assay within the laboratory (to include at least the lowest 0.5% of values). A confirmed level of <5 μmol/L (0.08 mg/dL) generated a request for a second sample. The median level of dried blood spot acetyl carnitine was 44 μmol/L (0.9 mg/dL), and the low cutoff was 10 μmol/L (0.2 mg/dL). Our cutoff levels for other acylcarnitine species have been published.[10]

RESULTS

Table II shows the free carnitine levels measured by MS/MS of dried blood samples from the newborn population collected on days 2, 3 (the preferred day), 4, or 5 to 8. The percentage of samples with a level of <5 and <10 μmol/L (<0.08 and <0.16 mg/dL) was similar for each group. The distribution was skewed to the right. The median levels of free carnitine increased slightly from day 2 to days 5 to 8, from

26.6 to 27.5 μmol/L (0.43-0.44 mg/dL). Babies weighing <2000 g had somewhat higher median free carnitine levels, and only 2 of 3268 had a free carnitine level of <5 μmol/L (0.08 mg/dL), which was not significantly different from the proportion in normal birth weight babies. The affected baby (patient 1), identified by newborn screening, had a level of free carnitine below the cutoff on day 3, when the routine screening sample was taken. The affected baby (patient 2), born during the pilot study before our cutoff levels had been established, had a free carnitine level of 10.2 μmol/L (0.16 mg/dL) and so would not have been identified with our current protocol. Neither of the two babies retrospectively tested would have been diagnosed with the current protocol, although minor adjustments to this would have identified one of them. Reassay of the dried blood spot of patient 2 after an interval of 30 months showed that although the acetyl carnitine value had decreased greatly during storage, the free carnitine value was similar (Table I).

The carnitine uptake assay in cultured skin fibroblasts from all our patients with the carnitine transporter defect demonstrated very low uptakes of carnitine, 5% to 8% of control values. The tritiated water release assay on cultured skin fibroblasts, a general fatty acid oxidation screening assay, also provided abnormal results for patient 1, with low oxidation rates of both oleate and myristate of 21% of intra-batch control values for each.

DISCUSSION

We have shown that some cases of the carnitine transporter defect can be detected by expanded newborn screening. However, at least one and perhaps two of our clinically identified patients could not have been identified by measurement of free carnitine alone with any reasonable protocol. We have provided data on free carnitine levels in the healthy neonatal population for days 2 to 8 and for low birth weight babies. The measurements are of whole blood levels determined by MS/MS and do not equate with levels measured in plasma by this or other methods. Shenai et al[13] measured total carnitines in whole cord blood and found levels approximately 23% higher than our median level of free carnitine on day 2. These authors also substantiate that total carnitine levels are higher in red blood cells than in plasma. Because published plasma levels of free carnitine are similar to those we found in whole blood,[13] the excess is likely to be largely acetylcarnitine. There is not a sufficient change in whole blood free carnitine levels from day 2 to day 8 that would dictate different action levels for newborn screening dependent on day of screening.

In neonates, low free carnitine levels are seen principally in those who are sick. We found, in intensive care units, an over-representation of babies with levels <10 μmol/L (0.16 mg/dL). In neonates with a variety of inborn errors of organic acid or fatty acid metabolism, low free carnitine levels are usual, but generally other individual acyl carnitines will be elevated, and these elevations are readily detectable by MS/MS in the same assay. When there is very low free carnitine with no other elevated acylcarnitines, a carnitine transporter defect must be considered.

Of the two affected babies tested prospectively, one (patient 2) had a level well above our current cutoff for action. Of the two retrospectively studied, both had levels above this action level, although one level (patient 3) was only marginally so. The other retrospectively studied baby (patient 4) had been identified prenatally as being affected. Her mother was receiving carnitine medication during the pregnancy, and the baby was breast fed. This may have influenced her carnitine levels on day 3, when her screening sample was taken. The timing of the samples in relation to feeding is not known, but feeding regimens are an important determinant of postnatal carnitine levels.[13] Moreover, carnitine is actively transported across the placenta,[14] and maternal levels influence the neonatal load.[15] Babies with a defect in the carnitine transporter fail to reabsorb carnitine in the proximal renal tubule, and blood levels fall rapidly after birth.[9] In stored blood samples, acetyl carnitine apparently degrades, and measured levels are low. This is not accompanied by a significant rise in free carnitine measured by MS/MS.

To have a better chance of identifying all babies with the carnitine transporter defect, the cutoff level would need to be increased, at the cost of having more false-positive cases. It is arguable to what extent this is acceptable, especially for such an apparently rare disorder. In our program a cutoff level of 5 μmol/L led to a recall rate of only 0.004%, whereas adopting a higher cutoff of 10 μmol/L (0.16 mg/dL) would have generated a recall of 0.65%. This in itself is modest, but screening by MS/MS involves the analysis of 20 or more analytes for potential detection of a large number of disorders.[10] Each analyte will have an individual cutoff level, and sometimes ratios of two analytes may be used in addition. Because of this, the overall recall rate could easily become unacceptably high. For detecting the carnitine transporter defect, there is a case for a modest increase in the cutoff level of free carnitine to 6 or 7 μmol/L (0.1-0.11 mg/dL). A cutoff of 7 μmol/L (0.11 mg/dL) would have generated a recall rate of 0.058% in our series and would probably have identified one of the clinically diagnosed patients in addition.

It is known that obligate heterozygotes for the carnitine transporter defect may have levels of plasma carnitine below the normal range.[6] This implies that carriers, particularly when the mother is also a carrier, might be identified in newborn screening programs. It is not yet known whether the false-positive cases, babies with initially low

levels of carnitine but who do not have the transporter defect, will turn out to be carriers. The prevalence of carriers is not known in our population. Extrapolating from all cases diagnosed in New South Wales from 1990 to 1999, we would anticipate a carrier rate of 1:240 (95% CI 1:150-1:480).

The detection of the carnitine transporter defect in a neonate, reported here, helps to affirm the usefulness of the new trend in newborn screening, use of MS/MS to detect in a single test a wide range of disorders that are too rare to merit individual testing procedures. The carnitine transporter defect is but one of many rare but treatable disorders that have been detected by this form of screening.[10,16]

REFERENCES

1. Stanley CA. New genetic defects in mitochondrial fatty acid oxidation and carnitine deficiency. Adv Pediatr 1987; 34:59-88.
2. Shaw RD, Li BU, Hamilton JW, Shug AL, Olsen WA. Carnitine transport in rat small intestine. Am J Physiol 1983; 245:G376-81.
3. Rebouche CJ, Mack DL. Sodium gradient-stimulated transport of L-carnitine into renal brush border membrane vesicles: kinetics, specificity, and regulation by dietary carnitine. Arch Biochem Biophys 1984;235:393-402.
4. Rebouche CJ, Engel AG. Carnitine transport in cultured muscle cells and skin fibroblasts from patients with primary systemic carnitine deficiency. In Vitro 1982;18:495-500.
5. Stanley CA, DeLeeuw S, Coates PM, Vianey-Liaud C, Divry P, Bonnefont JP, et al. Chronic cardiomyopathy and weakness or acute coma in children with a defect in carnitine uptake. Ann Neurol 1991;30:709-16.
6. Tamai I, Ohashi R, Nezu J, Yabuchi H, Oku A, Shimane M, et al. Molecular and functional identification of sodium ion dependent, high affinity human carnitine transporter OCTN2. J Biol Chem 1998;273:20378-82.
7. Nezu J-I, Tamai I, Oku A, Ohashi R, Yabuchi H, Hashimoto N, et al. Primary systemic carnitine deficiency is caused by mutations in a gene encoding sodium ion-dependent carnitine transporter. Nat Genet 1999;21:91-4.
8. Pierpont ME, Breningstall GN, Stanley CA, Singh A. Familial carnitine transporter defect: a treatable cause of cardiomyopathy in children. Am Heart J 2000;139:96-106.
9. Christodoulou J, Teo SK, Hammond J, Sim KG, Hsu BYL, Stanley CA, et al. First prenatal diagnosis of the carnitine transporter defect. Am J Med Genet 1996;66:21-4
10. Wiley V, Carpenter KH, Wilcken B. Tandem mass spectrometry in the New South Wales Newborn Screening Programme. Acta Pediatr 1999;(Suppl 432):48-51.
11. Stanley CA, Treem WR, Hale DE, Coates PM. A genetic defect in carnitine transport causing primary carnitine deficiency. Prog Clin Biol Res 1990;321:457-64.
12. Manning NJ, Olpin SE, Pollitt RJ, Webley J. A comparison of [9,10-^{3}H]palmitic and [9,10-^{3}H]myristic acids for the detection of defects of fatty acid oxidation in intact cultured fibroblasts. J Inherit Metab Dis 1990;13:58-68.
13. Shenai JP, Borum PR, Mohan P, Donlevy SC. Carnitine status at birth of newborn infants of varying gestation. Pediatr Res 1983;17:579-82.
14. Tamai I, Ohashi R, Nezu J, Yabuchi H, Shimane M, Sai Y, et al. Molecular and functional identification of sodium ion-dependent, high affinity human carnitine transporter. J Biol Chem 1998;273:20378-82.
15. Sachan DS, Smith RB, Plattsmier J, Lorch V. Maternal, cord, and neonatal carnitine correlations and lipid profiles of various birthweight infants. Am J Perinatol 1989;6:14-7.
16. Levy HL. Newborn screening by tandem mass spectrometry: a new era. Clin Chem 1998;44:2401-2.

J. Inherit. Metab. Dis. 24 (2001) 51–59

Carnitine palmitoyltransferase I deficiency in neonate identified by dried blood spot free carnitine and acylcarnitine profile

K. G. Sim, V. Wiley, K. Carpenter and B. Wilcken*
New South Wales Biochemical Genetics Services and Newborn Screening, The Children's Hospital at Westmead (Sydney), Locked Bag 4001, Westmead, NSW 2145, Australia

* *Correspondence: E-mail: bridgetw@chw.edu.au*

MS received 17.10.00 Accepted 9.11.00

Summary: A neonate at risk for hepatic carnitine palmitoyltransferase I (L-CPT I) deficiency was investigated from birth. The free carnitine and acylcarnitine profile in dried whole blood filter paper samples collected at ages 1 and 4 days showed a markedly elevated concentration of free carnitine (141 and 142 μmol/L, respectively), normal concentrations of acetyl- and propionylcarnitine, with the near absence of all other species. The diagnosis was confirmed by *in vitro* fatty acid oxidation screening assays and enzyme assay in cultured skin fibroblasts. Retrospective study of the newborn whole blood sample of the index case showed a similar profile (free carnitine 181 μmol/L). The newborn population distribution of free carnitine ($n = 143981$) showed that only three samples had free carnitine >140 μmol/L (>99.9th centile), two were from L-CPT I-deficient neonates and one from a baby with sepsis. While there are other conditions that can cause elevated concentrations of free carnitine, an isolated elevation of free carnitine only in an apparently healthy term neonate warrants further investigation to exclude L-CPT I deficiency.

Long-chain fatty acids serve as an important source of energy, and mitochondrial fatty acid β-oxidation is the major process of energy production, especially during conditions of fasting or metabolic stress. Carnitine palmitoyltransferase I (EC 2.3.1.21), is one of the enzymes of the carnitine cycle that facilitates the transport of long-chain fatty acids across the mitochondrial membrane for β-oxidation. It catalyses the esterification of fatty acyl-CoA to its corresponding acylcarnitine, requiring carnitine as a co-substrate, and its catalytic activity is sensitive to malonyl-CoA inhibition (McGarry et al 1978).

Two different isoforms of CPT I have been described, a liver type (L-CPT I) and a muscle type (M-CPT I). Their gene map loci have been identified (Britton et al 1997; Yamazaki et al 1996). Studies of rat tissues indicate L-CPT I is the primary isoform found in liver, kidney, lung, spleen, intestine and brain. (Brown et al 1997). In humans, both isoforms are found in the heart (McGarry and Brown 1997), and L-CPT I is expressed in liver and is the same isoform expressed in fibroblasts (Britton et al 1995).

L-CPT I deficiency (McKusick 255120) is a mitochondrial β-oxidation disorder with autosomal recessive inheritance, usually presenting in infancy and characterized by recurrent episodes of hypoketotic hypoglycaemia, hepatomegaly, hepatic dysfunction, seizure and coma, similar to Reye syndrome, and triggered by fasting or intercurrent illness. It was first reported in 1981 (Bougneres et al 1981), and 14 families have been described to date (Bonnefont et al 1999). The first delineation of the molecular defect was reported in 1998 (Ijlst et al 1998). Two L-CPT I disease-causing mutations have since been identified (Abadi et al 1999; Gobin et al 2000). The biochemical findings in presenting samples of affected patients have been quite subtle: organic acid profiles are usually unremarkable; hypoglycaemia is a common feature but could be absent (Vianey-Saban et al 1993). The plasma total and free carnitine concentrations may be in the high-normal range; however, normal levels were reported in two affected neonates in one family (Innes et al 2000). We report the identification of L-CPT I deficiency in a neonate (baby S.H.), who has an affected sibling J.H. (the index case), by free carnitine and acylcarnitine analysis of dried blood spot samples. Retrospective analysis of the newborn sample from the index case showed a similar acylcarnitine profile.

PATIENTS AND METHODS

Baby S.H. was the third child of unrelated Maori parents, born in New Zealand at full term after a normal pregnancy. He was breast fed, with supplementary feeding for the first two days until breast feeding was well established. Samples of whole blood dried on filter paper were collected soon after birth and on day 4 of life. His, brother J.H., the index case, was a second child in the family, and had been healthy until at 7 months he had a seizure and collapsed after an episode of fever and intermittent vomiting. Initial findings included a large, firm liver and elevated serum transaminases, creatine kinase and triglycerides, and normal plasma ammonium. Liver biopsy showed microvesicular and macrovesicular steatosis and excessive lipid inclusions in bone marrow. He suffered an ischaemic brain insult resulting in profound neurological deficit, seizure and impairment of vision and hearing, but has remained otherwise well on a low long-chain fat diet supplemented with medium-chain triglycerides, and a protocol designed to avoid fasting.

Unaffected infants were investigated as part of the routine testing in the New South Wales (Australia) Newborn Screening Programme (Wiley et al 1999).

Free carnitine and acylcarnitine analysis in dried blood samples

Sample preparation for the extraction of free carnitine and acylcarnitines from whole blood samples dried on filter paper was as described (Wiley et al 1999). Briefly, a 3 mm blood disk was delivered into a deep-well microtitre plate, and methanol containing deuterium-labelled internal standards (Cambridge Isotope Labs) was added, and shaken for 1 h. (Shaking for 1 h produced a more reliable analytical recovery compared with the 30 min shaking as used by others (Chace et al 1996)). The supernatant was transferred to another microtitre plate, dried and derivatized with butanolic HCl and the butyl esters of acylcarnitines were analysed by electrospray tandem mass spectrometer (Micromass Quattro II/LC), detecting the precursors of m/z 85.2. Data of the newborn population aged 2–4 days were acquired by multiple reaction-monitoring mode (MRM), which specifies which analytes will be detected, hence providing a better sensitivity. Data of the acylcarnitine profiling were analysed by scan mode, multichannel acquisition (MCA), in order to detect all the butylated acylcarnitine species in the samples.

Oxidation rate study

Skin fibroblasts were grown in Ham's F10 nutrient mixture supplemented with 10% fetal calf serum, 2 mmol/L glutamine, 1% penicillin, streptomycin and fungizone. Rates of fatty acid oxidation in fibroblasts were estimated by the tritium water release assay employing two substrates, [9, 10(n)-^{3}H]oleate and [9, 10(n)-^{3}H]-myristate, as described (Olpin et al 1997).

Fibroblast acylcarnitine profiling study

The *in vitro* analysis of the mitochondrial β-oxidation pathway to study the acylcarnitine profiles generated by cultured skin fibroblasts was as described (Nada et al 1996) with modification. Fibroblasts were incubated in reaction medium supplemented with palmitic acid (C_{16} fatty acid) and carnitine, and incubation was at 37°C for 72 h. The acylcarnitine profiles of the culture media were acquired by the MCA scan mode and processed using Neolynx software. The protein content of the cell monolayers was measured by the Pierce bicinchoninic acid (BCA) method on a COBAS BIO centrifugal analyser.

L-CPT I activity

L-CPT I activity in cultured skin fibroblasts was measured as described previously (Demaugre et al 1988) with modification. The assay measures the formation of palmitoyl-[*methyl*-^{14}C]carnitine from palmitoyl-CoA and [*methyl*-^{14}C]carnitine in cells permeabilized by digitonin, in reaction mixtures with and without the addition of malonyl-CoA.

RESULTS

Table 1 shows the dried whole blood sample levels of free carnitine in the New South Wales newborn population aged 2 to 4 days. The population distribution of free carnitine in 143 981 neonates had a median of 27 μmol/L whole blood (1st to 99th centile, 11–72 μmol/L). In the neonatal population only one baby (0.0007%) had free carnitine concentration above 140 μmol/L (>99.9th centile), a baby with sepsis whose free carnitine concentration was 165 μmol/L on sample collected on day 3. Baby S.H. had free carnitine of 141 and 142 μmol/L in samples collected on days 1 and 4, respectively. Retrospective study of the newborn whole blood sample from the index case, J.H., collected on day 3, showed a free carnitine concentration of 181 μmol/L.

The acylcarnitine profile of the dried whole blood sample of baby S.H. collected on day 4 showed a free carnitine concentration of 142 μmol/L, acetylcarnitine 12 μmol/L and propionylcarnitine 3 μmol/L, with the near absence of all other acylcarnitines. The median value for whole blood acetylcarnitine in the newborn population was 47 μmol/L (1st to 99th centile, 8–160). The median value for propionylcarnitine was 1.5 μmol/L, with a cut-off for elevated concentrations of 9 μmol/L. Cut-off values for other acylcarnitines have been published (Wiley et al 1999).

Oxidation rates of [9,10(*n*)-^{3}H]myristate (M) and [9,10(*n*)-^{3}H]oleate (O) on intact cultured skin fibroblasts from the index case were 10% and 5% of mean control values, respectively, and those of baby S.H. were reduced to 4% for both substrates, with O/M ratio of 1.

The *in vitro* fibroblast acylcarnitine profiling study for the cell line from baby S.H. showed reduced formation of acylcarnitine species of even-numbered chain lengths, those derived from the oxidation of fatty acids. Values of palmitoylcarnitine,

Table 1 Free carnitine in dried whole blood on filter paper measured by electrospray tandem mass spectrometry

Newborns aged 2–4 days	*Carnitine* (μmol/L whole blood)		*Number of cases (%) with free carnitine levels* (μmol/L) *of*					
	Median	*1st–99th centiles*	*100–<120*	*120–<130*	*130–<140*	*140–<150*	*150–<160*	*>160*
$n = 143\,981$	27	11–72	172 (0.120%)	18 (0.014%)	2 (0.001%)	Nil	Nil	1 (0.0007%)
CPT I cases Newborn sample (μmol/L whole blood)						S.H. (142)		J.H. (181)

Table 2 Acylcarnitine levels (nmol/mg protein) in culture medium accumulated after incubation for 72 h

Cell lines	C_2	C_3	C_4	C_5	C_6	C_8	C_{10}	C_{12}	C_{14}	C_{16}
Baby S.H.	28	7	1.5	0.9	0.3	0	0.1	0.1	0.0	0.2
Normals										
Mean ($n = 11$)	40	9	2.5	1.1	0.8	0.9	1.1	0.2	0.2	1.0
Mean±SD	26–53	5–13	1.5–3.5	0.8–1.5	0.5–1.2	0.3–1.4	0.5–1.7	0.1–0.3	0.1–0.2	0.6–1.4

C_2 to C_{16} represent the chain length of the respective acylcarnitine species Data for Baby S.H. are the mean of duplicate analysis

decanoylcarnitine, octanoylcarnitine and hexanoylcarnitine were below one SD of mean controls (n=11) (Table 2).

The L-CPT I specific activity in cultured fibroblasts for both patients, identified by its inhibition by malonyl-CoA, was reduced to less than 1% of normal control values (Table 3). The parent's fibroblasts were not available.

DISCUSSION

We investigated baby S.H., who was at 1 : 4 risk of L-CPT I deficiency from birth and found that the free carnitine concentration in dried whole blood samples was among the highest seen in a newborn population distribution, without any elevation of other acylcarnitine species. The cultured skin fibroblast oxidation rate and O/M ratio studies indicated a defect in long-chain fatty acid oxidation not related to the β-oxidation spiral (Olpin et al 1997, 1999). The *in vitro* fibroblast acylcarnitine profile revealed an overall reduced formation of acylcarnitines, especially palmitoylcarnitine, suggestive of a defect upstream of carnitine palmitoyltransferase II or carnitine acylcarnitine translocase (Roe and Roe 1999; Ventura et al 1999), and

Table 3 CPT activities were assayed in fibroblasts as described in Methods and are expressed in nmol labelled palmitoyl-L-carnitine produced per min per milligram protein

	Total activities (Sum of L-CPT I, CPT II carnitine acyltransferases)	*Malonyl-CoA-insensitive components (Sum of CPT II, carnitine acyltransferases)*	*L-CPT I (Difference of total and malonyl-CoA-insensitive components)*
Baby S.H.	0.10	0.09	0.01
Index case J.H.	0.07	0.05	0.02
Control cell lines (n=3)	0.99–2.12	0.09–0.26	0.90–1.91

Data are means of duplicates

CPT I deficiency seemed a possibility (Schaefer et al 1997). The results of the fibroblast CPT enzyme assay firmly established that baby S.H. and the index case suffered from L-CPT I deficiency.

The newborn population distribution of whole blood free carnitine is positively skewed (skewness 1.5), so that there is a continuum of values in the 'tail' from 100 to 130 μmol/L (Table 1) with 0.13% of the total population having a concentration above 100 μmol/L whole blood. The free carnitine concentrations of dried whole blood extracted from filter paper measured by tandem mass spectrometry (Wilcken et al 2001) are different from concentrations measured in plasma by other methods (Schmidt-Sommerfeld et al 1988). A greater proportion of carnitine is found in the erythrocytes compared to plasma (Shenai et al 1983). Hence, the concentration of free carnitine measured in dried whole blood would be expected to be higher than in plasma. Borum and colleagues measured total carnitines in cord blood and found concentrations approximately 23% higher than our median concentration of free carnitine (Shenai et al 1983). The concentrations in newborn screening samples from the L-CPT I affected patients, well at the time of sampling, were greatly elevated, and quite distinct from the newborn population as a whole. The other neonate who had an elevated free carnitine level (165 μmol/L) had sepsis when the sample was collected. Although elevated concentrations of free carnitine may be seen in a critically ill neonate as in this case, it seems rare, as there is an over-representation of sick neonates among those with low concentrations of circulating free carnitine. In fact, they are at risk of carnitine deficiency because of the immaturity of their carnitine-synthesizing enzymes and of their mechanisms devoted to carnitine conservation (Borum 1995). Perimortem or postmortem samples frequently exhibit grossly elevated free carnitine, but the concentrations of all other acylcarnitines are usually elevated concomitantly. This may be due to the autolysis of muscle tissue, which has a high intracellular carnitine concentration. Other conditions associated with an elevated concentration of carnitine include rhabdomyolysis and cardiomyopathy. Elevated concentrations of serum carnitine have been reported in patients with idiopathic hypertrophic cardiomyopathy, due to reduced carnitine uptake into the myocardium (Nakamura et al. 1999). A high concentration of circulating free carnitine has not been reported to cause any adverse clinical outcome as judged by those patients on carnitine therapy. Because of this, there is a danger that an elevated concentration of free carnitine in neonates might be considered an insignificant finding.

Fatty acid oxidation disorders are generally associated with a decreased concentration of plasma total and free carnitine, and the elevation of short-, medium- or long-chain acylcarnitines. In contrast, in L-CPT I deficiency, the plasma free carnitine concentrations are usually elevated, with the absence of other acylcarnitine species except for acetylcarnitine. These phenomena could be due to an unusually high renal threshold for free carnitine (Stanley et al 1992) and the defective esterification of acyl-CoA to carnitine. The dried blood spot acylcarnitine profile of baby S.H. showed an isolated, persistent, marked elevation of free carnitine, and near absence of long-chain acylcarnitines. This contrasts with a recent report,

where whole blood acylcarnitine profiles and serum total and free carnitine were normal in two siblings with L-CPT I deficiency (Innes et al 2000).

In almost all reported cases, the prominent organ dysfunction in clinically presenting L-CPT I deficiency is the liver. Heart involvement is generally absent. Although several cases with slight cardiomegaly or arrhythmias have been reported, most have recovered spontaneously (Bergman et al 1994; Schaefer et al 1997; Tein et al 1989; Vianey-Saban et al 1993). This may be related to the presence of CPT I isoforms and their distinct tissue distributions. Both L-CPT I and M-CPT I are present in human heart with M-CPT I being the predominant isoform. A study of the developing rat heart showed that L-CPT I contributed approximately 25% to total CPT I activity, which declined to 2–3% in the adult (Brown et al 1995). It is not known whether these findings are paralleled in man. Reported cases have usually suffered irreversible neurological impairment following an episode of metabolic crisis. This is presumably due to hypoketotic hypoglycaemia, since the brain does not directly use fatty acids for oxidative metabolism but oxidizes ketone bodies derived from acetyl-CoA and acetoacetyl-CoA produced by β-oxidation of fatty acids in the liver (Mitchell et al 1995).

In conclusion, the finding of an isolated elevation of free carnitine in an apparently healthy term neonate is most unusual and warrants further investigation to rule out L-CPT I deficiency. Early recognition is important because institution of preventive measures can be achieved by relatively simple dietary treatment, and without intervention there is significant morbidity and mortality.

REFERENCES

Abadi N, Thuillier L, Prasad C, et al (1999) Molecular resolution of carnitine palmitoyltransferase I deficiency in a Hutterite family. *Proceedings of the 49th Annual Meeting of the American Society of Human Genetics*, San Francisco (abstract).

Bergman AJ, Donckerwolcke RA, Duran M, et al (1994) Rate-dependent distal renal tubular acidosis and carnitine palmitoyltransferase I deficiency. *Pediatr Res* **36**: 582–588.

Bonnefont JP, Demaugre F, Prip-Buus C, et al (1999) Carnitine palmitoyltransferase deficiencies. *Mol Genet Metab* **68**: 424–440.

Borum PR (1995) Carnitine in neonatal nutrition. *J Child Neurol* **10**(supplement 2): S25–S31.

Bougneres PF, Saudubray JM, Marsac C, Bernard O, Odievre M, Girard J (1981) Fasting hypoglycemia resulting from hepatic carnitine palmitoyl transferase deficiency. *J Pediatr* **98**: 742–746.

Britton CH, Schultz RA, Zhang B, Esser V, Foster DW, McGarry JD (1995) Human liver mitochondrial carnitine palmitoyltransferase I: characterization of its cDNA and chromosomal localization and partial analysis of the gene. *Proc Nat Acad Sci USA* **92**: 1984–1988.

Britton CH, Mackey DW, Esser V, et al (1997) Fine chromosome mapping of the genes for human liver and muscle carnitine palmitoyltransferase I (CPT IA and CPT IB). *Genomics* **40**: 209–211.

Brown NF, Weis BC, Husti JE, Foster DW, McGarry JD (1995) Mitochondrial carnitine palmitoyltransferase I isoform switching in the developing rat heart. *J Biol Chem* **270**: 8952–8957.

Brown NF, Hill JK, Esser V, et al (1997) Mouse white adipocytes and 3T3-L1 cells display an anomalous pattern of carnitine palmitoyltransferase (CPT) I isoform expression during differentiation. Inter-tissue and inter-species expression of CPT I and CPT II enzymes. *Biochem J* **327**(Pt 1): 225–231.

Chace DH, Hillman SL, Millington DS, Kahler SG, Adam BW, Levy HL (1996) Rapid diagnosis of homocystinuria and other hypermethioninemias from newborns' blood spots by tandem mass spectrometry. *Clin Chem* **42**: 349–355.

Demaugre F, Bonnefont JP, Mitchell G, et al (1988) Hepatic and muscular presentations of carnitine palmitoyl transferase deficiency: two distinct entities. *Pediatr Res* **24**: 308–311.

Gobin S, Bonnefont J-P, Prip-Buus C, et al (2000) Characterizing the organization of the "liver-type" carnitine palmitoyltransferase 1 (L-CPT I) gene helps unravelling the molecular basis of L-CPT I deficiency in human. *Proceedings of the 50th Annual Meeting of the American Society for Human Genetics, Philadelphia*, Abstract 278.

Ijlst L, Mandel H, Oostheim W, Ruiter JP, Gutman A, Wanders RJ (1998) Molecular basis of hepatic carnitine palmitoyltransferase I deficiency. *J Clin Invest* **102**: 527–531.

Innes AM, Seargeant LE, Balachandra K, et al (2000) Hepatic carnitine palmitoyltransferase I deficiency presenting as maternal illness in pregnancy. *Pediatr Res* **47**: 43–45.

McGarry JD, Brown NF (1997) The mitochondrial carnitine palmitoyltransferase system. From concept to molecular analysis. *Eur J Biochem* **244**: 1–14.

McGarry JD, Leatherman GF, Foster DW (1978) Carnitine palmitoyltransferase I. The site of inhibition of hepatic fatty acid oxidation by malonyl-CoA. *J Biol Chem* **253**: 4128–4136.

Mitchell GA, Kassovska-Bratinova S, Boukaftane Y, et al (1995) Medical aspects of ketone body metabolism. *Clin Invest Med* **18**: 193–216.

Nada MA, Vianey-Saban C, Roe CR, et al (1996) Prenatal diagnosis of mitochondrial fatty acid oxidation defects. *Prenat Diagn* **16**: 117–124.

Nakamura T, Sugihara H, Kinoshita N, et al (1999) Serum carnitine concentrations in patients with idiopathic hypertrophic cardiomyopathy: relationship with impaired myocardial fatty acid metabolism. *Clin Sci (Colch)* **97**: 493–501.

Olpin SE, Manning NJ, Pollitt RJ, Clarke S (1997) Improved detection of long-chain fatty acid oxidation defects in intact cells using [9,10-^{3}H]oleic acid. *J Inherit Metab Dis* **20**: 415–419.

Olpin SE, Manning NJ, Pollitt RJ, Bonham JR, Downing M, Clark S (1999) The use of [9,10-^{3}H]myristate, [9,10-^{3}H]palmitate and [9,10-^{3}H]oleate for the detection and diagnosis of medium and long-chain fatty acid oxidation disorders in intact cultured fibroblasts. *Adv Exp Med Biol* **466**: 321–325.

Roe CR, Roe DS (1999) Recent developments in the investigation of inherited metabolic disorders using cultured human cells. *Mol Genet Metab* **68**: 243–257.

Schaefer J, Jackson S, Taroni F, Swift P, Turnbull DM (1997) Characterisation of carnitine palmitoyltransferases in patients with a carnitine palmitoyltransferase deficiency: implications for diagnosis and therapy. *J Neurol Neurosurg Psychiatry* **62**: 169–176.

Schmidt-Sommerfeld E, Werner D, Penn D (1988) Carnitine plasma concentrations in 353 metabolically healthy children. *Eur J Pediatr* **147**: 356–360.

Shenai JP, Borum, PR, Mohan P, Donlevy SC (1983) Carnitine status at birth of newborn infants of varying gestation. *Pediatr Res* **17**: 579–582.

Stanley CA, Sunaryo F, Hale DE, Bonnefont JP, Demaugre F, Saudubray JM (1992) Elevated plasma carnitine in the hepatic form of carnitine palmitoyltransferase-1 deficiency. *J Inherit Metab Dis* **15**: 785–789.

Tein I, Demaugre F, Bonnefont JP, Saudubray JM (1989) Normal muscle CPT I and CPT2 activities in hepatic presentation patients with CPT I deficiency in fibroblasts. Tissue specific isoforms of CPT I? *J Neurol Sci* **92**: 229–245.

Ventura FV, Costa CG, Struys EA, et al (1999) Quantitative acylcarnitine profiling in fibroblasts using [U-^{13}C] palmitic acid: an improved tool for the diagnosis of fatty acid oxidation defects. *Clin Chim Acta* **281**: 1–17.

Vianey-Saban C, Mousson B, Bertrand C, et al (1993) Carnitine palmitoyl transferase I deficiency presenting as a Reye-like syndrome without hypoglycaemia. *Eur J Pediatr* **152**: 334–338.

Wilcken B, Wiley V, Sim K, Carpenter K (2001) Carnitine transporter defect diagnosed by newborn screening using electrospray tandem mass spectrometry. *J Pediatr* (in press).

Wiley V, Carpenter K, Wilcken B (1999) Newborn screening with tandem mass spectrometry: 12 months' experience in NSW Australia. *Acta Paediatr Suppl* **88**: 48–51.

Yamazaki N, Shinohara Y, Shima A, Yamanaka Y, Terada H (1996) Isolation and characterization of cDNA and genomic clones encoding human muscle type carnitine palmitoyltransferase I. *Biochim Biophys Acta* **1307**: 157–161.

Available online at www.sciencedirect.com
ScienceDirect
Molecular Genetics and Metabolism 90 (2007) 446–448
Molecular Genetics and Metabolism
www.elsevier.com/locate/ymgme

Brief Communication

Non-Ketotic Hyperglycinemia is usually not detectable by tandem mass spectrometry newborn screening

E.S. Tan [a], V. Wiley [b], K. Carpenter [c], B. Wilcken [b,c,*]

[a] *Genetic Metabolic Disorders Service, The Children's Hospital at Westmead, Sydney, Australia*
[b] *NSW Newborn Screening Program, The Children's Hospital at Westmead, Corner Hawkesbury and Hainsworth Street, Locked Bag 4001, Westmead, NSW 2145, Sydney, Australia*
[c] *NSW Biochemical Genetics Service, The Children's Hospital at Westmead, Corner Hawkesbury and Hainsworth Street, Locked Bag 4001, Westmead, NSW 2145, Sydney, Australia*

Received 24 October 2006; received in revised form 27 November 2006; accepted 27 November 2006
Available online 4 January 2007

Abstract

Diagnosis of Non-Ketotic Hyperglycinemia by MSMS newborn screening might benefit patients with post-neonatal presentation. We screened 733,527 babies over eight years, and nine babies were subsequently diagnosed with NKHG. Two had newborn glycine levels above our cut-off and presented within 72 h. The remaining patients could not have been diagnosed by newborn screening without an unacceptably high recall rate. We conclude that babies with NKHG are not usually identifiable by current newborn screening strategies.

Keywords: Non-Ketotic Hyperglycinemia; Newborn screening; Glycine; Tandem mass spectrometry

Introduction

Non-Ketotic Hyperglycinemia (NKH) (OMIM 605899) is a devastating metabolic condition caused by a defect in the glycine cleavage system, a mitochondrial enzyme complex of four proteins: P-protein (pyridoxal phosphate containing), H-protein (lipoic acid containing), T-protein (tetrahydrofolate-requiring, amino methyltransferase) and L-protein (lipoamide dehydrogenase). A defect in this system leads to accumulation of glycine, particularly in the central nervous system. Glycine levels are increased in both plasma and cerebrospinal fluid with an increased CSF/plasma ratio [1]. However, in some cases, this ratio is not diagnostic and confirmatory diagnosis requires liver glycine cleavage complex assays [2].

The classical form of NKH presents early in the neonatal period with intractable seizures and encephalopathy. There is a high mortality rate and survivors have significant neurological sequelae including profound mental retardation [3]. Current treatment strategies involving the use of sodium benzoate and dextromethorphan have not been successful [4–6]. Atypical forms of NKH have been described with age of onset ranging from infancy to adulthood [7]. Clinical presentation is varied and includes seizures, motor and/or language developmental delay, intermittent or persistent ataxia and choreoathetosis, behavioural problems and attention deficit hyperactive disorder. It has been suggested that early institution of therapy may improve the neurological sequelae of these patients.

Since the advent of newborn screening for inborn errors of metabolism using tandem mass spectrometry (MSMS), early diagnosis is possible for many conditions. The range of disorders screened via MSMS includes inborn errors affecting the metabolism of the urea cycle, amino acids, organic acids and fatty-acid oxidation. Tandem mass spectrometry was introduced in New South Wales and the Australian Capital Territory in early 1998, in South Australia in 1999, and in Victoria in 2002 [8].

* Corresponding author. Fax: +61 2 9845 3121.
E-mail address: Bridgetw@chw.edu.au (B. Wilcken).

1096-7192/$ - see front matter © 2006 Elsevier Inc. All rights reserved.
doi:10.1016/j.ymgme.2006.11.010

Glycine levels can be measured by MSMS in the routine newborn screening blood samples. NKH is not one of the disorders recently recommended for newborn screening in the United States [9], but early diagnosis of NKH might be particularly significant for atypical forms where early intervention may prove beneficial. We have evaluated the use of current newborn screening techniques in diagnosing cases of NKH.

Methods

Newborn screening bloodspot samples were collected between 48 and 72 h of life and analysed using MSMS as previously described [10]. To minimise the false positive rate, action limits were initially set after analysing archived dried blood spot samples from babies proven to have NKH ($N=2$) plus 25,000 consecutive newborn screening samples. Action limits were reviewed at least annually with regard to true positive and known false negative cases, the recall rate, and the potential for benefit. We ceased recalling patients with elevated glycine levels in 2004 after screening 624,000 babies. We reviewed the newborn blood glycine levels in all newborn screening samples and known cases of NKH from April 1998 to March 2006

Patients with the confirmed diagnosis of NKH born between April 1998 and March 2006 were identified from our clinical database. We reviewed the clinical records and newborn glycine levels of these patients. All patients had plasma and cerebrospinal fluid glycine levels measured by our biochemical genetics laboratory at the time of diagnosis. Plasma and cerebrospinal fluid amino acid quantification was carried out using ion exchange chromatography with post column ninhydrin derivatisation. A CSF:plasma glycine ratio of >0.05 was considered diagnostic in our laboratory. Liver glycine cleavage enzyme assays were performed in cases with CSF:plasma glycine ratios of less than 0.05.

Results

From April 1998 to March 2006, 733,527 babies were screened by tandem mass spectrometry in our centre. The median, upper 95th and 99.9th centiles for dried blood spot glycine were 280, 549 and 1036 μmol/L, respectively. The observed distribution of newborn blood glycine levels is shown in Fig. 1. The cut-off value had been set at 1000 μmol/L.

From this cohort, 9 babies were subsequently diagnosed with NKH. The newborn blood glycine level in the patients with NKH ranged from 242 to 1210 μmol/L. Two patients

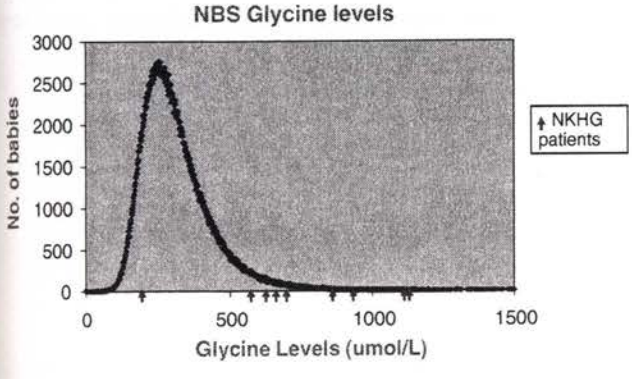


Fig. 1. Glycine levels (μmol/L) in 733,527 newborn screening samples, including 9 NKHG patients.

Table 1
Clinical and biochemical data

Patient	Age of presentation	Clinical presentation	Age at testing	Plasma glycine[a] (μmol/L)	CSF glycine[b] (μmol/L)	CSF:plasma glycine ratio	NBS glycine μmol/L (centile)	Enzyme confirmation	Progress
1	4.5 Months	Seizures, developmental delay	6 Months	488	43	0.09	671	N	Global development delay, dystonic
2	4 Months	Seizures, developmental delay	8 Months	359	25	0.07	242	N	Moved overseas
3	3 Days	Seizures, poor feeding	6 Days	1162	176	0.152	1210	N	Special school, marked spasticity
4	1 Day	Encephalopathy	5 Days	1347	262	0.2	1210	N	Died day 7
5	2.5 Years	Abnormal gait	2.5 Years	844	23	0.027	530	Y	Spastic diplegia, mild intellectual delay
6	2 Days	Recurrent apnea	7 Days	1760	405	0.23	565	N	Extubated on day 7 for conservative management
7	3 Days	Lethargy and poor suck	8 Days	852	289	0.339	810	N	Severe intellectual delay, dystonia, seizures
8	6 Months	Seizures, developmental delay	6 Months	961	52	0.054	909	Y	Died at 13 months
9	NA	Previously diagnosed sibling	3 Days	713	47	0.07	580	N	Mild motor delay

[a] Plasma glycine reference ranges: 0–28 days, 160–527; >28 days, 119–368 μmol/L.
[b] CSF glycine reference ranges: 0–90 days, 0–14; >90 days, 3–9 μmol/L.

had levels above 1000 μmol/L, but the remaining 7 could not have been diagnosed by newborn screening without a recall rate for glycine alone of 0.3–72% of babies born (Table 1).

Of the 9 patients, 4 presented with classical NKH and were symptomatic within the first 72 h of life. These included the 2 cases with newborn blood glycine levels above our cut off value of 1000 μmol/L. However, as these patients presented with severe neurological symptoms within 72 h of life, they were clinically diagnosed and treatment was instituted within the first week. Newborn screening would not have been useful as a tool of early diagnosis. Four of the 9 patients presented later in life, ranging from 4 months to 2.5 years of age. None of them had significantly raised newborn blood glycine levels. Another patient was diagnosed while he was asymptomatic as he had a sibling previously diagnosed with NKH. The clinical and biochemical features of these patients are described in Table 1.

In total, 35 babies (among 624,000) had repeat samples collected due to an isolated increase of glycine and only 2 were subsequently confirmed to have NKH. Therefore, the positive predictive value of the newborn glycine level for NKH is 5.7%, but the false negative rate was 78%.

Discussion

The only cases identifiable by newborn screening in our cohort of patients were severe classical forms of NKH. In both cases, the diagnosis of NKH was evident clinically and confirmatory diagnosis with plasma and CSF glycine levels was performed quickly. Newborn screening results had no clinical impact in the management of the patients. Despite early therapeutic intervention, such cases have a rapidly deteriorating course with severe neurological sequelae or high mortality. The 2 other cases with severe, classical presentations had newborn blood glycine levels within the normal range.

The cases that may have benefited from early diagnosis were those that presented beyond the neonatal period and with less severe symptoms. This is especially true with the expanding clinical spectrum of atypical NKH. Clinical suspicion may not be high and diagnosis may be delayed. Early diagnosis and intervention might help to improve the neurological outcome of these patients. However, all our cases with atypical NKH had newborn glycine levels within the normal range. To adjust our cut-off value for glycine to include these cases would result in an unacceptable recall rate. We ceased recalling patients for elevated glycine levels after reviewing the recall rates that would have been necessary to identify patients with a post-neonatal presentation.

In conclusion, babies with NKH do not usually have a sufficiently elevated blood glycine level at 48–72 h of life for identification by current newborn screening strategies. None of those with a post-neonatal presentation could have been identified by our programme without an unacceptable recall rate. Newborn screening for NKH is unlikely to be beneficial, as those babies who might benefit from early treatment could not be identified.

References

[1] D.A. Applegarth, J.R. Toone, Nonketotic Hyperglycinemia (glycine encephalopathy): laboratory diagnosis, Mol. Genet. Metab. 74 (2001) 139–146.

[2] K. Hayasaka, K. Tada, G. Kikuchi, S. Winter, W.L. Nyhan, Nonketotic Hyperglycinemia: two patients with primary defects of P-protein and T-protein, respectively, in the glycine cleavage system, Pediatr. Res. 17 (1983) 967–970.

[3] J.E. Hoover-Fong, S. Shah, J.L.K. Van Hove, D. Applegarth, J. Toone, A. Hamosh, Natural history of Nonketotic Hyperglycinemia in 65 patients, Neurology 63 (2004) 1847–1853.

[4] S.H. Korman, I.D. Wexler, A. Gutman, M.O. Rolland, J. Kanno, S. Kure, Treatment from birth of Nonketotic Hyperglycinemia due to a novel GLDC mutation, Ann. Neurol. 59 (2) (2006) 411–415.

[5] A. Hamosh, M. Johnston, Nonketotic Hyperglycinemia, in: C.R. Scriver, A.L. Beaudet, W.S. Sly, D. Valle (Eds.), The Metabolic and Molecular Bases of Inherited Disease, eighth ed., McGraw-Hill, New York, 2001, pp. 2065–2076.

[6] E. Zammarchi, M.A. Donati, F. Ciani, E. Pasiquini, I. Pela, P. Fiorini, Failure of early dextromethorphan and sodium benzoate therapy in an infant with Nonketotic Hyperglycinemia, Neuropaediatrics 25 (1994) 274–276.

[7] A. Dinopoulos, Y. Matsubara, S. Kure, Atypical variants of Nonketotic Hyperglycinemia, Mol. Genet. Metab. 86 (2005) 61–69.

[8] B. Wilcken, V. Wiley, J. Hammond, K. Carpenter, Screening newborns for inborn errors of metabolism by tandem mass spectrometry, NEJM 348 (2003) 2301–2312.

[9] Maternal and Child Health Bureau, Newborn screening: toward a uniform screening panel and system. <http://mchb.hrsa.gov/screening/>, 2006 (accessed 27.11.06.).

[10] V. Wiley, K. Carpenter, B. Wilcken, Newborn screening with tandem mass spectrometry: 12 months' experience in NSW, Australia Acta. Paediatr. Suppl. 432 (1999) 48–51.

Acylcarnitine Profiles in Fibroblasts From Patients With Respiratory Chain Defects Can Resemble Those From Patients With Mitochondrial Fatty Acid β-Oxidation Disorders

Keow Giak Sim, Kevin Carpenter, Judith Hammond, John Christodoulou, and Bridget Wilcken

Mitochondrial fatty acid β-oxidation (FAO) is coupled to the respiratory chain (RC). Functional defects of one pathway may lead to secondary alteration in flux through the other. We investigated the acylcarnitine profiles in cultured fibroblasts obtained from 14 healthy subjects, 31 patients with 8 different primary enzyme deficiencies of FAO, and 16 patients with primary RC defects including both isolated and multiple enzyme complex defects. Intact cells were incubated in media containing deuterium-labeled hexadecanoic acid and L-carnitine, and the acylcarnitines analysed using an electrospray tandem mass spectrometer. All FAO-deficient cell lines revealed disease-specific acylcarnitine profiles related to the sites of defects. Some cell lines from patients with RC defects showed profiles similar to those of controls, whereas others had abnormal profiles mimicking those found in FAO disorders. The acylcarnitine profiles of patients with RC enzyme defects were not predictable, and in some patients defects caused by mutations in either nuclear-encoded or mitochondrial DNA were associated with acylcarnitine abnormalities. While in vitro acylcarnitine profiling is useful for the diagnosis of FAO deficiencies, abnormal profiles do not exclusively indicate these disorders, and primary defects of the RC remain a possibility. Awareness of this diagnostic pitfall will aid in the selection of subsequent confirmatory tests and therapeutic options.

MITOCHONDRIAL fatty acid β-oxidation (FAO) and respiratory chain (RC) oxidative phosphorylation play pivotal roles in energy production. Both pathways are linked via the ubiquinone pool at 2 stages,[1] and also via the activity of the citric acid cycle. Inborn errors of mitochondrial fatty acid β-oxidation,[2-4] and the respiratory chain[5,6] are clinically and biochemically heterogenous. Whereas FAO defects usually involve single enzymes, defects in the RC may occur as multiple defects of 2 or more of the 5 enzyme complexes. These 2 groups of disorders have some clinical and biochemical features in common, including muscle weakness, cardiomyopathy, encephalopathy, hepatopathy, metabolic decompensation during catabolic stress, hypoglycemia, and lactic acidosis. Patients with primary RC defects can exhibit organic acidurias mimicking FAO disorders[7-9] Functional enzyme analyses have revealed concomitantly reduced activities of some enzymes in both pathways.[10-13] Histopathological findings[11,14] and in vitro oxidation rate studies[15] may not discriminate between defects of these 2 groups of disorders. It is apparent that impairment of one pathway may lead to secondary alteration in flux through the other.

Incubating cells from patients with documented FAO defects in media enriched with fatty acids and L-carnitine revealed disease-specific acylcarnitine profiles,[16-21] and thus in vitro acylcarnitine profiling is a useful screening tool for the diagnosis of patients suspected of having one of the numerous enzyme defects in this group of disorders. It is not clear whether acylcarnitine profiling is entirely specific for FAO disorders, as fibroblasts from patients with RC defects have not been studied in depth. The current report examines this question.

From the New South Wales Biochemical Genetics Service, Western Sydney Genetics Program, The Children's Hospital at Westmead, Sydney; and the Department of Paediatrics and Child Health, University of Sydney, Sydney, Australia.

Submitted May 15, 2001; accepted September 14, 2001.

Supported in part by the Small Grants Scheme of The Children's Hospital at Westmead, Sydney, Australia.

Address reprint requests to Clinical Associate Professor Bridget Wilcken, The Children's Hospital at Westmead, Locked Bag 4001, Westmead, NSW 2145, Australia.

0026-0495/02/5103-0012$35.00/0

doi:10.1053/meta.2002.30521

MATERIALS AND METHODS

Cell Lines and Cell Culture

Controls. Normal control cultured skin fibroblasts were from 14 healthy laboratory workers.

Mitochondrial FAO-deficient cell lines. These were obtained from 31 patients with 8 different enzyme deficiencies confirmed by specific enzyme assay and/or mutational analysis. The FAO enzyme deficiencies included hepatic carnitine palmitoyltransferase I (CPT IA; n = 1), carnitine palmitoyltransferase II (CPT II; n = 5), carnitine acylcarnitine translocase (CACT; n = 2), very-long-chain acyl-CoA dehydrogenase (VLCAD; n = 3), long-chain L-3-hydroxyacyl-CoA dehydrogenase (LCHAD; n = 5), medium-chain acyl-CoA dehydrogenase (MCAD; n = 9), multiple acyl-CoA dehydrogenase (MAD; n = 5), and short-chain acyl-CoA dehydrogenase (SCAD; n = 1). Table 1 shows the clinical phenotype and gene defects where known.

Respiratory chain defects. Cultured skin fibroblasts were available from 16 patients with RC defects confirmed by measurement of the activities of the enzyme complexes in one or more tissues including skeletal muscle, heart muscle, liver, and cultured skin fibroblasts, with DNA mutation analysis performed in some cases. The defects included isolated complex I (n = 5); complex II (n = 1); complex IV (n = 3); combined deficiencies of complexes I, III, and IV (n = 5); and combined complexes I, II + III, III, and IV deficiencies (n = 2). Cells between passage 4 and 13 were used for the assays, except for one, R10, at passage 19. A summary of clinical phenotype and gene defects (if identified) for each patient are presented in Table 2.

Skin fibroblasts were grown in Ham F10 nutrient mixture supplemented with 10% fetal calf serum, 2 mmol/L glutamine, 1% penicillin, streptomycin, and fungizone in a 25-cm^2 culture flask until 100% confluent. The cells were checked for microbial contamination using the Hoechst 33258 stain (Hoechst, St Louis, MO) whenever they were trypsinized.

Table 1. Incubation of Fibroblasts From Controls and Patients With Mitochondrial Fatty Acid β-Oxidation Deficiencies With $^2H_5$-Hexadecanoate and L-Carnitine for 72 Hours

Enzyme Deficiencies	Gene Defect	Clinical Phenotype		Acylcarnitine (nmol/mg protein) in Media at 72 h								
				C_{4}-	C_{5}-	C_{6}-	C_{8}-	C_{10}-	C_{12}-	C_{14}-	C_{16}-	C_{16}-OH
CPT IA (n = 1)	NK	Liver dysfunction	Average	0.3	0.9	<0.1	<0.1	<0.1	<0.1	<0.1	0.3	<0.1
CPT II early (n = 2)	NK	Neonates, died before 2 m	Average	<0.1	0.9	0.1	0.1	0.1	0.6	0.2	**13.4**	<0.1
			Limits§	<0.1	0.5-1.8	<0.2	<0.2	<0.2	0.1-1.9	0.1-0.3	**6.7-22.6**	<0.1
CPT II late (n = 3)	NK	Childhood and adult onset;	Average	1.2	0.9	0.5	0.6	0.8	0.2	0.1	**3.3**	<0.1
		rhabdomyolysis	Limits§	0.6-2.4	0.6-1.0	0.3-0.9	0.3-0.8	0.3-1.3	0.1-0.3	<0.2	**2.3-4.9**	<0.1
CACT (n = 2)	NK	1 neonatal death, 1 with mild	Average	0.3	1.5	0.3	0.3	0.3	0.7	0.1	**10.2**	<0.1
		symptoms	Limits§	0.1-0.6	1.4-1.6	0.1-0.4	0.3-0.4	0.2-0.5	0.2-1.3	<0.2	**9.2-11.2**	<0.1
MAD: profile 1 (n = 3)	NK	All neonatal deaths	Average	0.1	1.8	0.2	0.2	0.5	0.6	**1.5**	**10.5**	<0.1
			Limits§	<0.3	0.9-5.6	0.1-0.6	0.1-0.8	0.1-1.3	0.1-1.2	**0.3-2.1**	**5.5-19.8**	<0.1
MAD: profile 2 (n = 2)	NK	Proximal muscle weakness	Average	0.7	1.2	0.9	1.5	**2.3**	**2.7**	**2.3**	3.1	<0.1
			Limits§	0.2-1.3	0.9-1.7	0.5-1.1	1.0-1.8	**1.8-3.1**	**1.8-3.8**	**1.9-2.8**	2.1-4.9	<0.1
VLCAD (n = 3)	(1) R429W/R429W	(1) Fatal neonatal onset	Average	1.4	0.9	0.2	0.3	0.5	**1.0**	**2.4**	4.0	<0.1
	(2) G869A/G881A	(2) Reye-like at 2 yr	Limits§	0.4-5.2	0.5-1.7	0.1-0.3	0.2-0.5	0.3-0.8	**0.4-2.2**	**0.5-4.9**	0.9-8.1	<0.1
	(3) NK	(3) Asymptomatic¶										
LCHAD (n = 5)	G1528C/G1528C (n = 3)	All died before 6 mo of age	Average	0.7	1.3	0.3	0.4	0.6	**0.9**	**0.6**	**5.7**	**1.7**
	G1528C/NK (n = 2)	Both are well on low-fat diet	Limits§	0.2-2.1	0.7-2.2	0.1-0.6	0.2-0.8	0.2-1.5	**0.4-2.1**	**0.2-1.0**	**3.0-10.4**	**0.9-2.4**
MCAD (n = 9)	G985A/G985A (n = 4)	Included 4 clinical and 5	Average	1.4	1.0	2.4	**6.5**	1.4	0.2	0.1	1.2	<0.1
	G985A/Others (n = 5)	asymptomatic¶ patients	Limits§	0.4-3.9	0.4-2.9	1.0-5.5	**2.5-12.3**	0.4-5.4	0.1-0.4	<0.2	0.4-4.0	<0.1
SCAD (n = 1)	A57V/Q374X	Asymptomatic¶	Average	**10.2**	0.8	0.3	0.5	0.6	0.2	0.1	0.8	<0.1
Controls (n = 14)			Mean of 72 observations	0.9*	1.1	0.5*	0.6	0.8	0.2*	0.1*	0.9*	<0.1†
			Reference range‡	0.3-2.2*	0.5-1.7	0.2-1.4*	0.1-1.1	0.1-1.6	<0.4*	<0.2*	0.4-2.2*	0.1†

NOTE. Each cell line was analyzed in duplicate in 2 to 6 separate assays. For the FAO-deficient cell lines, the average of each analyte was calculated from all cell lines from different patients with the same enzyme defect.

Abbreviations: NK, not known; CPT IA, hepatic carnitine palmitoyltransferase I; CPT II, carnitine palmitoyltransferase II; CACT, carnitine acylcarnitine translocase; VLCAD, very-long-chain acyl-CoA dehydrogenase; LCHAD, long-chain L-3-hydroxyacyl-CoA dehydrogenase; MCAD, medium-chain acyl-CoA dehydrogenase; MAD, multiple acyl-CoA dehydrogenase; SCAD, short-chain acyl-CoA dehydrogenase.

*Logarithmic transformation of data to determine normal distribution.

†Median and upper limit of observed range.

‡Mean ± 1.96 SD.

§Observed lower and upper limits.

¶Detected by newborn screening program.

Table 2. Incubation of Fibroblasts From Patients With Respiratory Chain Defects With $^2H_5$-Hexadecanoate and L-Carnitine

Complex Defects (patient)	Gene Defect	Clinical Phenotype	Acylcarnitine (nmol/mg protein) in Media at 72 h								
			C_4-	C_5-	C_6-	C_8-	C_{10}-	C_{12}-	C_{14}-	C_{16}-	C_{16}-OH
Complex I											
R1	NK	Dysmorphic, LA	2.0	1.5	0.3	0.3	0.4	0.2	0.1	1.6	<0.1
R2	NK	Cardiomyopathy	1.7	1.3	0.6	0.4	0.6	0.2	0.1	1.4	<0.1
R3	mt $tRNA^{Leu(UUR)}$	MELAS	**2.8**	0.8	0.3	0.4	0.5	0.1	0.1	0.9	<0.1
R4	mt $tRNA^{Leu(UUR)}$	Skeletal myopathy	**3.0**	1.1	0.5	0.6	0.7	0.2	0.1	1.1	<0.1
R5	NK	Leigh syndrome	1.3	1.0	0.5	0.5	0.6	0.2	0.1	1.1	<0.1
Complex II											
R6	NK	Hepatic failure	1.5	1.1	**1.9**	**2.8**	**1.9**	0.3	0.1	1.5	<0.1
Complex IV											
R7	NK	Leigh disease	0.9	1.9	0.7	0.8	1.1	0.3	0.1	1.4	<0.1
R8	SURF-1	Dysmorphic, LA	0.9	1.0	1.5	**2.2**	**2.6**	**1.3**	**0.6**	2.0	<0.1
R9	NK	Neonatal LA	1.1	1.0	0.7	0.7	0.6	0.2	0.1	1.9	<0.1
Complex I, III, and IV											
R10	NK	Hepatic failure	0.4	0.9	0.1	0.2	0.2	0.2	0.3	**6.7**	<0.1
R11	NK	LA	1.3	1.3	1.0	1.0	1.3	0.3	0.1	1.4	<0.1
R12	mt $tRNA^{Lys}$	MERRF	0.7	1.2	0.3	0.4	0.8	0.4	0.3	3.1	**0.4**
R13	NK	Hypoglycemia, LA	0.8	0.9	0.2	0.2	0.4	0.1	0.1	1.6	<0.1
R14	NK	Hepatic failure	1.3	1.3	0.3	0.4	0.5	0.5	0.3	**7.5**	<0.1
Complex I, II + III, III, and IV											
R15	NK	LA	**3.1**	1.3	0.5	0.5	0.4	0.2	0.1	**4.8**	<0.1
R16	NK	LA	2.2	1.1	0.6	0.7	1.1	0.2	0.1	1.2	<0.1
Controls (N = 14)		Mean of 72 observations	0.9*	1.1	0.5*	0.6	0.8	0.2*	0.1*	0.9*	<0.1†
		Reference range‡	0.3-2.2*	0.5-1.7	0.2-1.4*	0.1-1.1	0.1-1.6	<0.4*	<0.2*	0.4-2.2*	0.1†

NOTE. Patient results are the average of 4 to 8 observations obtained in 2 to 4 separate assays.

Abbreviations: NK, not known; LA, lactic acidosis; MELAS, mitochondrial encephalomyopathy, lactic acidosis, and stroke-like episodes; MERRF, myoclonic epilepsy ragged red fiber syndrome.

*Logarithmic transformation of data to determine normal distribution.

†Median and upper limit of observed range.

‡Mean ± 1.96 SD.

Culture Media and Chemicals

All cell culture media, antibiotics, trypsin-EDTA were from Gibco BRL (Life Technologies, Grand Island, NY). Stable-isotope labeled palmitate, [15,15,16,16,16]-$^2H_5$ hexadecanoic acid was obtained from C/D/N Isotopes (Quebec, Point-Claire, Canada). Bovine serum albumin (essentially fatty acid free), L-carnitine, L-isovalerylcarnitine (C_5-), DL-hexanoylcarnitine (C_6-), DL-octanoylcarnitine (C_8-), DL-decanoylcarnitine (C_{10}-), and L-palmitoylcarnitine (C_{16}-) were obtained from Sigma (St Louis, MO). Internal standards (IS) containing a mixture of $^2H_3$-labeled butyrylcarnitine (C_4-), C_8-, C_{16}-acylcarnitine, and $^2H_9$-labeled C_5- and tetradecanoylcarnitine (C_{14}-) were purchased from NeoGen (Cambridge Isotope Laboratory, Andover, MA).

The In Vitro Acylcarnitine Assay in Intact Fibroblasts

The in vitro acylcarnitine assay in intact skin fibroblasts was performed as described[19] with some modifications. At 100% confluency, the fibroblasts were detached with trypsin, the cells counted, 10^5 cells (50 to 95 μg protein) subcultured into a well of a 24-well culture dish (Costa Products, Corning, NY), and settled for 24 hours to form a cell monolayer. Duplicate wells were set up for every cell line in each batch. The culture media was then replaced with 0.5 mL of reaction media containing $^2H_5$-labeled hexadecanoic acid 0.11 mmol/L complexed to bovine serum albumin 0.5 mg/mL, L-carnitine 0.4 mmol/L, and fetal calf serum 10% in Ham F10 nutrient mixture, prepared just before use. After incubation at 37°C in a water jacketed incubator for 72 hours, the reaction media in each well was transferred to a tube for acylcarnitine analysis, and the cell monolayer retained for protein quantitation.

Quantitative Profiling of Acylcarnitine Using Electrospray Tandem Mass Spectrometer

IS prepared in ethanol were added to each tube of reaction media, mixed, and centrifuged at 28,000 × g min to remove the precipitate. The supernatant was evaporated to dryness under a stream of air, derivatized with 300 μL butanolic-hydrochloric acid 10% (vol/vol) with heating at 60°C for 15 minutes, and dried again. The dried butylated sample was reconstituted with 100 μL of 50% acetonitrile: water (vol/vol) and transferred to a 96-well plate for automated injection.

The butyl esters of acylcarnitines were detected using a Micromass Quattro LC electrospray ionization tandem mass spectrometer (EIS-MS/MS, Micromass, Manchester, UK), detecting the precursors of m/z of 85.2 by the multichannel acquisition mode scanning in the range m/z 250 to 500. Following data acquisition, the concentrations of acylcarnitines were processed using Neolynx software (Micromass), measuring the intensities of analyte/internal standard ratio and concentration data from externally calibrated analytes including C_5-, C_6-, C_8-, C_{10}-, and C_{16}-acylcarnitine. Concentrations of C_4-, C_{12}- C_{14}-, and C_{16}-OH-acylcarnitine were calculated using the responses of the nearest counterparts, C_5-, C_{10}-, and C_{16}-acylcarnitine.

Protein Determination

The cell monolayer was washed twice with phosphate-buffered saline, hydrolyzed with 2.0 mol/L sodium hydroxide, and then neutralized with 1.0 mol/L hydrochloric acid. The protein was measured by the Pierce (Rockford, IL) bicinchoninic acid (BCA) method using bovine serum albumin in saline as standard on a Cobas BIO analyzer (Roche Diagnostics, Basel, Switzerland).

Data Analysis

Control cell lines (n = 14) were analyzed repeatedly with increasing passage numbers until a total of 72 observations were obtained. The data distribution of the levels of each $^2H_5$-labeled acylcarnitine species was checked with a test of normality, the Kolmogorov-Smirnov test, after logarithmic transformation of data for some analytes where required (SPSS version 9, SPSS Inc, Chicago, IL). Reference ranges were then derived, based on ±1.96 SD for allanalytes except for C_{16}-OH-acylcarnitine, which was expressed as the median and the upper limit of the observed range.

All patient cell lines were analyzed in duplicate in each batch and batches were repeated 2 to 6 times following successive subculture. For the data obtained from the FAO-deficient cells, the average for each analyte was calculated from all cell lines from different patients with the same enzyme defect, and the lower and upper limit of all these observations recorded (Table 1).

Fibroblast acylcarnitine levels in individual patients with RC defects were calculated as the average of 4 to 8 observations obtained from 2 to 4 separate assays (Table 2).

RESULTS

Butyl esters of $^2H_5$-labeled acylcarnitines derived from the precursor $^2H_5$-palmitate, and unlabeled acylcarnitines originating from the endogenous source of fatty acids and amino acids that had accumulated in the reaction media after 72 hours incubation were detected. Results shown in Tables 1 and 2 are the butyl esters of deuterium labelled acylcarnitines, except for C_5-acylcarnitine, which is the unlabelled species derived from branched-chain amino acids in the fetal calf serum.

The possibility of effect from the passage number (long-term subcultures) was investigated in fibroblasts from an RC affected patient with complex IV deficiency demonstrated in skeletal muscle, transformed lymphoblasts, and fibroblasts, who had homozygous mutation in SURF-1 gene (R8). Retrieved cell line cultured over a period of 6 weeks representing passages 5 to 8 following successive subcultures was analyzed in 4 separate assays. The average (observed lower-upper limit) of $^2H_5$-labeled C_8-, C_{10}-, C_{12}-, C_{14}-acylcarnitine were 2.2 (1.5-3.6), 2.6 (1.9-4.1), 1.3 (0.7-2.3), 0.6 (0.4-1.1) nmol/mg protein, respectively. These analytes were consistently elevated above the reference range of normal controls, and there were no significant changes in pattern or indication of a trend of change in concentration of any analytes with increasing passages.

Table 1 shows the results of quantitative profiling of acylcarnitines generated by fibroblasts from healthy subjects and those from patients with various inherited deficiencies of FAO (data not overlapping with the reference range are in bold type). The control cell lines showed the presence of low concentrations of acylcarnitines with carbon chain lengths corresponding to C_4-, C_5-, C_6-, C_8-, C_{10}-, C_{16}-, even lower levels of C_{12}-, C_{14}-, and near absence of C_{16}-OH-acylcarnitine. CPT IA-deficient cells displayed generally reduced levels of all $^2H_5$-labeled acylcarnitines, but normal level of C_5-acylcarnitine compared to the controls. All cell lines from patients with CPT II and CACT defects showed an elevated concentration of $^2H_5$-C_{16}-acylcarnitine. Two distinct profiles were exhibited by the MAD-deficient cell lines: one with massive accumulation of $^2H_5$-C_{16}-acylcarnitine and variable elevation of C_5-acylcarnitine (n = 3), and another with markedly increased concentrations of a number of medium- and long-chain species, notably $^2H_5$-labeled C_{10}-, C_{12}-, and C_{14}-acylcarnitine (n = 2). Cell lines with VLCAD deficiency were associated with elevation of $^2H_5$-labeled C_{12}-, C_{14}-, and C_{16}-acylcarnitine, with $^2H_5$-C_{14}-acylcarnitine being the predominant species. All of the LCHAD-deficient cell lines revealed accumulation of deuterium-labeled long-chain species and the distinctive C_{16}-OH-acylcarnitine. The MCAD-deficient cell lines were characterized by the predominant elevation of $^2H_5$-C_8-acylcarnitine, and the SCAD deficiency showed increased concentration of $^2H_5$-C_4-acylcarnitine only.

Fibroblast acylcarnitine profiles in individual patients with RC defects are shown in Table 2. Of the 5 cell lines with isolated complex I deficiency, 3 showed a profile similar to those of controls, while 2 revealed increased concentration of $^2H_5$-C_4-acylcarnitine (R3 and R4), suggestive of SCAD dysfunction. The cell line from a patient with complex II defect (R6) exhibited elevated concentrations of $^2H_5$-labeled C_6-, C_8-, and C_{10}-acylcarnitine, with C_8-acylcarnitine being the dominant species, strongly resembling those with MCAD deficiency. One of the 3 cell lines with isolated complex IV defect (R8) showed accumulation of a number of medium- and long-chain species suggestive of MAD deficiency. Of the 5 cell lines with multiple deficiencies of complexes I, III, and IV, profiles similar to controls (R11 and R13), or mimicking those with CPT II/CACT (R10 and R14), or LCHAD (R12) deficiencies were observed. One of the 2 cell lines with combined deficiencies of complexes I, II + III, III, and IV revealed marked accumulation of $^2H_5$-C_{16}- and $^2H_5$-C_4-acylcarnitine (R15), a profile suggestive of combined dysfunctions of short- and long-chain FAO.

DISCUSSION

In this study, only the acylcarnitines accumulating in the reaction media were analyzed, in contrast to previous studies where cells plus media were used.[21,22] The acylcarnitines are derived from the acyl-CoA esters, catalyzed by the carnitine acyltransferases in the presence of excess L-carnitine. These acylcarnitines leave mitochondria via the reverse action of carnitine acylcarnitine translocase,[23] and exit the cell cytosol possibly with the aid of the plasma membrane carnitine transporter,[24] as indicated by the appearance of acylcarnitines in the blood of patients with FAO defects during metabolic stress. Substantial levels of intracellular acylcarnitines have been found upon in vitro analysis of rat hepatocytes[25] and fibroblasts of normal controls.[26] Analyzing the acylcarnitines in the reaction media without the cells would appear to exclude these intracellular "background levels" of acylcarnitines. Our approach may improve the diagnostic sensitivity of the assay, as indicated by the unique profile of CPT IA deficiency,[27] which has been previously reported to be indistinguishable from controls.[22]

The acylcarnitine profile observed in CPT IA deficiency is different from that of other FAO defects in that all $^2H_5$-labeled acylcarnitines were reduced instead of elevated. All other FAO-deficient cell lines revealed elevation of acylcarnitines related to the site(s) of metabolic block, findings similar to other reports.[16,19-22] The same enzyme defects exhibited disease-specific acylcarnitine profiles (analyte concentrations >1.96 SD of control mean), irrespective of severity of symptoms (CPT II, CACT, VLCAD) or different mutations (LCHAD, VLCAD, MCAD). Hence, in vitro acylcarnitine profiling is a useful test for selected patients suspected with FAO when in vivo metabolite findings are not conclusive.

Of the 16 cell lines from patients with RC defects (Table 2), 8 showed acylcarnitine profiles similar to controls, and the other 8 exhibited abnormal profiles mimicking several enzyme defects in FAO. There was no clear association of a characteristic acylcarnitine profile with a specific respiratory chain defect (R10 to R14). Acylcarnitine abnormalities were seen in some cell lines with defects caused by mutations in the nuclear-encoded DNA (R8) or mitochondrial DNA (R3, R4, and R12).

The principal function of FAO is generation of acetyl-CoA, and that of the RC is the production of adenosine triphosphate (ATP). Functional defects in RC should not be associated with acylcarnitine abnormalities theoretically, as acyl-CoA esters are not the intermediate substrates. The observation of abnormal acylcarnitine profiles could be related to the linkages of L-3-hydroxyacyl-CoA dehydrogenases to complex I via NADH, and the acyl-CoA dehydrogenases to ubiquinone via electron transfer flavoprotein (ETF) and ETF dehydrogenase, and the reversible reactions between them. A primary functional defect of the RC resulting in dysfunction of these dehydrogenases in FAO is evident by accumulation of long-chain acyl-CoA esters, their L-3-hydroxyacyl derivatives and acylcarnitines as has been demonstrated in rat heart mitochondria, although high levels were not reached.[28] However, the abnormal metabolites thus accumulated could further impair oxidative phosphorylation,[29] and exacerbate the FAO dysfunction by the inhibitory effects of acylcarnitines on CACT,[30] and L-3-hydroxyacyl-CoA on enoyl-CoA hydratases.[31] The latter reaction could lead to accumulation of 2-enoyl-CoA, which has been shown to inhibit the acyl-CoA dehydrogenases.[32] These and other genetic factors may explain the variable and unpredictable acylcarnitine profiles observed in fibroblasts from patients with primary RC defects.

The profile of the complex IV-deficient cell line (R8) indicating general dysfunction of numerous enzymes which act on medium- and long-chain acyl-CoA is probably not surprising. However, it is not clear how a deficiency in complex I may be related to elevation of $^2H_5$-C_4-acylcarnitine (R3, R4), instead of the hydroxy-acylcarnitines. Complex II performs a key step in the citric acid cycle, in which succinate is dehydrogenated to fumarate. In turn, the citric acid cycle is linked to FAO via acetyl-CoA, the end product of thiolytic cleavage. The predominant accumulation of C_8-acylcarnitine suggesting a specific inhibition of MCAD in the fibroblasts of this complex II-affected patient (R6) is unexpected, especially since the complex II activity was not abnormal in skin fibroblasts, but deficient in liver. There was in fact no association between the expression of the RC defect in skin fibroblasts and a normal or abnormal acylcarnitine profile (data not shown), but our profile results were reproducible. It is possible that these cell lines harbor 2 genetic defects,[33] but the occurrence of abnormal findings in 8 of 16 patients studied seemed too high for this possibility.

Mitochondrial β-oxidation disorders described to date show autosomal-recessive inheritance (nuclear encoded), and may be potentially fatal if undetected. However, in most diagnosed cases the prognosis is generally favorable,[34] even in patients with 2 null mutations.[35] Moreover, treatment is relatively simple, involving avoidance of fasting together with dietary therapy and L-carnitine in some disorders. In contrast, defects in RC can be caused by mutations in nuclear or mitochondrial DNA, and the inheritance patterns may be autosomal (recessive or dominant), or maternal (mitochondrial). Treatment in most individuals has generally been ineffective[34,36] and the prognosis especially in early onset defects is poor, making prenatal diagnosis and genetic counseling particularly important.

In summary, in vitro acylcarnitine profiling in fibroblasts is a useful test for the detection of patients suspected with FAO disorders. However, there are limitations. Neither defects in plasma membrane cellular uptake of long-chain fatty acids[37] nor in ketogenesis not involving acyl-CoA intermediates would be detected. Tissue-specific enzymes that are not expressed in fibroblasts would also be missed. Moreover, abnormal acylcarnitine profiles may also be detected in fibroblasts from patients with RC defects. Whereas the clinical and biochemical abnormalities of FAO and RC defects may be similar, patient management, prognosis and genetic counseling are vastly different. Awareness of this diagnostic pitfall enables the appropriate selection of follow-up confirmatory tests and assists in targeted therapeutic strategies. Integrated evaluation of clinical information, in vivo and in vitro findings, histopathology, enzymology, and molecular studies are required for accurate diagnosis.

ACKNOWLEDGMENT

The authors are grateful to Dr David Thorburn for analyzing the enzyme complex activities of respiratory chain; Dr Ronald Wanders, Dr Christine Vianey-Saban, and Dr Michelle Brivet for performing some of the confirmatory mitochondrial fatty acid β-oxidation enzyme assays; Dr Brage Andresen for the mutational analyses on the β-oxidation–deficient fibroblasts; and all staff of the NSW Biochemical Genetics Service and Newborn Screening Programme for their valuable help in different aspects of the study.

REFERENCES

1. Eaton S, Bartlett K, Pourfarzam M: Mammalian mitochondrial beta-oxidation. Biochem J 320:345-357, 1996

2. Roe CR, Ding J: Mitochondrial fatty acid oxidation disorders, in Scriver CR, Beaudet AL, Valle D, et al (eds): The Metabolic and Molecular Bases of Inherited Disease. New York, NY, McGraw-Hill, 2001, pp 2297-2326

3. Pollitt RJ: Disorders of mitochondrial long-chain fatty acid oxidation. J Inherit Metab Dis 18:473-490, 1995

4. Bennett MJ, Rinaldo P, Strauss AW: Inborn errors of mitochondrial fatty acid oxidation. Crit Rev Clin Lab Sci 37:1-44, 2000

5. Shoffner JM: Oxidative phosphorylation diseases, in Scriver CR, Beaudet AL, Valle D, et al (eds): The Metabolic and Molecular Bases of Inherited Disease. New York, NY, McGraw-Hill, 2001, pp 2367-2423

6. Christodoulou J: Genetic defects causing mitochondrial respiratory chain disorders and disease. Hum Reprod 15:28-43, 2000 (suppl 2)

7. Bennett MJ, Weinberger MJ, Sherwood WG, et al: Secondary 3-hydroxydicarboxylic aciduria mimicking long-chain 3-hydroxyacyl-CoA dehydrogenase deficiency. J Inherit Metab Dis 17:283-286, 1994

8. Christensen E, Brandt NJ, Schmalbruch H, et al: Muscle cytochrome c oxidase deficiency accompanied by a urinary organic acid pattern mimicking multiple acyl-CoA dehydrogenase deficiency. J Inherit Metab Dis 16:553-556, 1993

9. Enns GM, Bennett MJ, Hoppel CL, et al: Mitochondrial respiratory chain complex I deficiency with clinical and biochemical features of long-chain 3-hydroxyacyl-coenzyme A dehydrogenase deficiency J Pediatr 136:251-254, 2000

10. Reichmann H, Scheel H, Bier B, et al: Cytochrome c oxidase deficiency and long-chain acyl coenzyme A dehydrogenase deficiency with Leigh's subacute necrotizing encephalomyelopathy. Ann Neurol 31:107-109, 1992

11. Tyni T, Majander A, Kalimo H, et al: Pathology of skeletal muscle and impaired respiratory chain function in long-chain 3-hydroxyacyl-CoA dehydrogenase deficiency with the G1528C mutation. Neuromusc Disord 6:327-337, 1996

12. Das AM, Fingerhut R, Wanders RJ, et al: Secondary respiratory chain defect in a boy with long-chain 3-hydroxyacyl-CoA dehydrogenase deficiency: possible diagnostic pitfalls. Eur J Pediatr 159:243-246, 2000

13. Diogo L, Garcia P, Rebelo O, et al: Clinical heterogeneity in MAD—Report of two cases. J Inherit Metab Dis 23:134-134, 2000 (suppl 1, abstr)

14. Watmough NJ, Bindoff LA, Birch-Machin MA, et al: Impaired mitochondrial beta-oxidation in a patient with an abnormality of the respiratory chain. Studies in skeletal muscle mitochondria. J Clin Invest 85:177-184, 1990

15. Venizelos N, von Dobeln U, Hagenfeldt L: Fatty acid oxidation in fibroblasts from patients with defects in beta-oxidation and in the respiratory chain. J Inherit Metab Dis 21:409-415, 1998

16. Kler RS, Jackson S, Bartlett K, et al: Quantitation of acyl-CoA and acylcarnitine esters accumulated during abnormal mitochondrial fatty acid oxidation. J Biol Chem 266:22932-22938, 1991

17. Pourfarzam M, Schaefer J, Turnbull DM, et al: Analysis of fatty acid oxidation intermediates in cultured fibroblasts to detect mitochondrial oxidation disorders. Clin Chem 40:2267-2275, 1994

18. Schaefer J, Pourfarzam M, Bartlett K, et al: Fatty acid oxidation in peripheral blood cells: characterization and use for the diagnosis of defects of fatty acid oxidation. Pediatr Res 37:354-360, 1995

19. Nada MA, Rhead WJ, Sprecher H, et al: Evidence for intermediate channeling in mitochondrial beta-oxidation. J Biol Chem 270: 530-535, 1995

20. Schmidt-Sommerfeld E, Bobrowski PJ, Penn D, et al: Analysis of carnitine esters by radio-high performance liquid chromatography in cultured skin fibroblasts from patients with mitochondrial fatty acid oxidation disorders. Pediatr Res 44:210-214, 1998

21. Ventura FV, Costa CG, Struys EA, et al: Quantitative acylcarnitine profiling in fibroblasts using [U-^{13}C] palmitic acid: An improved tool for the diagnosis of fatty acid oxidation defects. Clin Chim Acta 281:1-17, 1999

22. Roe CR, Roe DS: Recent developments in the investigation of inherited metabolic disorders using cultured human cells. Mol Genet Metab 68:243-257, 1999

23. Pande SV: A mitochondrial carnitine acylcarnitine translocase system. Proc Natl Acad Sci USA 72:883-887, 1975

24. Wu X, Huang W, Prasad PD, et al: Functional characteristics and tissue distribution pattern of organic cation transporter 2 (OCTN2), an organic cation/carnitine transporter. J Pharmacol Exp Ther 290: 1482-1492, 1999

25. Eaton S, Bartlett K, Pourfarzam M, et al: Production and export of acylcarnitine esters by neonatal rat hepatocytes. Adv Exp Med Biol 466:155-159, 1999

26. Shen JJ, Matern D, Millington DS, et al: Acylcarnitines in fibroblasts of patients with long-chain 3-hydroxyacyl-CoA dehydrogenase deficiency and other fatty acid oxidation disorders. J Inherit Metab Dis 23:27-44, 2000

27. Sim KG, Wiley V, Carpenter K, et al: Carnitine palmitoyltransferase I deficiency in neonate identified by dried blood spot free carnitine and acylcarnitine profile. J Inherit Metab Dis 24:51-59, 2001

28. Eaton S, Pourfarzam M, Bartlett K: The effect of respiratory chain impairment of beta-oxidation in rat heart mitochondria. Biochem J 319:633-640, 1996

29. Ventura FV, Ruiter JP, Ijlst L, et al: Inhibitory effect of 3-hydroxyacyl-CoAs and other long-chain fatty acid beta-oxidation intermediates on mitochondrial oxidative phosphorylation. J Inherit Metab Dis 19:161-164, 1996

30. Baillet L, Mullur RS, Esser V, et al: Elucidation of the mechanism by which (+)-acylcarnitines inhibit mitochondrial fatty acid transport. J Biol Chem 275:36766-367688, 2000

31. He XY, Yang SY, Schulz H: Inhibition of enoyl-CoA hydratase by long-chain L-3-hydroxyacyl-CoA and its possible effect on fatty acid oxidation. Archiv Biochem Biophys 298:527-531, 1992

32. Powell PJ, Lau SM, Killian D, et al: Interaction of acyl coenzyme A substrates and analogues with pig kidney medium-chain acyl-coA dehydrogenase. Biochemistry 26:3704-3710, 1987

33. Vockley J, Rinaldo P, Bennett MJ, et al: Synergistic heterozygosity: Disease resulting from multiple partial defects in one or more metabolic pathways. Mol Genet Metab 71:10-18, 2000

34. Pons R, De Vivo DC: Mitochondrial disease. Curr Treat Opt Neurol 3:271-288, 2001

35. Touma EH, Rashed MS, Vianey-Saban C, et al: A severe genotype with favourable outcome in very long chain acyl-CoA dehydrogenase deficiency. Arch Dis Child 84:58-60, 2001

36. Walker UA, Byrne E: The therapy of respiratory chain encephalomyopathy: A critical review of the past and current perspective. Acta Neurol Scand 92:273-280, 1995

37. Odaib AA, Shneider BL, Bennett MJ, et al: A defect in the transport of long-chain fatty acids associated with acute liver failure. N Engl J Med 339:1752-1757, 1998

JIMD Short Report #015 (2006) Online
DOI 10.1007/s10545-006-0366-0

SHORT REPORT

Newborn screening may fail to identify intermediate forms of maple syrup urine disease

K. Bhattacharya · V. Khalili · V. Wiley · K. Carpenter · B. Wilcken

Received: 11 March 2006 / Accepted: 23 May 2006

Summary The New South Wales state-wide newborn screening programme has offered comprehensive screening for inborn errors of metabolism, including MSUD, using electrospray tandem mass spectrometry since 1998. Over this period, a number of patients with classic MSUD have been identified with subsequent good neurological outcome. We describe two patients with an intermediate form of MSUD who presented later in childhood. Retrospective review of their newborn screening results demonstrates that the diagnosis could not have been made by current newborn screening. Their neurological outcome is much less satisfactory. Despite the usefulness of expanded newborn screening programmes in detecting severe neonatal presentations of inborn errors of metabolism, partial enzyme deficiencies may not be detected. Metabolic diseases still need to be considered in appropriate clinical situations later in life.

Abbreviations

BCAA	branched-chain amino acids
BCKD	α-ketoacid dehydrogenase
ES-TMS	electrospray tandem mass spectrometry
MSUD	mple syrup urine disease

Communicating editor: Rodney Pollitt

Competing interests: None declared

K. Bhattacharya · B. Wilcken (✉)
Genetic Metabolic Service, The Children's Hospital at Westmead, Westmead, New South Wales, Australia; University of Sydney, Sydney, New South Wales, Australia
e-mail:BridgetW@chw.edu.au

V. Wiley · B. Wilcken
NSW Newborn Screening Programme, The Children's Hospital at Westmead, Westmead, New South Wales, Australia

V. Khalili · K. Carpenter · B. Wilcken
Biochemical Genetics Laboratory, Western Sydney Genetics Program, The Children's Hospital at Westmead, Westmead, New South Wales, Australia

Introduction

Maple syrup urine disease (MSUD) (OMIM 248600) is an autosomal recessive disorder caused by decreased activity of branched-chain α-ketoacid dehydrogenase (BCKD) (EC 1.2.4.4). The mitochondrial BCKD complex is involved in the second step of the degradative pathway for the branched-chain amino acids (BCAA) leucine, isoleucine and valine. Classic MSUD usually presents in the first week of life with nonspecific symptoms that rapidly progress to encephalopathy. If the condition is not treated effectively at this stage, the prognosis is usually poor. Conversely, patients identified by population-based newborn screening programmes or targeted screening in high-incidence areas and who are effectively treated before irreversible complications ensue can have a very good neurological outcome (Schulze et al 2003; Heldt et al 2005; Morton et al 2002).

The molecular basis of BCKD deficiency is complicated. The BCKD enzyme complex comprises several subunits denoted E1α, E1β, E2 and E3. In addition the integral regulatory BCKD kinase and phosphatase form the complete complex. The spatial arrangement of BCKD within the mitochondrion is dependent on extensive posttranslational modification, transport and assembly of the sub-units (Chuang et al 2001). It is not surprising, therefore, that a broad range of phenotypes exist encompassing the classic phenotype to intermittent and intermediate forms. These nonclassical forms of MSUD have a subacute course in infancy and can often present a diagnostic challenge. It would be predicted that if

newborn screening programmes were able to identify these patients, significant morbidity could be avoided.

We present two patients who participated in the New South Wales extended screening programme using electrospray tandem mass spectrometry (ES-TMS) but had normal results in the newborn period. They subsequently presented early in childhood with acute neurological problems and background developmental delay. A diagnosis of intermittent MSUD was subsequently made in each patient.

Case 1

Preliminary findings from this boy have been described previously (Khalili et al 2003). This child had a normal newborn screening result of a combined leucine/isoleucine of 325 μmol/L (cut off <400 μmol/L) with a ratio to phenylalanine of 6.5 (Fig. 1) and a ratio to alanine of 0.78 (Fig. 2). He was the third child of consanguineous Lebanese parents. He had a history of recurrent asthma with 15 attendances to the emergency department in the first 2 years prior to a presentation after his second birthday with a 2-week history of ataxia. He had elevated urinary BCAA and decreased alanine. Urine organic acids demonstrated elevation in 2-hydroxyisovalerate, 2-hydroxy-3-methylvalerate and 2-hydroxyisocaproate. A subsequent plasma amino acid profile showed elevation of BCAA typical of MSUD (Table 1). Prior to presentation, he had reached developmental milestones late: he had crawled at 11 months, walked independently at 20 months and had his first meaningful words at 20 months. On developmental assessment using the Bayley Scales of Infant Development (2nd edition) at the age of 27 months he had a mental scale ability at the 15-month level and a motor scale ability at the 18-month level. Subsequent reassessment of development at 44 months of age using the same scale demonstrated motor ability at 36 months and cognitive skills at the 26–30 month range. He has been managed on a mildly protein-restricted diet with the addition of a BCAA-free supplement. Since diagnosis, plasma leucine levels have been largely within the reference range, with occasional values elevated, all less than 550 μmol/L, and there has been no clinical evidence of metabolic decompensation.

Case 2

This child had a normal newborn screening result: combined leucine/isoleucine of 209 μmol/L (cut-off <400 μmol/L) with a ratio to phenylalanine of 3.7 (Fig. 1) and a ratio to alanine of 0.39 (Fig. 2). She was the second of three children born to nonconsanguineous Lebanese parents. She presented to the emergency department at the age of 27 months with a focal right-sided seizure that became sustained and generalized. Urinary BCAA were elevated, and there were elevations of 2-hydroxyisovalerate, 2-hydroxy-3-methylvalerate and 2-hydroxyisocaproate on urine organic acid analysis. The plasma amino acid profile demonstrated elevation of BCAA typical of MSUD (Table 1). She attained many developmental landmarks late: she started rolling at 9 months, sat unsupported at 1 year, crawled at 13 months and walked independently at 2 years of age. She had her first meaningful single words at 18 months of age. Many investigations had been performed for developmental delay prior to presentation with a seizure, but urine amino acids, organic acids and plasma amino acids had not been done. She has been managed on a protein-restricted diet with a BCAA-free amino acid formula since the diagnosis was made. Her plasma leucine concentrations have all been either in or below the normal range since diagnosis, except during one mild upper respiratory tract infection when the leucine concentration was 550 μmol/L.

Cases detected by newborn screening

Between 1 April 1998 and 31 March 2005, three cases of classical MSUD were identified, from a population of 638 000 babies screened by our New South Wales newborn screening laboratory. Two patients were identified and assessed on day 7 of life and the other on day 8. By the time of assessment, the three patients had clinical signs compatible with MSUD but were not under medical care. Their peak plasma leucine concentrations were 2850, 3900 and 4300 μmol/L. All three patients required treatment in paediatric intensive care with continuous veno-venous haemofiltration at this first presentation to our team, in addition to standard therapy comprising protein restriction, high-energy

Table 1 Branched-chain amino acid concentrations (μmol/L) in two patients diagnosed with intermittent/intermediate MSUD at newborn screening and during decompensation at time of first assessment by the Genetic Metabolic Team

DBS, dried blood spot; ND, not detected

Amino acid	Patient 1	Patient 2	Reference range
At screening			
Leucine + isoleucine (DBS)	325	209	≤400 (cut-off)
At decompensation			
Leucine (plasma)	1078	1209	56–178
Valine (plasma)	1013	1094	108–314
Isoleucine (plasma)	552	527	34–106
Alloisoleucine (plasma)	133	118	ND

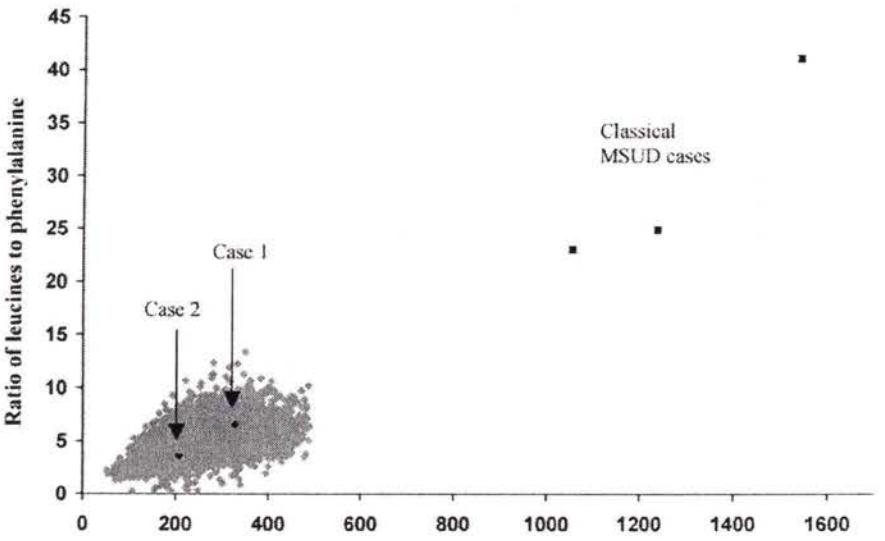


Fig. 1 Combined leucine + isoleucine to phenylalanine ratio versus leucine + isoleucine in 5 MSUD patients, 2 intermittent and 3 classical, within 30 000 newborn screening samples

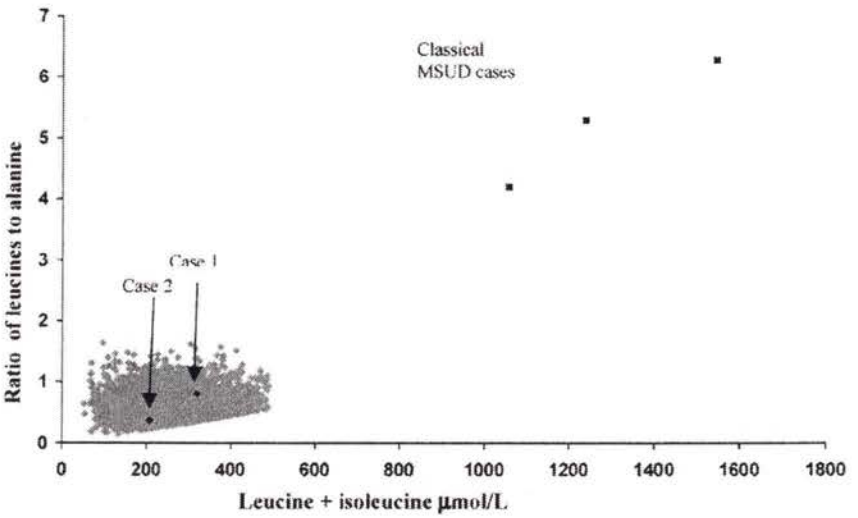


Fig. 2 Combined leucine + isoleucine to alanine ratio versus isoleucine + leucine in 5 MSUD patients, 2 intermittent and 3 classical, within 30 000 newborn screening samples

parenteral nutrition and supplementation with non-BCAA formula. The patients were well enough to be discharged home after 15, 14 and 18 days. They have each subsequently been managed with a low-protein diet, supplementation of a nutritionally complete non-BCAA amino acid formula, regular dietetic and clinical surveillance and regular monitoring of plasma BCAA. Enteral hypercaloric supplementation has been implemented during intercurrent illness at home and each has required inpatient intravenous therapy, when plasma leucine levels have been elevated during intercurrent illness. The three patients, who are now aged 1, 3 and 6 years, have attained normal developmental milestones without any concerns about their neurological progress, the oldest child being in normal mainstream education without additional support.

Newborn screening methods

Newborn screening bloodspot samples were collected between 48 h and 72 h of age as previously described (Wilcken et al 2003). Amino acids were analysed as their butyl esters using multiple reaction monitoring and quantified against dried blood spot calibrators with stable-isotope internal standards; results were initially interpreted using Neolynx software (Micromass Ltd, UK). The results were transferred to the Labmaster database (Wiley Associates Pty Ltd, Sydney, Australia), where each was checked against predefined algorithms and reference ranges (previously described; Wiley et al 1999) including reference ranges for ratios, and age-matched ranges for second samples where requested. For the detection of MSUD, a positive result is currently defined

as a combined leucine and isoleucine peak greater than 400 μmol/L, and ratios to phenylalanine or alanine are not routinely used.

Confirmatory diagnosis and subsequent monitoring was performed on fresh plasma samples. Plasma amino acid quantification was carried out using ion-exchange chromatography with postcolumn ninhydrin detection.

Discussion

It is reported that partial deficiency of BCKD can cause one of two clinical scenarios: namely intermittent MSUD, in which patients have episodes of decompensation but are otherwise well, or intermediate MSUD, which results in patients developing chronic moderate neurological disability without episodes of decompensation. In our experience, there is overlap between the two scenarios. These forms of MSUD have been regarded as more 'mild' than classic MSUD but they are still disabling diseases.

Current newborn screening protocols identify patients with MSUD either by demonstrating elevation of leucine, often combined with isoleucine above an established 'cut-off' value, or by an elevation of the ratio of these BCAA to phenylalanine or alanine (Chace et al 1995; Morton et al 2002). Although there is a clear rationale for the use of these ratios in classical MSUD (Korein et al 1994), it is less clear whether they are predictive in individuals with intermediate forms of MSUD.

Our screening protocol urgently recalls individuals who have a combined leucine and isoleucine value greater than 400 μmol/L. We were able to analyse retrospectively the ratios of these BCAA to the blood spot alanine and phenylalanine, which were measured at the same time in the newborn period. Our three patients with classical MSUD had clear elevation of combined isoleucine and leucine from the newborn screening card, clinical symptoms at the time of recall and gross elevation of plasma leucine when re-tested. The diagnosis of MSUD was consequently made early in life. The two patients with intermediate/intermittent MSUD had normal combined isoleucine and leucine values with the ratio to phenylalanine also retrospectively found to be normal in the newborn period. One patient had a normal BCAA to alanine ratio, while the other had a value of 0.786 (95.8th centile). This value would still be below any 'cut-off' level that we would consider, as it would imply a recall rate of 5% (approximately 36 000 babies) for detection of one case in the last 8 years. Given that these nonclassical forms may also not always demonstrate typical biochemical features, there is also no guarantee that the diagnosis would even be made at recall. We conclude that our ES-TMS-based newborn screening programme has been unable to detect cases of intermittent/intermediate MSUD in its eight-year history.

We demonstrate, ironically, that 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. It is important, therefore, to consider the diagnosis of MSUD in the appropriate clinical setting irrespective of the newborn screening result.

References

Chace DH, Hillman SL, Millington DS, et al (1995) Rapid diagnosis of maple syrup urine disease in blood spots from newborns by tandem mass spectrometry. *Clin. Chem* **41**: 62–68.

Chuang DT, Shih VE (2001) Maple syrup urine disease (branched-chain ketoaciduria). In: Scriver CR, Beaudet AI, Sly WS, Valle D, eds; Child B, Kinzler KW, Vogelstein B, assoc eds. *The Metabolic and Molecular Bases of Inherited Disease*, 8th edn. New York: McGraw-Hill, 1971–2005.

Heldt K., Schwahn B, Marquardt I, Grotzke M, Wendel U (2005) Diagnosis of MSUD by newborn screening allows early intervention without extraneous detoxification. *Mol Genet Metab* **84**: 313–316.

Khalili V, Regaglia G, Carpenter K, et al (2003) Mild maple syrup urine disease is unlikely to be detected by newborn screening but has serious consequences. *J Inherit Metab Dis* **26**: S2–6.

Korein J, Sansaricq C, Kalmijn M, et al (1994) Maple syrup urine disease: clinical, EEG, and plasma amino acid correlations with a theoretical mechanism of acute neurotoxicity *Int J Neurosci* **79**: 21–45.

Morton DH, Strauss KA, Robinson DL, Puffenberger EG, Kelley RI (2002) Diagnosis and treatment of maple syrup disease: a study of 36 patients. *Pediatrics* **109**: 999–1008.

Schulze A, Lindner M, Kohlmuller D, et al (2003) Expanded newborn screening for inborn errors of metabolism by electrospray ionization-tandem mass spectrometry: results, outcome, and implications. *Pediatrics* **111**: 1399–1406.

Wilcken B, Wiley V, Hammond J, Carpenter K (2003) Screening newborns for inborn errors of metabolism by tandem mass spectrometry. *N Engl J Med* **348**: 2304–2312.

Wiley V, Carpenter K, Wilcken B (1999) Newborn screening with tandem mass spectrometry: 12 months' experience in NSW Australia. *Acta Paediatr Suppl* **88**: 48–51.

ORIGINAL ARTICLE

Screening Newborns for Inborn Errors of Metabolism by Tandem Mass Spectrometry

Bridget Wilcken, M.B., Ch.B., Veronica Wiley, Ph.D., Judith Hammond, Ph.D., and Kevin Carpenter, Ph.D.

ABSTRACT

BACKGROUND

The recent development of electrospray tandem mass spectrometry makes it possible to screen newborns for many rare inborn errors of metabolism, but the efficacy and outcomes of screening remain unknown. We examined the effect of the screening of newborns by tandem mass spectrometry on the rates of diagnosis of 31 disorders.

METHODS

We compared the rates of detection of 31 inborn errors affecting the metabolism of the urea cycle, amino acids, and organic acids and fatty-acid oxidation among 362,000 newborns screened by tandem mass spectrometry over a four-year period (April 1998 through March 2002) with the rates in six preceding four-year birth cohorts in New South Wales and the Australian Capital Territory, Australia, where screening, diagnostic, and clinical services were centralized.

RESULTS

The overall prevalence of disorders during the periods when clinical diagnosis was used did not vary between 1982 and 1998. In the cohort screened with tandem mass spectrometry, the prevalence of inborn errors, excluding phenylketonuria, was 15.7 per 100,000 births (95 percent confidence interval, 11.9 to 20.4), as compared with adjusted rates of 8.6 to 9.5 per 100,000 births in the four preceding four-year cohorts. Of the 57 cases diagnosed after the introduction of newborn screening, 15 were diagnosed clinically; 7 of the 15 newborns had a normal result on screening. The rate of detection was increased specifically for medium-chain acyl-coenzyme A dehydrogenase deficiency ($P<0.001$) and other disorders of fatty-acid oxidation ($P=0.007$), as compared with the 16-year period before the implementation of neonatal screening for these disorders.

CONCLUSIONS

More cases of inborn errors of metabolism are diagnosed by screening with tandem mass spectrometry than are diagnosed clinically. It is not yet clear which patients with disorders diagnosed by such screening would have become symptomatic if screening had not been performed.

From the New South Wales Newborn Screening Programme (B.W., V.W.); the New South Wales Biochemical Genetics Service (B.W., J.H., K.C.); the Children's Hospital at Westmead (B.W., V.W., J.H., K.C.); and the Discipline of Paediatrics and Child Health, University of Sydney (B.W.) — all in Sydney, Australia. Address reprint requests to Professor Wilcken at the Children's Hospital at Westmead, Locked Bag 4001, Westmead, NSW 2145, Australia, or at bridgetw@chw.edu.au.

N Engl J Med 2003;348:2304-12.

IN MANY COUNTRIES, IT IS A ROUTINE COMponent of neonatal care to screen infants for congenital hypothyroidism, phenylketonuria, and a variable number of other disorders. The primary aim is the early detection and treatment of clinically important disorders in order to minimize morbidity and mortality in early childhood. Recently, with the development of electrospray tandem mass spectrometry, it has become possible to use a single test to screen for a wide range of very rare disorders that have not been screened for previously.

Formal evidence of the clinical effectiveness of newborn screening is lacking. The only randomized, controlled trials have been of screening for cystic fibrosis in the United States and the United Kingdom.[1,2] The clinical effectiveness of screening for phenylketonuria and hypothyroidism is generally accepted,[3] although no formal trials were ever conducted; there is also evidence to support screening for sickle cell disease in regions where it is prevalent.[4] One reason for the lack of randomized, controlled trials of screening is the rarity of these disorders; a very large number of infants would need to be enrolled in a trial for it to have sufficient power to assess the benefits of screening.[5] Another factor that has discouraged researchers from conducting randomized, controlled trials of screening for some disorders has been a strong conviction, based on clinical experience, that there is a benefit from early diagnosis. As a result, the usefulness of many screening tests remains uncertain.

Tandem mass spectrometry is used in many screening programs to analyze amino acids and acylcarnitines in blood to detect disorders of amino acids, organic acids, and fatty-acid metabolism. The ability to select which analytes to detect effectively permits screeners to choose which disorders to screen for. In Australia, the screening of newborns by tandem mass spectrometry was introduced in New South Wales and the Australian Capital Territory in early 1998, in South Australia in 1999, and in Victoria in 2002. The technology is being introduced widely in the United States, some European countries, and elsewhere.[6-13]

Among the disorders that may be diagnosed, some cause severe illness or death within the first few days of life, and newborn screening may serve only to suggest a diagnosis that might otherwise have been missed. However, most of the disorders are treatable if they are diagnosed early. With early diagnosis and appropriate treatment, some problems can be avoided; these include biochemical disturbances such as hyperammonemia in patients with urea-cycle disorders that present after the newborn period, severe metabolic acidosis in patients with disorders of organic acids, or hypoketotic hypoglycemia, cardiomyopathy, or rhabdomyolysis in patients with disorders of fatty-acid oxidation; if left untreated, these disorders may lead to brain damage, other organ damage, or death.

We compared the rates and profile of diagnoses identified through newborn screening by tandem mass spectrometry with those during periods preceding the use of this technology in the population of New South Wales and the Australian Capital Territory (population, 6 million).

METHODS

DISORDERS

We studied 31 disorders that may be able to be diagnosed by tandem mass spectrometry in blood samples from newborns (Table 1). Phenylketonuria and pterin disorders were excluded from consideration, since they had been screened for by another method for many years. Also excluded were disorders known to be benign, maternal disorders such as maternal 3-methylcrotonyl-coenzyme A (CoA) carboxylase deficiency, and errors that are not inborn, such as vitamin B_{12} deficiency.

CLINICAL DIAGNOSIS

Since 1974, whenever patients in New South Wales and the Australian Capital Territory have had symptoms suggestive of an inborn error of metabolism, their cases have been investigated in our biochemical genetics laboratory, the only laboratory in the state that provides diagnostic testing for defects of amino acids, organic acids, and fatty-acid metabolism. Diagnostic tests have included analysis of urinary organic acids by gas chromatography (until 1991) or gas chromatography–mass spectrometry; analysis of urinary amino acids, initially by high-voltage electrophoresis; analysis of plasma amino acids by quantitative amino acid analysis; and analysis of plasma acylcarnitines by tandem mass spectrometry. Other biochemical genetic tests, including enzyme and molecular analyses, were performed as indicated. Since April 1998, the laboratory has also evaluated patients identified by routine newborn screening. We have maintained a data base of all patients with a confirmed diagnosis of an inborn error of metabolism.

Table 1. **Numbers of Patients in the Birth Cohorts with Inborn Errors of Metabolism Diagnosed after Clinical Presentation (between April 1974 and March 1998) or during the Newborn-Screening Period (April 1998 to March 2002).***

Disorder	Screening Year							Total
	1974–1978	1978–1982	1982–1986	1986–1990	1990–1994	1994–1998	1998–2002	
	number of patients							
Urea cycle								
Carbamyl phosphate synthetase deficiency	1	1	0	1	1	0	1	5
Ornithine transcarbamylase deficiency	2	3	5	4	4	3	3	24
Argininosuccinate synthase deficiency	2	0	0	2	2	1	0	7
Argininosuccinate lyase deficiency	2	1	1	1	0	0	2	7
Arginase deficiency	0	0	0	0	0	0	0	0
Citrullinemia, type II (citrin deficiency)	0	0	0	0	0	0	1	1
Amino acid								
Nonketotic hyperglycinemia	3	1	1	3	3	4	3	18
Cystathionine β-synthase deficiency (homocystinuria)	5	0	2	5	1	0	2	15
Maple syrup urine disease	1	1	4	3	3	2	1	15
Tyrosinemia								
Type I	0	1	1	1	0	1	2	6
Type II	1	2	2	0	0	1	1	7
Organic acid								
Propionyl-CoA carboxylase deficiency	0	0	1	0	0	1	1	3
Methylmalonyl-CoA mutase deficiency plus cobalamin A and B defects	2	3	1	2	4	3	0	15
Cobalamin C defect	0	1	0	0	0	3	1	5
Isovalericacidemia	0	1	1	1	0	0	1	4
Glutaryl-CoA dehydrogenase deficiency	1	1	0	0	0	0	2	4
Holocarboxylase synthase deficiency	0	1	1	1	3	0	0	6
Biotinidase deficiency	1	0	1	0	0	1	0	3
Hydroxymethylglutaryl-CoA lyase deficiency	0	0	1	0	0	1	0	2
Methylglutaconicaciduria	1	1	0	0	1	0	1	4
3-Methylcrotonyl CoA carboxylase deficiency	0	0	0	0	0	0	3	3
3-Ketothiolase deficiency	0	0	0	0	1	0	3	4

We identified patients with any of the 31 disorders who received a diagnosis after clinical referral and who were born during the six four-year periods from April 1974 through March 1998 or during the first four years when screening was being performed, from April 1, 1998, through March 31, 2002. The diagnosis was confirmed in almost all of these patients by enzymatic or molecular genetic tests. In cases of argininosuccinate lyase deficiency and three of five cases of short-chain acyl-CoA dehydrogenase deficiency detected by newborn screening, we relied on biochemical variables. To take into account a possible late age at the time of clinical diagnosis, we examined the age at diagnosis in all patients born between 1974 and 1998. For disorders diagnosed in any child at ages beyond 4.5 years (the interval between the end of the last period before screening began and the time of writing),

Table 1. **(Continued.)**

Disorder	Screening Year							Total
	1974–1978	1978–1982	1982–1986	1986–1990	1990–1994	1994–1998	1998–2002	
	number of patients							
Fatty acid								
Short-chain acyl-CoA dehydrogenase deficiency	0	0	1	0	0	0	5†	6
Medium-chain acyl-CoA dehydrogenase deficiency	2	4	3	3	7+1	1+2	17	37
Very-long-chain acyl-CoA dehydrogenase deficiency	0	0	1	2	0	0	3	6
Long-chain 3-hydroxyacyl-CoA dehydrogenase deficiency	0	0	1	1	3	2	0	7
Multiple acyl-CoA dehydrogenase deficiency	0	0	1	0	0	0	0	1
Carnitine transporter defect	0	0	0	2	1	2	3	8
Carnitine palmitoyl transferase deficiency								
Type I	0	0	0	0	0	0	0	0
Type II	1	0	0	0	0	0	0	1
Carnitine acylcarnitine translocase deficiency	0	0	0	0	0	0	1‡	1
Unadjusted total	25	22	29	32	34+1	26+2	57	228
Adjustment for estimated missing data due to late age at diagnosis	1	1	1	1	2	6		
Adjusted total	26	23	30	33	36+1	32+2	57	

* All periods began on April 1 and ended on March 31. A second number (following a plus sign) represents the number of patients given a diagnosis of medium-chain acyl-coenzyme A (CoA) dehydrogenase deficiency after the disorder had been diagnosed in a younger sibling through newborn screening. These cases are included in the unadjusted total for all the disorders combined but not in the total for the particular disorder.
† Three patients with probable cases of short-chain acyl-CoA dehydrogenase deficiency had elevated urinary ethylmalonic acid and elevated plasma butyrylcarnitine levels (3.53 to 5.06 μmol per liter [reference range, 0.08 to 0.24]), but skin-fibroblast studies were declined.
‡ This patient had a presumed diagnosis only: there was a substantial elevation of the palmitoylcarnitine concentration (20.0 μmol per liter); the patient subsequently died suddenly on day 3, but no tissue was available for testing.

we calculated the likely number of patients in the 1994–1998 birth cohort whose disorder remained undiagnosed. Similarly, for the 1990–1994 cohort, we calculated the likely number of patients whose disorder remained undiagnosed for 8.5 years, and so forth. For patients with cystathionine β-synthase deficiency, diagnoses were in some cases made in adulthood, and we estimated the expected number of diagnoses per four-year birth period on the basis of our data base of 51 patients with this disorder.

DIAGNOSIS DURING THE NEWBORN-SCREENING PERIOD

Between April 1998 and March 2002, blood samples obtained at 48 to 72 hours of life from all infants born in New South Wales or the Australian Capital Territory were tested by tandem mass spectrometry, as previously described.[11] Samples were butylated. Spectra were initially interpreted with the use of NeoLynx software (Micromass). Amino acids and acylcarnitines were quantitated against dried-blood-spot calibrators. Only selected compounds were analyzed, so as to avoid the identification of benign disorders. The results were transferred to a central data base, where each result was checked against predefined algorithms and reference ranges (described previously[11]), including reference ranges for ratios and for second samples, when requested. Results were generally available within 24 hours. Confirmatory testing was performed by our biochemical genetics laboratory, as was the investigation of patients in this cohort who presented with suggestive symptoms but had not been identified by newborn screening as having a disorder.

BIRTH COHORTS

The numbers of infants born during the four-year periods were taken from data of the Australian Bureau of Statistics on birth registrations. The number of first tests performed in any year was more than the number of births registered, despite double checks for duplication. Routine periodic checks according to name and hospital of birth indicate that coverage is greater than 99 percent.

STATISTICAL ANALYSIS

Fisher's exact test was used to compare the number of patients in the 1998–2002 birth cohort in whom a given disorder was diagnosed with the number of patients with a diagnosis of that disorder during previous years. Chi-square tests of linear trend were performed with the use of Epi Info software, version 6.[14]

RESULTS

The numbers of patients with the 31 target disorders born during the four-year periods before screening by tandem mass spectrometry was implemented and during the first four years of screening with this technique are shown in Table 1. The total number of births and the prevalence of each class of disorder are shown in Table 2. During the six four-year periods preceding the implementation of screening with mass spectrometry, 22 to 34 cases were diagnosed per period, resulting in rates of 6.6 to 9.0 cases per 100,000 births. When a possible late age at diagnosis was taken into account, the expected final range for these cohorts is estimated to be 23 to 36 cases per period, with rates of 6.9 to 9.5 cases per 100,000 births. There was no trend toward increased overall rates of diagnosis of these disorders between 1982 and 1998 (Table 2). Specifically, there was no increase in the rate of diagnosis of disorders of fatty-acid metabolism over this period, although many of these disorders were described for the first time during these years.

During the four years in which tandem mass spectrometry was used routinely for newborn screening, we tested 362,000 newborns, of whom 560 (0.15 percent) required a second test or urgent follow-up because of an abnormal test result. During the screening period, 57 newborns were given a diagnosis of 1 of the 31 inborn errors of metabolism (15.7 diagnoses per 100,000 births [95 percent confidence interval, 11.9 to 20.4]). Of these cases, 48 were diagnosed by newborn screening, and 6 of those diagnosed by screening were also diagnosed clinically before or at the same time as the screening result became available. Two patients, siblings with ornithine transcarbamylase deficiency born to a mother with known risk, did not undergo newborn screening, since both screening and

Table 2. **Incidence of the 31 Disorders, According to Type of Disorder and Birth Cohort.***

Variable	Screening Year							P Value for Comparison of 1982–1998 with 1998–2002
	1974–1978	1978–1982	1982–1986	1986–1990	1990–1994	1994–1998	1998–2002	
Urea-cycle disorders	7	5	6	8	7	4	7	0.70
Amino-acid disorders	10	5	10	12	7	8	9	0.15
Organic-acid disorders	5	8	6	4	9	9	12	0.11
Medium-chain acyl-CoA dehydrogenase deficiency	2	4	3	3	7+1	1+2	17	<0.001
Other fatty-acid–oxidation disorders	1	0	4	5	4	4	12	0.007
No. of births	336,000	331,000	349,000	360,000	378,000	367,000	362,000	
Incidence per 100,000 births	7.4	6.6	8.3	8.9	9.0	7.0	15.7	
Adjusted incidence per 100,000 births†	7.9	6.9	8.6	9.2	9.5	8.7		

* A second number (following a plus sign) represents the number of patients given a diagnosis of medium-chain acyl-CoA dehydrogenase deficiency after the disorder had been diagnosed in a younger sibling through newborn screening.
† Data are adjusted for estimated missing data due to late age at diagnosis.

treatment for the affected sons were declined. Seven patients in whom disorders were later diagnosed clinically had a negative result on newborn screening.

The greatest increase in the rate of diagnosis was for medium-chain acyl-CoA dehydrogenase deficiency (Table 1). The overall positive predictive value of an abnormal screening test was 10 percent, but the positive predictive value varied among analytes. For example, an abnormal level of tyrosine had a positive predictive value of only 2 percent for the detection of type I or type II tyrosinemia. The overall specificity was high, with a false positive rate of only 0.15 percent. In some patients with either a positive or a negative result on newborn screening the diagnosis was made on clinical grounds (Table 3).

The incremental cost of tandem mass spectrometry within the laboratories where newborn screening and biochemical genetic testing were conducted was approximately $0.70 (1.17 Australian dollars; amounts are reported here in U.S. dollars) per newborn screened. This amount included the cost of reagents, microtiter plates and other consumables, maintenance and depreciation of instruments, staffing, and confirmatory tests. The mean cost of confirmatory testing for infants who actually required it was $217. The cost per relevant disorder detected (excluding phenylketonuria) was $3,939. Including phenylketonuria, the cost was $2,519. Data on costs incurred outside of the newborn-screening and biochemical genetics laboratories are not yet available.

DISCUSSION

It is now possible to screen rapidly, simultaneously, and inexpensively for a number of very rare disorders with the use of tandem mass spectrometry, but the yield and usefulness of testing have not been clear. It is difficult to conduct a randomized, controlled trial of such screening because of the rarity of the individual disorders (necessitating the use of a very large sample) and because the strong belief of many proponents in the benefit of early diagnosis by tandem mass spectrometry has led to a public campaign in the United States and elsewhere for universal screening.[16]

Our study was designed to assess the diagnostic potential of this technology. To evaluate the technology fully, studies of the clinical effectiveness of screening and a detailed cost analysis will be necessary, but we do not yet have data to address these questions.

Our study design has potential drawbacks that we have attempted to counter. Although we used historical controls, the rate of clinical diagnosis for the 31 disorders had remained stable between 1982 and 1998, before tandem mass spectrometry was introduced for routine screening. Our centralized services and a high level of awareness of inborn errors of metabolism among pediatricians make it likely that we would be aware of patients in our region who have these disorders. Because of a high level of cooperation among state referral centers, children with such disorders diagnosed at any of the other five diagnostic centers in Australia would be routinely brought to our attention, and diagnosis outside of Australia would be very unlikely.

We have tried to account for the presence of children in the four-year birth cohorts from 1974 to 1998 that probably have undiagnosed cases by correcting for the expected number of cases remaining undiagnosed at the end of each four-year period. Although it is difficult to apply the same methods to the screened cohort, and this could affect the apparent sensitivity of the screening method, any such correction would not alter the overall conclusion that the target disorders have been diagnosed in more patients during the screening period than during the periods before screening. We cannot systematically account for patients who die with undiagnosed disorders. Although all coroners' offices have a protocol for the collection and investigation of samples in cases of possible metabolic disease, we do not have data on whether such investigations were routinely performed.

The two disorders that were clearly more frequently diagnosed by screening than clinically were medium-chain and short-chain acyl-CoA dehydrogenase deficiencies. Probably the rates of diagnosis of very-long-chain acyl-CoA dehydrogenase, 3-ketothiolase, and 3-methylcrotonyl CoA carboxylase deficiencies were also increased, although none of these increases was individually significant. During the past 24 years, only one case each of short-chain acyl-CoA dehydrogenase deficiency and 3-ketothiolase deficiency, three cases of very-long-chain acyl-CoA dehydrogenase deficiency, and no cases of 3-methylcrotonyl CoA carboxylase deficiency had been diagnosed clinically. It appears likely that some patients with these disorders could be at very low risk of ever having symptoms,[17] and short-chain acyl-CoA dehydrogenase deficiency and 3-methylcrotonyl CoA carboxylase deficiency may be largely benign disorders. There is no clear evidence yet to indicate whether early detection of

Table 3. **Presentation and Results on Newborn Screening in Patients with a Clinically Diagnosed Disorder.***

Patient No.	Disorder	Presentation	Result on Newborn Screening†	Comments
Positive result on newborn screening				
1	Glutaryl-CoA dehydrogenase deficiency	Hypotonia at 10 mo; MRI results suggestive of disorder	Retrospectively, glutarylcarnitine, 3.2 μmol/liter on day 3 (cutoff, 0.5 μmol/liter)	First 2 wk of screening and analyte not included in panel
2	Propionyl-CoA carboxylase deficiency	Metabolic acidosis on day 4	Propionylcarnitine, 16.0 μmol/liter (cutoff, 10 μmol/liter); result available on day 7	
3	Ornithine transcarbamylase deficiency	Prenatal diagnosis; treatment for 12 hr	Citrulline, 4 μmol/liter before treatment started (lower limit, <5 μmol/liter)	
4	Nonketotic hyperglycinemia	Hypotonia, seizures; diagnosis at 24 hr	Glycine, 1210 μmol/liter (cutoff, 1100 μmol/liter)	
5	Nonketotic hyperglycinemia	Hypotonia, seizures; diagnosis at 72 hr	Glycine, 1230 μmol/liter (cutoff, 1100 μmol/liter)	
6	Argininosuccinate synthase deficiency	Hyperammonemia; diagnosed at 72 hr	Citrulline, 230 μmol/liter (upper cutoff, 75 μmol/liter)	
Negative result on newborn screening				
7	Tyrosinemia, type I	Liver failure at 2 mo	Tyrosine 226 at 95th percentile (cutoff, 500 μmol/liter)	No alteration of the cutoff planned
8	Tyrosinemia, type II	Diagnosed interstate at 9 mo	Tyrosine 220 at 94th percentile (cutoff, 500 μmol/liter)	No alteration of the cutoff planned
9	Glutaryl-CoA dehydrogenase deficiency	Hypotonia from birth; gradually increasing dystonia; diagnosed at 10 mo	Glutarylcarnitine, 1.2 μmol/liter; repeated sample, 0.6 μmol/liter (cutoff at the time, 0.8 μmol/liter)	Protocol altered to recommend immediate clinical and biochemical review, without request for a second sample; cutoff lowered to 0.5 μmol/liter
10	Cobalamin C defect	Hypotonia and failure to thrive at 5 wk	Propionylcarnitine, 7.5 μmol/liter, 95th percentile; acetylcarnitine, 24.6 μmol/liter; methionine, 19 μmol/liter (cutoff, 5 μmol/liter)	Retrospectively found the ratio of propionylcarnitine to acetylcarnitine elevated (0.3; cutoff, 0.25); now this ratio used as primary variable for diagnosis‡
11	Very-long-chain acyl-CoA dehydrogenase deficiency	Hypoglycemia at 12 mo	Initial sample: tetradecanoylcarnitine, 1.3 μmol/liter; tetradecenoylcarnitine, 2.0 μmol/liter (cutoff for both at the time, 2.0 μmol/liter); repeated sample on day 14: tetradecanoylcarnitine, 0.4 μmol/liter; tetradecenoylcarnitine, 1.3 μmol/liter	Cutoff levels have been modified to 1.5 μmol/liter for both and to 0.9 μmol/liter for repeated samples
12	Nonketotic hyperglycinemia	Hypotonia, seizures on day 1	Glycine, 560 μmol/liter of whole blood (cutoff, 1100 μmol/liter)	No alteration of the cutoff planned
13	3-Ketothiolase deficiency	Severe ketosis at 15 mo	2- or 3-Hydroxy-3-methylbutyrylcarnitine, 1.20 μmol/liter of whole blood (cutoff, 1.50 μmol/liter)	No alteration of the cutoff planned

* MRI denotes magnetic resonance imaging.
† Values are in micromoles per liter of whole blood.
‡ Change in protocol is according to Chace et al.[15]

short-chain acyl-CoA dehydrogenase deficiency would be clinically useful. All programs of newborn screening by tandem mass spectrometry have resulted in the diagnosis of cases of 3-methylcrotonyl CoA carboxylase deficiency, which was previously reported extremely rarely,[12] and have led to the diagnosis of maternal cases of this deficiency.[18]

After phenylketonuria, medium-chain acyl-CoA dehydrogenase deficiency was the most common disorder detected, with a prevalence of 1 per 10,000 births to 1 per 20,000 births; it is potentially lethal. Before screening was implemented, 25 percent of children with diagnosed cases died, usually during the first episode of decompensation, and an additional 30 to 40 percent had some developmental delay.[19] Some patients died during the newborn period.[20] From family studies, it was known that many patients with medium-chain acyl-CoA dehydrogenase deficiency remain healthy,[21] and population studies of the common mutation have indicated that a substantial proportion of cases remain undiagnosed. However, in a recent study in the United Kingdom, symptoms were common in patients whose disorder was diagnosed by retrospective screening of stored newborn samples but had escaped clinical diagnosis.[22]

In another report, some of the cases of medium-chain acyl-CoA dehydrogenase deficiency that had been diagnosed through newborn screening by tandem mass spectrometry in the United States have involved an unusual mutation that had not previously been seen in clinically diagnosed cases.[23] Two of our patients carried this mutation and may not be at risk for the development of symptoms. However, it was not possible to determine with certainty whether any of our patients were indeed at low or no risk of a dangerous decompensation, since they all had a demonstrable functional (biochemical) abnormality.[24] All such patients require careful follow-up and a management plan, including avoidance of fasting and maintenance of caloric intake during episodes of infection. With appropriate management, there appears to be a negligible risk of serious decompensation or death after diagnosis,[19] although formal studies are lacking.

For certain disorders, false negative test results, obtained primarily during the first two years of the screening program, have led us to alter the cutoff point used in our protocols for the definition of the disorder (Table 3). These alterations have resulted in an overall increase in the false positive rate of approximately 0.01 percent, but a review of the results obtained before the change in cutoff value did not reveal any other likely missed cases.

The ultimate sensitivity of testing by tandem mass spectrometry will depend on what false positive rate will be tolerated, both overall and for each disorder. This tolerable level may, in turn, be predicated on the perceived seriousness of each disorder and the urgency of early treatment. Our study so far indicates a high sensitivity for most disorders in the target group but also defines some disorders in which sensitivity is likely to be low. Sensitivity was apparently 100 percent for cases of phenylketonuria and pterin disorders (data not shown) and is likely to be close to 100 percent for medium-chain acyl-CoA dehydrogenase deficiency.[24] There are clearly diagnostic problems for type I tyrosinemia unless another assay for succinylacetone is added[25]; given the rarity of this disorder in our population, we have elected not to add such an assay. It is unlikely that pyridoxine-responsive homocystinuria can be detected without a high false positive rate by the current strategy of measuring methionine.[26]

New South Wales has a mixed population, largely derived from the United Kingdom and Ireland, but with substantial contribution from southern Europe, the Middle East, and Asia. We believe that these results are likely to be mirrored in many other parts of the world.

The screening of newborns by tandem mass spectrometry detects more cases than are diagnosed after clinical presentation, but the excess cases seem to be confined to a small number of disorders. The long-term outcomes and costs associated with neonatal screening for these disorders require further study. However, the diagnosis of cases that might never come to clinical attention should not be used as an argument against this expanded program of newborn screening. With the possible exceptions of short-chain acyl-CoA dehydrogenase deficiency and 3-methylcrotonyl CoA carboxylase deficiency, the disorders in question can all lead to substantial morbidity and mortality. It is important that follow-up tests be performed by experienced biochemical genetics laboratories and that clinical care be provided by physicians experienced in the management of inborn errors of metabolism, so that the benefit of early detection can be maximized and the risk of harm and unnecessary worry minimized.

We are indebted to Professor Jennifer Peat and Dr. Katrina Williams for advice on statistics.

REFERENCES

1. Chatfield S, Owen G, Ryley HC, et al. Neonatal screening for cystic fibrosis in Wales and the West Midlands: clinical assessment after five years of screening. Arch Dis Child 1991;66:29-33.
2. Farrell PM, Kosorok MR, Rock MJ, et al. Early diagnosis of cystic fibrosis through neonatal screening prevents severe malnutrition and improves long-term growth. Pediatrics 2001;107:1-13.
3. U.S. Congress, Office of Technology Assessment. Healthy children: investing in the future. Washington, D.C.: Government Printing Office, 1988. (Publication no. OTA-H-345.)
4. Lees CM, Davies S, Dezateux C. Neonatal screening for sickle cell disease. Cochrane Database Syst Rev 2000;2:CD001913.
5. Wilcken B. Rare diseases and the assessment of intervention: what sorts of clinical trials can we use? J Inherit Metab Dis 2001; 24:291-8.
6. Albers S, Marsden D, Quackenbush E, Stark AR, Levy HL, Irons M. Detection of neonatal carnitine palmitoyltransferase II deficiency by expanded newborn screening with tandem mass spectrometry. Pediatrics 2001;107:1417. abstract.
7. Ranieri E, Gerace R, Bartlett B, Barnard K, Fletcher JM. The introduction of tandem mass spectrometry into the South Australian Neonatal Screening Programme: benefits and costs. J Inherit Metab Dis 2000;23:Suppl 1:3. abstract.
8. Roscher A, Liebl B, Fingerhut R, Olgemöller B. Prospective study of MS-MS newborn screening in Bavaria, Germany: interim results. J Inherit Metab Dis 2000;23: Suppl 1:4. abstract.
9. Wilcken B, Wiley V, Carpenter K. Two years of routine newborn screening by tandem mass spectrometry (MSMS) in New South Wales, Australia. J Inherit Metab Dis 2000;23:Suppl 1:4. abstract.
10. Wilcken B, Wiley V, Sim KG, Carpenter K. Carnitine transporter defect diagnosed by newborn screening with electrospray tandem mass spectrometry. J Pediatr 2001;138: 581-4.
11. Wiley V, Carpenter K, Wilcken B. Newborn screening with tandem mass spectrometry: 12 months' experience in NSW Australia. Acta Paediatr Suppl 1999;88:45-51.
12. Zytkovicz TH, Fitzgerald EF, Marsden D, et al. Tandem mass spectrometric analysis for amino, organic, and fatty acid disorders in newborn dried blood spots: a two-year summary from the New England Newborn Screening Program. Clin Chem 2001;47: 1945-55.
13. Wood JC, Magera MJ, Rinaldo P, Seashore MR, Strauss AW, Friedman A. Diagnosis of very long chain acyl-dehydrogenase deficiency from an infant's newborn screening card. Pediatrics 2001;108:173-4. abstract.
14. Dean AG, Dean JA, Coulombier D, et al. Epi Info, version 6: a word processing, database, and statistics program for epidemiology on microcomputers. Atlanta: Centers for Disease Control and Prevention, 1994.
15. Chace DH, DiPerna JC, Kalas TA, Johnson RW, Naylor EW. Rapid diagnosis of methylmalonic and propionic acidemias: quantitative tandem mass spectrometric analysis of propionylcarnitine in filter-paper blood specimens obtained from newborns. Clin Chem 2001;47:2040-4.
16. Marshall E. Fast technology drives new world of newborn screening. Science 2001; 294:2272-4. [Erratum, Science 2002;295: 2370.]
17. Rhead WJ, Allain D, Van Calcar S, et al. Short chain acyl-CoA dehydrogenase (SCAD) and 3-methylcrotonyl-CoA carboxylase (MCC) deficiencies: tandem mass spectrometry newborn screening detects many clinically benign cases. J Inherit Metab Dis 2002;25: Suppl 1:4. abstract.
18. Gibson KM, Bennett MJ, Naylor EW, Morton DH. 3-Methylcrotonyl-coenzyme A carboxylase deficiency in Amish/Mennonite adults identified by detection of increased acylcarnitines in blood spots of their children. J Pediatr 1998;132:519-23.
19. Iafolla AK, Thompson RJ Jr, Roe CR. Medium-chain acyl-coenzyme A dehydrogenase deficiency: clinical course in 120 affected children. J Pediatr 1994;124:409-15.
20. Leung KC, Hammond JW, Chabra S, Carpenter KH, Potter M, Wilcken B. A fatal neonatal case of medium-chain acyl-coenzyme A dehydrogenase deficiency with homozygous A→G985 transition. J Pediatr 1992;121: 965-8.
21. Heptinstall LE, Till J, Wraith JE, Besley GT. Common MCAD mutation in a healthy parent of two affected siblings. J Inherit Metab Dis 1995;18:638-9.
22. Pourfarzam M, Morris A, Appleton M, Craft A, Bartlett K. Neonatal screening for medium-chain acyl-CoA dehydrogenase deficiency. Lancet 2001;358:1063-4.
23. Andresen BS, Dobrowolski SF, O'Reilly L, et al. The mutational spectrum in the MCAD gene of newborns identified by prospective tandem MS screening for "diagnostic" acyl-carnitines in blood spots differs from that observed in clinically affected patients. J Inherit Metab Dis 2000;23:Suppl 1:12. abstract.
24. Carpenter K, Wiley V, Sim KG, Heath D, Wilcken B. Evaluation of newborn screening for medium chain acyl-CoA dehydrogenase deficiency in 275 000 babies. Arch Dis Child Fetal Neonatal Ed 2001;85:F105-F109.
25. Schulze A, Frommhold D, Hoffmann GF, Mayatepek E. Spectrophotometric microassay for delta-aminolevulinate dehydratase in dried-blood spots as confirmation for hereditary tyrosinemia type I. Clin Chem 2001;47:1424-9.
26. Peterschmitt MJ, Simmons JR, Levy HL. Reduction of false negative results in screening of newborns for homocystinuria. N Engl J Med 1999;341:1572-6.

Evaluation of newborn screening for medium chain acyl-CoA dehydrogenase deficiency in 275 000 babies

K Carpenter, V Wiley, K G Sim, D Heath, B Wilcken

Abstract

***Objective*—To evaluate newborn screening by tandem mass spectrometry for detection of medium chain acyl-CoA dehydrogenase (MCAD) deficiency, a fatty acid oxidation disorder with significant mortality in undiagnosed patients.**

***Design*—The following were studied: (*a*) 13 clinically detected MCAD deficient subjects, most homozygous for the common A985G mutation, whose newborn screening sample was available; (*b*) 275 653 consecutive neonates undergoing routine newborn screening. Screened infants with blood octanoylcarnitine levels ⩾ 1 μmol/l were analysed for the A985G mutation, had analysis of plasma and repeat blood spot acylcarnitines and urinary organic acids, and had fibroblast fatty acid oxidation or acylcarnitine studies.**

***Result*—Twelve of the 13 patients later diagnosed clinically had newborn octanoylcarnitine levels > 2.3 μmol/l. Twenty three screened babies had initial octanoylcarnitine levels ⩾ 1 μmol/l. Eleven of 12 babies with persistent abnormalities had metabolite and/or enzyme studies indicating MCAD deficiency. Only four were homozygous for the A985G mutation, the remainder carrying one copy.**

***Conclusions*—Most patients with symptomatic MCAD deficiency could be detected by newborn screening. Infants actually detected had a lower frequency of A985G alleles than clinically diagnosed cases and may have a lower risk of becoming symptomatic.**

(*Arch Dis Child Fetal Neonatal Ed* 2001;**85**:F105–F109)

Keywords: acylcarnitines; octanoylcarnitine; fatty acid oxidation; acylglycines; Reye syndrome

New South Wales Newborn Screening Programme, The Children's Hospital at Westmead, Sydney, Australia
V Wiley
D Heath
B Wilcken

Biochemical Genetics Service, The Children's Hospital at Westmead
K Carpenter
K G Sim
B Wilcken

Correspondence to:
Dr Wilcken, The Children's Hospital at Westmead, Locked Bag 4001, Westmead, NSW 2145, Australia
Bridgetw@chw.edu.au

Fatty acid oxidation in the mitochondria is an essential source of cellular energy. Medium chain acyl-CoA dehydrogenase (MCAD) deficiency is the commonest disorder of mitochondrial fatty acid oxidation. It is recessively inherited, and is particularly common in populations of northern European origin.[1] Several studies have shown that about 80% of clinically ascertained cases are homozygous for an A to G transition at position 985, resulting in a lysine to glutamate substitution,[2–4] with almost all of the remainder having one copy of this mutation together with some other rare mutation.

MCAD deficiency may result in episodic hypoketotic hypoglycaemia and hepatic encephalopathy triggered by periods of catabolic stress, suggestive of Reye syndrome. Episodes occur most often between 3 months and 6 years of age, but a proportion of the diagnosed patients have symptoms in the neonatal period, sometimes with a fatal outcome.[5] An overall mortality of about 25% has been reported,[6–8] but it is clear that some people remain asymptomatic throughout life.[9] Almost all deaths have been in previously undiagnosed patients. After diagnosis, institution of management to avoid fasting stress, especially during intercurrent illnesses, appears effective in avoiding serious morbidity and mortality.[6] Patients with MCAD deficiency have raised urinary acylglycines, even when well, and have impaired oxidation of medium chain fatty acids, which can be detected in cultured skin fibroblasts.[10]

Mass screening for MCAD deficiency (and other disorders) has become practicable through the use of tandem mass spectrometry to analyse acylcarnitines in dried blood spots.[11] In the United Kingdom, two health technology assessments have examined the implications of expanded newborn screening, including screening for MCAD deficiency.[12 13] The second study suggested that among many disorders potentially detectable in neonates by tandem mass spectrometry, a case could only be made for introducing screening for MCAD deficiency and glutaric aciduria type 1, and recommended monitoring of the clinical effectiveness and cost, although such an evaluation has not yet taken place.[14] The recommendations for MCAD deficiency were based in part on the clear benefit from early detection and management and also on a calculated incidence of MCAD deficiency derived from A985G carrier frequency. However, early results from such programmes have indicated a much lower frequency of homozygotes for the A985G mutation than expected, and a higher frequency of compound heterozygotes.[15 16]

The NSW Newborn Screening Programme screens all babies born in New South Wales and the Australian Capital Territory (about 95 000 births per year), and has used tandem mass spectrometry since April 1998. Parents may dissent from the screening tests as a whole, are handed a descriptive brochure, and are informed by nursing staff about what tests are to be carried out. Randomised controlled trials of screening for such individually rare disorders are not practicable because of the enormous numbers that would be needed for adequate power.[17] Other levels of evidence must be taken into account. In this study, we report the results of screening for MCAD deficiency in over a quarter of a million neonates.

Methods

SUBJECTS

Newborn screening samples were collected from > 99% of babies born in New South Wales and the Australian Capital Territory between April 1998 and March 2001.The total number of babies tested was 275 653. Median age at sampling was 3 days, and over 99% of babies were sampled before day 6. Samples collected before 48 hours of life because of early discharge (1.9%) were deemed inappropriate, and second samples were collected from those babies. However, the initial samples were assayed. Stored newborn screening samples were also available from 13 babies born between January 1981 and June 1997 who were subsequently diagnosed as suffering from MCAD deficiency. Twelve of the samples were taken on days 4–6, with one taken on day 10. This comprised all the available newborn samples from babies with MCAD deficiency.

LABORATORY METHODS

Acylcarnitines were analysed as butyl esters on a Micromass Quattro II tandem mass spectrometer as previously described.[18] Mutational analysis for A985G was based on the method of Gregersen.[19] Urinary organic acids were extracted with ethyl acetate and analysed as their trimethylsilyl esters on a Hewlett-Packard 5890 GC and 5971 MSD. Assays of the rate of fatty acid oxidation in cultured skin fibroblasts monitored tritium release with oleate and myristate as substrates.[20]

Fibroblast acylcarnitine accumulation in culture medium was analysed by tandem mass spectrometry after incubation with palmitate and carnitine for 72 hours. Results are expressed as means of replicate (between two and six) analyses.[21] In two patients, a specific electron transfer flavoprotein (ETF) based assay for MCAD activity was performed by Dr C Vianey-Saban (Hôpital Debrousse, Lyons, France).

NEWBORN SCREENING PROTOCOL

All samples with octanoylcarnitine ≥ 0.8 μmol/l were assayed again in duplicate in the laboratory. If the octanoylcarnitine level in the final assay was ≥ 1 μmol/l, a second dried blood sample together with a plasma and urine sample were requested, and analysis for the common A985G mutation performed. At clinical follow up, a skin biopsy was performed for culture of fibroblasts.

CRITERIA FOR ASSIGNMENT OF DIAGNOSIS OF MCAD

Patients were diagnosed with MCAD deficiency if one or more of the following criteria were met: homozygous for A985G mutation; raised hexanoylglycine and suberylglycine in urine, increased hexanoylcarnitine, octanoylcarnitine, or decenoylcarnitine in plasma; studies of fibroblast fatty acid oxidation rate with reduced myristate oxidation and an oleate to myristate oxidation ratio > 1.3; or accumulation of medium chain acylcarnitines in culture medium from skin fibroblasts, with octanoyl carnitine > 1.1 nmol/mg protein/72 hours. In two cases in which ETF assays were performed, the results were taken as the definitive measure and the diagnosis of affected or carrier assigned.

Results

RETROSPECTIVE ANALYSIS OF NEWBORN SCREENING SAMPLES FROM PATIENTS LATER DIAGNOSED WITH MCAD DEFICIENCY, AND DEVELOPMENT OF A CUT OFF VALUE

We performed retrospective analysis of newborn screening cards from 13 patients diagnosed clinically with MCAD deficiency. Eleven of the 13 were homozygous for the common MCAD mutation A985G, and two were heterozygous. Twelve had octanoylcarnitine concentrations well in excess of 1 μmol/l (range 2.4–6.8). The remaining baby,[22] who was homozygous for the common mutation, developed hypoglycaemia and hepatic encephalopathy on day 2, leading to carnitine depletion, and had an octanoylcarnitine level of 0.27 μmol/l. Prospective analysis of newborn screening samples from 24 000 newborns showed 99.2% to have octanoylcarnitine concentrations of < 0.3 μmol/l, all the remainder falling between 0.3 and 0.7 μmol/l. None had a level greater than 1 μmol/l, and this level was chosen as the cut off value for prospective screening.[18] The level of octanoylcarnitine did not alter appreciably in newborn screening samples stored for up to three years, and in no case did it increase (data not shown).

PROSPECTIVE NEWBORN SCREENING RESULTS

Distribution of octanoylcarnitine concentrations in dried blood spots did not vary greatly with birth weight or day of sampling.

Samples from 23 babies had octanoylcarnitine values greater than 1 μmol/l in the initial dried blood spot sample. All were tested for the common A985G mutation, and a repeat dried blood spot sample requested. Eleven of the 23 did not carry A985G on either allele, and in six of these the octanoylcarnitine had normalised on repeat sampling. One baby had a blood octanoylcarnitine concentration of 1.5 μmol/l on retesting, but further investigation of plasma carnitines and urinary acylglycines did not indicate MCAD deficiency (data not shown). The remaining four infants died from a variety of causes before a second sample could be collected. One had hydrops foetalis. MCAD deficiency was eliminated on enzyme analysis in cultured skin fibroblasts. Three others died before further investigations were possible, but information obtained from the responsible clinicians, including postmortem findings which did not show any increased hepatic fat, did not suggest that MCAD deficiency could have been the cause of death.

Twelve infants had persistently abnormal acylcarnitines and carried at least one copy of A985G. These were investigated further by analysis of urinary organic acids and/or quantitation of plasma carnitine and acylcarnitine. Six patients had a skin biopsy, and studies of oxidation rate were performed on cultured skin fibroblasts. Table 1 summarises the results. Overall 11 patients had results that were clearly

Table 1 Initial dried blood spot and follow up assays in 12 babies with persistently increased octanoylcarnitine and at least one copy of A985G

	Newborn screening		*Biochemical genetics follow up*					*Fibroblast results*				
			Urinary acylglycines		*Plasma carnitines (μmol/l)*				*Acylcarnitine profile (nmol/mg/72 h)*			*ETF assays MCAD activity (Dr Vianey-Saban) (2.42 (0.61) n=50)*
Patient	*Octanoyl-carnitine (μmol/l) (cut off <1)*	*A985G mutational analysis*	*Hexanoyl*	*Suberyl*	*C6 (<0.1)*	*C8 (<0.3)*	*C10:1 (<0.3)*	*Fatty acid oxidation rate assay % of intrabatch controls O:M ratio (0.7–1.3)*	*C6 (0.2–1.4)*	*C8 (0.1–1.1)*	*C10 (0.6–1.6)*	
1	24.6	+/−	+++	+++	0.6	2.0	0.3	NP	0.81	2.06	0.34	NP
2	14.2	+/+	NP	NP	1.3	5.6	0.5	NP	NP	NP	NP	NP
3	13.0	+/−	+++	+++	1.0	3.8	0.6	Ole 63%, Myr 10% O:M 6.4	2.90	6.97	1.25	NP
4	8.6	+/−	+++	+++	0.9	3.8	0.3	NP	1.56	4.07	0.66	NP
5	8.0	+/+	+++	+++	0.5	2.6	0.3	NP	NP			NP
6	8.0	+/+	+++	+++	1.3	8.4	2.5	NP	NP			NP
7	8.0	+/+	+++	+++	1.1	5.7	0.7	NP	NP			NP
8	4.9	+/−	++	++	0.2	1.1	0.4	Ole 51–87%, Myr 31–52% O:M 1.6–1.7	1.60	3.07	1.43	NP
9	3.5	+/−	+	++	1.1	3.1	0.4	NP	NP			NP
10	3.3	+/−	+	+	0.2	0.7	0.2	Ole 107%, Myr 38% O:M 2.8	2.35	7.81	4.01	NP
11	1.1	+/−	TR, ND	TR, ND	0.7	1.9	1.0	Ole 41–59%, Myr 42–44% O:M 1.0–1.3	0.92	1.03	0.67	0.44, 0.31
12	1.1	+/−	ND	ND	NP	NP	NP	Ole 32%, Myr 22% O:M 1.5	1.45	1.52	1.11	1.28, 1.21

Patients 1–11 were assigned a diagnosis of medium chain acyl-CoA dehydrogenase (MCAD) deficiency, and patient 12 was considered to be a carrier at low risk of developing symptoms. +/+, homozygous for A985G; +/−, heterozygous for A985G; Ole, oleate; Myr, myristate; O:M, oleate to myristate ratio; C4, butyrylcarnitine; C6, hexanoylcarnitine; C8, octanoylcarnitine; C10, decanoylcarnitine; C10:1, decenoylcarnitine; ND, not detected; NP, not performed; NI, not increased; +++, grossly elevated; ++, moderately increased; +, slightly increased; TR, trace increase.

consistent with MCAD deficiency. Four were homozygous for the common mutation, ten had clearly increased urinary acylglycines (one had not had urine collected), and in six who had skin biopsies, there was decreased oxidation of myristate and an increased oleate to myristate oxidation ratio in cultured skin fibroblasts or pronounced accumulation of octanoylcarnitine on fibroblast acylcarnitine profiling. One further patient (patient 11) was classified as MCAD deficient on the basis of decreased MCAD activity in cultured skin fibroblasts using a specific ETF based assay performed by Dr C Vianey-Saban. This patient had intermittently raised urinary acylglycines and decreased oxidation of myristate in cultured skin fibroblasts, but the rate of oleate oxidation was also decreased, and the oleate to myristate ratio was not typical of MCAD deficiency (table 1).

PREVALENCE OF THE A985G MUTATION IN MCAD DEFICIENT PATIENTS

The 11 babies classified as MCAD deficient in the newborn screening cohort carried the common mutation on 15 of the 22 alleles. There were 17 MCAD deficient patients clinically diagnosed in New South Wales before screening began. Sixteen of these had had DNA analysis for the common mutation, which was carried on 29 of the 32 alleles, a non-significant difference (Fisher's Exact Test).

Discussion

MCAD deficiency has been considered appropriate for inclusion in newborn screening programmes because it is a disorder with significant morbidity and mortality, and estimates of incidence based on mutation testing suggest the defect remains underdiagnosed.[4] Several studies have shown that around 80% of patients presenting clinically are homozygous for the AG985 mutation, with about 18% being compound heterozygotes.[23] These figures are in keeping with our own experience in New South Wales, Australia where 29 of 32 alleles from 16 of our clinically presenting patients carried the A985G mutation. Combined figures from three neonatal population studies in Australia (one each in New South Wales, Victoria, and Queensland) have shown the prevalence of A985G carriers to be 1:86 (17 of 1461). This implies a birth prevalence of homozygous A985G MCAD deficiency of 1:29 500 (95% confidence interval 1:11 500 to 1:87 000) and an overall birth prevalence of MCAD deficiency of 1:19 000.[24 25]

As in most other centres, before introduction of newborn screening, our rate of diagnosis of MCAD deficiency was considerably lower than that predicted by the carrier frequency for A985G and the frequency of this mutation in MCAD alleles of clinically diagnosed patients. From 1980 to 1997 we would have expected 86 cases of MCAD deficiency, but only 15 were clinically ascertained. Seven of these were born between 1990 and 1997, the years immediately before screening began, when 38 would have been expected. The assumption was that the missing patients had had no or very mild symptoms or had died without a diagnosis being made. The introduction of newborn screening has led to the discovery of 11 new patients in 36 months, but only four of these are homozygous for the common mutation. Several groups have reported this unexpected mutational spectrum from patients identified by newborn screening.[15 16 26–28] Our finding of 4:275 653 (1:68 913) patients homozygous for A985G is considerably lower than the predicted prevalence from Australian figures but is still within the rather wide 95% confidence interval.

Two important questions arise from these findings. If the carrier frequencies are correct, and there is considerable agreement from many studies of populations derived from Northern Europe, could there be some homozygous A985G individuals who have normal acylcarnitine concentrations during the neonatal period and are thus missed on newborn screening? Secondly, do all those presumed compound heterozygotes detected by screening have the same risk of death and disability as homozygotes, or are we detecting and treating some patients who have very little chance of becoming symptomatic?

Our retrospective analysis of newborn screening cards from previously diagnosed patients suggests that patients with later symptoms are very unlikely to have had normal acylcarnitine profiles in the newborn period. The one baby who did have a normal octanoylcarnitine concentration was homozygous for the A985G mutation, had severe hepatic encephalopathy as a neonate, and had carnitine depletion.[22] No other acylcarnitine measure or ratio would have distinguished her from healthy neonates. It remains possible that not all patients homozygous for the A985G mutation will show raised octanoylcarnitine in the neonatal period. Maybe other polymorphisms, perhaps affecting chaperonins, could offer protection by stabilising the mutant protein and normalising enzyme activity, thus accounting for the "missing" cases.

The non-A985G mutations in our compound heterozygotes are probably rare and diverse. The question of whether patients with certain mutations are likely to be mildly affected, with high residual enzyme activities, has been investigated by expression studies, but a correlation of genotype to phenotype is not straightforward.[29] It is important to remember, however, that these individuals are being detected because of a raised octanoylcarnitine, which in itself is a clear indication of a functional impairment of their ability to metabolise fatty acids. It is very likely that environmental or other genetic factors will determine if an affected individual becomes symptomatic. One of our compound heterozygote neonates (patient 1, table 1) did indeed become hypoglycaemic, requiring transfer to a neonatal intensive care unit, but has remained well since on conservative treatment. She had very high octanoylcarnitine concentration in her newborn screening sample.

Two cases illustrate the problems of atypical follow up results in heterozygous A985G infants detected by screening. Patient 12 had considerably reduced rates of oxidation of both medium and long chain fatty acids in cultured skin fibroblasts, but a specific assay for MCAD activity showed results in the heterozygous MCAD range (table 1). His risk of developing symptoms is unknown, but he may well carry a defect in another part of the fatty acid oxidation pathway that results in his biochemical findings. Patient 11 had urine, plasma, and fibroblast results atypical for MCAD deficiency, but her specific ETF based assay of MCAD activity suggests that she does have MCAD deficiency. However, mutational analysis of all exons in the gene failed to identify another mutation (Professor B Andreson, Aarhus University, Denmark).

It is clear from our retrospective study that newborn screening using tandem mass spectrometry can detect almost all patients with MCAD deficiency who would later have presented with symptoms, and some of these will certainly be saved from death or morbidity with simple and appropriate management. Some other babies are detected who have abnormal MCAD activity, but may have a benign disorder. At present it is not possible to distinguish these cases from those at risk of decompensation, and all will require counselling and a stringent management plan. Thus, decisions on expansion of newborn screening programmes to include MCAD deficiency should be taken in the knowledge that this is a complex disorder, and that not all cases discovered by newborn screening may be at risk. Screening programmes uncover unexpected findings, and these are not always predictable from theoretical projections.[12 13 30]

Professor Brage Andresen, Aarhus, sequenced the exons of the MCAD gene in one of the patients. Dr Christine Vianey-Saban, Lyons, performed an ETF based enzyme assay on skin fibroblasts from two of the patients. All staff of the NSW Newborn Screening Programme and the staff of the NSW Biochemical Genetics Service gave valuable help in different aspects of the study.

1 Tanaka K, Gregersen N, Ribes A, *et al*. A survey of the newborn populations in Belgium, Germany, Poland, Czech Republic, Hungary, Bulgaria, Spain, Turkey, and Japan for the G985 variant allele with haplotype analysis at the medium chain acyl-CoA dehydrogenase gene locus: clinical and evolutionary consideration. *Pediatr Res* 1997;**41**:201–9.
2 Blakemore AI, Singleton H, Pollitt RJ, *et al*. Frequency of the G985 MCAD mutation in the general population [letter]. *Lancet* 1991;**337**:298–9.
3 Curtis D, Blakemore AI, Engel PC, *et al*. Heterogeneity for mutations in medium chain acyl-CoA dehydrogenase deficiency in the UK population. *Clin Genet* 1991;**40**:283–6.
4 Pollitt RJ, Leonard JV. Prospective surveillance study of medium chain acyl-CoA dehydrogenase deficiency in the UK. *Arch Dis Child* 1998;**79**:116–19.
5 Wilcken B, Carpenter KH, Hammond J. Neonatal symptoms in medium chain acyl coenzyme A dehydrogenase deficiency. *Arch Dis Child* 1993;**69**:292–4.
6 Iafolla AK, Thompson RJJ, Roe CR. Medium-chain acyl-coenzyme A dehydrogenase deficiency: clinical course in 120 affected children. *J Pediatr* 1994;**124**:409–15.
7 Touma EH, Charpentier C. Medium chain acyl-CoA dehydrogenase deficiency. [Review] *Arch Dis Child* 1992;**67**:142–5.
8 Wilcken B, Hammond J, Silink M. Morbidity and mortality in medium chain acyl coenzyme A dehydrogenase deficiency. *Arch Dis Child* 1994;**70**:410–12.
9 Heptinstall LE, Till J, Wraith JE, *et al*. Common MCAD mutation in a healthy parent of two affected siblings. *J Inherit Metab Dis* 1995;**18**:638–9.
10 Stanley C. Disorders of fatty acid oxidation. In: Fernandez J, Saudubray JM, van den Berghe G, eds. *Inborn metabolic disease*. Berlin: Springer-Verlag, 2000;139–50.
11 Naylor EW, Chace DH. Automated tandem mass spectrometry for mass newborn screening for disorders in fatty acid, organic acid, and amino acid metabolism. *J Child Neurol* 1999;**14**(suppl 1):S4–8.
12 Pollitt RJ, Green A, McGill J, *et al*. Neonatal screening for inborn errors of metabolism: cost, yield and outcome. *Health Technol Assess* 1997;**1**:1–202.
13 Seymour CA, Thomason MJ, Chalmers RA, *et al*. Newborn screening for inborn errors of metabolism: a systematic review. [Review] *Health Technol Assess* 1997;**1**:i–iv.
14 Tanner S, Sharrard M, Cleary M, *et al*. Screening for medium chain acyl-CoA dehydrogenase deficiency has still not been evaluated. *BMJ* 2001;**322**:112.
15 Marsden D, Zykovicz T, Larson C, *et al*. Spectrum of expression of MCAD and SCAD detected by newborn screening. *J Inherit Metab Dis* 2000;**23**(suppl 1):14.
16 Ziadeh R, Hoffman EP, Finegold DN, *et al*. Medium chain acyl-CoA dehydrogenase deficiency in Pennsylvania: neonatal screening shows high incidence and unexpected mutation frequencies. *Pediatr Res* 1995;**37**:675–8.
17 Wilcken B. Rare diseases and the assessment of intervention. *J Inherit Metab Dis* 2001;in press.

18 Wiley V, Carpenter K, Wilcken B. Newborn screening with tandem mass spectrometry: 12 months' experience in NSW Australia. *Acta Paediatr* 1999;**88**(suppl):48–51.
19 Gregersen N, Blakemore AI, Winter V, *et al.* Specific diagnosis of medium-chain acyl-CoA dehydrogenase (MCAD) deficiency in dried blood spots by a polymerase chain reaction (PCR) assay detecting a point-mutation (G985) in the MCAD gene. *Clin Chim Acta* 1991;**203**:23–34.
20 Olpin SE, Manning NJ, Pollitt RJ, *et al.* The use of [9,10-^{3}H]myristate, [9,10-^{3}H]palmitate and [9,10-^{3}H]oleate for the detection and diagnosis of medium and long-chain fatty acid oxidation disorders in intact cultured fibroblasts. *Adv Exp Med Biol* 1999;**466**:321–5.
21 Sim KG, Wiley V, Carpenter K, *et al.* Carnitine palmitoyltransferase I deficiency in a neonate identified by dried blood spot free carnitine and acylcarnitine profile. *J Inherit Metab Dis* 2001;**24**:51–9.
22 Christodoulou J, Hoare J, Hammond J, *et al.* Neonatal onset of medium-chain acyl-coenzyme A dehydrogenase deficiency with confusing biochemical features. *J Pediatr* 1995;**126**:65–8.
23 Coates PM, Tanaka K. Workshop on molecular aspects of MCAD deficiency. Mutations causing medium-chain acyl-CoA dehydrogenase deficiency: a collaborative compilation of the data from 172 patients. In: Coates PM, Tanaka K, eds. *New developments in fatty acid oxidation.* New York: Wiley-Liss, 1992;499–506.
24 Heath, D. Neonatal hypoglycaemia and heterozygosity for the 985A to G common mutation of medium chain acyl-CoA dehydrogenase (MCAD) deficiency. Thesis, University of Western Sydney, 1996.
25 Matsubara Y, Narisawa K, Tada K. Medium-chain acyl-CoA dehydrogenase deficiency: molecular aspects. [Review] *Eur J Pediatr* 1992;**151**:154–9.
26 Andresen B, Dobrowolski SF, O'Reilly L, *et al.* The mutational spectrum in the MCAD gene of newborns identified by prospective tandem MS screening for "diagnostic" acyl-carnitnes in blood spots differs from that observed in clinically affected patients. *J Inherit Metab Dis* 2000;**23**:12.
27 Marsden D, Zykovicz T, Larson C, *et al.* Prevalence of fatty acid oxidation disorders and organic acidaemias in New England newborns screened by tandem mass spectrometry. *J Inherit Metab Dis* 2000;**23**:15.
28 Carpenter K, Wiley V, Sim KG, *et al.* Newborn screening for MCAD deficiency: the New South Wales experience. *J Inherit Metab Dis* 2000;**23**:16.
29 Andresen BS, Bross P, Udvari S, *et al.* The molecular basis of medium-chain acyl-CoA dehydrogenase (MCAD) deficiency in compound heterozygous patients: is there correlation between genotype and phenotype? *Hum Mol Genet* 1997;**6**:695–707.
30 Clayton PT, Doig M, Ghafari S, *et al.* Screening for medium chain acyl-CoA dehydrogenase deficiency using electrospray ionisation tandem mass spectrometry. *Arch Dis Child* 1998;**79**:109–15.

Available online at www.sciencedirect.com
SCIENCE DIRECT®
Molecular Genetics and Metabolism 87 (2006) 32–39
Molecular Genetics and Metabolism
www.elsevier.com/locate/ymgme

Medium-chain acyl-CoA dehydrogenase deficiency: Genotype–biochemical phenotype correlations

Leigh Waddell [a,d], Veronica Wiley [a], Kevin Carpenter [b], Bruce Bennetts [c], Lyn Angel [d], Brage S. Andresen [e,g], Bridget Wilcken [a,b,f,*]

[a] *The NSW Newborn Screening Programme, Children's Hospital at Westmead, Sydney, Australia*
[b] *The NSW Biochemical Genetics Laboratory, Children's Hospital at Westmead, Sydney, Australia*
[c] *The Molecular Genetics Laboratory, Children's Hospital at Westmead, Sydney, Australia*
[d] *School of Biomedical Sciences, Charles Sturt University, Australia*
[e] *Research Unit for Molecular Medicine, Aarhus University Hospital, Denmark*
[f] *Discipline of Paediatrics, University of Sydney, New South Wales, Australia*
[g] *Institute of Human Genetics, Aarhus University, Denmark*

Received 19 July 2005; received in revised form 27 September 2005; accepted 27 September 2005
Available online 15 November 2005

Abstract

The fatty acid oxidation disorder most commonly identified by tandem mass spectrometry newborn screening is the potentially fatal medium-chain acyl-CoA dehydrogenase deficiency (MCAD). In clinically presenting cases, 80% are homozygous for the common mutation, c.985A > G and 18% heterozygous. We screened 592,785 babies and identified 34 with MCAD, 17 homozygous for c.985A > G, 14 with one copy, and 3 with no copy. We sequenced the exons of 19 patients, the 17 carrying one or no copy of c.985A > G, and two with marginal findings, and examined correlations between groups of mutations and biochemical markers. We found two known or putative pathogenic mutations in 18 of the 19 patients. Two mutations appeared more than once: c.199T > C, not recorded in clinically presenting cases ($n = 4$), and c.583G > A ($n = 2$). Patients homozygous for c.985A > G had the highest levels of neonatal octanoylcarnitine, plasma octanoylcarnitine when asymptomatic, and urinary acylglycines. Compound heterozygotes of c.985A > G and other mutations had intermediate levels, and those without c.985A > G, or heterozygous for that and c.199T > C had the lowest levels of these analytes. There was overlap in all values. The c.985A > G and c.583G > A mutations appear to have functional effects towards the severe end of the spectrum, and the c.199T > C mutation a smaller effect, as has been previously postulated. If these results are confirmed and extended, this could influence the advice given to parents of babies with MCAD detected by newborn screening, and make management more specific. In the meantime, all MCAD patients identified by newborn screening have, by definition, a functional defect and require careful clinical management.

Keywords: Newborn screening; Fatty acid oxidation; c.583G > A; c.199T > C; Octanoylcarnitine; Tandem mass spectrometry; MCAD

Introduction

Medium-chain acyl-CoA dehydrogenase deficiency (MCAD) is a potentially fatal disorder of the β-oxidation of fatty acids. Symptoms of the disorder are induced by catabolic stress, and while some patients die or suffer catastrophic neurological damage during episodes of decompensation with hypoketotic hypoglycaemia, other patients with the same genotype may never have symptoms [1]. After diagnosis, a clinical management plan to avoid or rapidly treat catabolic stress is successful in preventing adverse outcomes [2]. Newborn screening for MCAD by tandem mass spectrometry (MSMS) has good sensitivity and specificity [3–5], but far more cases of MCAD are being detected than were found clinically [6]. In cases detected by newborn screening, the frequency of the prevalent

* Corresponding author. Fax: +612 9845 3121.
E-mail address: bridgetw@chw.edu.au (B. Wilcken).

1096-7192/$ - see front matter © 2005 Elsevier Inc. All rights reserved.
doi:10.1016/j.ymgme.2005.09.020

c.985A > G mutant allele has been lower than that observed in clinically affected patients, and some have had mutations not previously recorded [3,7]. It has been suggested that some genotypes in screening-detected cases are "mild" and unlikely to be associated with clinical disease [8].

Between April 1998 and October 2004, we screened 592,785 babies by MSMS and identified 36 with biochemical evidence of MCAD. All were tested for the c.985A > G mutation. The patients biochemically ascertained to have MCAD but not homozygous for c.985A > G had further analysis to determine their genotype. We correlated genotype with biochemical phenotype as demonstrated by newborn screening results (samples taken at 48–72 h), plasma acylcarnitine, and urinary hexanoylglycine levels of samples taken during a clinically stable phase, and, in a limited subset, the acylcarnitine profile obtained with cultured skin fibroblasts.

Methods

MSMS screening and diagnosis of MCAD

Newborn screening by MSMS has been carried out in NSW since April 1998, as has been previously described, using dried blood samples obtained on all babies at 48–72 h of age [6,9]. MCAD deficiency was suspected in babies whose initial blood octanoylcarnitine level was 1.0 μmol/L or greater, and whose octanoyl:decanoylcarnitine ratio was >1.0. To confirm MCAD deficiency, these patients were further investigated by analysis of urinary organic acids, including acylglycines by gas chromatography–mass spectrometry (GC–MS), plasma acylcarnitines by MSMS, and DNA extraction and analysis for the common mutation was carried out on the initial dried blood spot. As previously described [9], the cut-off value for octanoylcarnitine of 1.0 μmol/L was arrived at after a pilot study of 25,000 neonates, and retrospective analysis of stored neonatal dried blood spot samples from patients with MCAD clinically diagnosed, and following an early period where further investigation was initiated at 0.8 μmol/L. Urinary hexanoylglycine was semi-quantified as part of routine organic acid analysis and scored as slight (approximately 5–10 μmol/mmol creatinine), moderate (10–20) or gross (>20) increase.

Further DNA analysis

The exons, known mutations, and single nucleotide polymorphisms (SNPs) were mapped and highlighted on the *ACADM* gene sequence (NM_000016.2). Fragments covering the exons, splice sites, and all known mutations were PCR-amplified with the product assessed by polyacrylamide gel electrophoresis and subsequently sequenced to identify variations. These variations were assessed for pathogenicity. Missense mutations were assessed using the PolyPhen website: http://www.bork.enbl-heidelberg.de/PolyPhen/ which scores variations using multiple sequence alignments. Sequence variants potentially involving splice sites were assessed using a splice-site prediction programme: http://www.fruitfly.org/seq_tools/splice-instrucs.html and compared against the consensus motifs (GTRAGT for the 5′ splice site and YYYYYYYYYYYN CAG for the 3′ splice site). Correlations were then made with respect to their biochemical and clinical phenotypes.

Genotype–biochemical phenotype correlations

The patients were classified into the following mutational categories: c.985A > G/c.985A > G ($N = 17$); c.985A > G/c.199T > C($N = 4$); c.985A > G/c.583G > A($N = 2$); c.98 5 A > G/other, i.e., compound heterozygotes ($N = 9$); and other/other, i.e., no c.985A > G mutation ($N = 4$). These categories were then correlated with the initial dried blood spot octanoylcarnitine concentrations (during the catabolic stress of the newborn period, day 2–3 of life), the plasma acylcarnitines, and urine acylglycines obtained in most cases before 21 days of age (median 15; mean 10 days), when the baby was stable, to identify the relationship between the genotype and neonatal and later biochemical phenotype.

Results

Patients identified by screening were considered to have MCAD deficiency if they had two previously described disease-causing mutations and/or increased urinary hexanoylglycine. Among 592,785 screened neonates, 60 babies had octanoylcarnitine levels of 1 μmol/L or greater, and among these we confirmed 34 MCAD patients, a positive predictive value for the test of 0.57. The birth prevalence of MCAD was 1 in 17,435. There were two additional patients fulfiling some but not all of the diagnostic criteria (see below). All 34 confirmed MCAD patients had increased urinary hexanoylglycine and typical elevations of medium-chain plasma acylcarnitines, with increased octanoyl/decanoylcarnitine ratios (data not shown). Seventeen of the 34 (50%) were homozygous for the common mutation c.985A > G. Fourteen (41%) had one copy of this mutation. A further three (9%) had no copy of the common mutation. Thus, 71% of alleles carried the common c.985A > G mutation. Of the two biochemically atypical unconfirmed cases, one had one copy of the common mutation.

The fragment containing exon 1 proved difficult to amplify, possibly due to its high GC content. Therefore, we were only able to sequence exon 1 in 9 of the 19 patients. This was not pursued as there had been no mutations reported in the literature in this exon. The results obtained in our patients are shown in Table 1.

Seven different previously reported mutations were found in 11 of the patients, six of which were heterozygous with the common mutation (c.985A > G). Four of these mutations (c.199T > C; c.799G > A; c.1012insTAGAATGAGTTAC; and c.583G > A) are known to be associated with reduced enzyme activity and have been reported in the Human Gene

Table 1
Biochemical and genotype data on 36 patients with biochemical evidence of MCAD detected on newborn screening

Patient	Age at newborn screening sample (days)	Newborn screening octanoylcarnitine (μmol/L)	Urinary hexanoylglycine	Plasma octanoylcarnitine (μmol/L)	Mutation results
1	2	35.0	+++	7.8	c.985A > G/c.985A > G
2	3	26.0	NP	NP	c.985A > G/c.985A > G
3	3	22.5	+++	0.5	c.985A > G/c.985A > G
4	3	21.0	+++	6.5	c.985A > G/c.985A > G
5	2	17.3	+++		c.985A > G/c.985A > G
6	3	16.5	++	1.5	c.985A > G/c.985A > G
7	2	16.0	+++	7.6	c.985A > G/c.985A > G
8	1	16.0	+++	2.4	c.985A > G/c.985A > G
9	3	14.2	NP	5.6	c.985A > G/c.985A > G
10	3	13.5	++	4.5	c.985A > G/c.985A > G
11	3	10.0	+++	5.9	c.985A > G/c.985A > G
12	3	9.0	+++	8	c.985A > G/c.985A > G
13	3	8.1	++	4.1	c.985A > G/c.985A > G
14	4	8.0	+++	2.6	c.985A > G/c.985A > G
15	4	8.0	+++	NP	c.985A > G/c.985A > G
16	3	7.1	++	5	c.985A > G/c.985A > G
17	4	5.0	+++	2.7	c.985A > G/c.985A > G
18	3	37.0	+++	4.9	c.985A > G/c.583G > A
19	3	24..6	+++	2	c.985A > G/c.1012insTAGAATGAGTTAC
20	0	18.5	+++	3.6	c.985A > G/c.387delG*
21	3	15.0	+++	3.1	c.985A > G/c.583G > A
22	3	13.0	+++	3.8	c.985A > G/c.311A > G
23	3	8.6	+++	3.8	c.985A > G/IVS10 – 2A > C
24	3	4.9	++	1.1	c.985A > G/c.199T > C
25	3	4.5	+	3.3	c.985A > G/c.199T > C
26	2	3.9	++	1.4	c.985A > G/c.617G > A
27	4	3.6	+++	9.3	c.985A > G/IVS6 + 7A > G
28	3	3.5	+	3.1	c.985A > G/c.799G > A
29	4	3.3	+	0.7	c.985A > G/c.253G > T
30	3	1.2	+	1	c.985A > G/c.199T > C
31	3	1.2	++	0.4	c.985A > G/c.199T > C
32	3	5.9	++	2.1	c.250C > T/c.694C > T
33	3	4.0	+++	2.9	IVS3 + 5G > T/c.616C > T
34	3	3.1	+++	2.4	–/–
35	2	4.8	TR	0.3	c.890A > G/IVS10 + 4delAinsGC
36	3	1.1	TR,ND	1.9	c.985A > G/c.127G > A

ND, not detected; NP, not performed; TR, trace increase; +, slight increase; ++, moderate increase; +++, gross increase.

Mutation Database http://uwcmml1s.uwcm.ac.uk/uwcm/mg/search/118958.html. The c.387delG mutation and the c.616C > T mutation have recently been reported in other patients found by newborn screening [7]. The c.387delG leads to a shifted reading frame and thus has severe consequences for the encoded protein. The c.616C > T causes a change in amino acid from arginine (CGT), a basic amino acid, to the polar, uncharged cysteine (TGT) within the middle domains β-sheet 6. Prediction using the PolyPhen website ranked this variation as potentially damaging with an alignment score of 2.420.

The c.127G > A mutation (patient 36) is located in the B α-helix of the N-terminal domain. It changes glutamic acid (GAA), an acidic amino acid, to lysine (AAA), a basic amino acid. Such a change in the polarity of the amino acid would be expected to cause a significant change in the formation of this region of the protein and affect the overall structure and function of the protein. However, this patient, who carried one copy of the common mutation, had a plasma acylcarnitine profile in which the C8/C10 ratio was less than 1, and did not fully fulfil our criteria for MCAD deficiency, with hexanoylglycine only marginally and intermittently increased. The c.127G > A mutation has previously been reported by Zytkovicz et al. [5] in a baby detected by newborn screening who did not have increased urinary acylglycines.

Two genotypes occurred more than once in independent patients. The genotype c.985A > G/c.199T > C that has been noted to be common in other newborn screening populations [3,7] was identified in four of the total of 34 MCAD patients, accounting for 6% of alleles, and present in 12% of the NSW MCAD population. The only other genotype found more than once in the cohort was c.985A > G/c.583G > A, found in two of the patients.

Novel sequence variations

Seven of the remaining eight patients had among them nine sequence variations of uncertain pathogenicity. None

of the nine variations identified were found in homozygous form. Five were heterozygous with the common c.985A > G mutation. The IVS10 − 2A > C changes the invariant A of the AG at the splice site, so that the normal 3′ splice site will not be used. This mutant splice site does not conform to the consensus motifs for splice sites and it is therefore most likely that this mutation is deleterious.

The c.253G > T mutation changes glycine at position 60 to a cysteine. The importance of glycine at this position is underscored by the fact that it is conserved across species and also in the other acyl-CoA dehydrogenases with a PolyPhen score of 3.040. The c.253G > T mutation has been detected in a newborn identified in another screening program (unpublished data), indicating its pathogenic potential.

The c.311A > G mutation changes aspartic acid at position 79 to a glycine. This is a drastic change and would be expected to have a severe effect. Aspartic acid at this position is conserved in MCAD from other species, but not in other acyl-CoA dehydrogenases. A test with the PolyPhen program gives a score of 0.948 and suggests that this change would be benign. It is therefore not clear what the effect of this mutation is.

The IVS3 + 5G > T variant (patient 33) changes the consensus G to a non-consensus T at the +5 position. This will compromise this splice site. Mutation of this position has been shown to cause miss-splicing in other genes [10]. If this variant does destroy the splice site, the gene would encode the intron as well, adding an extra 37 codons and causing a frameshift (V48fsX84).

Exon 4 contains the c.250C > T mutation (patient 32) and the resultant leucine (CTT) to phenylalanine (TTT) change is in the D α-helix of the N-terminal. Both of these amino acids are non-polar and hydrophobic, though their structures differ with the benzene ring in phenylalanine. The fact that we have found this mutation in three other newborns and in one patient with clinically manifest disease (Unpublished data) strongly suggests that this mutation is deleterious.

The pathogenicity of the IVS6+7G > A mutation is not easily resolved. The splice-site prediction program identified the 5′ splice donor site of intron 6 with a low score of 0.27. The normal 5′ splice donor site of intron 6 in the IVS6 + 7A > G variant (patient 27) was not identified though another 5′ splice donor site was predicted 50 nucleotides into the intron (AACAAAGgtgctatt) with a score of 0.49. This would suggest that a putative splice junction 50 bp into the intron may be used, adding 16 codons and causing a frameshift resulting in a stop codon at X489. IVS6 + 7A is conserved in all species except zebrafish. RNA studies would be needed for confirmation of this putative splicing error.

The c.617G > A mutation from exon 8 (patient 26) will cause arginine (CGT) to change to histidine (CAT). These are both basic amino acids, although the histidine has an imidazole ring in its structure, compared with the long tail of arginine. Prediction using the PolyPhen website is ranked as potentially damaging with an alignment score of 2.214. A mutation of the same nucleotide (c.617G > T) has been reported to cause disease in two patients [11].

The c.694C > T mutation (patient 32) will cause the polar, uncharged glutamine (CAG) to change to a premature termination codon (TAG). The last 190 amino acids would not be coded for, producing a truncated protein (completely lacking the C-terminal domain).

Patient 34 appears to have a number of polymorphisms but no clear-cut variations thought to be pathogenic were identified. The polymorphisms are almost identical to those seen in patients 27 and 28 (both of whom had other disease-causing mutations), except that two of them, both common polymorphisms (IVS3 + 10T > C and IVS5 + 32C > G) [3], are homozygous in patient 34 whereas they are only found on one allele in patients 27 and 28. All the above patients had undoubted MCAD deficiency biochemically.

In addition to patient 36, who had a previously reported genotype discussed above, one other patient did not fulfil our criteria. This was patient 35, who had two novel sequence variations, but marginal findings in confirmatory tests for MCAD deficiency. The c.890A > G mutation causes the acidic aspartic acid (GAT) to change to the neutral glycine (GGT), located on the same helix (H) from the C terminal as the common c.985A > G mutation. The IVS10 + 4delAinsGC mutation is located within the 5′ splice donor site region of intron 10. The mutant splice site has mismatches at the +4, +5, and +6 positions, which probably would inhibit stable base pairing with the U1 splice factor and thus cause miss-splicing.

Genotype–biochemical phenotype correlations

Genotype vs newborn screening octanoylcarnitine

Patients homozygous for c.985A > G had the highest median concentrations of octanoylcarnitine in the dried blood spot newborn screening samples (median 14 μmol/L; range 5–35). The genotypic group of c.985A > G/other compound heterozygotes was associated with the next highest group of octanoylcarnitine concentrations, but the range was very wide (median 4.9–8.6 μmol/L; range 3–37). The c.583G > A mutation also appeared to have a relatively severe biochemical effect, as judged by the newborn screening DBS octanoylcarnitine levels. The four patients with the c.985A > G/c.199T > C genotype had lower concentrations of octanoylcarnitine, 1.2–4.9 μmol/L, and samples with no c.985A > G mutation had octanoylcarnitine concentrations at the lower end of the scale. Fig. 1 shows the individual results among the mutation groups. The same trend was apparent when the genotype groups were related to octanoylcarnitine/decanoylcarnitine ratio (Fig. 2).

Genotype vs plasma acylcarnitine

The plasma concentrations of octanoylcarnitine, measured when the babies were stable, displayed the same overall trends but were less differentiated amongst the genotype groups, compared with the dried blood spot results (Fig. 3). The timing of the samples in relation to feeding was not recorded.

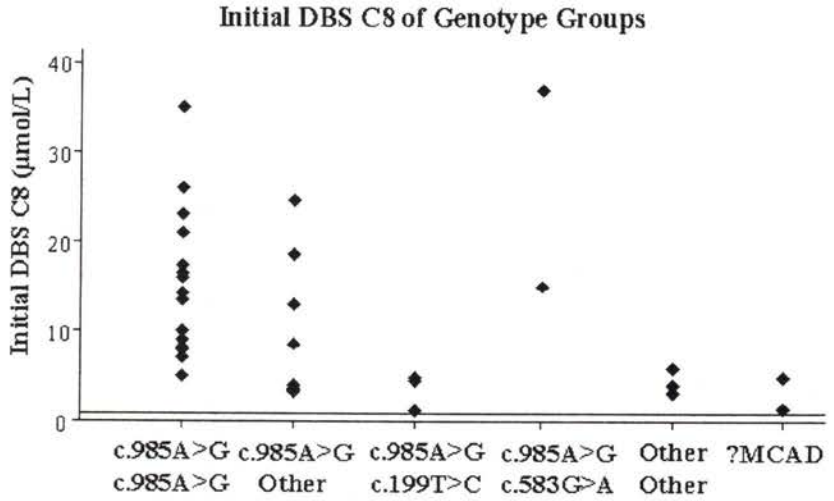


Fig. 1. Newborn screening octanoylcarnitine concentrations in different genotype groups. Solid line indicates cutoff for repeat testing (1 µmol/L) ?MCAD, this category of patient did not fulfil all our criteria for a diagnosis of MCAD (see text).

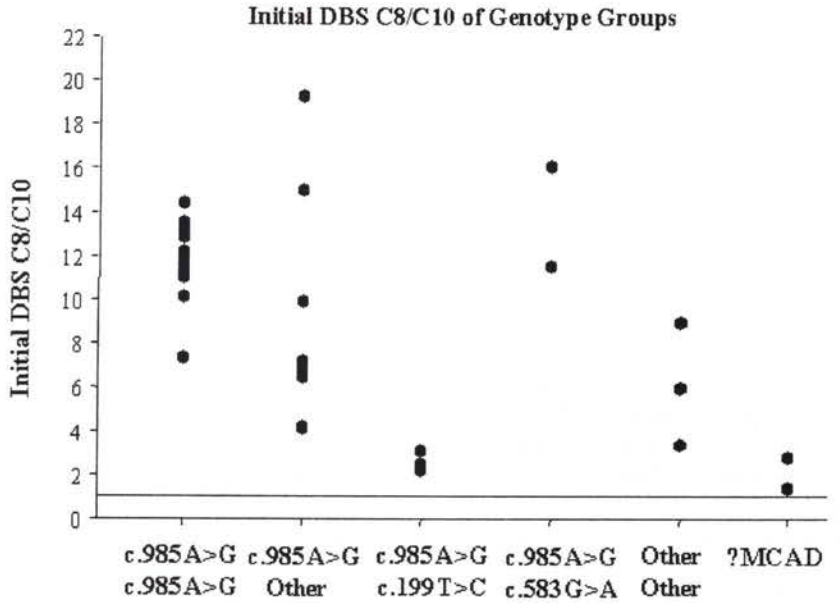


Fig. 2. Octanoyl to decanoylcarnitine ratio in initial newborn screening sample in different genotype groups. Solid line represents ratio of 1. ?MCAD, this category of patient did not fulfil all our criteria for a diagnosis of MCAD (see text).

Genotype vs urinary hexanoylglycine

The urinary hexanoylglycine concentrations were assessed semi-quantitatively with normal levels being below the limit of detection. An increased urinary hexanoylglycine level is required to assign a diagnosis of MCAD. Urinary hexanoylglycines were grouped according to the previously established genotype categories. All of the c.985A > G homozygous samples had moderately to grossly increased levels of urinary hexanoylglycine and most had similar increases in suberylglycine (data not shown), while the c.985A > G/c.199T > C genotype as a group had lower levels. As would be expected at this age, none of the infants excreted phenylpropionylglycine.

Clinical correlation

Two patients within the cohort had neonatal symptoms (patients 2 and 19) and patient 2 died as a consequence. He was homozygous for the common c.985A > G mutation, confirming its clinical severity, whereas subject 19, who became significantly ill with an episode of hypoglycaemia but did not die, was found to have c.985A > G and c.1012insTAGAATGAGTTAC. This suggests that this mutation falls into the "severe" category. Patient 22 also experienced a significant episode of decompensation, despite preventive measures. He was found to have c.985A > G and c.311A > G, also associating this genotype with a clinically severe phenotype.

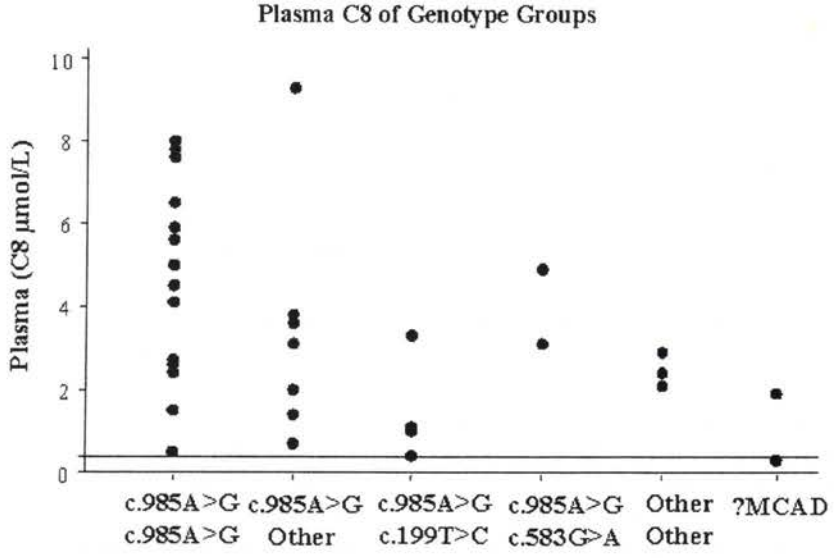


Fig. 3. Plasma octanoylcarnitine concentrations during a clinically stable phase in different genotype groups. Solid line represents upper limit of normal (0.4 µmol/L). ?MCAD, this category of patient did not fulfil all our criteria for a diagnosis of MCAD (see text).

Discussion

Screening for MCAD deficiency has been regarded as desirable because this is a potentially fatal disorder, and adverse outcomes can be avoided by appropriate management after diagnosis [12]. However, screening programmes diagnose many more cases than those that come to light clinically [13,14,6] and this poses a dilemma of whether too many cases are being diagnosed, and whether it may become possible to identify those at very low risk of decompensation. In this it is similar to other screening initiatives, with phenylketonuria screening also identifying benign hyperphenylalaninaemia, and cystic fibrosis screening identifying patients with clearly mild mutations [15]. In this study, we have investigated the mutations found in our MCAD patients detected by screening and correlated the mutations to biochemical indices of disease severity.

Our cut-off value of 1.0 µmol/L for octanoylcarnitine is higher than other published figures [5]. While it is not possible to compare levels exactly among different laboratories, our cut offs were established with regard to values obtained from retrospective analysis of newborn screening cards from clinically diagnosed patients, as well as pilot study data, and were optimised to maximise positive predictive value whilst maintaining sensitivity for the disorder in question.

Our study confirms the findings of previous reports about mutations in MCAD patients diagnosed by newborn screening, and extends these. The previous reports showed that in the US and in Bavaria, Germany, fewer screening-detected patients were homozygous for the common Northern European mutation c.985A>G, and the overall frequency of this mutation among the alleles of screened patients was much lower [3,7]. Tables 2 and 3 compare our findings with published newborn screening findings and those in clinically diagnosed patients. Among clinically presenting MCAD cases recorded by Tanaka et al. [16], 89% of ACADM alleles harboured the common c.985A>G mutation. We found this mutation to be present in 71% of alleles, a frequency falling somewhere between that found by Andresen et al. [3] and that found in Bavaria by Maier et al. [7] in screened newborns. While our percentage does not significantly differ from those of the other newborn screening series, there is a very significant difference between frequencies of the common mutation among the alleles of all screened babies when compared with clinically presenting cases ($\chi^2 = 15.8, p < 0.001$). The other prevalent mutation, c.199T>C, occurred in a similar small percentage of our population when compared with other screened populations. This mutation had never been reported (but may not have been excluded) from clinically presenting cases. However, recently, a patient with the 799G>A/199T>C genotype was reported as requiring hospital admis-

Table 2
Patient genotypes in newborn screening identified and clinically presenting MCAD patients

Reference	Andresen et al. [3]	Maier et al. [7]	Australia (SA and NSW 1998–2004)[a]	This study (NSW only)	All screened	Clinical Tanaka et al. [16]
Patients	59	57	46	34	162	172
c.985A > G homozygous (%)	66	47	54	50	56	80
c.985 A > G heterozygous (%)	30	32	35	41	32	18
No 985A > G (%)	3	21	11	9	11	2

[a] Combined data for NSW and South Australia unpublished. SA, South Australia; NSW, New South Wales.

Table 3
Independent allele frequencies in newborn screening identified and clinically presenting MCAD patients

Reference	Andresen et al. [3]	Maier et al. [7]	Australia (SA and NSW 1998–2004)[a]	This study (NSW only)	All screened	Clinical Tanaka et al. [16]
Alleles: no	118	114	94	68	326	344
c.985A > G (%)	81	63	74	71	73	89
c.199C > T (%)	6	8	4	6	6	?
Other (%)	13	29	21	24	21	11

[a] Combined data for NSW and South Australia unpublished. SA, South Australia; NSW, New South Wales.

sion during an episode of vomiting [17]. These results from our patients are similar to a larger series (including the patients reported here) representing patients screened in two states of Australia, from 1998 to 2004 inclusive (see Tables 2 and 3).

This general similarity of prevalent mutation frequency among screened cohorts is interesting, as the populations are somewhat different. Our population is quite mixed, with 13% of births born to parents themselves born outside Australasia, UK or Europe [18]. At least 10% of our newborns are of Asian origin. Maier and colleagues found that all but one of the MCAD patients not harbouring any copy of the common mutation were non-German, with six of the nine being Turkish, while Turkish births accounted for only about 4.6% of population births. In our series, those with no copy of the common mutation had parents whose origins were Indian (1), Italian (1), and European and Indonesian (1). While MCAD is thought to be very uncommon in Asia, there are certainly some reported cases [19]. Population differences probably account for the lower birth-incidence of MCAD we found. However, the definition of MCAD has not been clear in previous publications.

The common mutation c.985A > G was shown to be associated with the most pronounced biochemical markers. This mutation can be regarded as "severe," in that it is often associated with an adverse clinical phenotype. But since homozygotes for this mutation can be asymptomatic, the interpretation of severity is quite complex. We have confirmed that the c.199T > C mutation seems to account for a somewhat milder biochemical phenotype. Octanoylcarnitine levels in the newborn period in compound heterozygotes for this mutation and the "severe" common mutation were at the lower end of the distribution. However, there was overlap in values with other compound heterozygotes, although no overlap with values obtained for the group homozygous for the common mutation. The findings were similar but less marked in plasma of patients aged 1–3 weeks who were well and clinically stable. Even with this information, it is difficult to be sure that a patient with compound heterozygosity which includes c.199T > C would be at minimal risk for a significant adverse episode. Functional testing of c.199T > C (Y42H) mutation showed that it is temperature sensitive and would be likely to lose its residual enzyme activity if the patients experienced fever [17].

Elevation of hexanoylglycine has long been used as a definitive diagnostic marker for MCAD deficiency (although it is not specific, occurring also in multiple acyl-CoA dehydrogenase deficiency). Virtually all patients with a clinical presentation of MCAD deficiency would have increased urinary hexanoylglycine in order to be assigned this diagnosis. The two babies in whom hexanoylglycine was intermittently marginally increased or not detected, patients 35 and 36, included the baby previously reported in whom diagnosis of MCAD deficiency was somewhat uncertain [4].

Ideally, we would have liked to correlate our genotype with the biochemical phenotype revealed by the fibroblast acylcarnitine profile [20]. Unfortunately, patients diagnosed by newborn screening often did not have a skin biopsy for fibroblast culture, and of the 19 patients not homozygous for the common ACADM mutation, only seven had a skin fibroblast line—too few for any overall correlations. However, the two patients tested with the c.985A > G/c.199T > C mutation had generally lower accumulation of acylcarnitines in fibroblasts than those with other genotypes [20]. Interestingly, patient 19 had the lowest fibroblast acylcarnitine levels even though the newborn screening octanoylcarnitine was amongst the highest observed.

Evaluation of a correlation between clinical phenotype and genotype was usually not possible. In all cases, once babies were suspected of a diagnosis of MCAD a regimen to minimise fasting was implemented. Only three patients suffered episodes of decompensation: two as neonates, and one at age 3 years. The mutations they had (the common mutation and two others) are clearly potentially severe, but while there is certainly some correlation between genotype and phenotype, it is not at all straightforward, as previously noted [21]. Both environmental and epigenetic influences must be involved in the variations seen in clinical effect. In the meantime, we cannot easily dismiss from careful clinical management any of the MCAD patients identified by newborn screening, as all have a functional deficit, and all could possibly be at risk of clinically significant decompensation events.

Acknowledgment

B.S.A. was supported by a grant from the March of Dimes Foundation (Grant No. 1-FY-2003-688).

References

[1] L.E. Heptinstall, J. Till, J.E. Wraith, G.T. Besley, Common MCAD mutation in a healthy parent of two affected siblings, J. Inherit. Metab. Dis. 18 (1995) 638–639.

[2] C.J. Wilson, M.P. Champion, J.E. Collins, P.T. Clayton, J.V. Leonard, Outcome of medium chain acyl-CoA dehydrogenase deficiency after diagnosis, Arch. Dis. Child. 80 (1999) 459–462.

[3] B.S. Andresen, S.F. Dobrowolski, L. O'Reilly, J. Muenzer, S.E. McCandless, D.M. Frazier, S. Udvari, P. Bross, I. Knudsen, R. Banas, D.H. Chace, P. Engel, E.W. Naylor, N. Gregersen, Medium-chain acyl-CoA dehydrogenase (MCAD) mutations identified by MS/MS-based prospective screening of newborns differ from those observed in patients with clinical symptoms: identification and characterization of a new, prevalent mutation that results in mild MCAD deficiency, Am. J. Hum. Genet. 68 (2001) 1408–1418.

[4] K. Carpenter, V. Wiley, K.G. Sim, D. Heath, B. Wilcken, Evaluation of newborn screening for medium chain acyl-CoA dehydrogenase deficiency in 275 000 babies, Arch. Dis. Child. Fetal Neonatal Ed. 85 (2001) F105–F109.

[5] T.H. Zytkovicz, E.F. Fitzgerald, D. Marsden, C.A. Larson, V.E. Shih, D.M. Johnson, A.W. Strauss, A.M. Comeau, R.B. Eaton, G.F. Grady, Tandem mass spectrometric analysis for amino, organic, and fatty acid disorders in newborn dried blood spots: a two-year summary from the New England Newborn Screening Program, Clin. Chem. 47 (2001) 1945–1955.

[6] B. Wilcken, V. Wiley, J. Hammond, K. Carpenter, Screening newborns for inborn errors of metabolism by tandem mass spectrometry, N. Engl. J. Med. 348 (2003) 2304–2312.

[7] E.M. Maier, B. Liebl, W. Roschinger, U. Nennstiel-Ratzel, R. Fingerhut, B. Olgemoller, U. Busch, N. Krone, R. Kries, A.A. Roscher, Population spectrum of ACADM genotypes correlated to biochemical phenotypes in newborn screening for medium-chain acyl-CoA dehydrogenase deficiency, Hum. Mutat. 25 (2005) 443–452.

[8] J. Zschocke, A. Schulze, M. Lindner, S. Fiesel, K. Olgemoller, G.F. Hoffmann, J. Penzien, J.P. Ruiter, R.J. Wanders, E. Mayatepek, Molecular and functional characterisation of mild MCAD deficiency, Hum. Genet. 108 (2001) 404–408.

[9] V. Wiley, K. Carpenter, B. Wilcken, Newborn screening with tandem mass spectrometry: 12 months' experience in NSW Australia, Acta Paediatr. Suppl. 88 (1999) 48–51.

[10] K. Fujii, Y. Kohno, K. Sugita, M. Nakamura, Y. Moroi, K. Urabe, M. Furue, M. Yamada, T. Miyashita, Mutations in the human homologue of *Drosophila* patched in Japanese nevoid basal cell carcinoma syndrome patients, Hum. Mutat. 21 (2003) 451–452.

[11] B.Z. Yang, J.H. Ding, C. Zhou, M.M. Dimachkie, L. Sweetman, M.J. Dasouki, J. Wilkinson, C.R. Roe, Identification of a novel mutation in patients with medium-chain acyl-CoA dehydrogenase deficiency, Mol. Genet. Metab. 69 (2000) 259–262.

[12] R.J. Pollitt, A. Green, J. McGill, C.J. McCabe, A. Booth, N.J. Cooper, J.V. Leonard, J. Nicholl, P. Nicholson, J.R. Tunaley, N.K. Virdi, Neonatal screening for inborn errors of metabolism: cost, yield and outcome, Health Technol. Assess. (South Hampton, NY) 1 (1997) 1–202.

[13] T.G. Derks, M. Duran, H.R. Waterham, D.J. Reijngoud, L.P. Ten Kate, G.P. Smit, The difference between observed and expected prevalence of MCAD deficiency in The Netherlands: a genetic epidemiological study, Eur. J. Hum. Genet. 13 (2005) 947–952.

[14] M. Pourfarzam, A. Morris, M. Appleton, A. Craft, K. Bartlett, M. Pourfarzam, A. Morris, M. Appleton, A. Craft, K. Bartlett, Neonatal screening for medium-chain acyl-CoA dehydrogenase deficiency, Lancet 358 (2001) 1063–1064.

[15] R.J. Massie, B. Wilcken, P. Van Asperen, S. Dorney, M. Gruca, V. Wiley, K. Gaskin, Pancreatic function and extended mutation analysis in DeltaF508 heterozygous infants with an elevated immunoreactive trypsinogen but normal sweat electrolyte levels, J. Pediatr. 137 (2000) 214–220.

[16] K. Tanaka, I. Yokota, P.M. Coates, A.W. Strauss, D.P. Kelly, Z. Zhang, N. Gregersen, B.S. Andresen, Y. Matsubara, D. Curtis, Mutations in the medium chain acyl-CoA dehydrogenase (MCAD) gene, Hum. Mutat. 1 (1992) 271–279.

[17] L. O'Reilly, P. Bross, T.J. Corydon, S.E. Olpin, J. Hansen, J.M. Kenney, S.E. McCandless, D.M. Frazier, V. Winter, N. Gregersen, P.C. Engel, B.S. Andresen, The Y42H mutation in medium-chain acyl-CoA dehydrogenase, which is prevalent in babies identified by MS/MS-based newborn screening, is temperature sensitive, Eur. J. Biochem. 271 (2004) 4053–4063.

[18] Australian Bureau of Statistics, Table 8.14 Births Cat #3301.0 (2002) 56–57.

[19] R. Ensenauer, J.L. Winters, P.A. Parton, D.F. Kronn, J.W. Kim, D. Matern, P. Rinaldo, S.H. Hahn, Genotypic differences of MCAD deficiency in the Asian population: novel genotype and clinical symptoms preceding newborn screening notification, Genet. Med. 7 (2005) 339–343.

[20] K.G. Sim, K. Carpenter, J. Hammond, J. Christodoulou, B. Wilcken, Quantitative fibroblast acylcarnitine profiles in mitochondrial fatty acid beta-oxidation defects: phenotype/metabolite correlations, Mol. Genet. Metab. 76 (2002) 327–334.

[21] B.S. Andresen, P. Bross, S. Udvari, J. Kirk, G. Gray, S. Kmoch, N. Chamoles, I. Knudsen, V. Winter, B. Wilcken, I. Yokota, K. Hart, S. Packman, J.P. Harpey, J.M. Saudubray, D.E. Hale, L. Bolund, S. Kolvraa, N. Gregersen, The molecular basis of medium-chain acyl-CoA dehydrogenase (MCAD) deficiency in compound heterozygous patients: is there correlation between genotype and phenotype? Hum. Mol. Genet. 6 (1997) 695–707.

J. Inherit. Metab. Dis. 24 (2001) 291–298

Rare diseases and the assessment of intervention: What sorts of clinical trials can we use?

B. WILCKEN
The Children's Hospital at Westmead, Locked Bag 4001, Westmead, New South Wales 2145, (Sydney), Australia. E-mail: bridgetw@chw.edu.au

Summary: There is increasing emphasis on the importance of practising evidence-based medicine. Randomized controlled trials are the standard way to assess the benefits of an intervention, and observational studies are not usually accorded much weight; the results are likely to be considered misleading. For rare diseases, there are great difficulties in obtaining adequate evidence for interventions or for the benefits of early diagnosis. This is because the disorders are not only very rare but also have variable expression, may have very long courses, and have incompletely known late effects; and surrogate end-points often have to be used. Randomized controlled trials are usually impossible because of inadequate power, and because there are preconceived notions of the effects of treatments already in use. The adoption of the best possible design for observational trials, formation of a central registry of such trials, and a greater general appreciation of the problems that rare diseases pose will help in obtaining the best possible evidence for the effects of interventions.

Evidence-based medicine was one of the catch-cries of the 1990s, and still is. From just 18 published articles found in Medline to address this subject in the early 1990s, the numbers rose inexorably year by year: 77 in 1995, then 239, 661, 1079, and most recently a massive 1557 in 1999. A randomized controlled trial, and better still a systematic review of randomized controlled trials, is the accepted standard for assessing clinical effectiveness, and anything that falls too far below this standard is in danger of being discounted. This leaves those of us who diagnose and care for patients with rare diseases in a double difficulty. Not only is it difficult to get evidence of diagnostic or therapeutic efficacy, but such evidence as can be obtained is often regarded as too unreliable to be taken into account at all.

One of the first trials of treatment in a rare inborn error of metabolism, phenylketonuria, (McKusick 261600) was undertaken by Professor Horst Bickel and his colleagues, and reported in a preliminary communication in the *Lancet*

(Bickel et al 1953). Their patient was 2 years old, 'an idiot, unable to stand, walk, or talk'. She was placed on a specially prepared low-protein diet and over a few months improved markedly. Then (without the mother's knowledge) 5 g per day of phenylalanine was added back into the diet. Within 6 h she started to bang her head as formerly, and within days she had lost all the ground previously gained. To test this further, she was admitted to hospital, where the experiment was repeated (with her mother's permission), with similar results. Professor Bickel had performed a study with single-blind and open-label phases. The conclusion was that 'In this child at least, the benefits of a low-phenylalanine intake seem unequivocal'. There have been no randomized trials of treatment (versus no treatment) of phenylketonuria.

A recent editorial in the *New England Journal of Medicine* was entitled 'Randomized trials or observational tribulations' (Pocock and Elbourne, 2000) and commented on two articles showing no difference in estimated treatment effects between randomized and nonrandomized trials, a finding disputed in the editorial on several grounds. This was by no means the first time that such a question has been investigated (e.g. Sacks et al 1982). Because of the many biases that can arise in observational studies, most people would agree that, where possible, a randomized trial is the preferred model for clinical trials. Although different questions require different trial methodologies, the hierarchy of evidence is generally agreed to be

- Randomized controlled trials, and their derivatives (systematic reviews of RCTs)
- Controlled observational studies
- Uncontrolled studies
- Expert opinion.

It is unfortunate that scientists and clinicians dealing with the very rare diseases often seem to be locked into the bottom rung of this hierarchy.

The problems with assessing intervention in inborn errors of metabolism are several. The disorders are usually very rare and, despite being largely monogenic, are in reality complex diseases with very variable expression, which complicates the use of historical controls. There is often a very long course of the illness, with long-term rather than short-term complications. On top of all this, monitoring must often be by surrogate measures, and use surrogate end-points. It is important not only to be able to assess current treatment options but also to be able to assess the effectiveness of early or presymptomatic treatment. How do we achieve the levels of evidence we need for these endeavours?

Very often there has been treatment that is believed to work, albeit not as obviously as is the case with phenylketonuria. If we believe in the efficacy of the treatment, then collaborating in a randomized controlled trial becomes ethically difficult or impossible. Then, too, there is the problem of desperation on the part of parents. There may be no known treatment for a progressive disease. When something of promise comes along, it is very hard to persuade parents to agree to randomization, especially when the time course is likely to be long. This arose with X-linked adrenoleukodystrophy (McKusick 300100) and the use of the 'Lorenzo's oil' (glyceryl trioleate and glyceryl trierucate), which was thought to pre-

vent fatty acid chain lengthening and thus reduce the accumulation of very long-chain fatty acids (Rizzo et al 1989). No randomized trial could be undertaken. The regimen with a low-fat diet and Lorenzo's oil did lower the circulating levels of very long-chain fatty acids, but data available now from observational studies suggest that the therapy is probably ineffective in preventing progression. However, that still remains uncertain (Alger et al 2000). The same scenario may occur with the newer treatments for lysosomal storage disorders.

Statistical power is a major difficulty. Underpowered trials are poorly regarded. One publication suggests that they usually have, 'poor design, poor randomization, ill-defined end-points, poor supervision, inexperienced researchers, ...' (Griffiths 1997). Yet here is a major problem for most of us. This can be illustrated by considering one newborn screening topic: is early diagnosis of congenital adrenal hyperplasia (McKusick 201910) beneficial? One aim of such screening is to prevent death in male babies with a severe salt-losing phenotype during an adrenal crisis. Detection of a 50% increase in deaths in the unscreened, when compared with screened babies (and surely the percentage would be less), would require 2 500 000 in each arm of the trial—a trial unlikely to be funded (Figure 1). There are of course other proposed benefits of such screening, but a firm proof of efficacy is bedevilled by the comparative rarity (Edwards et al 1997). But she also notes that, theoretically, publication of the results of a small underpowered trial with no likelihood of a statistically significant result may disturb the previous equipoise—that is, the

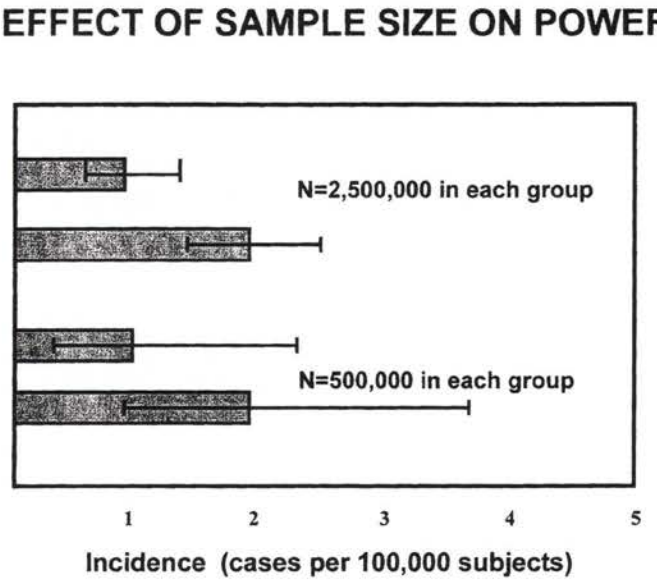


Figure 1 In a disorder with a frequency of 1 : 50 000, demonstration of a 50% reduction in the occurrence of an end-point with 95% confidence would require 2 500 000 subjects in each arm of a study. With 'only' 500 000 in each arm, there is an overlap in the confidence intervals

uncertainty of the value of an intervention—so that further trials would be unethical. However, a small trial may be all that can be managed with available numbers, and others have argued that some information is better than none (Lilford et al 1995).

The necessity of using surrogate end-points in chronic diseases is a common problem, not specific to inborn errors of metabolism. In some cases we know a great deal. For example, we have a reasonable idea about plasma levels of phenylalanine and the outcome in treated phenylketonuria. One of the best studies (Smith et al 1990) showed that mean IQ standard deviation scores in patients were similar to those of controls if mean blood phenylalanine levels were maintained below 400 μmol/L during the first four years of life, but declined progressively with higher average phenylalanine levels. This makes it possible to use blood phenylalanine levels when assessing interventions in children. Even so, we are aware now that measures of brain phenylalanine might be more relevant (Koch et al 2000). But we have not progressed very far with other disorders. We do not know the levels of plasma leucine that ensure a good outcome in maple syrup urine disease (McKusick 248600). Nor do we know the level of homocysteine, either free or bound, that we should be aiming for to maintain the lowest possible risk of thromboembolism in cystathionine β-synthase deficiency (McKusick 236200), or indeed whether some other analyte would be a more important modifier of risk. We suspect that plasma very long-chain fatty acids, elevated in X-linked adrenoleukodystrophy, are not useful analytes to measure for monitoring the effects of treatment. Treatments that reduce the elevated levels do not appear to modify outcome, although, as alluded to earlier, uncertainty remains (Alger et al 2000).

Many inborn errors are chronic diseases with unpredictable but often long courses. The effects of therapy may only be evident after many years. Examples of inborn errors where the start of intervention may precede a clinical end-point by up to 20 years or even more include childhood familial hypercholesterolaemia (McKusick 143890), homocystinuria (cystathionine β-synthase deficiency), and any mild organic acidaemia. This sort of time course makes it hard to embrace enthusiastically a suitably designed trial, and especially a randomized controlled trial. And the course may be quite unpredictable. Once again we turn to X-linked adrenoleukodystrophy as an example. Here, about half the affected boys will have a devastating course with severe physical disability, dementia, and death during the first to second decade. The other half will have a much more prolonged and benign course, with more or less intact survival to adulthood, the final phenotype still not being fully known (Moser et al 2000). It has not so far proved possible to distinguish between these phenotypes at an early age. This is perhaps an extreme example, but many inborn errors have sufficiently unpredictable courses to make the evaluation of a trial of treatment very difficult.

A further problem is that the late effects of an inborn error are likely to remain unknown for a prolonged period. For many disorders, the long-term outlook is no doubt still unknown. Methylmalonic aciduria, (methylmalonyl-CoA mutase deficiency; McKusick 251000), perhaps the most frequent of the classical organic acidurias, was first described in the late 1960s (Morrow and Barness 1968). Thus the early and life-threatening symptoms have been well known for over 30 years;

but the first publication revealing that renal failure was likely to occur in survivors of infantile presentations came only recently (Walter et al 1989). It is still not clear whether this is universal in serve cases, and whether mild congenital methylmalonic aciduria could lead to renal failure in the long term. Dozens of other examples could be given. When the long-term outlook is still uncertain, and some complications as yet uncovered, the effects of therapy are hard to judge. Perhaps the relative newness of our specialty rivals the rarity of our disorders in making it difficult to use the best tools of evidence-based medicine.

What sorts of studies could be used with these unpredictable and rare inborn errors of metabolism? The problems with lesser-order studies (i.e. those that do not incorporate double-blinded randomization of subjects and controls) have been well-ventilated. Observational studies require some control subjects, but without randomization it is difficult to control for confounding and bias. For example, researchers are interested in the outcome, and are not blinded. Subjects are also interested in the outcome and self-selection may be an especial problem. Use of historical controls presents particular difficulties because not only will management have altered and presumably improved over time, but more importantly, we very often do not know the natural history accurately. One excellent study of natural history was that of cystathionine β-synthase deficiency (Mudd et al 1985). This study was conducted by postal questionnaire to physicians caring for patients with inborn errors of metabolism. Data were collected on 629 patients from 114 individual physicians. Most of the physician responders were caring for only one to three such patients. Awareness of the disorder was at that time not generally high, and there might well be a bias towards the severe end of the spectrum, with only the most obvious cases being diagnosed in some regions. Another very useful study of natural history also had inbuilt and unavoidable bias towards the severe. Pitt and Danks (1991) studied the outcome of 51 never-treated adults with PKU. This showed, *inter alia*, that 6% had an IQ of 68 or greater; but of course it tells us nothing of the whole spectrum at the mild end of the range, as PKU patients with a relatively normal IQ may well go unrecognized throughout life. To improve the knowledge of natural history by retrospective study there need to be especially strict diagnostic criteria, and all known cases from a centre should be reported if they fulfil the criteria, even though some data may be missing. Data from regional centres with good diagnostic facilities may be preferred over those from referral centres, or at least should be analysed separately, as the patient base may be more comprehensive. These precautions will not ensure lack of bias, but will reduce it to a minimum.

Screening can provide otherwise hidden information. Natural history may be illuminated, although newborn screening, for example, is usually only undertaken when some form of intervention is intended. But it can enable the study of mild variants, and expose ascertainment bias, such as occurs when patients being investigated because of symptoms are found to have a rare genetic disorder, which is then thought to be the cause of the symptom. This was exemplified in the case of histidinaemia (McKusick 235800), which was initially thought to result in developmental delay and speech defects (Ghadimi and Partington, 1967). Only with the availability of both newborn screening and family studies did it become clear

that this enzyme deficiency is likely to be benign (Coulombe et al 1983). A new chapter is being written now with the advent of tandem mass spectrometry. Already, many screening programmes have found atypical cases of medium-chain acyl-CoA dehydrogenase deficiency (Andresen et al 2000; Carpenter et al 2000; Lindner et al 2000). In addition, an unexpectedly large incidence of two other disorders previously though extremely rare—short-chain acyl-CoA dehydrogenase deficiency and 3-methylcrotonyl-CoA carboxylase deficiency—has been found (Roscher et al 2000; Wilcken et al 2000). The clinical significance of these sorts of cases is at present unknown.

Despite all these difficulties, good studies of intervention are possible in rare diseases. While a randomized controlled trial could probably still be performed, for instance, to investigate the continued use of diet in adult males with PKU, many centres would nowadays find that unethical. There are few other inborn errors where randomization, for all the reasons discussed above, would be feasible. Observational studies with historical controls could be performed in a number of instances. Other designs are possible, depending on the question to be answered. What is needed is that trials, either multicentre or small trials, have the very best design possible, that protocols are well reviewed at the outset, and that there is some central body with which the trial can be registered. Review of trial design by an experienced clinical trials centre would ensure that the most reliable information was obtained. Registration of a trial would ensure first that, where agreeable to the investigators, others could know of a trial planned or in progress, and second that at the conclusion all of the data that resulted could be accessed, whether or not publication was achieved. To answer treatment questions for individual patients, '*N*-of-1' randomized trials are easily designed. These are indicated when effectiveness of the treatment is in doubt, when there is quick onset and offset of treatment effects, where there is a measurable treatment target and, of course, when the patient is keen to do the trial. A good example of the helpfulness of such a trial was a report of the effectiveness of benzoate and imipramine in a patient with late-onset nonketotic hyperglycinaemia (Wiltshire et al 2000).

There are many studies waiting to be done. An obvious one already mentioned is a study of the effectiveness and efficiency of tandem mass spectrometry in newborn screening. Does early identification improve outcome, and for which disorders? It is not realistic to wait for the answers before some screening programmes embark on testing, as without large numbers of participating programmes no answers can possibly be forthcoming. But it is surprising that no coherent plan has yet emerged to study this. The establishment of some central resource would perhaps encourage cooperative action. Many questions are being asked repeatedly: for example, the place of carnitine therapy in medium-chain acyl-CoA dehydrogenase deficiency (McKusick 201450) and other fatty acid oxidation defects, the severity of galactose restriction needed in galactosaemia (McKusick 230400), and so forth. Most could be answered. Is there a possibility of a central body to encourage appropriate trials, and to register them? Who could take this on? A web-site for the registry of trials is not such a difficult concept, and could easily be linked to the web-sites of the various societies—the SSIEM, SIMD, JSIMD, ASIEM. Possibly the Cochrane Col-

laboration will move more rapidly towards dealing with the problem of rare diseases and trials. Certainly, publicizing the problems can only help to dispel the idea that the only questions that ought to be addressed are those that can be answered with randomized controlled trials.

ACKNOWLEDGEMENT

I am most grateful to Dr Jennifer Peat and Dr Katrina Williams for helpful discussion and advice.

REFERENCES

Andresen BS, Dobrowoloski SF, O'Reilly L, et al (2000) The mutational spectrum in the MCAD gene of newborns identified by prospective tandem MS screening for 'diagnostic' acyl-carnitines in blood spots differs from that observed in clinically affected patients. *J Inherit Metab Dis* **23**(supplement 1): 12.

Alger S, Green A, Kohler W, Sokolwski P, Moser H (2000) Proceedings of the 4th International Workshop of the Adrenolekodystrophy International Research Group, September 1998. *J. Inherit Metab Dis* **23**: 449–452.

Bickel H, Gerrard J, Hickmans EM (1953) Influence of phenylalanine intake on phenylketonuria *Lancet* **ii**: 812–813.

Carpenter KH, Wiley V, Sim KS, Hammond J, Heath D, Wilcken B (2000) Newborn screening for MCAD deficiency: the New South Walse experience. *J Inherit Metab Dis* **23**(supplement 1):16.

Coulombe JT, Kammerer BL, Levy HL, Hirsch BZ, Scriver CR (1983) Histidinaemia part III: Impact: a prospective study. *J Inher Metab Dis* **6**: 58–61.

Edwards SJ, Lilford RJ, Braunholtz D, Jackson J (1997) Why underpowered trials are not necessarily unethical. *Lancet* **350**: 804–847.

Ghadimi H, Partington MW (1967) Salient features of histidinemia. *Am J Dis Child* **113**: 83–87.

Griffiths M (1987) 'Underpowered' trials. *Lancet* **350**: 1406–1407.

Koch R, Moats R, Guttler F, Guldberg P, Nelson M Jr (2000) Blood–brain phenylalanine relationships in persons with phenylketonuria. *Pediatrics* **106**: 1093–1096.

Lilford RJ, Thornton JG, Braunholtz, D (1995) Clinical trials and rare diseases: a way out of the conundrum *Br Med J* **311**: 1621–1625.

Lindner M, Zschocke J, Schulze A, et al (2000) Tandem mass spectrometry detects mild MCAD deficiency with negative phenylpropionic acid test. *J Inherit Metab Dis* **23**(supplement 1): 125.

Morrow G, Barness LA (1968) Methylmalonic aciduria. A newly discovered inborn error. *Ann Intern Med* **69**: 633–635.

Moser HW, Bexman L, Lu SE, Raymond GV (2000) Therapy of X-linked adrenoleukodystrophy: prognosis based upon age and MRI abnormality and plans for placebo-controlled trials. *J Inherit Metab Dis* **23**: 273–277.

Mudd SH, Skovby F, Levy HL, et al (1985) The natural history of homocystinuria due to cystathonine β-synthase deficiency. *Am J Hum Genet* **37**: 1–31.

Pitt D, Danks D (1991) The natural history of untreated phenylketonuria over 20 years. *J Paediatr Child Health* **27**: 189–190.

Pocock SJ, Elbourne DR (2000) Randomized trials or observational tribulations? *N Engl J Med* **342**: 1907–1909.

Rizzo WB, Leshner RT, Odone A, et al (1989) Dietary erucic acid therapy for X-linked adrenoleukodystrophy. *Neurology* **39**: 1415–1422.

Roscher A, Liebl B, Fingerhut, Olgemoller B (2000) Prospective study of MS-MS newborn screening in Bavaria, Germany. Interim results. *J Inherit Metab Dis* **23**(Suppl. 1): 4.
Sacks H, Chalmers TC, Smith H Jr (1982) Randomized versus historical controls for clinical trials. *Am J Med* **72**: 233–240.
Smith I, Beasley MG, Ades AE (1990) Intelligence and quality of dietary treatment in phenylketonuria. *Arch Dis Child* **65**: 472–478.
Walter JH, Michalski A, Wilson WM, Leonard JV, Barrett TM, Dillon MJ (1989) Chronic renal failure in methylamalonic acidaemia. *Eur J Pediatr* **148**: 344–348.
Wilcken B (2000) Two years of routine newborn screening by tandem mass spectrometry. *J Inherit Metab Dis* **23** (Suppl. 1): 4.
Wiltshire EJ, Poplawski NK, Harrison JR, Fletch JM (2000) Treatment of late-onset non-ketotic hperglycinaemia: effectiveness of imipramine and benzoate. *J Inherit Metab Dis* **23**: 15–21.

J Inherit Metab Dis (2006) 29:366–369
DOI 10.1007/s10545-005-0254-z

Mini-Symposium: Newborn screening for inborn errors of metabolism—Clinical effectiveness

Bridget Wilcken

Received: 27 October 2005 / Accepted: 23 January 2005

Summary With the application of tandem mass spectrometry, newborn screening has become an important topic in inborn metabolic disease. The aim of newborn screening is to produce an improved clinical outcome by early detection of disease, but it has been difficult to measure clinical effectiveness. Good evidence of clinical effectiveness has been hard to obtain because of the rarity of individual disorders, often precluding randomized controlled trials, the increase in diagnosis of individual disorders by screening, compared with clinical diagnosis, variable definitions of what constitutes a case, uncertainty about completeness of ascertainment, and differences in treatment in different geographical areas or at different times. Multiplex testing has introduced some new problems. There have been recent attempts to standardize screening in several countries, which have taken different approaches. Public pressure has driven the introduction of screening for inborn errors in some areas. Since it seems inevitable that screening may often be implemented ahead of hard evidence of benefit, ongoing evaluation of clinical effectiveness is a necessary part of any screening programme.

Introduction

Biochemical newborn screening has been a quiet, and at times almost hidden, endeavour since phenylketonuria screening was introduced in the early 1960s (Guthrie and Susi 1963). Testing for other inborn errors was slow to take off, and endocrine disorders became important targets. Now, the application of tandem mass spectrometry to newborn screening has dramatically changed the scenery and has propelled newborn screening squarely into the inborn errors arena. The first meeting involving the Society for the Study of Inborn Errors of Metabolism at which newborn screening was given a specific category in the abstracts was in 2000. At the Paris meeting in 2005, 35 abstracts (7%) were exclusively about newborn screening, which was also the subject of this mini-symposium. It is still true, however, that many screening programmes are not allied to paediatric clinical services, let alone to clinical or laboratory services dealing with inborn errors of metabolism. Newborn screening is a medical intervention that is now extremely widespread, with a substantial proportion of the world's newborns having some biochemical screening. The major gaps are in Africa and much of Asia, where in any case there may well be other more pressing priorities. While it is inexpensive to screen an individual baby, newborn screening programmes are expensive overall. As the core business of screening is to improve clinical outcome by early intervention, it is important to be able to measure clinical effectiveness.

Good evidence of benefit

Good evidence in newborn screening has been hard to come by. There have only been two randomized clinical trials of newborn screening, both of cystic fibrosis (Chatfield et al 1991; Farrell and Mischler 1992). For some other disorders there is clear-cut evidence of benefit (and cost-effectiveness) without any trials having been conducted. For example, there is very little doubt that screening is clinically effective for

Communicating editor: Jean-Marie Saudubray

Competing interests: None declared

Presented at the 42nd Annual Meeting of the SSIEM, Paris, 6–9 September 2005

B. Wilcken
The Children's Hospital at Westmead, and the Discipline of Paediatrics, Locked Bag 4001, Westmead NSW 2145, University of Sydney, Sydney, Australia

PKU and the same holds true for hypothyroidism. Clear-cut evidence of *lack* of benefit was certainly the case for histidinaemia, which was found to be a benign disorder, and for neuroblastoma, at least for the approaches used (Scriver and Levy 1983; Woods et al 2002). For most other disorders currently detected by screening, there is little hard evidence. Clinical experience points to likely benefit (it is usually better to know about a treatable disorder early), but the magnitude of the benefit and the magnitude of possible harms have usually not been well investigated. In the hierarchy of study designs, meta-analysis of randomized control trials, the highest in order, seems especially unlikely to be achieved in newborn screening. Nevertheless, there should be an advance on 'expert opinion', the lowest form of evidence. Regrettably, sometimes this has been the only evidence supporting newborn screening for certain disorders. At least controlled observational studies could be aimed at.

Even using controlled observational studies, there are well-known problems in measuring the clinical effectiveness of screening for rare disorders.

- *Power:* With very rare disorders, there is a major problem with power for any study, and the population base needed to show major change will be of several millions. This in itself will usually preclude a randomized controlled trial, but also makes clear-cut findings in observational studies equally hard to achieve.
- *Overdiagnosis by screening:* It is usual that more cases are found by screening than by clinical diagnosis, and in some instances mild disease, and even nondisease, may be hard to distinguish from the target affected population. Finding more cases has certainly been true in medium-chain acyl-CoA dehydrogenase (MCAD) deficiency, where almost double the number of cases are found in a screened population (Wilcken et al 2003), and histidinaemia is a good example of detecting what was later found to be nondisease (Scriver and Levy 1983). Of course, without screening there may be not only delayed diagnosis, but also underdiagnosis.
- *Definition of a case:* There has been an understandable difficulty for many disorders in providing definitions of what should be classified as a case. Here even mutation analysis is of little help; the finding of unusual mutations in a screen-detected population might lead to functional studies, but our incomplete knowledge of modifying genes is complicating this approach (Drumm et al 2005). We all acknowledge that for most inborn errors of metabolism there is a wide and graduated spectrum in any one disorder from severely affected individuals to the functionally normal. We need to decide on meaningful case-definitions: clinical definitions are clearly not useful for comparing screen-diagnosed and clinically ascertained cases if useful preventive strategies have been put in place. As alluded to above, molecular definitions may well not be sufficient, although they are very useful sometimes; for example in both cystic fibrosis and MCAD deficiency there is a common mutation of known severe effect. Biochemical definitions have a certain advantage obvious to aficionados of inborn errors, in that abnormalities do signal an *in vivo* functional derangement.
- *Completeness of ascertainment:* A further problem is the necessity for complete ascertainment. Where clinical and metabolic services are not centralized, this may be difficult, and for some disorders the diagnosis of more cases by screening than by clinical diagnosis implies likely incomplete ascertainment in any available control groups. Systematic searching for 'missed' cases is vital for the screened cohort, but is not possible for the controls.
- *Comparability of treatment:* While it is obvious that control and screened populations must have comparable and effective treatment before differences in outcome can be shown, this is not always something taken into account.
- *Control groups:* Because of all these factors, it is difficult to define what will be adequate control groups for defining clinical effectiveness of screening, and a striving for comparability is necessary. Control groups could be contemporaneous, from a different geographical area, or historical, from the same area. In both cases there is the risk that for one group there may have been less skilled or practised diagnosis, different treatment, or a different population demographic. Another approach was employed in Sweden (Alm et al 1984) and by the group in Newcastle UK (Pourfarzam et al 2001), who conducted retrospective screening of newborn blood spots. This, however, has the some of the same potential biases of historical controls, could be expensive, and also introduces the complexity of recoverability of analytes from a stored blood spot.

Multiplex testing: new problems?

Problems we face now in tandem mass spectrometry screening are largely not new: they were all seen in different guises after newborn screening was initially introduced. Screening for PKU started in the very early 1960s, following Dr Bob Guthrie's seminal paper (Guthrie and Susi 1963). Not surprisingly, the birth incidence was higher than expected, presumably largely due to previous underascertainment. Interesting new problems were quickly identified. A goodly number of babies were diagnosed as having PKU who in retrospect did not need any treatment. (Menkes and Holtzman 1970). They had a mild, clinically insignificant reduction in phenylalanine hydroxylase activity. Even now there is not full agreement on what levels of blood phenylalanine indicate the need for

treatment. Then new disorders were discovered. 'Malignant phenylketonuria' comprised previously unknown disorders of pterin synthesis and recycling, which were diagnosed by newborn screening programmes because of elevated blood phenylalanine levels(Tada et al 1970). These pterin disorders needed different treatment. New complications were delineated as the risks of maternal PKU were clearly recognized (Komrower et al 1979)

Approaches

All of these types of problems can be seen with the current introduction of screening by tandem mass spectrometry. But some aspects *are* new. With multiplex testing, disorders can be piggybacked onto a testing programme, with virtually no added cost. Thus it is tempting to screen for a wide range of disorders, with the danger of a loosening of the criteria for screening. Different countries have adopted different approaches to deal with this: Australia, with closely integrated clinical and laboratory services in each state, has opted at present to screen for a wide range of disorders, it being reasoned that only an experienced team of metabolic physicians will be dealing with babies identified by newborn screening. The United States has recommended a panel of 29 disorders detected by current MS/MS screening to be included, following a structured enquiry from a wide range of interested parties, and an expert review of the results (Maternal and Child Health Bureau 2005).

The United Kingdom has approached the issue cautiously, and at present is including only MCAD deficiency in a pilot programme (Goddard 2004). Germany has decided on the inclusion of a number of disorders in screening programmes, but incidental findings that could indicate any disorder outside this panel are not allowed to be disclosed (Guidelines of the German Ministry of Health and Social Affairs 2004). Of necessity, none of these models has been based on really good evidence.

With a widening range of disorders able to be detected, and with new treatments becoming available, it will be difficult to resist screening for certain disorders before the long-term or medium-term outcome is known. Thus already, only months after the first published report of favourable short-term outcome in Krabbe disease (globoid cell leukodystrophy) following presymptomatic umbilical cord-blood transplantation in affected siblings of known cases (Escolar et al 2005), New York State will start a pilot study (Wall Street Journal 2005). Clearly this is the sort of thing that the public wants. In the United States there has been very strong pressure from the public on screening programmes for treatable disorders. There is evidence from the United Kingdom that the public is moving towards a position of wanting to know as early as possible about a disorder, even if it is untreatable (Campbell and Ross 2003). It is inevitable that public pressure will drive screening programmes to some extent, even though this seems undesirable, so it is doubly important that outcomes are rigorously monitored. Suitable measures of outcome may be disease-specific. They will include death, medical complications, neuropsychological development and, inevitably, costs. While the problems outlined above in assessing clinical effectiveness will remain, the evaluation of this is most easily done when screening, diagnostic and clinical services are as closely integrated as possible.

It is never likely to be possible to collect firm evidence of the clinical effectiveness of newborn screening *before* implementing a screening programme, but as programmes are implemented ahead of hard evidence of benefit, ongoing evaluation of clinical effectiveness is a necessary part of the whole.

References

Alm J, Hagenfeldt L, Larsson A, Lundberg K (1984) Incidence of congenital hypothyroidism: retrospective study of neonatal laboratory screening versus clinical symptoms as indicators leading to diagnosis. *BMJ Clin Res Ed* **289**(6453): 1171–1175.

Campbell E, Ross LF (2003) Parental attitudes regarding newborn screening of PKU and DMD. *Am J Med Genet Pt A* **120**(2): 209–214.

Chatfield S, Owen G, Ryley HC, et al (1991) Neonatal screening for cystic fibrosis in Wales and the West Midlands: clinical assessment after five years of screening. *Arch Dis Child* **66**: 29–33.

Drumm ML, Konstan MW, Schluchter MD, et al (2005) Genetic modifiers of lung disease in cystic fibrosis. *N Engl J Med* **353**(14): 1443–1453.

Escolar ML, Poe MD, Provenzale JM, et al (2005) Transplantation of umbilical-cord blood in babies with infantile Krabbe's disease. *N Engl J Med* **352**(20): 2069–2081.

Farrell PM, Mischler EH (1992) Newborn screening for cystic fibrosis. The Cystic Fibrosis Neonatal Screening Study Group. *Adv Pediatr* **39**: 35–70.

Goddard P (2004) Newborn screening for medium chain acyl-CoA dehydrogenase deficiency (MCADD) in the UK. *J Fam Health Care* **14**(4): 90–92.

Guidelines of the German Ministry of Health and Social Affaires (2004) 21st December 2004 Internet references. Retreived from: http://www.aerzteblatt.de/v4/archiv/artikel.asp?id = 46488 28/10/2005.

Guthrie R, Susi A (1963) A simple phenylalanine method for detecting phenylketonuria in large populations of newborn infants. *Pediatrics* **32**: 338–343.

Komrower GM, Sardharwalla IB, Coutts JM, Ingham D (1979) Management of maternal phenylketonuria: an emerging clinical problem. *Br Med J* **1**(6175):1383–1387.

Maternal and Child Health Bureau (2005) Internet references. Retrieved from: http://mchb.hrsa.gov/screening/ 26/10/2005.

Menkes JH, Holtzman NA (1970) Phenylalaninemia or classical phenylketonuria. *Neuropadiatrie* **2**(1): 121–122.

Pourfarzam M, Morris A, Appleton M, Craft A, Bartlett K (2001) Neonatal screening for medium-chain acyl-CoA dehydrogenase deficiency. *Lancet* **358**: 1063–1064.

Scriver CR, Levy HL (1983) Histidinaemia. Part I: Reconciling retrospective and prospective findings. *J Inherit Metab Dis* **6**(2): 51–53.

Tada K, Yoshida T, Mochizuki K, Konno T, Nakagawa H (1970) Two siblings of hyperphenylalaninemia: suggestion to a genetic variant of phenylketonuria. *Tohoku J Exp Med* **100**(3): 249–253.

Wall Street Journal (2005) Internet references. Retrieved from http://www.post-gazette.com/pg/05139/507278.stm 28/10/2005.

Wilcken B, Wiley V, Hammond J, Carpenter K (2003) Screening newborns for inborn errors of metabolism by tandem mass spectrometry. *N Engl J Med* **348**: 2304–2312.

Woods WG, Gao RN, Shuster JJ, et al (2002) Screening of infants and mortality due to neuroblastoma. *N Engl J Med* **346**: 1041–1046.

J Inherit Metab Dis (2006) 29:383–389
DOI 10.1007/s10545-006-0278-z

'Classical' organic acidurias, propionic aciduria, methylmalonic aciduria and isovaleric aciduria: Long-term outcome and effects of expanded newborn screening using tandem mass spectrometry

Carlo Dionisi-Vici · Federica Deodato · Wulf Röschinger · William Rhead · Bridget Wilcken

Received: 17 November 2005 / Accepted: 27 January 2005

Summary 'Classical organic acidurias' comprise isovaleric aciduria, propionic aciduria and methylmalonic aciduria. Available data from the literature suggest that the use of 'new' therapeutic strategies has improved survival but has not modified neurodevelopment. Progressive neurocognitive deterioration is almost invariably present in propionic and methylmalonic acidurias, while large-scale studies on the long-term outcome of patients with isovaleric aciduria are still lacking. In order to answer to some of the questions suggested by Wilson and Jungner in 1968 about the criteria of disease screening, we compared the natural history of patients with 'classical' organic acidurias diagnosed on clinical bases to those diagnosed through neonatal mass screening using tandem mass spectrometry. Decreased early mortality, less severe symptoms at diagnosis, and more favourable short-term neurodevelopmental outcome were recorded in patients identified through expanded newborn screening. The short duration of follow-up so far does not allow us to draw final conclusions about the effects of newborn screening on long-term outcome. The evaluation of the effect of neonatal screening on the detection rate of these three diseases showed that the incidence of isovaleric aciduria was significantly higher in the screening population than in clinically detected cases, with no changes for propionic and methylmalonic acidurias. Further multicentre longitudinal studies are needed to assess the usefulness of expanded newborn screening for 'classical' organic acidurias and to better understand the clinical spectrum of these diseases. This paper describes the long-term outcome and the impact of expanded newborn screening on the so-called 'classical' organic acidurias (propionic aciduria, methylmalonic aciduria and isovaleric aciduria).

Communicating editor: Jean-Marie Saudubray

Competing interests: None declared

Presented at the 42nd Annual Meeting of the SSIEM, Paris, 6–9 September 2005

C. Dionisi-Vici (✉) · F. Deodato
Division of Metabolism, Bambino Gesù Children's Hospital, Piazza S. Onofrio 4, I-00165 Rome, Italy
e-mail: dionisi@opbg.net

W. Röschinger
Department of Metabolic Diseases, Ludwig-Maximilians University, München, Germany

W. Rhead
Genetic Center, Children's Hospital of Wisconsin, Milwaukee, Wisconsin, USA

B. Wilcken
The Children's Hospital at Westmead, Westmead, Australia

Introduction

The term 'classical organic acidurias' conventionally defines three different types of inherited disorders of branched-chain amino acids: isovaleric aciduria (IVA, McKusick 243500), propionic aciduria (PA, McKusick 606054) and methylmalonic aciduria (MMA, McKusick 251000) (Ogier de Baulny and Saudubray 2000). IVA is caused by a deficiency of isovaleryl-CoA dehydrogenase (EC 1.3.99.10), which catalyses the third step in the catabolism of leucine. PA is caused by a deficiency of propionyl-CoA carboxylase (EC 6.4.1.3), which converts propionyl-CoA, derived from the catabolism of isoleucine, valine, methionine, threonine and odd-chain fatty acids, to methylmalonyl-CoA. MMA is caused by a defect in the conversion of methylmalonyl-CoA to succinyl-CoA, a reaction catalysed by methylmalonyl-CoA mutase (EC 5.4.99.2). Patients with isolated methylmalonic acidaemia may have a defect of the

apoenzyme (mut^0 and mut^-) or of the synthesis of its coenzyme adenosylcobalamin (cbl A, cbl B).

All these disorders are characterized by an acute or progressive neurological involvement caused by the accumulation of toxic compounds proximal to the metabolic block. In most cases the onset of the disease is in the neonatal period, but a later-onset presentation can occur with appearance of symptoms during the firsts years of life.

The neonates present acute neurological deterioration after an initial symptom-free period, ranging from hours to days from birth. Typically the progression of symptoms moves from feeding refusal, vomiting, progressive weight loss, generalized hypotonia, and abnormal posturing and movements, through lethargy, seizures and coma, leading to death within a few days or to severe brain damage, if not promptly treated.

In the later-onset form, the clinical picture is more variable, ranging from acute life-threatening encephalopathy to intermittent or chronic symptoms of various degrees. These include intermittent ataxia, abnormal behaviour, and poor feeding with selective refusal of protein-rich foods, recurrent vomiting, etc. Failure to thrive is often present as well as neurodevelopmental delay.

Acidosis, ketonuria, hyperammonaemia, leukothrombocytopenia, anaemia and hyperuricemia are the most common laboratory abnormalities. In IVA only, urine and sweat have a characteristic smell of 'sweaty feet'.

The diagnosis is based on urinary organic acid analysis and blood acylcarnitine profile, showing the characteristic metabolic compounds for each disorder. Further confirmation can be obtained by enzymatic and molecular studies.

Despite the improvement in our understanding of the biochemistry of these diseases, their management remains difficult. Treatment can be schematically divided into two main steps: emergency treatment and long-term management. The acute phase, either at the onset or during a relapse of metabolic decompensation, represents a true medical emergency and its management is usually performed in the intensive care units. Emergency treatment is based on exogenous protein restriction, inhibition of endogenous catabolism by adequate parenteral energy supply, high dose of carnitine, vitamin supplementation and, in patients unresponsive to medical therapy within the firsts few hours, on extracorporeal removal of toxic metabolites through dialysis (Picca et al 2001). The use of sodium benzoate to increase ammonia detoxification in organic acidurias is still controversial. More recently, carbamyl glutamate has been successfully used for this purpose in MMA and PA (Gebhardt et al 2003, 2005).

The mainstay of the long-term treatment is a low-protein and high-energy diet, supplemented with a specific amino acid mixture free of precursor amino acids. Nasogastric and gastrostomy feeds are often used to maintain a satisfactory nutritional status. Carnitine is supplemented to prevent deficiency, metronidazole is used to reduce gut propionate produced by intestinal bacteria, and vitamin B_{12} is used in the responsive forms of MMA. In IVA, peroral glycine supplementation increases the excretion of isovaleric acid.

The aim of treatment is to prevent brain damage and the selective organ involvement (i.e. kidney, heart, pancreas) while maintaining normal development and nutritional status. However, despite intense medical therapy most patients, especially those with the severe form with neonatal onset, have a high risk of relapsing episodes of acute metabolic decompensation triggered by intercurrent stressing events such as fever, gastroenteritis and vaccinations.

More recently, as an alternative therapy to conventional medical treatment, liver transplantation has been attempted to cure the underlying metabolic defect in PA and MMA. However, the few scattered experiences reported to date did not clearly demonstrate the effectiveness of this therapy either to prevent further deterioration or to improve survival and quality of life (Chakrapani et al 2002; Kayler et al 2002; Leonard et al 2001; Morioka et al 2005; Nyhan et al 2002; Saudubray et al 1999; van't Hoff et al 1999).

What we know from the literature about the long-term outcome of patients with 'classical' organic acidurias

There is great difficulty in establishing the natural history and predicting the prognosis and long-term outcome in patients with 'classical' organic acidurias. From a careful review studies reporting longitudinal evaluation on large series of patients as well as selective organ complications (Baumgartner and Viardot 1995; Burlina et al 1995; Deodato et al 2004; Lehnert et al 1994; Mardach et al 2005; Massoud and Leonard 1993; Nagarajan et al 2005; Nicolaides et al 1998; North et al 1995; Ogier de Baulny et al 2005; Rousson and Guibaud 1984; Sass et al 2004; Saudubray et al 1999; Surtees et al 1992; Van Calcar et al 1998; van der Meer et al 1994, 1996; van't Hoff et al 1999), we can at least conclude the following.

- The late-onset forms have a better prognosis compared to the neonatal/early onset ones.
- The progresses made in treatment with the use of the 'new' therapeutic strategies has improved survival but has not modified the neurodevelopmental outcome.
- Progressive neurocognitive deterioration is almost invariably present.
- Relapsing episodes of acute metabolic decompensation are associated with a high risk of basal ganglia stroke, responsible for severe motor disabilities.
- Long-term complications with selective organ impairment are frequent. In particular, cardiomyopathy and pancreatitis can represent severe life-threatening events, and

progressive renal failure is typically observed in vitamin B_{12}-unresponsive MMA.
- Failure to thrive and poor nutritional status are frequently observed.
- The most severe forms with neonatal onset are more frequently observed in patients with PA and IVA (75%) compared to those with MMA (56%).
- There are no uniform criteria with respect to the day of life for the definition of what is a neonatal/early-onset form.
- There are no standardized measures for evaluating the long-term cognitive outcome.

Surprisingly, a careful retrospective analysis of the literature revealed the complete lack of large-scale studies describing the long-term outcome of patients with IVA, with only one abstract reporting the follow-up of 22 patients (Ensenauer et al 2003).

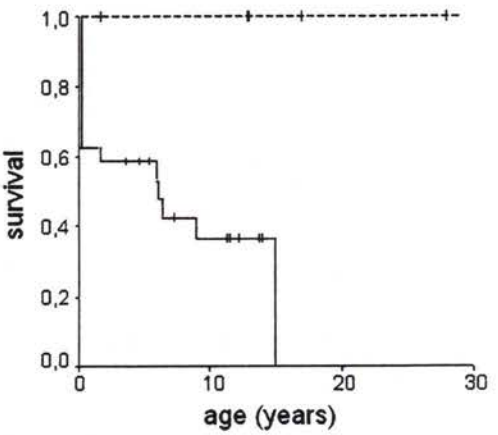


Fig. 1 Kaplan–Mayer actuarial survival curves comparing patients with neonatal onset (solid line) and those with late onset (dashed line): $p < 0.01$.

The impact of expanded newborn screening for 'classical' organic acidurias

In recent years, the application of electrospray tandem mass spectrometry (MS/MS) to newborn screening for inborn errors of metabolism has radically changed the scenario. In contrast t conventional programmes based on the use of a specific test for each disease, the simultaneous analysis of amino acids and acylcarnitines in blood by MS/MS allows the identification of more than 30 different diseases in the one test. However, the introduction of this new technology into routine clinical practice raises several questions about the real advantage that may be achieved by diagnosing newborns with 'classical' organic acidurias at a presymptomatic or early symptomatic stage. According to the criteria established more than 35 years ago (Wilson and Jungner 1968) and still valid, many issues remain to be clarified. To answer some of these questions, in this review we compared the natural history in two groups of patients with 'classical' organic acidurias: those diagnosed on clinical bases and those diagnosed through neonatal screening. The study evaluated mortality, neurocognitive development and organ complications in each group.

Patients identified by clinical symptoms

The group of symptomatic patients comprised 29 individuals diagnosed and followed at the Bambino Gesù Children's Hospital in Rome from 1983. Fifteen patients had MMA, 13 had PA and 1 had IVA. In 24 patients (82%) symptoms appeared within the first 28 days of life (neonatal-onset) and in the remaining 5 cases symptoms appeared later (late-onset). In the neonatal-onset group the mean age at diagnosis was 11 days. This value decreased to 7 days among patients diagnosed after 1995.

Overall, the mortality rate was 51%, with a significantly higher mortality ($p < 0.01$) in the neonatal-onset group compared to the later-onset cases, in which no deaths were recorded (Fig. 1). About 40% of patients died during the initial phase of the disease or before the second year of life, a finding very similar to the report by Surtees and colleagues in a series of patients with PA (Surtees et al 1992). The 'new' therapies developed throughout the years, based on the use of standardized emergency protocols combined with the improvement of dialysis techniques, influenced early survival of patients by reducing neonatal mortality.

In all patients, cognitive development was longitudinally assessed with conventional neuropsychological tests, and according to the IQ/DQ score patients were classified as normal (IQ > 79, DQ > 74), mildly retarded (IQ50–79/DQ60–74) or severely retarded (IQ < 49/DQ < 59). For patients with neonatal onset, the neurocognitive outcome was evaluated in two sequential periods, considering the psychomotor development in the firsts two years of life (short-term outcome) and the final long-term outcome.

At the first evaluation before the 2nd year of life, normal neurodevelopment was observed in 60% of surviving patients. As we previously reported in newborns affected both by branched-chain organic acidurias and by urea cycle defects (Picca et al 2001), the only parameter that significantly influenced early survival and the short-term outcome was the duration of coma, and in particular the duration of coma preceding the beginning of therapy. Through the following years, most of the affected children showed progressive cognitive deterioration or died. The mortality rate increased from 36% to 62%, and a normal neurocognitive outcome was recorded in only 27% of patients, with a complete reversal of the proportions observed at the early evaluation. As is evident

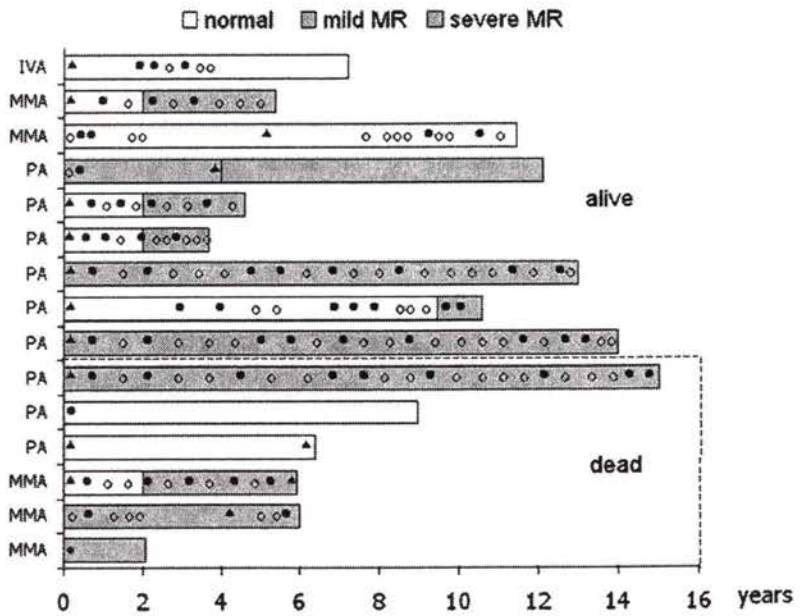


Fig. 2 Cognitive development of 15 patients with neonatal-onset 'classical' organic acidurias. Deceased patients are illustrated in the lower part, living patients in the upper part. Episodes of metabolic decompensation, mild (○) and severe (●), as well as episodes of coma (▲) are indicated for each patient. IVA, isovaleric aciduria; MMA, methylmalonic aciduria; PA, propionic aciduria; MR, mental retardation

from Fig. 2, the course of these diseases is characterized by frequent and severe relapses of metabolic decompensation, which undoubtedly influences the progression of neurological symptoms and the poor long-term prognosis. The findings of our study are very similar to the longitudinal evaluations previously reported in PA and MMA (Nicolaides et al 1998; North et al 1995).

As shown in Fig. 3, if we combine the risk of death with the risk of developing neurocognitive impairment we observe an event free 'window' corresponding to the first two years in which patients surviving the neonatal period show a relatively stable metabolic condition. This suggests that the option of liver transplantation as a potentially radical treatment should preferably be considered within this time period before the development of an irreversible neurological damage.

Unlike in neonatal-onset cases, the natural history in the late-onset cases clearly appears more favourable. Although our study was limited by the small number of patients, not only survival but also neurocognitive outcome was better, with 4 out 5 patients showing a normal/borderline development. Besides the absence of progressive cognitive deterioration, patients have a more stable clinical course with less frequent and less severe relapsing episodes of metabolic decompensation. One of our patients with mut^{t} MMA had two successful pregnancies with an uneventful postpartum period (Deodato et al 2002).

Regarding specific organ complications, we observed cardiomyopathy in 3 patients with PA, basal ganglia stroke in 2 patients with MMA, and acute pancreatitis in 1 patient with PA. One patient with neonatal-onset PA and normal cognitive level died suddenly at school from cardiomyopathy at the age

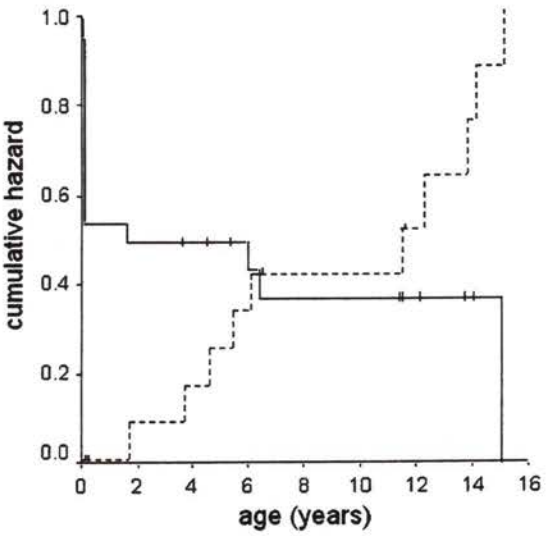


Fig. 3 Hazard plot showing the actuarial survival (solid line) and the developing of mental retardation (dashed line) in 15 patients with neonatal-onset 'classical' organic acidurias surviving neonatal period

of 9 years. Signs of chronic renal failure were detectable in all MMA patients after the sixth year of life.

Patients identified by newborn screening

We evaluated the clinical characteristics at diagnosis, the neurocognitive outcome and the mortality rate in a series of 18 patients (7 IVA, 7 PA and 4 MMA) diagnosed by neonatal screening in Australia (11 patients) and in Munich, Germany (7 patients, 4 of whom were prospectively diagnosed because

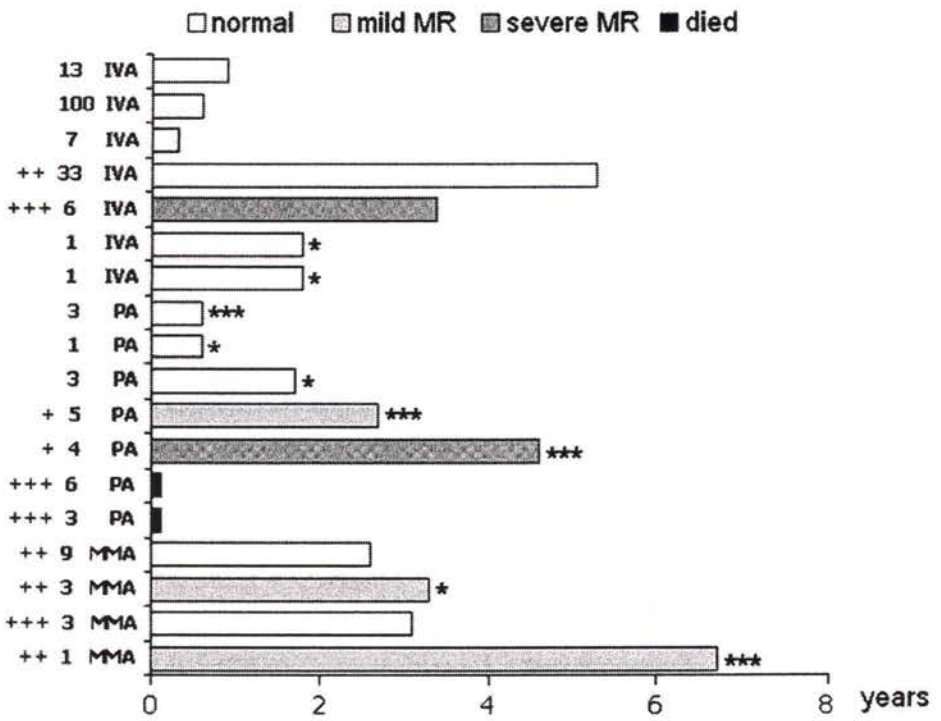


Fig. 4 Outcome of 18 patients with 'classical' organic acidurias detected by expanded newborn screening. The presence and severity of symptoms before the result of screening (+) as well as metabolic instability (*, mild; **, moderate; ***, severe) are indicated. IVA, isovaleric aciduria; MMA, methylmalonic aciduria; PA, propionic aciduria; MR, mental retardation.

of an affected elder sibling). The median age at diagnosis was 4 days (range 1–33 days) and more than 50% of patients were already symptomatic at the time of diagnosis. The clinical signs included tachypnoea in 3 patients, and metabolic decompensation with ketoacidosis and/or hyperammonaemia in 7 patients. Two of these 7 patients received treatment prior to the result of neonatal screening being available.

The neonatal mortality in this group of patients was significantly lower than in symptomatic patients ($p < 0.03$): only two patients with PA died, at the age of 5 and 13 days, respectively.

Early neurocognitive outcome was normal in 69% of cases, with a median age at follow-up in these patients of 1.7 years (range 0.3–5.3 years). The remaining 31% with neurocognitive impairment (severe in two patients,1 IVA and 1 PA; mild in three patients, 1 PA and 2 MMA) had a median age at follow-up of 3.4 years (range 2.7–6.7 years).

The clinical course was more stable than that observed in the series of symptomatic patients, with less-frequent relapsing episodes of metabolic decompensation; this was even more evident in IVA, in which metabolic crises never occurred in 5 out 7 cases (Fig. 4).

Only one patient with MMA, aged 40 months at last evaluation, shows mild renal dysfunction.

Impact of newborn screening on the detection rate

We analysed the effect of expanded neonatal screening by MS/MS on the frequency of these three diseases in the general population and compared the results with the available epidemiological studies based on clinical diagnosis. Combining the results of the different studies, the mean incidence of the three diseases in the symptomatic population is 1.19 ± 0.32 cases per 100 000 births (range 0.88–1.58) (Applegarth et al 2000; Dionisi-Vici et al 2002; Klose et al 2002; Wilcken et al 2003), compared to 1.94 ± 0.38 per 100 000 births in the neonatal screening group (range 1.58–2.4) (Hoffmann et al 2004; Rhead this study; Schulze et al 2003; Wilcken this study). Despite an apparent positive trend, these figures did not reach statistical significance ($p < 0.19$).

If we analyse the three diseases separately, figures for MMA and PA did not show significant differences between the two groups, with an overall incidence for these two diseases of 0.99 ± 0.35 per 100 000 births in the clinically diagnosed group compared to 1.04 ± 0.36 in the screening group. Remarkably, the incidence of IVA was significantly higher in the screened population, 0.89 ± 0.49 per 100 000 births versus 0.20 ± 0.13 ($p < 0.0001$) (Fig. 5).

Combining all cases reported in clinical studies ($n = 308$), the proportion of IVA in this group of patients is 10.4%, with an IVA/PA+MMA ratio of 1/9.6 (Applegarth et al 2000; Dionisi-Vici et al 2002; Hori et al 2005; Klose et al 2002; Rousson and Guibaud 1984; Waisbren et al 2003; Wilcken et al 2003). Whereas in the screened population ($n = 44$) this proportion increases to 40.9% of cases, with an IVA/PA+MMA ratio of 1/1.4 (Rhead this study; Röschinger this study; Schulze et al 2003; Wilcken this study).

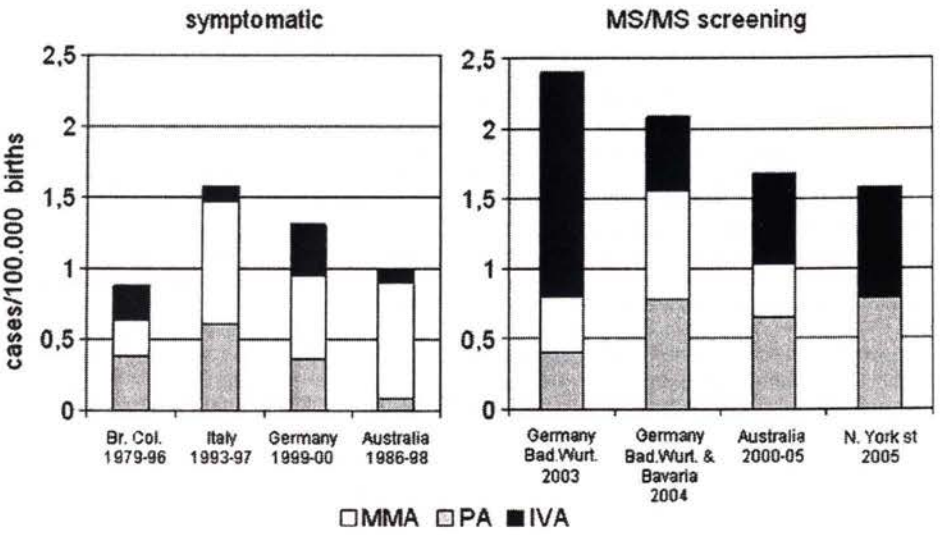


Fig. 5 Incidence (number of cases per 100 000 births) of 'classical' organic acidurias in symptomatic and neonatal mass screening populations. Different bars correspond to references (Applegarth et al 2000; Dionisi-Vici et al 2002; Hoffmann et al 2004; Klose et al 2002; Rhead this study; Schulze et al 2003; Wilcken et al 2003, this study)

Conclusions

Our study based on the comparison of the natural history of patients with 'classical' organic acidurias diagnosed on clinical bases with those diagnosed through expanded newborn screening may help to indicate whether these conditions fulfil the still valid principles of early disease detection established in 1968 (Wilson and Jungner 1968). In our series of cases, expanded newborn screening decreased early mortality and symptoms at diagnosis were less severe. Moreover, the short-term neurodevelopmental outcome seems to be more favourable in patients detected by newborn screening However, the duration of follow-up in these patients is too short and therefore the long-term outcome seems to be still unpredictable. A less compromised condition might be of further benefit when deciding the option of liver transplantation in selected cases.

Since neonatal screening could not discriminate the early-onset patients from those with a later onset, the major advantage from screening would probably be obtained in patients with milder phenotypes, facilitating appropriate treatment in a presymptomatic stage of the disease to prevent the appearance of severe clinical signs.

Regarding the potential psychological and social consequences of neonatal screening on patients and their families, the markedly increased detection rate of patients with IVA clearly requires a particular awareness. This increase would be the result either of underdetection by conventional diagnosis or of overdetection of mild cases by MS/MS screening. The recent identification of a common mutation of the isovaleryl-CoA dehydrogenase gene (932C > T) associated with a mild phenotype or even asymptomatic IVA (Ensenauer et al 2004), further supported by the observation of six cases of IVA diagnosed by neonatal screening with a fully normal phenotype (Knerr et al 2005), may explain this finding. Clinicians should as far as possible avoid inducing the so-called 'vulnerable' child syndrome by treating a benign condition as a serious disease (Waisbren et al 2003).

Neonatal screening does not modify the detection rate of MMA and PA, unlike for IVA, underlining the existence of a more severe phenotype in these two diseases, as confirmed by the unfavourable outcome at follow-up.

In conclusion, further multicentre longitudinal studies are needed to assess the usefulness of expanded newborn screening for 'classical' organic acidurias and to allow better understanding of the clinical spectrum of these diseases. In addition, the complete lack of large scale studies on the outcome of IVA should stimulate clinicians to collaborate in establishing the natural history of this disease. For all these reasons, the opportunity to include 'classical' organic acidurias in the expanded neonatal screening is still being debated.

Acknowledgements This work was supported by grants from the Italian Ministry of Health. Ricerca Corrente 2005, Ricerca Finalizzata 2003 and Ricerca Finalizzata 2005. Special thanks to S. Picca, A. Bartuli and S. Caviglia for their contribution in the clinical management of patients at the Bambino Gesù Children's Hospital.

References

Applegarth DA, Toone JR, Lowry RB (2000) Incidence of inborn errors of metabolism in British Columbia, 1969–1996. *Pediatrics* **105**: e10.

Baumgarter ER, Viardot C (1995) Long-term follow-up of 77 patients with isolated methylmalonic acidaemia. *J Inherit Metab Dis* **18**: 138–142.

Burlina AB, Dionisi-Vici C, Piovan S, et al (1995) Acute pancreatitis in propionic acidaemia. *J Inherit Metab Dis* **18**: 169–172.

Chakrapani A, Sivakumar P, McKiernan PJ, Leonard JV (2002) Metabolic stroke in methylmalonic acidemia five years after liver transplantation. *J Pediatr* **140**: 261–263.

Deodato F, Rizzo C, Boenzi S, Baiocco F, Sabetta G, Dionisi-Vici C (2002) Successful pregnancy in a woman with mut⁰ methylmalonic acidaemia. *J Inherit Metab Dis* **25**: 133–134.

Deodato F, Boenzi S, Rizzo C, et al (2004) Inborn errors of metabolism: an update on epidemiology and on neonatal-onset hyperammonemia. *Acta Paediatr Suppl* **93**: 18–21.

Dionisi-Vici C, Rizzo C, Burlina AB, et al (2002) Inborn errors of metabolism in the Italian pediatric population: a national retrospective survey. *J Pediatr* **140**: 321–327.

Ensenauer R, Gruenert S, Willard J, et al (2003) Natural history of isovaleric acidemia. *J Inherit Metab Dis* **26**: 38.

Ensenauer R, Vockley J, Willard JM, et al (2004) A common mutation is associated with a mild, potentially asymptomatic phenotype in patients with isovaleric acidemia diagnosed by newborn screening. *Am J Hum Genet* **75**: 1136–1142.

Gebhardt B, Vlaho S, Fischer D, Sewell A, Bohles H (2003) *N*-Carbamylglutamate enhances ammonia detoxification in a patient with decompensated methylmalonic aciduria. *Mol Genet Metab* **79**: 303–304.

Gebhardt B, Dittrich S, Parbel S, Vlaho S, Matsika O, Bohles H (2005) *N*-Carbamylglutamate protects patients with decompensated propionic aciduria from hyperammonaemia. *J Inherit Metab Dis* **28**: 241–244.

Hoffmann GF, von Kries R, Klose D, et al (2004) Frequencies of inherited organic acidurias and disorders of mitochondrial fatty acid transport and oxidation in Germany. *Eur J Pediatr* **163**: 76–80.

Hori D, Hasegawa Y, Kimura M, Yang Y, Verma IC, Yamaguchi S (2005) Clinical onset and prognosis of Asian children with organic acidemias, as detected by analysis of urinary organic acids using GC/MS, instead of mass screening. *Brain Dev* **27**: 39–45.

Kayler LK, Merion RM, Lee S, et al (2002) Long-term survival after liver transplantation in children with metabolic disorders. *Pediatr Transplant* **6**: 295–300.

Klose DA, Kolker S, Heinrich B, et al (2002) Incidence and short-term outcome of children with symptomatic presentation of organic acid and fatty acid oxidation disorders in Germany. *Pediatrics* **110**: 1204–1211.

Knerr I, Weiglel C, Rauh M, et al (2005) 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* **28**: 12.

Lehnert W, Sperl W, Suormala T, Baumgartner ER (1994) Propionic acidaemia: clinical, biochemical and therapeutic aspects. Experience in 30 patients. *Eur J Pediatr* **153**(Supplement 1): 68–80.

Leonard JV, Walter JH, McKiernan PJ (2001) The management of organic acidaemias: the role of transplantation. *J Inherit Metab Dis* **24**: 309–311.

Mardach R, Verity MA, Cederbaum SD (2005) Clinical, pathological, and biochemical studies in a patient with propionic acidemia and fatal cardiomyopathy. *Mol Genet Metab* **85**: 286–290.

Massoud AF, Leonard JV (1993) Cardiomyopathy in propionic acidaemia. *Eur J Pediatr* **152**: 441–445.

Morioka D, Takada Y, Kasahara M, et al (2005) Living donor liver transplantation for noncirrhotic inheritable metabolic liver diseases: impact of the use of heterozygous donors. *Transplantation* **80**: 623–628.

Nagarajan S, Enns GM, Millan MT, Winter S, Sarwal MM (2005) Management of methylmalonic acidaemia by combined liver–kidney transplantation. *J Inherit Metab Dis* **28**: 517–524.

Nicolaides P, Leonard J, Surtees R (1998) Neurological outcome of methylmalonic acidaemia. *Arch Dis Child* **78**: 508–512.

North KN, Korson MS, Gopal YR, et al (1995) Neonatal-onset propionic acidemia: neurologic and developmental profiles, and implications for management. *J Pediatr* **126**: 916–922.

Nyhan WL, Gargus JJ, Boyle K, Selby R, Koch R (2002) Progressive neurologic disability in methylmalonic acidemia despite transplantation of the liver. *Eur J Pediatr* **161**: 377–379.

Ogier de Baulny H, Saudubray JM. (2000) Branched-chain organic acidurias. In: Fenrnades J, Saudubray JM, van den Berge G, eds. *Inborn Metabolic Diseases*. Springer-Verlag, Berlin, 197–212.

Ogier de Baulny H, Benoist JF, Rigal O, Touati G, Rabier D, Saudubray JM (2005) Methylmalonic and propionic acidaemias: management and outcome. *J Inherit Metab Dis* **28**: 415–423.

Picca S, Dionisi-Vici C, Abeni D, et al (2001) Extracorporeal dialysis in neonatal hyperammonemia: modalities and prognostic indicators. *Pediatr Nephrol* **16**: 862–867.

Rousson R, Guibaud P (1984) Long term outcome of organic acidurias: survey of 105 French cases (1967–1983). *J Inherit Metab Dis* **7**(Supplement 1): 10–12.

Sass JO, Hofmann M, Skladal D, Mayatepek E, Schwahn B, Sperl W (2004) Propionic acidemia revisited: a workshop report. *Clin Pediatr (Phila)* **43**: 837–843.

Saudubray JM, Touati G, Delonlay P, et al (1999) Liver transplantation in propionic acidaemia. *Eur J Pediatr* **158**(Supplement 2): 65–69.

Schulze A, Lindner M, Kohlmuller D, Olgemoller K, Mayatepek E, Hoffmann GF (2003) Expanded newborn screening for inborn errors of metabolism by electrospray ionization-tandem mass spectrometry: results, outcome, and implications. *Pediatrics* **111**: 1399–1406.

Surtees RA, Matthews EE, Leonard JV (1992) Neurologic outcome of propionic acidemia. *Pediatr Neurol* **8**: 333–337.

Van Calcar SC, Harding CO, Lyne P, et al (1998) Renal transplantation in a patient with methylmalonic acidaemia. *J Inherit Metab Dis* **21**: 729–737.

van der Meer SB, Poggi F, Spada M, et al (1994) Clinical outcome of long-term management of patients with vitamin B_{12}-unresponsive methylmalonic acidemia. *J Pediatr* **125**: 903–908.

van der Meer SB, Poggi F, Spada M, et al (1996) Clinical outcome and long-term management of 17 patients with propionic acidaemia. *Eur J Pediatr* **155**: 205–210.

van't Hoff W, McKiernan PJ, Surtees RA, Leonard JV (1999) Liver transplantation for methylmalonic acidaemia. *Eur J Pediatr* **158**(Supplement 2): 70–74.

Waisbren SE, Albers S, Amato S, et al (2003) Effect of expanded newborn screening for biochemical genetic disorders on child outcomes and parental stress. *JAMA* **290**: 2564–2572.

Wilcken B, Wiley V, Hammond J, Carpenter K (2003) Screening newborns for inborn errors of metabolism by tandem mass spectrometry. *N Engl J Med* **348**: 2304–2312.

Wilson JMG, Jungner G (1968) Principles and practice of screening for disease. *Public Health Papers WHO* **34**: 26–39.

Outcome of neonatal screening for medium-chain acyl-CoA dehydrogenase deficiency in Australia: a cohort study

Bridget Wilcken, Marion Haas, Pamela Joy, Veronica Wiley, Meredyth Chaplin, Carly Black, Janice Fletcher, Jim McGill, Avihu Boneh

Summary

Background Medium-chain acyl-CoA dehydrogenase (MCAD) deficiency is the disorder thought most to justify neonatal screening by tandem-mass spectrometry because, without screening, there seems to be substantial morbidity and mortality. Our aim was to assess the overall effectiveness of neonatal screening for MCAD deficiency in Australia.

Methods We identified MCAD-deficient patients from a total population of 2 495 000 Australian neonates (810 000 screened) born between April 1, 1994, and March 31, 2004. Those from a cohort of 1 995 000 (460 000 screened) were followed up for at least 4 years, and we recorded number of deaths and severe episodes, medical and neuropsychological outcome, and hospital admissions within the screened and unscreened groups.

Findings In cohorts aged at least 4 years there were 35 MCAD-deficient patients in those not screened (2·28 per 100 000 total population) and 24 in the screened population (5·2 per 100 000). We estimated that patients with this disorder in the unscreened cohort remained undiagnosed. Before 4 years of age, three screened patients had an episode of severe decompensation (including one neonatal death) versus 23 unscreened patients (including five deaths). At the most conservative estimate, relative risk of an adverse event was 0·44 (95% CI 0·13–1·45). In the larger cohort the relative risk (screened *vs* unscreened) of an adverse event by age 2 years was 0·26 (95% CI 0·07–0·97), also a conservative estimate. 38 of 52 living patients had neuropsychological testing, with no suggestions of significant differences in general cognitive outcome between the groups.

Interpretation Screening is effective in patients with MCAD deficiency since early diagnosis reduces deaths and severe adverse events in children up to the age of 4 years.

Lancet 2007; 369: 37–42

See Comment page 5

Children's Hospital at Westmead, Sydney, NSW 2145, Australia (Prof B Wilcken MBChB, P Joy PhD, V Wiley PhD, C Black BPsych); University of Sydney, Sydney, Australia (B Wilcken); Centre for Health Economics, Research and Evaluation, University of Technology, Sydney, Australia (M Haas PhD, M Chaplin BAppSc); Women's and Children's Hospital, Adelaide, Australia (J Fletcher MD), Royal Hospital for Children, Brisbane, Australia (J McGill MBBS); and Genetic Health Services Victoria, Murdoch Childrens Research Institute and University of Melbourne, Melbourne, Australia (A Boneh MD)

Correspondence to: Prof Bridget Wilcken, Children's Hospital at Westmead, Locked Bag 4001, Westmead, NSW 2145, Australia. bridgetw@chw.edu.au

Introduction

Neonatal screening by tandem-mass spectrometry has burgeoned in the past decade. This type of screening is universal in Australia and widespread in many other countries, but despite preliminary studies,[1,2] the overall effectiveness of this expanded screening has not been clearly shown.

Medium-chain acyl-CoA dehydrogenase (MCAD) deficiency is the disorder most frequently diagnosed by neonatal screening with tandem-mass spectrometry, together with phenylketonuria.[3–6] The disorder results in decreased ability to withstand catabolic stress. Patients risk hypoketotic hypoglycaemia, which leads to coma or death during intercurrent illness. In reports of patients diagnosed clinically with the disorder, 20–25% died, usually during a first episode, and a further 20% sustained neurological damage.[7–11] A retrospective neonatal-screening study of 100 000 stored samples reported one in eight patients had died and only one survivor had learning difficulties.[12] Some individuals with MCAD deficiency could remain healthy, with no episodes of decompensation.[12,13] After diagnosis, management plans to avoid catabolism during fasting or illness generally prevent adverse episodes.[14] This success with intervention implies the probable advantage of newborn baby screening for MCAD deficiency, which is the disorder thought to most justify neonatal screening by tandem mass spectrometry.[15–19] However, there have been only very few reports of outcome after screening.[1,2]

Data for patients' outcome after screening for rare disorders are scarce and difficult to interpret because of small numbers, variable definitions and phenotypes of individual disorders, increased detection by screening,[4] different mutation range in screened and unscreened patients,[20–22] and often insufficient follow-up. We have therefore done a nationwide study in Australia of the overall effectiveness of neonatal screening by tandem-mass spectrometry.

Methods

Patients

We obtained data from all five newborn baby screening laboratories, all six biochemical genetics laboratories, and all five genetic metabolic clinical services in Australia, to identify patients with MCAD deficiency born between April 1, 1994 and March 31, 2004. The institutional ethics committees of all six centres approved this study, and written informed consent was obtained from parents of the patients for data access and additional testing.

There were three main patient groups: clinically diagnosed historical patients born between April 1, 1994 and March 31, 1998; clinically diagnosed contemporaneous patients born between April 1, 1998 and March 31, 2002 in states not screening at the time; and babies born between April 1, 1998 and March 31, 2002 who were tested by neonatal screening. Additionally, we recorded deaths and severe episodes of metabolic decompensation

in all screened and unscreened babies with MCAD deficiency, born between April, 2002 and March, 2004.

Procedures

Neonatal screening by tandem-mass spectrometry was done with dried blood samples obtained between 48 and 72 h of age.[4,23] MCAD deficiency was suspected in babies whose initial blood octanoylcarnitine (C8) concentration was greater than the concentration predetermined by individual laboratories, (≥1 μmol/L) and whose C8: decanoylcarnitine ratio was greater than one. These samples also had analysis of the common disease causing 985A→G mutation.[24] Low free-carnitine concentrations triggered further investigation in all patients irrespective of C8 concentrations. MCAD deficiency was confirmed by an isolated increase in straight medium-chain acylcarnitines in plasma and increased urinary hexanoylglycine concentrations (both criteria were present in all living patients) or two copies of the common *ACADM* (acyl-coenzyme A dehydrogenase, C-4 to C-12 straight chain) gene mutation. In clinically diagnosed patients, diagnosis was principally by enzymatic analysis and increased urinary hexanoylglycine concentration.

Parents of babies diagnosed with MCAD deficiency were given a written management plan, which included maximum length of overnight fasts by age, a sick-day regimen for minor illness, and instructions to telephone the metabolic physician for advice during illness. All parents had a letter for emergency departments, to expedite treatment and administration of intravenous glucose. Carnitine treatment was not routine.

Hospital admissions, episodes of decompensation, other complications, and death were recorded from the medical case-notes and after discussion with physicians treating genetic metabolic disease. A questionnaire to families recorded admission to other centres. A severe episode was defined as an emergency admission, with need for intravenous therapy, and not a prophylactic admission. For young patients (born April, 2002–March, 2004) only severe episodes and death were recorded.

The Woodcock-Johnson III (WJ III) tests of cognitive abilities and achievement[25] were used to assess cognitive outcome for patients aged 4 years or more. This validated test measures general intellectual ability, specific cognitive abilities, and academic achievement, across a wide range of ages (2–90 years). Vineland adaptive behaviour scales (VABS)[26] were used to measure the child's adaptive functioning across communication, daily living, and socialisation domains.

Statistical analysis

To estimate effectiveness of screening on rates of death and severe episodes before 4 years of age, we assumed that the birth incidence in the unscreened cohorts would be the same as that seen in the screened group. The undiagnosed patients in the unscreened cohorts could have been never symptomatic, could have been as symptomatic as those diagnosed, or might have occupied some midpoint. We did not consider the extreme position that all undiagnosed patients might have died or had severe episodes. The relative risk of adverse events was analysed, with 95% CIs and an absolute risk of adverse events provided.

We used a bigger sample (from April, 1994, to March, 2004) to compare actual rates of death or a severe episode by 2 years of age, and the relative risk was analysed with 95% CIs. A series of independent samples *t* tests were done to compare neuropsychological outcome variables in the screened group with those in the unscreened group and fully tested with partly tested groups.

Because admissions data were not normally distributed, we used non-parametric tests (Mann-Whitney *U* tests) to compare admission rates and length of stay per admission in the screened and unscreened groups.

Role of the funding source

This study was supported by a grant from The National Health and Medical Research Council, Australia, which played no part in study design, collection, analysis or interpretation of data, or writing of the report. The corresponding author had full access to all the data and had final responsibility for the decision to submit for publication.

Results

For those with at least 4 years follow-up, (1994–2002 cohorts) the median age at diagnosis was substantially later in the unscreened patients than in those who were screened (table 1). Five (14%) unscreened patients clinically

	Number of births	Clinical diagnoses	Family screening diagnoses	Diagnosis by age 4 years (total diagnoses at any age)	Total diagnoses per 100 000 population	Number of deaths by age 4 years (total at any age)	Number who had severe episode by age 4 years (total at any age)	Median age at diagnosis of proband (range) months
Unscreened 1994–98	1 002 000	13	6	10 (19)	1·9	2 (3)	9 (11)	19 (0·1–93)
Unscreened 1998–2002	533 000	12	4	16 (16)	3·0	3* (3)	9 (9)	10† (0·1–57)
All unscreened	1 535 000	25	10	26 (35)	2·3	5 (6)	18 (20)	16† (0·1–93)
Screened 1998–2002	460 000	..	..	24 (24)	5·2	1‡ (1)	2 (2)	0·5† (0·1–3)†

*2/3 neonatal deaths. †Excluding neonatal deaths. ‡1/1 neonatal death.

Table 1: **Medium-chain acyl-CoA dehydrogenase deficiency cases in screened and unscreened cohorts, age at diagnosis, and numbers of cases with severe adverse events by 4 years of age**

diagnosed and four (11%) siblings diagnosed by family screening were older than 4 years of age at diagnosis. The unscreened cohorts included three (9%) patients diagnosed after a younger sibling had been detected by newborn screening. The total diagnoses in the screened cohort was just over double that in unscreened babies; no false-negative screening results are known. The 1998–2002 overall false-positive rate was 0·01%, and the positive predictive value was 42%. Based on MCAD deficiency frequency in the screened population, the expected number of patients with the disorder in the unscreened group was 80 (95% CI 52–121).

The results of analysis for the common 985A→G mutation in the 1998–2002 cohort study were available for 21 (88%) of 24 screened patients (12 homozygous for 985A→G, allele frequency 76%) and for 20 (80%) of the 25 unscreened patients (16 homozygous 985A→G, allele frequency 90%). Only one patient had no copy of the common mutation. (Data from one state, New South Wales, were reported in a study of genotype and biochemical phenotype correlations. This cohort included all 17 patients from New South Wales included in the main study).[22] Two patients carried one copy of the mild mutation 199C→T,[20] neither of whom had any hospital admissions or events attributable to MCAD deficiency.

In the unscreened 1994–2002 cohorts, there were six (17%) deaths, at age 3 days (two babies), and at 17, 19, 47, and 93 months, all apparently secondary to MCAD deficiency. In the 24 patients diagnosed by screening there was one death, at 3 days of age, after poor feeding, then irrecoverable coma. The newborn screening sample taken post mortem was diagnostic (octanoylcarnitine concentration 34 μmol/L) and the baby was homozygous for the common 985A→G mutation. A 6-year-old sibling, who had been admitted twice for prostration during intercurrent illness, was reported to be affected. In all deaths, the diagnosis was made post mortem, and only one child had had a previous admission.

In the unscreened 1994–2002 cohorts, 20 (57%) patients had a severe episode of decompensation that led to diagnosis, 18 (90%) of whom were less than 4 years of age. In the screened patients, two had severe episodes. One (4%) patient had hypoglycaemia on day 3, prompting transfer to neonatal intensive care, and the other was a boy aged 3 years who lost consciousness early one morning and was admitted, requiring intravenous glucose, 2 days after recovering from gastroenteritis and seemingly in good health.

Patients who died or were known to have had a severe decompensation in unscreened cohorts were $1·5\times10^{-5}$ of the population, versus $0·65\times10^{-5}$ in the screened cohorts (relative risk [RR] 0·44; 95% CI 0·13–1·45). This result is the most conservative estimate of the true picture, and assumes that all patients never diagnosed were completely asymptomatic. A liberal estimate (same rate of severe adverse episodes in undiagnosed as diagnosed patients, including siblings diagnosed by family screening) and a midpoint estimate between these two extremes would both show significantly fewer severe adverse events in the screened population than in the unscreened—three versus 53 events (0·19; 0·06–0·60) for the liberal estimate, and three versus 38 events (0·26, 0·08–0·85) for the midpoint estimate.

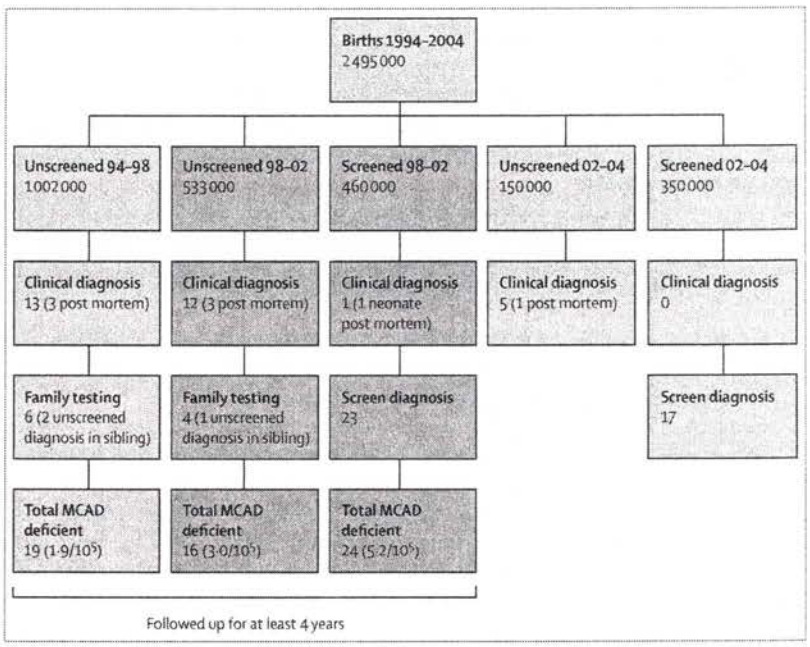


Figure: **Birth cohorts from April, 1994 to March, 2004, unscreened and screened, and the number of babies diagnosed with medium-chain acyl-CoA dehydrogenase (MCAD) deficiency in each cohort**

We identified 17 further MCAD-deficient patients by screening 350 000 patients between April, 2002, and March, 2004 (more states in Australia were by then screening) and these patients were all healthy between 2 and 4 years of age (figure). In the same 2 years, five patients were identified among 150 000 unscreened patients, one of whom died aged 2 days, and three others had severe events. Table 2 shows the total number of patients studied to 2 years of age (born from April, 1994 to March, 2004), and number of deaths plus severe events. The relative risk of death or a severe event by 2 years of age in screened patients (the most conservative estimate) was 0·26 (95% CI 0·07–0·97). The absolute risk of death or a severe event by 2 years of age was 5% (two of 41) for screened patients, and 55% (22 of 40) for unscreened patients diagnosed clinically or by family testing.

Of the 24 children diagnosed by screening between 1998, and 2002, ten (42%) were admitted in their first 4 years, generating 43 admissions, compared with 22 (63%) of the 35 unscreened children, who generated 74 admissions (table 3). Overall, the rate of admission for the screened group was 1·8 per child and for the unscreened group 2·1 per child. On a population basis the total admissions were, however, significantly higher for the screened cohort (p=0·038), and all but one of these were prophylactic admissions during intercurrent

	Total	Unscreened	Screened	Relative risk for screened vs unscreened
Births 1994–2004	2 495 000	1 685 000	810 000	
MCAD deficiency diagnoses by 2 years of age (total diagnosed in the cohort)	69 (81)	28 (40)	41 (41)	
Death by 2 years of age (neonatal deaths)	5 (4)	4 (3)	1 (1)	
Severe episode by 2 years of age	19	18	1	0·16 (95% CI 0·02–1·09)
Death or severe episode by 2 years	24	22	2	0·26 (95% CI 0·07–0·97)

Table 2: **Medium-chain acyl-CoA dehydrogenase (MCAD) deficiency diagnosed by the age of 2 years in babies born in Australia from April, 1994 to March, 2004**

illness or poor feeding. The length of admission was shorter for screened patients (mean 2·35 *vs* 2·95 days) but this was not significant (U=1456, p=0·44).

We tested 38 (73%) of 52 living patients for neuropsychological outcomes. Of these, 25 (13 screened) had full testing and a further 13 (six screened) were tested only by VABS. 14 could not be tested—four patients were geographically distant, four refused, and six were not able to be contacted for this part of the study. Mean general intellectual ability scores by the WJ III testing were 103·6 (SD 11·7) for screened and 104·9 (14·8) for unscreened patients (p=0·809). Only two of 25 patients (one screened) had a score greater than one SD below the mean. Adaptive behaviour composite scores were also similar—19 screened patients had a mean score of 101·6 (17·8), versus 98·4 (14·6) for 19 unscreened patients (p=0·547). Patients without cognitive testing had similar adaptive behaviour scores to those tested (data not shown). The detailed results of neuropsychological testing will be presented elsewhere.

Discussion

We have identified all patients diagnosed with MCAD deficiency in Australia, who are now at least 2 years of age. From a population of almost 2·5 million newborn babies, 32% were screened for the disorder. One patient in the screened cohort died on day 3 (before screening), and all remaining patients are well. No cases are known to have been missed by screening. More patients in the unscreened cohort died or had a severe adverse event by 2 years of age than in the screened cohort, but we detected no difference in cognitive outcome in surviving screened or unscreened children in the older cohorts.

There are several distinct advantages to our study. First, we are confident about complete ascertainment of our patient population, because all laboratories that provide neonatal screening and biochemical genetic services, and all genetic metabolic clinical services in Australia collaborated in this study. Second, there was an agreed definition for MCAD deficiency, on the basis of our findings in clinically diagnosed cases over the past 20 years, and published reports. In some other reports[1,2] there have not been clear definitions of what was included as MCAD deficiency, which could in part be responsible for the lower birth rate we noted compared with the rate others have reported.[17] Additionally, our study adds in that we report outcomes in patients up to 4 years of age in a screened cohort. Previous studies reported results for children of mixed ages, many of whom had shorter follow-up. No other study has documented numbers of hospital admissions, nor have outcomes been reported with use of a validated measure of cognitive ability.

There were two main factors that complicated the interpretation of our results. First, the higher frequency of MCAD deficiency recorded by screening, as we previously reported,[4] makes comparison between the clinical phenotype of screened and unscreened patients somewhat difficult. For the analysis we modelled different clinical situations for the missing patients in the unscreened groups—ie, those who were never diagnosed with the disorder. The first model assumes that all were asymptomatic, which is a very conservative and unlikely

	Screened (n=24)					Unscreened (n=35)				
	Number of children ever admitted*	Number of admissions (range per child)	Total length of stay (days)	Total admissions (per 100 000 population screened)	Mean length of stay per admission days	Number of children ever admitted*	Number of admissions (range per child)	Total length of stay (days)	Total admissions (per 100 000 population unscreened)	Mean length of stay per admission days
<1 year	5 (21%)	18 (1–8)	37	..	..	13 (37%)	29 (1–6)	101	..	..
1 to <2 years	6 (25%)	17 (1–9)	45	..	..	12 (34%)	21 (1–4)	59	..	..
2 to <3 years	3 (13%)	3 (1–1)	7	..	..	5 (14%)	16 (1–6)	35	..	..
3 to <4 years	2 (8%)	5 (1–4)	12	..	..	6 (17%)	8 (1–3)	23	..	..
Total	10 (42%)	43	101	9·3	2·35 (SD 2·6)	22 (63%)	74	218	4·1	2·95 (SD 3·4)

*The proportion of children alive at the beginning of a period who were admitted to hospital.

Table 3: **Number of hospital admissions and length of stay for cohorts born April, 1994 to March, 2002**

estimate. In this case there was no statistically significant difference in the occurrence of death or a severe decompensation in screened and unscreened patients to the age of 4 years. In reality, we are certain that some unscreened patients experienced severe events, but remained undiagnosed. This notion was proven in one of the three older siblings, who had been very symptomatic, but was diagnosed only because of a neonatal screening diagnosis in a younger brother. There could also have been patients with MCAD deficiency who died undiagnosed. Although pathologists were alerted to the implications of hepatic steatosis, appropriate further investigations are only now becoming routine.

The second model assumes the rate of adverse events to be similar to those in the unscreened cohorts diagnosed clinically or by family study, and those never diagnosed. The third model assumes a midpoint between these two models. In both of these two models there were significant differences seen, which favour newborn screening. When we analysed the larger cohort, which included infants with MCAD deficiency aged 2–4 years, there were substantially more patients who had died or had a severe adverse event in the unscreened than the screened group.

A second complication was a consideration of how comparable were patients diagnosed clinically or by screening in terms of genotype and biochemical phenotype. Patients diagnosed by neonatal screening had a different profile of mutations in the *ACADM* gene,[22] which could have included a group with milder disease. Before neonatal screening, around 80% of patients with symptoms were homozygous for the common mutation 985A→G and a further 18% were heterozygous for this mutation and another.[28] In our screened patients the allele frequency for this mutation was only 76%, although all but one screened patients diagnosed with MCAD deficiency carried at least one copy.

A different mutation profile has also been documented in other neonatal screening programmes in the USA and Germany.[20,21] All our patients had altered biochemistry, which indicates changed function, and biochemical hallmarks of pronounced enzyme deficiency. Our definition of the disorder for diagnosis by neonatal screening is based on our findings in symptomatic, clinically diagnosed patients. However, there is no doubt that MCAD deficient patients diagnosed by newborn baby screening do include some patients with more marginal biochemical changes.[22] With further screening programmes worldwide the diagnosis of MCAD deficiency could be better refined.

Population demographics are probably an important factor that contributes to the varying rate of deficiency reported from screening centres and the lower rate we recorded. In some 10% of births in our study, both parents were from east Asia, where the disorder is believed to be rare.[29] Only one patient in our study had a parent from east Asia (Indonesia); he was a compound heterozygote for uncommon mutations, which Waddell and colleagues[22] have reported.

Previous studies of neonatal screening outcome have not taken into account the probable missing patients, who were never diagnosed, in unscreened cohorts[7] or have used a retrospectively screened group from another country as a control group,[1] although this study did confine the comparison to patients homozygous for the common 985A→G mutation. Otherwise there has been no clear diagnosis of MCAD deficiency, and management protocols might have differed from ours, so we are unclear how comparable other studies are. In the Bavarian study[1] there were two non-neonatal deaths, both at 10 months, in the screened cohort. One of these patients also had congenital adrenal hyperplasia, which could have contributed to outcome. Non-neonatal deaths have also been reported in two patients diagnosed by screening in Pennsylvania, USA[30] whereas previously deaths after diagnosis were assumed to be very rare.[8] Our patients and their primary-care physicians were given a written management plan, and encouraged to telephone a metabolic physician for advice at any time, which could have helped to avoid severe events. However, four of our 81 patients (5%) died in the first 72 h. Such deaths cannot be avoided by neonatal screening, since the process includes an overnight assay at a central laboratory, with samples not valid in Australia unless taken after 24 h.[4]

Previously, we and others[7–10] reported that around 20% of patients who survived had some neuropsychological impairment, although a major review on this topic suggested fewer than this proportion.[25] By contrast, in our study both unscreened and screened patients were progressing well. Since fully tested children were as young as 4 years of age, some could later develop subtle problems such as specific learning difficulties, although emergency care is now more readily available and treatment more timely and skilled than previously.

In summary, we have shown that screening for MCAD deficiency provides a benefit, with reduction in mortality and morbidity in screened patients up to 4 years of age, compared with unscreened patients, but no differences in neuropsychological functioning in survivors.

Contributors

B Wilcken conceived the study, participated in data collection, analysis and interpretation, wrote the first draft of the report, and approved the final version. M Haas, P Joy, and V Wiley participated in designing the study, and in data collection, analysis and interpretation, in writing the report, and approved the final version. M Chaplin, C Black, J McGill, J Fletcher, and A Boneh participated in data collection, analysis and interpretation, and approved the final version of the report. A Boneh also participated in writing the report.

Conflict of interest statement

We declare that we have no conflict of interest

Acknowledgments

This study was supported by a grant from The National Health and Medical Research Council of Australia. We thank our associate investigators, Kevin Carpenter, John Christodoulou, Caroline Ellaway, Judith Hammond, and Edwin Kirk (New South Wales); Frank Bowling and David Cowley (Queensland); Enzo Ranieri (South Australia); the late

Paul Carman, Barry Lewis, and Geoff Knight (West Australia); and James Pitt and Heidi Peters (Victoria) for clinical and laboratory support and sharing their data, and to the many other administrative and laboratory staff who have given help; and Carol Dezateux for very helpful criticism.

References

1 Nennstiel-Ratzel U, Arenz S, Maier EM, et al. Reduced incidence of severe metabolic crisis or death in children with medium chain acyl-CoA dehydrogenase deficiency homozygous for c.985A>G identified by neonatal screening. *Mol Genet Metab* 2005; **85:** 157–59.

2 Waisbren SE, Read CY, Ampola M, et al. Newborn screening compared to clinical identification of biochemical genetic disorders. *J Inherit Metab Dis* 2002; **25:** 599–600.

3 Schulze A, Lindner M, Kohlmuller D, Olgemoller 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:** 1399–406.

4 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:** 2304–12.

5 Zytkovicz TH, Fitzgerald EF, Marsden D, et al. Tandem mass spectrometric analysis for amino, organic, and fatty acid disorders in newborn dried blood spots: a two-year summary from the New England Newborn Screening Program. *Clin Chem* 2001; **47:** 1945–55.

6 Frazier DM, Millington DS, McCandless SE, et al. The tandem mass spectrometry newborn screening experience in North Carolina: 1997–2005. *J Inherit Metab Dis* 2006; **29:** 76–85.

7 Iafolla AK, Thompson RJ Jr, Roe CR. Medium-chain acyl-coenzyme A dehydrogenase deficiency: clinical course in 120 affected children. *J Pediatr* 1994; **124:** 409–15.

8 Pollitt RJ, Leonard JV. Prospective surveillance study of medium chain acyl-CoA dehydrogenase deficiency in the UK. *Arch Dis Child* 1998; **79:** 116–19.

9 Touma EH, Charpentier C. Medium chain acyl-CoA dehydrogenase deficiency. *Arch Dis Child* 1992; **67:** 142–45.

10 Wilcken B, Hammond J, Silink M. Morbidity and mortality in medium chain acyl coenzyme A dehydrogenase deficiency. *Arch Dis Child* 1994; **70:** 410–12.

11 Derks TG, Reijngoud DJ, Waterham HR, et al. The natural history of medium-chain acyl CoA dehydrogenase deficiency in the Netherlands: clinical presentation and outcome. *J Pediatr* 2006; **148:** 665–70.

12 Pourfarzam M, Morris A, Appleton M, Craft A, Bartlett K. Neonatal screening for medium-chain acyl-CoA dehydrogenase deficiency. *Lancet* 2001; **358:** 1063–64.

13 Roe CR, Ding JH. Mitochondrial fatty acid oxidation disorders. In: Scriver CR, Beaudet AL, Sly WS, et al, eds. The metabolic and molecular bases of inherited disease, 8th edn. New York: McGraw Hill, 2005: 2297–2326.

14 Wilson CJ, Champion MP, Collins JE, Clayton PT, Leonard JV. Outcome of medium chain acyl-CoA dehydrogenase deficiency after diagnosis. *Arch Dis Child* 1999; **80:** 459–62.

15 Maternal and Child Health Bureau. Newborn screening: toward a uniform screening panel and system; 2005. Available from http://mchb.hrsa.gov/screening/ (accessed Oct 31, 2006).

16 Dezateux C. Newborn screening for medium chain acyl-CoA dehydrogenase deficiency: evaluating the effects on outcome. *Eur J Pediatr* 2003; **162** (suppl 1): S25–28.

17 Pollitt RJ, Green A, McCabe CJ, et al. Neonatal screening for inborn errors of metabolism: cost, yield and outcome. *Health Technol Assess* 1997; **1:** 1–202.

18 Seymour CA, Thomason MJ, Chalmers RA, et al. Newborn screening for inborn errors of metabolism: a systematic review. *Health Technol Assess* 2001; **1:** 23–26.

19 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:** 1–121.

20 Andresen BS, Dobrowolski SF, O'Reilly L, et al. Medium-chain acyl-CoA dehydrogenase (MCAD) mutations identified by MS/MS-based prospective screening of newborns differ from those observed in patients with clinical symptoms: identification and characterization of a new, prevalent mutation that results in mild MCAD deficiency. *Am J Hum Genet* 2001; **68:** 1408–18.

21 Maier EM, Liebl B, Roschinger W, et al. Population spectrum of ACADM genotypes correlated to biochemical phenotypes in newborn screening for medium-chain acyl-CoA dehydrogenase deficiency. *Hum Mutat* 2005; **25:** 443–52.

22 Waddell L, Wiley V, Carpenter KH, et al. Medium-chain acyl-CoA dehydrogenase deficiency: genotype-biochemical phenotype correlations. *Mol Genet Metab* 2005; **87:** 32–39.

23 Wiley V, Carpenter K, Wilcken B. Newborn screening with tandem mass spectrometry: 12 months' experience in NSW Australia. *Acta Paediatrica Suppl* 1999; **88:** 48–51.

24 Matsubara Y, Narisawa K, Ye-Qi Y, et al. Prevalence of K329E mutation in medium-chain acyl-CoA dehydrogenase gene determined from Guthrie cards. *Lancet* 1991; **338:** 552–53.

25 Woodcock RW, McGrew KS, Mather N. Woodcock-Johnson III. Itasca, IL: Riverside Publishing, 2001.

26 Sparrow SS, Balla DA, Cicchetti DV. Vineland adaptive behaviour scales. Interview edition survey form manual. Circle pines, MN: American Guidance Service Inc. 1984.

27 Grosse SD, Khoury MJ, Greene CL, Crider KS, Pollitt RJ. The epidemiology of medium chain acyl-CoA dehydrogenase deficiency: an update. *Genet Med* 2006; **8:** 205–12.

28 Tanaka K, Yokota I, Coates PM, Strauss AW, Kelly DP, Zhang Z, et al. Mutations in the medium chain acyl-CoA dehydrogenase (MCAD) gene. *Hum Mutat* 1992; **1:** 271–79.

29 Ensenauer R, Winters JL, Parton PA, et al. Genotypic differences of MCAD deficiency in the Asian population: novel genotype and clinical symptoms preceding newborn screening notification. *Genet Med* 2005; **7:** 339–43.

30 Ziadeh R, Hoffman EP, Finegold DN, et al. Medium chain acyl-CoA dehydrogenase deficiency in Pennsylvania: neonatal screening shows high incidence and unexpected mutation frequencies. *Pediatr Res* 1995; **37:** 675–78.

Healthcare Use and Costs of Medium-chain Acyl-Coa Dehydrogenase Deficiency in Australia: Screening Versus No Screening

Marion Haas, PhD, Meredyth Chaplin, BAppSc, Pamela Joy, PhD, Veronica Wiley, PhD, Carly Black, Bpsych, and Bridget Wilcken, MBChB

Objective To describe and analyze the use and costs of hospital services for children diagnosed with medium-chain acyl-CoA dehydrogenase (MCAD) deficiency either with newborn screening or clinical diagnosis in Australia between 1994 and 2002. MCAD deficiency is a potentially lethal disorder of fatty-acid oxidation.

Study design We conducted a retrospective audit of medical records supplemented by a parental survey.

Results A total of 59 children with MCAD deficiency were identified, 24 by using newborn screening. In the first 4 years of life, screening children cost an average of $A1676 (US$1297) per year for inpatient, emergency department, and outpatient visits, compared with $A1796 (US$1390) for children in whom a clinical diagnosis was made. Forty-two percent of the children who underwent screening were admitted to the hospital, compared with 71% of children who did not undergo screening. Children who did not undergo screening used significantly more inpatient services and cost significantly more in emergency services. There were also some significant differences in use on a year-by-year basis.

Conclusions Children who do not undergo screening may be more likely to be admitted to the hospital and to incur higher emergency department costs than children who underwent screening, and children seem more likely to attend hospital outpatient clinics. Screening does not result in higher costs from a hospital perspective. *(J Pediatr 2007;151:121-6)*

Tandem mass spectrometry (MS/MS) screening for disorders of amino acid, organic acid, and fatty acid metabolism has become widespread. However, increasing the scale and scope of screening programs requires understanding of costs, not just of screening but also downstream costs, use of healthcare, and longer term outcomes. The most common disorder newly detectable with MS/MS is medium-chain acyl-CoA dehydrogenase (MCAD) deficiency, a disorder of fatty acid oxidation. Hypoketotic hypoglycemia with intercurrent illness develops in children with MCAD deficiency, and there is a substantial risk of death in patients in whom it is not diagnosed.[1]

Several economic evaluations and systematic reviews of newborn screening with MS/MS have been published,[1-7] but only 2 specifically have considered MCAD deficiency.[2,7] Venditti et al reported estimated costs for patients who did not undergo screening using a 30-year retrospective chart review of 32 patients.[7] These data were used to construct a "typical" pattern of care for patients with MCAD deficiency diagnosed with screening. Pandor et al carried out a systematic review and modeling exercise.[2] The study reported here compares the actual costs and use of children with MCAD deficiency, both those who underwent screening and those who did not.

The evaluation of costs and outcomes is complicated because more children with MCAD deficiency are detected with screening, and they have a different mutation profile from children in whom MCAD is diagnosed clinically.[8] Some children in whom MCAD deficiency is detected with screening might not be at risk of adverse outcomes. The implications of these findings may be: 1) increased use and costs—families with early diagnosis may more readily seek medical assistance during intercurrent illness and undergo more routine monitoring; 2) decreased use and costs—because a higher proportion of children in the screened group may have a less severe variant of MCAD deficiency[8]; or

See editorial, p 108, and related article, p 115

ED	Emergency department	MCAD	Medium-chain acyl-CoA dehydrogenase
LOS	Length of stay	MS/MS	Tandem mass spectrometry

From the Centre for Health Economics Research and Evaluation, University of Technology, Sydney, Australia (M.H., M.C.); Children's Hospital at Westmead, Sydney, Australia (P.J., V.W., C.B., B.W.); University of Sydney, Sydney, Australia (B.W.).

Supported by a project grant from the National Health and Medical Research Council (249406).

Submitted for publication Aug 29, 2006; last revision received Jan 30, 2007; accepted Mar 5, 2007.

Reprint requests: Associate Professor Marion Haas, CHERE, UTS, PO Box 123, Broadway NSW 2007. E-mail: marion.haas@chere.uts.edu.au.

0022-3476/$ - see front matter

10.1016/j.jpeds.2007.03.011

3) increased use and decreased costs—with early detection, careful monitoring, and earlier intervention an increased rate of use per patient may result, but the cost, per patient or total, may be lower with cheaper interventions and shorter length of stay.

In Australia, newborn screening with MS/MS began in 2 of 6 states in 1998 and 1999 and a third at the end of 2001. We captured relevant data on all children born between April 1994 and March 2002 and examined the differences in costs, healthcare use, and outcomes between children who underwent screening and children who did not. The human ethics committee of each tertiary children's hospital separately approved the conduct of this study.

METHODS

Patients

PATIENT GROUPS: 1) children with MCAD deficiency diagnosed clinically from April 1994 to March 1998 in all states in Australia; 2) children with MCAD deficiency diagnosed clinically from April 1998 to March 2002 in states not screening for MCAD deficiency, and 3) patients born from April 1998 to March 2002 in whom MCAD deficiency was diagnosed with newborn screening. Because only 5 laboratories in Australia perform diagnostic testing for MCAD deficiency, we believe we have virtually complete ascertainment.[9]

Data Collection

USE. All patients with MCAD deficiency were seen at tertiary children's hospitals. Each of the hospitals in the Australian states of New South Wales (2), Queensland (2), Victoria, South Australia, and Western Australia (1 each) were requested to provide details of emergency department (ED), inpatient, and outpatient (ie, hospital-based ambulatory care) episodes of care for all children in whom MCAD deficiency was diagnosed between April 1994 and March 2002. Details requested included number of episodes, year of episode, hospital length of stay (LOS), whether the patient was admitted from the ED, principal and secondary diagnoses, and procedures, including diagnostic tests, performed. In addition, because some children may have gone to hospitals other than the tertiary facility, we surveyed all the parents of children with MCAD deficiency to inquire about this aspect of use.

COSTS. The 3 States in which screening was established during this period provided information about the cost of the MS/MS test, including consumables, equipment, overhead, and staffing. The equipment was figured at 1 instrument per screening laboratory, with depreciation over 7 years. The cost estimated for this study assumes that a minimum of 50,000 MS/MS tests per machine per year will be performed. Therefore, the cost per test is likely to be higher than average in smaller states and lower in larger states. In the smallest state, the cost of the screening test per child was $A2.24 (US$1.73), and in the largest state it was $A1.02 (US$0.79), compared with an average cost of $A1.63 (US$1.26). Similarly, the total costs of the screening program varied by state depending on the cost per child screened and the number of tests performed.

The perspective taken is that of the state public hospital systems. A request was made to each hospital to provide information on available costs (eg, bed-day costs, costs of tests, costs of procedures, etc) for each admission for each child. Hospitals were generally able to provide detailed information about the costs of individual inpatient admissions. When such costs were not available (eg, in some hospitals, costs were not available for admissions before 2000), a cost was assigned to each inpatient episode on the basis of the average cost of a child in the study of the same age and sex with the same principal diagnosis, adjusted for LOS. Outpatient clinic and ED costs were based on the NSW Health Costs of Care Standards.[10] Visits to the ED were assigned a cost of $A450 (US$348) for patients who were admitted and $A373 (US$289) for patients who were not admitted, which were the benchmark average costs per ED weighted episode for triage 2 category patients for the specialist hospitals located in NSW. Similarly, a cost of $A254 (US$197) was assigned to each outpatient clinic visit on the basis of the cost calculated for Pediatric Medicine outpatient clinics.[11] Additional costs were assigned to outpatient allied health visits (eg, dietician, physiotherapy) on the basis of the costs calculated for the same set of standards. All costs were converted to 2002 Australian dollars.[11] No attempt was made to account for out-of-pocket costs for this group of children and their families, although this has been done for the larger study.

Because the youngest of the children who underwent screening was aged 4 years, only years 1 to 4 have been compared across the groups.

Data Analysis

USE. The number of inpatient, outpatient, and ED visits were collected for each individual in the study. Admissions in which the principal diagnosis was clearly unrelated to MCAD deficiency were not counted (eg, trauma, acute appendicitis). Because the age distribution varied between the screened and unscreened groups (children who underwent screening were born between 1998 and 2002; children who did not undergo screening were born between 1994 and 2002), it was necessary to take this into account when calculating use rates. The average number and range of inpatient, ED, and outpatient services and the rates of use for both the screened and unscreened groups were calculated. Within each group, the use rates per age group were calculated. Differences in the groups were tested by using the Fisher exact test.

COSTS. Total costs for each individual were calculated by multiplying the costs per ED, inpatient, and outpatient episode by the number of episodes used, and an annual cost per child was calculated. Cost data for individuals, plus the number of services used, were used to calculate average costs per year in the screened and unscreened groups. The t test was

Table I. Health care use in the first 4 years of life: comparison of screened group with unscreened group (total group), contemporaneous unscreened group, and historical unscreened group

	Screened (n = 24)	All unscreened (n = 35)	P value	Contemporaneous unscreened (n = 16)	P value	Historical unscreened (n = 19)	P value
Inpatient							
% of patients using	42	71	.03	63	.33	79 (63)‡	.03
Mean #pa* if used (SD)	1.1 (1.4)	0.9 (0.9)		0.7 (0.5)		1.1 (1.1)	
Mean los pa† if used (SD)	2.5 (2.6)	2.6 (3.1)		1.7 (1.3)		3.2 (3.8)	
ED							
% using	21	34	.38	50	.09	21 (21)‡	1.0
Mean #pa if used (SD)	0.5 (0.2)	1.3 (1.3)		1.2 (1.3)		1.6 (1.6)	
Outpatient							
% using	63	60	1.00	69	.75	53 (37)‡	.55
Mean #pa if used (SD)	1.5 (1.0)	1.6 (1.4)		1.4 (1.2)		1.9 (1.7)	

All *P* values with Fisher exact test.
*Mean number of admissions per annum as an inpatient in patients admitted at least once.
†Mean length of stay per annum among those who were admitted to hospital.
‡The percent of patients in this group from states that had introduced screening.

used to test for differences in costs. The 2-sample *t* test for unequal variances was used to compare the mean cost of the health services used by the screened group with those used by the unscreened groups. Because the cost data were skewed, a 95% CI for the mean cost difference between groups was calculated with the nonparametric BCa bootstrap method.[12]

Seven deaths, all apparently related to to MCAD deficiency, were recorded in children with MCAD deficiency, (6 unscreened, 1 screened). Of the 7 deceased children, 2 had used services (and hence incurred costs). However, because both children lived for at least 4 years, their deaths have not had any influence on the data presented here.

RESULTS

A total of 59 children with MCAD deficiency who were born between 1994 and 2002 were identified, 24 in the screened group and 35 in the unscreened group. Within the unscreened group, 16 children were born contemporaneously with children in the screened group (1998-2002), and 19 were classified as belonging to the historical unscreened group (ie, born between 1994 and 1998).

Cost of the MS/MS Test

Three states were able to provide detailed information on the additional costs of running the MS/MS testing system. Costs were based on testing 189,540 children per year in the 3 states. Among the 3 programs, equipment and overhead costs did not vary significantly, but the cost of consumables and staffing did. The cost of consumables depended on methodological variation, in particular the extent of use of labeled standards,* and staffing depended on contextual issues, such as the degree to which computerization was used to download and analyze data and the extent to which the screening services were able to justify the employment of additional staff to implement MS/MS screening. An average cost was estimated by using these assumptions:

- 1 additional hospital scientist
- 70% samples tested with a computerized laboratory interpretation management system
- A mid-range cost for labeled standards.

With these assumptions, the average cost of the MS/MS test per child screened was estimated at $A1.67 (US$1.29). This cost did not include confirmatory testing, which, for MCAD deficiency was estimated to add an additional $A0.04 (US$0.03) per child screened.

* (Labeled standards are isotopically labeled amino acids [AAs] or acyl carnitines [ACs]. All levels of AAs or ACs in samples are determined relative to these. Labeled standards can be made "in-house" or purchased, either in bulk or in "kit" mixes. Kit mixes are the most expensive, followed by bulk purchase, and then in-house manufacture.)

Use

Table I summarizes the annual use of the 3 groups. The proportion of children using inpatient, ED, and outpatient services is shown, as is the mean number of visits in the children who used the services and the mean LOS for those children who used inpatient services. It shows that significantly fewer children in the screened group used inpatient services compared with the children in the unscreened group (42% versus 71%; $P = .03$). Of the children who were admitted to the hospital, the mean number of admissions per year was approximately 1 for both groups (1.1 screened; 0.9 unscreened), and the mean LOS was 2.5 days (2.5 screened; 2.6 unscreened). There were significantly more admissions per year in children in the historical unscreened group (79%; $P = .03$) and, on average, they stayed for a longer time (3.2 days). The most common reasons for admission were gastrointestinal illness or infection (eg, respiratory), and management involved the intravenous replacement of fluids, parenteral infusion of nutrition, or antibiotic therapy. Approximately 32% of

Table II. Total health care use by year of life, percent using (number using): comparison of screened group with unscreened group (total group), contemporaneous unscreened group, and historical unscreened group

	Screened (n = 24)	All unscreened (n = 35)	P value	Contemporaneous unscreened (n = 16)	P value	Historical unscreened (n = 19†)	P value
Inpatient							
Year 1	13 (3)	46 (16)	.01*	50 (8)	.01	42 (8)	.08
Year 2	25 (6)	40 (14)	.27	25 (4)	1.00	53 (10)	.20
Year 3	13 (3)	20 (7)	.51	13 (2)	1.00	26 (5)	.68
Year 4	8 (2)	26 (6)	.17	19 (2)	.37	32 (4)	.21
ED							
Year 1	8 (2)	26 (9)	.17	50 (8)	.007*	5 (1)	1.0
Year 2	8 (2)	17 (7)	.45	19 (3)	.37	16 (4)	.64
Year 3	13 (3)	14 (5)	1.00	25 (4)	.41	5 (1)	.62
Year 4	0	14 (5)	.07	19 (3)	.06	11 (2)	.44
Outpatient							
Year 1	63 (15)	34 (12)	.04*	38 (7)	.20*	32 (5)	.06
Year 2	50 (12)	46 (16)	.80	38 (6)	.53	53 (10)	.76
Year 3	38 (9)	37 (12)	1.00	38 (5)	1.00	37 (7)	1.0
Year 4	17 (4)	26 (9)	.53	25 (4)	.69	26 (5)	.43

All *P* values with Fisher exact test.

*$P < .05$.

†The number (%) of children in the historical unscreened group from screening states: inpatient 5 (20%); emergency 2 (33%); outpatient 1 (4%).

admissions in the screened group and 23% of those in the unscreened group occurred in hospitals other than specialist children's hospitals (data not shown in Table I).

A higher proportion of children who did not undergo screening (34%) than children who did (21%) went to the ED, although, overall, similar proportions went to hospital outpatient clinics (63% and 60%, respectively). However, the differences in the groups for ED and outpatient visits were not significant, although there are marked differences in the rates of attendance between the contemporaneous and historical unscreened groups. Of those who went to the ED, the average number of visits per year varied from 0.5 for the screened group to 1.3 for the contemporaneous unscreened groups and 1.6 for the historical unscreened group. The mean number of outpatient visits per year was less variable. Any differences were not significant.

Table II illustrates the year-by-year use of inpatient, ED, and outpatient services. In the first year of life, the screened group used significantly fewer inpatient services than the combined unscreened group ($P = .01$) and each of the contemporaneous and historical screened groups ($P = .01$; $P = .04$), but the screened group used significantly more outpatient services than the combined unscreened group ($P = .04$). No other significant differences emerged.

Costs of Care

A summary of the costs for inpatient, ED, and outpatient hospital services is shown in Table III. At $A1676 (US$1297), the average total costs per year for the screened group were higher than for the contemporaneous unscreened group ($A1211; US$937), but lower than the historical group ($A2288; US$1770). However, in all groups the range of costs was large. The average cost for ED admissions was significantly lower in the screened group than in the combined unscreened group ($P = .05$), but there were no other significant differences in costs.

DISCUSSION

We collected data on the costs of MS/MS testing and combined these with actual costs and use of hospital-based services to compare children with MCAD deficiency who underwent screening and those who did not. Overall, there were significant differences between children in the screened group and children in the unscreened group in the cost per patient of ED care and in the rate of use of inpatient hospital care. However, these results should be treated with caution because of the wide confidence intervals around the differences in costs. Some significant differences emerged when use was analyzed on a year-by-year basis. It is not clear what impact the deaths that occurred had on costs. If they had not died, the deceased children would have used health services, thus adding to the total costs of care, but if some deaths were caused by treatment for decompensation being initiated later in the course of an intercurrent illness rather than a more severe case of the condition, rates of use and the average cost per patient would not have been affected.

Despite the potentially serious impact of decompensation episodes on children's development, there were no differences in the groups in either physical development such as height and weight or neuropsychological function. However, they may have produced more subtle differences in cognitive function, which were not observable in such young children.

Children in the contemporaneous and historical unscreened groups were significantly more likely than children in the screened group to visit the ED and to be admitted as an inpatient during the first year of life. This pattern may be

Table III. Average cost per annum ($Aus) in the first 4 years of life: comparison of screened group with unscreened group (total group), contemporaneous unscreened group, and historical unscreened group

	Screened (n = 24)	All unscreened (n = 35)	P value*	Contemporaneous unscreened (n = 16)	P value*	Historical unscreened (n = 19)	P value*
Inpatient							
Mean	1427	1392	.96	739	.31	1942	−.56
Difference screened-unscreened (95% CI)†		35 (−1124-1798)		687 (−338-2390)		−515 (−2026–1402)	
ED							
Mean	42	184	.05	235	.08	141	.30
Difference screened-unscreened (95% CI)†		−142 (−274-−12)		−193 (−383-0)		−99 (−286-48)	
Outpatient							
Mean	207	220	.85	237	.71	206	.99
Difference screened-unscreened (95% CI)†		−13 (−143-119)		−30 (−176-131)		1 (−178-163)	
Total cost							
Mean	1676	1796	.88	1211	.53	2288	.55
Difference screened-unscreened (95% CI)†		−120 (−1367-1737)		465 (−694-2247)		−612 (−2355-1503)	

*All *P* values with *t* test.
†BCa bootstrap confidence intervals.[12]

caused by repeated visits and admissions that eventually resulted in a diagnosis. Children who underwent screening were significantly more likely to visit outpatient clinics during their first year of life than children in the combined unscreened group. Such a pattern of use certainly reflects the age at which children were when MCAD deficiency was diagnosed. The more intensive level of treatment in the first and second years of life provided to children who underwent screening (avoidance of fasting and sick-day regimens) may also have contributed to their lower rates of admission in the first year of life.

The main cost associated with the introduction of screening is that of the test itself, including confirmatory testing, and at $A2.06 (US$1.59) per child, it does not appear to be excessive. Two issues should be noted. First, during the time of the study, the states that used screening relied on 1 machine each but had back-up arrangements in place with another state in case of breakdowns. Since the time of the study, each state has acquired a second machine. Second, the cost of the test covers all the conditions for which a child is screened, not only MCAD deficiency. With the notion of opportunity cost, therefore, the cost of the test for MCAD deficiency could be $0, because the MS/MS technology would still be used to detect other conditions (such as phenylketonuria). However, it is possible to assign a cost per MCAD deficiency test on the basis of actual numbers of tests. However, the total costs of the screening program will vary by state depending on the cost per child undergoing screening, the number of tests performed, and the "throughput" of each machine (ie, the extent to which each machine is used at it maximum rate). Because the birth rate in Australia is not expected to change substantially and the rate of detection is unlikely to rise, large increases in costs are not expected. Indeed, as the technology improves, an enhanced capacity to computerize the downloading and analyzing of tests may decrease costs in the long term.

The study has some important strengths. Because services are centralized and there are only 5 collaborating biochemical genetics laboratories serving Australia, there is almost certain to have been complete ascertainment of children with MCAD deficiency. Patients are treated by only 9 metabolic physicians, all of whom meet regularly and collaborated on this study, so management strategies are similar throughout Australia. Care is mainly provided through the specialist children's hospitals, through scheduled ambulatory attendances and telephone contact with the treating physicians in these institutions. Also, we were able to gain information about the use of hospital services and costs for all identified patients.

However, the study also has limitations. Although we were able to identify all known cases of MCAD deficiency in Australia, the small numbers of children in both groups make comparisons difficult. This will inevitably lead to problems of power in identifying significant differences. However, there is no reason to believe that this group of children with MCAD deficiency is different in illness experience (and therefore use) than those in other parts of the developed world. Although there are no missing children, there may be some missing or misreported data for numbers of episodes of care. This is particularly likely to apply to ED visits and inpatient episodes of care that occurred outside the major pediatric hospitals, because this information was sought from parents and clinicians but not by investigation of case notes.

We have only included the costs from a hospital perspective. Such costs are likely to be different from those in

other countries for reasons connected with both the structure of the health system and management practices. For example, practices adopted in Australia, such as regular monitoring of children and the provision of allied health care (eg, nutrition, physiotherapy, etc) via hospital-based ambulatory clinics may not be a feature of other health systems (eg. the United States). Finally, we have no way of knowing what healthcare costs are incurred by children with MCAD deficiency in whom it remains undiagnosed.

It seems as though the third scenario suggested in the introduction has occurred. That is, early detection has resulted in an increased rate of use per patient of outpatient services. However, as ED visits and inpatient admissions were prevented, costs were lower in the screened group. The key question is whether patterns of management affect physical and neuropsychological development, because this would certainly affect long-term costs. Answering such a question would require further research, preferably a randomized controlled trial of different management strategies.

Acknowledgments available at www.jpeds.com.

REFERENCES

1. Insinga R, Laessig R, Hoffman G. Newborn screening with tandem mass spectrometry: examining its cost-effectiveness in the Wisconsin Newborn Screening Panel. J Pediatr 2002;141:524-31.

2. Pandor A, Eastham J, Chilchott J, Paisley S, Beverley C. Economics of tandem mass spectrometry screening of neonatal inherited disorders. Int J Technol Assess Health Care 2006;22:321-6.

3. Carroll A, Downs S. Comprehensive cost-utility analysis of newborn screening strategies. Pediatrics 2006;117:S287-95.

4. 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 2001;1:i-iv.

5. Schoen E, Baker J, Colby C, To T. Cost-benefit analysis of universal tandem mass spectrometry for newborn screening. Pediatrics 2002;110:781-6.

6. 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 2001;1:i-iv.

7. Vendetti L, Venditti C, Berry G, Kaplan P, Kaye E, Glick H, et al. Newborn screening by tandem mass spectrometry for medium-chain cayl-CoA dehydrogenase deficiency: a cost-effectiveness analysis. Pediatrics 2003;112:1005-105.

8. Andresen BS, Dobrowolski SF, O'Reilly L, Muenzer J, McCandless SE, Frazier DM, et al. Medium-chain acyl-CoA dehydrogenase (MCADD) mutations identified by MS/MS-based prospective screening of newborns differ from those observed in patients with clinical symptoms: identification and characterization of a new, prevalent mutation that results in mild MCADD deficiency. Am J Hum Genet 2001;68:1408-18.

9. Wilcken B, Haas M, Joy P, Wiley V, Chaplin M, Black C, et al. Neonatal screening for medium-chain acyl-CoA dehydrogenase deficiency: outcome at four years. Lancet 2007;369:37-42.

10. NSW Health Department. NSW cost of care standards 2003/04. Sydney: NSW Department of Health; 2004.

11. 6401.0 Consumer Price Index, Australia TABLE 1 CPI: All Groups, Index Numbers (Quarterly)(a). Accessed April 28, 2005 at: www.abs.gov.au/ausstats/abs@.nsf/lookupresponses/dc7c73d0cda3e34cca2570a500802675?opendocument.

12. Efron B, Tibshirani RJ. An introduction to the bootstrap. Boca Raton, FL: Chapman & Hall/CRC; 1993.

We are grateful to our associate investigators, Dr Kevin Carpenter, Professor John Christodoulou, Dr Caroline Ellaway, Dr Judith Hammond, and Dr Edwin Kirk (New South Wales), Professor Frank Bowling and Dr David Cowley (Queensland), Dr Janice Fletcher and Dr Enzo Ranieri (South Australia), the late Dr Paul Carman, Dr Barry Lewis, and Dr Geoff Knight (West Australia), and Dr Avihu Boneh, Dr James Pitt, and Dr Heidi Peters (Victoria) for clinical and laboratory support and for advice and assistance in costing. For vital assistance in extracting and organizing the data used in this project, we thank Dr Ralph Hanson, Lee Jeffries, and Linda Hobson (The Children's Hospital at Westmead, Sydney), Jenny Smith (The Sydney Children's Hospital), Michael Moore-Crouch (Princess Margaret Hospital for Children, Perth), Dr Jim McGill and Tracey Guan, (Royal Children's Hospital, Brisbane), Avihu Boneh, Cheryl Apperly, and Cameron Barnes (Royal Hospital for Children, Melbourne), Dr Janice Fletcher and Meg Phythian (Women's and Children's Hospital, Adelaide), Loretta Seamer, Anna Porfyriou, and Anne-maree Schneider (Mater Children's Hospital, Brisbane). Patsy Kenny provided invaluable assistance in the statistical analysis of the data.

Child Neuropsychology, *iFirst*: 1–13, 2008
http://www.psypress.com/childneuropsych
ISSN: 0929-7049 print / 1744-4136 online
DOI: 10.1080/09297040701864570

Psychology Press
Taylor & Francis Group

NEUROPSYCHOLOGICAL FUNCTIONING IN CHILDREN WITH MEDIUM CHAIN ACYL COENZYME A DEHYDROGENASE DEFICIENCY (MCADD): THE IMPACT OF EARLY DIAGNOSIS AND SCREENING ON OUTCOME

Pamela Joy,[1] Carly Black,[2] Antonella Rocca,[3] Marian Haas,[4] and Bridget Wilcken[2]

[1]The Children's Hospital at Westmead, Developmental Cognitive Neuropsychology Research Unit, Sydney, Australia, [2]The Children's Hospital Westmead, Sydney, Australia, [3]Rivendell Adolescent Unit, Sydney, Australia, and [4]Centre for Health Economics, Research and Evaluation, Sydney, Australia

Children with medium chain acyl coenzyme A dehydrogenase deficiency (MCADD) have been reported to be at high risk for neurocognitive deficits. However this has not been systematically studied and little is known about the exact nature of neuropsychological sequelae or of the impact of early diagnosis and screening on outcome. We examined cognitive and adaptive outcome in children with MCADD (N = 38, age range: 2 years, 2 months – 10 years, 3 months) diagnosed either through a newborn screening program (tandem mass spectrometry/MSMS) or upon clinical presentation. There was no evidence of overall intellectual impairment in either groups but there was some suggestion of poorer verbal and specific executive functioning (i.e., planning) abilities in the unscreened cohorts. Adaptive functioning was relatively intact with the exception of reduced Daily Living Skills in both our screened and unscreened groups. Early diagnosis and greater number of hospitalizations were related to higher verbal, communication, and socialization skills. Overall, our results highlight the importance of early diagnosis and management for children with MCADD.

Keywords: *Neuropsychology; Pediatric; Medium chain acyl coenzyme A dehydrogenase deficiency; Cognitive outcome; Early diagnosis*

Medium chain acyl coenzyme A dehydrogenase deficiency (MCADD) is the most common of the disorders of fatty acid oxidation (Carpenter, Wiley, Sim, Heath, & Wilcken, 2001; Roe & Coates, 1997; Wilcken, Carpenter, & Hammond, 1993; Wilcken, Hammond, & Silink, 1994; Wilson, Champion, Collins, Clayton, & Leonard, 1999). This defect in fatty acid oxidation causes a decreased ability to withstand catabolic stress. With illness or prolonged fasting, babies and children are at risk of episodes of hypoketotic hypoglycemia, which can lead to coma or death. With the introduction of tandem mass spectrometry (MSMS) newborn screening programs, early diagnosis of MCADD has

This study was supported by a grant from The National Health and Medical Research Council of Australia.

Address correspondence to Pamela Joy, The Children's Hospital at Westmead, Developmental Cognitive Neuropsychology Research Unit, Locked Bag 4001, Westmead, 2145, Sydney, Australia. E-mail: pamelaj@chw.edu.au

become possible and preventative measures can be taken to avoid damaging episodes of metabolic instability. Previously, literature suggested that adverse outcomes were not uncommon; however much of the investigation into MCADD focused on medical outcome with little or no neurocognitive follow up. Studies that have reported neuropsychological follow-up show adverse long-term outcomes of varying degree (ranging from mild to severe) in some patients (Iafolla, Thompson, & Roe, 1994; Pollitt & Leonard, 1998; Wilcken et al., 1994; Wilson et al., 1999).

Reports of adverse outcomes are not surprising as prior to and indeed following diagnosis patients are at risk of episodes of metabolic instability that include hypoglycemia and can lead to neurocognitive impairment (Iafolla et al., 1994; Leonard & Dezateux, 2002; Pollitt & Leonard, 1998; Wilson et al., 1999). In an unscreened population the first indication of MCADD is usually associated with serious illness requiring hospitalization. Generally babies or young children are typically diagnosed when they present with symptoms such as somnolence, lethargy, and even coma (Touma & Charpentier, 1992). If it is unknown that a child has the condition then any fasting associated with intercurrent illness can lead to a metabolic crisis (Leonard & Dezateux, 2002; Pollitt & Leonard, 1998; Touma & Charpentier, 1992; Wilcken et al., 1994; Wilson et al., 1999). It is possible that once a child has become sick enough to present clinically some irreversible insult to the brain may have already occurred.

The degree and nature of adverse outcome for MCADD deficiency remains unclear. There is some indication that there is an increased risk of global development delay in children with MCADD. Wilson et al. (1999) and Wilcken et al. (1994) found that 14.6% and 33.3% respectively of each of their cohorts exhibited cognitive impairment ranging from mild to severe, and Pollitt and Leonard (1998) reported that 10% of their sample exhibited neurological impairment/developmental delay. However, no formal neuropsychological testing was undertaken in any of these studies and the findings were based largely on clinical reports with short-term follow-up.

To date only one study has attempted to formally document the presence of global and specific cognitive deficits. Iafolla et al. (1994) conducted a retrospective analysis and reported behavioral and learning problems in a sample of MCADD patients. However only a small portion (24.2%) of the sample had formal testing and there was considerable variability in the test measures used. With the exception of this study, formal neuropsychological assessment has not previously been conducted on children with MCADD. Reported results are based on clinical observations only and most follow-up is limited to younger children. There is a clear need for more comprehensive neuropsychological assessment of children with MCADD to delineate their neuropsychological profile.

Developmental issues are also relevant. Occasionally, in cases of early brain disease or injury there is sparing of function as the brain develops. Neural plasticity, however, tends to be the exception rather than the rule and much of the literature in this area indicates that the younger a child is when their brain insult is sustained, the more likely it is that their developmental outcome will be negatively affected (Taylor & Alden, 1997). This reflects not only the extent to which early damage to the brain can immediately affect outcome but also how it interrupts the further development of the brain and acquisition of skills. In addition, developmental constraints limit the range of skills that can be reliably tested in young children, with many higher level executive functions not fully developed until late adolescence. Subtle or specific problems may be effectively hidden during testing at a younger age and only become apparent in

older children when we expect to see acquisition of higher level processing. Thus long-term follow-up is required to determine the full impact of an early brain insult on later cognitive outcome.

Hypoglycemic episodes are cited frequently in the MCADD population and these may have a specific impact on neuropsychological function. MCADD is a disorder of beta (β) oxidation of fatty acids (Roe & Coates, 1997). Ketones, an alternative to cerebral fuels, e.g., glucose, are a product of β oxidation of fatty acids. Hypoglycemia occurs when blood glucose levels drop, for example during fasting. Healthy babies and children are able to produce ketone bodies when blood glucose concentrations are low, thus avoiding serious consequences of hypoglycemia. The inability of MCADD patients to mount a response to hypoglycemia places them at greater risk than normal children to suffer the serious neurocognitive impairments that can result from severe episodes of hypoglycemia (Hawdon, 1999). Follow-up studies of patients with diabetes who also have an increased risk of hypoglycemia have found memory deficits (Anderson, Northam, Wrennall, & Hendy, 2001; Chalmers et al., 1991; Hannonen, Tupola, Ahonen, & Riikonen, 2003; Kovacs, Ryan, & Obrosky, 1994), problems with processing speed, and executive function deficits (Anderson et al.). Whether these cognitive deficits are also present in MCADD has not yet been explored.

In this study, we have compared outcomes in children who have MCADD, diagnosed either by newborn screening or clinically. We wished to examine whether the condition has any implications for cognitive functioning and, if so, if that impact is global or specific. We then aimed to explore whether there is a difference in cognitive outcome for the screened versus unscreened groups. We investigated general intellectual functioning and adaptive behavior and also explored whether the groups differed on more specific aspects of cognition and adaptive functioning. It was also of interest to determine whether there is a relationship between age at diagnosis of MCADD and cognitive outcome or a relationship between episodes of hospitalization and outcome. In addition, the extent to which parental attitudes and coping might influence hospital admission and outcome were examined. A very brief report of the neuropsychological outcome of the cohort we studied has been published as part of a report on medical outcomes (Wilcken et al., 2007) and here we present the detailed findings.

METHODS

Participants

All children born in Australia between the years 1994 and 2002 who had been diagnosed with MCADD deficiency were identified from a list of participants in a larger study evaluating the effectiveness of expanded newborn screening. Parents of potential participants were mailed information packages inviting their children to participate in the study. Consent for participation was obtained by a follow–up telephone call and arrangements were made for clinic and assessment appointments.

Fifty-nine children with MCADD were identified: 35 unscreened (25 diagnosed clinically, 10 by family study), and 24 diagnosed by newborn screening. Seven of the children had died, 4 parents declined to have their children tested, and 10 could not be tested as they lived at a great distance or could not be contacted.

Clearance to conduct this study was obtained from The Children's Hospital at Westmead's Human Research Ethics Committee.

Materials

Woodcock Johnson-III. The Woodcock Johnson-III (WJIII; Woodcock, McGrew, & Mather, 2001) consists of two batteries: the WJIII Tests of Cognitive Abilities and the WJIII Tests of Achievement. The tests of cognitive abilities are specifically designed to measure a wide range of narrow and broad cognitive functions and to provide a valid system for evaluating domain specific skills. The normative sample is large (8818 participants) and the test has excellent reliability ($r = .97$ for the measure General Intellectual Ability [GIA]). The WJIII can be administered to a wide age range, from 2 to 90 years allowing participants of different ages to be assessed on the same test, thereby eliminating the confounding issue of different test measures.

The test protocol for this study consisted of 13 subtests from the WJIII Tests of Cognitive Abilities and 4 subtests from the WJIII Tests of Achievement. Each subtest provides a measure of a specific narrow ability, which can then be combined to form factor scores or more broad ability clusters. Participants aged 4 and above were administered the WJIII. Not all subtests can be administered to all ages: 2 cognitive subtests (Numbers Reversed and Pair Cancellation) were omitted for the 4-year-olds and achievement tests were only administered to children 5 years and above. As less than half the screened group completed the achievement tests only the results of the cognitive tests will be reported in this paper.

Vineland Adaptive Behaviour Scale – Interview Edition. The Vineland Adaptive Behavior Scale (VABS; Sparrow, Balla, & Cicchetti, 1984) is a well-recognized standardized scale utilized to measure adaptive functioning in children and adolescents in a variety of settings. It is administered to parents or caregivers using a semistructured interview format and is normed for individuals from birth to 18 years 11 months. It has high reliability (test-retest and interrater reliability coefficients .99 and .98, respectively) and good construct validity. The content and scales are organized into four domains of adaptive functioning: communication skills (expressive, receptive, and written), daily living skills (personal, domestic, and community), socialization skills (interpersonal relationships, play and leisure time, and coping skills), and motor skills (gross and fine motor). Standard scores (mean = 100, standard deviations = 15) are derived for each scale. As the motor skills domain is only administered to children under the age of 6 years, only the results of the three domains that are common across all age groups will be discussed.

Parenting Stress Index, 3rd Edition, Short Form. The Parenting Stress Index/Short Form (PSI/SF) is a derivative of the full-length Parenting Stress Index (PSI, 3rd Edition) (Abidin, 1995). The PSI is an inventory completed by parents that examines stressful areas in parent-child interactions. It was standardized to use with parents of children ranging in age from 1 month to 12 years (in fathers only from 6 months to 6 years). The PSI/SF contains 36 items that are divided into three subscales (Parental Distress, Difficult Child, and Parent-Child Dysfunctional Interaction) and all contribute to a Total Stress score. Percentile ranks are given for subscale and total scores. The PSI/SF has demonstrated both test-retest and internal consistency reliability and is likely to share the validity of the full-length PSI, which has been extensively validated.

Procedure

Recruited participants aged 4 years and above were individually administered the WJIII at the Children's Hospital at Westmead, their local hospital, or at home. Parents

were then interviewed using the Vineland Adaptive Behavior Scale and completed the PSI/SF along with other questionnaires. Data from the WJIII, VABS, and PSI/SF are reported here. Several participants were unable to attend a testing session or were too young to be administered the WJIII. In these cases a phone interview was conducted to complete the Vineland Adaptive Behavior Scales. Medical records were reviewed for each participant and data obtained on number of hospitalizations.

Data Analysis

Descriptive statistics were used to document age and gender of the study sample.[1] Cognitive and adaptive outcomes for the total, screened, and unscreened groups were examined with a series of independent samples *t*-tests. As multiple tests were carried out, we utilized the False Discovery Rate (FDR; Benjamini & Hochberg, 1995) to control the Type 1 error rate. MCADD is a rare and therefore low incident disorder, and thus only small numbers of participants with the condition were available for study. This significantly limited the power of the study to detect group differences and increased the risk of making a Type 2 error. To address this we calculated effect sizes for all comparisons to determine the extent and size of differences between the groups. Based on Cohen's (1988) classification we took medium and large effect sizes as an indication of meaningful differences. In interpreting these results however, it should be borne in mind that the present sample of all children available for testing represents a significant proportion of the surviving population (25/52) of all children born in Australia with this condition from April 1994–April 2002. Therefore we feel more weight can be placed on obtained differences than the absolute size of the sample would suggest. McNemar's test of correlated proportions was used to examine the variability of performance for overall cognitive outcome and adaptive behavior for the whole group. The relationship between age at diagnosis and cognitive and adaptive behavior outcome was investigated using a Pearson product-moment correlation coefficient. Independent samples *t*-tests were also used to examine the difference in means between an early- and a later-diagnosis group. The effect of hospital admissions on outcome was first analyzed using a one-way between groups ANOVA with three groups: no admissions, 1–2 admissions, and greater than 2 admissions. Due to the large variation in the greater than two admissions group, hospital admission data was also analyzed using Pearson correlations. Lastly, in order to examine the possibility that hospital admissions may be related to parenting practices we used Pearson correlations to examine the relationship between PSI scores and number of admissions and MANOVA to check for admission group differences in PSI scores.

RESULTS

The sample was comprised of 38 children who had been diagnosed with MCADD. Five participants (3 unscreened) did not have mutation analysis. Of the remainder, all

[1]Socioeconomic status (SES) measures were not included. The study sample represents a significant proportion of the surviving population (25/52) of all children born in Australia with this condition from April 1994 to April 2002. MSMS screening is a population-based screening program available universally to all newborns, and any SES differences in the unscreened group are unlikely due to the fact that children who are diagnosed with MCADD clinically present with severe, if not life-threatening, illness.

Table 1 Number and Age of Participants and Available Test Data for Total, Screened, and Unscreened Groups

	Screened	Unscreened	Total
Cognitive Data			
N	13	12	25
Age in Months			
M	55	87.75	70.72
SD	8.40	22.67	23.44
Gender (M/F)	9/4	4/8	13/12
Adaptive Data			
N	19	19	38
Age in Months			
M	53.26	88	70.63
SD	9.62	23.86	25.13
Gender (M/F)	11/8	6/13	17/21

carried at least one copy of the common ACADM (acyl-coenzyme A dehydrogenase, C-4 to C-12 straight chain) mutation c.A>G985, which accounted for 76% of alleles in the screened and 90% of alleles in the unscreened groups, but this difference was not significant. Therefore, further investigation of the relationship between the ACADM mutation and cognitive outcome was not undertaken.

The ages of the children ranged from 2 years 2 months to 10 years 3 months. Of these children 19 were clinically diagnosed (unscreened group) and 19 were diagnosed by MSMS (screened group). Table 1 shows descriptive information for the two groups and whole sample. An independent samples *t*-test indicated that for the total sample tested the groups differed significantly on age, $t(36) = 5.886$, $p < .001$, but not gender, $t(36) = 1.65$, $p = .108$. Cognitive testing was not completed for 13 patients. Twelve lived several hundred kilometers from the testing center or were too young for formal testing and one refused to speak. Comparison of the number of children with and without formal cognitive testing revealed no significant differences with respect to age, age at diagnosis, or adaptive ratings.

Cognitive Outcome for Total, Screened, and Unscreened Groups

Means and standard deviations for WJIII factor scores and supplemental tests are presented in Table 2. It should be noted that Cognitive Efficiency was omitted from factor score analysis. This factor is comprised of two subtests, Visual Matching and Numbers Reversed. Only a small number ($n = 5$) of our young participants were able to complete Numbers Reversed, making calculation of the factor score inappropriate. Results of Visual Matching along with the six supplement tests were analyzed separately. Scores on GIA, verbal ability, and thinking ability for the total participant group were in the average range. Examination of individual scores revealed that only two children (8%) returned scores in the range lower than one standard deviation below the mean on the GIA, Verbal Ability, and Thinking Ability measures; a comparatively smaller percentage than we would expect to see based on population norms.

Results of independent *t*-test analysis indicated no significant differences between screened and unscreened groups on GIA, $t(23) = .245$, $p = .809$, Verbal Ability, $t(23) = -1.302$, $p = .206$, and Thinking Ability, $t(23) = .155$, $p = .878$. Despite

Table 2 WJIII Outcome Scores for Factor Composites and Supplementary Tests for Total Sample, Screened and Unscreened Participants

	Total (n = 25)	Screened (n = 13)	Unscreened (n = 12)	Cohen's d
Factor				
GIA (IQ)				
M	104.24	103.62	104.92	0.10
SD	13.03	11.75	14.79	
Verbal Ability				
M	107.36	111.00	103.42	0.54*
SD	14.76	12.19	16.75	
Thinking Ability				
M	103.8	103.38	104.25	0.06
SD	13.69	12.65	15.29	
Supplementary Tests				
Visual Matching				
M	101.00[a]	104.75[b]	97.25	0.75**
SD	10.96	10.41	10.6	
Numbers Reversed				
M	110.25[c]	118.80[d]	106.36[e]	0.65*
SD	19.16	22.01	17.43	
Auditory Working Memory				
M	105.32[f]	104.91[e]	105.73[e]	0.05
SD	17.15	13.90	20.59	
Retrieval Fluency				
M	101.92	101.46	102.42	0.10
SD	9.81	9.43	10.60	
Decision Speed				
M	104.52[g]	102.90[h]	106.00[e]	0.27
SD	11.97	14.60	9.48	
Rapid Picture Naming				
M	95.56	99.38	91.42	.66*
SD	12.88	13.99	10.62	
Planning				
M	116.74[i]	124.63[j]	111.00[e]	1.7**
SD	10.66	7.65	8.8	
Pair Cancellation				
M	103.54[k]	107.00[l]	102.50[h]	0.38
SD	10.56	12.29	10.49	

[a]n = 24. [b]n = 12. [c]n = 16. [d]n = 5. [e]n = 11. [f]n = 22. [g]n = 2. [h]n = 10. [i]n = 19. [j]n = 8. [k]n = 13. [l]n = 3.
*moderate effect size. **large effect size.

the fact that significance was not reached for verbal ability the screened group's scores were higher than the unscreened group with moderate effect size found (Cohen's d = 0.54). This contrasted with small effect sizes (Cohen's d = 0.10 and 0.06) for GIA and Thinking Ability, respectively.

A significant difference was found between the screened and unscreened groups on the Planning subtest, $t(17) = -3.514$, $p = .003$, using the FDR adjusted alpha value of .05, with the screened group having a significantly higher mean score than the unscreened

group. Group differences between both the Visual Matching, $t(22) = -1.749$, $p = .094$, and Rapid Picture Naming, $t(23) = -1.594$, $p = .125$, did not reach a level of statistical significance. Nonetheless, the effect sizes calculated were large (Cohen's $d = 0.75$) and moderate (Cohen's $d = 0.66$) respectively, indicating that differences may exist that were undetectable due to our small sample size. All other subtest comparisons yielded nonsignificant findings and small effect sizes.

Adaptive Behavior Outcome for Total, Screened, and Unscreened Groups

Mean scores and standard deviations for adaptive behavior measures for the total, screened, and unscreened groups are reported in Table 3. Mean scores for the total group were all in the average range; however 10.5% of our sample (4/38) returned scores at or below one standard deviation below the mean for the Communication domain, 5.3% (2/38) for Socialization, 23.7% (9/38) for Daily Living Skills, and 18.4% (7/38) for the Adaptive Behavior Composite. The proportion of participants scoring below one standard deviation on Daily Living Skills and the Adaptive Behavior Composite is larger than would be predicted from the normal population.

Between-Group Comparison for Adaptive Behavior

No significant differences were found between the screened and unscreened groups on any of the Vineland Adaptive Behavior Scales variables.

Cognitive Outcome and Adaptive Behavior Performance Variability

McNemar's test of correlated proportions was used to analyze the spread of performance for the total group on cognitive outcome and on adaptive behavior. A nonsignificant result was obtained ($p = .63$) that indicates that there is no difference in the proportion of participants whose scores are more than one standard deviation below the mean on WJIII GIA and VABS Adaptive Behavior.

Table 3 VABS Outcome Scores for Total Sample, Screened and Unscreened Participants

	Total ($n = 38$)	Screened ($n = 19$)	Unscreened ($n = 19$)	d
Communication				
M	103.13	103.37	102.89	0.03
SD	13.92	15.05	13.10	
Daily Living Skills				
M	93.24	94.58	91.89	0.18
SD	15.45	17.64	13.27	
Socialization				
M	102.47	102.68	102.26	0.04
SD	11.37	13.93	8.48	
Adaptive Behavior Composite				
M	99.97	101.58	98.37	0.20
SD	16.14	17.80	14.62	

Relationship Between Age at Diagnosis and Cognitive and Adaptive Behavior Outcome

We investigated the possibility of a relationship between age at diagnosis of MCADD and cognitive outcome (general intellectual ability, verbal ability, and thinking ability) and age at diagnosis and adaptive behavior outcome (communication, daily living skills, socialization, and overall adaptive behavior) using Pearson correlations. For the whole group there were no significant correlations found between age at diagnosis and any of the WJIII index or VABS measures. Analysis was also performed using only the unscreened group who had a larger spread of age at diagnosis, which also produced no significant results. Additionally, we compared those with early diagnosis (within the first month of life) to those whose diagnoses came later in life. No significant differences were found between the early- and late-diagnosis groups on any of the VABS outcome measures. In terms of cognitive outcome, there were no significant differences between the groups; however the difference in Verbal Ability approached significance, $t(23) = 1.878$, $p = .073$, with those diagnosed at less than 1 month of age having a higher mean score than those diagnosed at greater than 1 month of age. A large effect size was found (Cohen's $d = 0.78$) indicating an outcome that is worthy of consideration despite not reaching significance.

Relationship Between Cognitive and Adaptive Behaviour Outcome and Number of Hospital Admissions

A one-way between-groups analysis of variance was conducted to explore the impact of number of admissions to hospital in the first 2 years of life on WJIII factors and VABS domains. Hospital admission data was available for 22 of the 25 participants who were tested on the WJIII and for 31 of the 38 participants for whom VABS data were collected. Participants were divided into three groups according to the number of admissions (Group 1: no admissions; Group 2: 1–2 admissions; Group 3: 2+ admissions). There were no significant differences between the groups on cognitive outcomes. There was a statistically significant difference between groups on VABS Socialization scores, $F(2, 28) = 4.17$, $p = .026$. Post-hoc comparisons using the Tukey HSD test indicated that for VABS Socialization the mean score for children with more than two admissions during the first 2 years of life was significantly higher than the scores of children with no admissions ($p = .034$) and one or two admissions only ($p = .031$). Groups 1 and 2 were not significantly different from each other. Table 4 shows the mean and standard deviations of the VABS Socialization scores in each of these groups.

The relationship between continuous number of admissions in the first 2 years of life and main WJIII outcome variables and VABS domains was examined using Pearson product-moment correlation coefficients. A significant positive correlation was found between the number of admissions a child has in the first 2 years of life and score on the Socialization domain of the VABS ($r = .420$, $n = 31$, $p = .019$).

Table 4 Mean Scores for VABS Socialization by Number of Hospital Admissions

	Group 1 (no admissions)		Group 2 (1–2 admissions)		Group 3 (2 + admissions)	
	M	*SD*	*M*	*SD*	*M*	*SD*
VABS Socialization	101.59	9.23	99.67	16.77	113.00	16.69

Table 5 Mean Scores for PSI by Number of Hospital Admissions

	Number of Admissions		
	None	1–2	2 or more
Parental Distress			
M	20.93	26.29	22.80
SD	4.99	7.37	4.76
Total Stress			
M	70.33	78.00	71.00
SD	16.79	27.99	19.95

We used data gathered from parents on the PSI to investigate the possibility that children with more admissions may have a parent who reports higher levels of parenting-related stress. There was no significant correlation between either the number of admissions in the first 2 years of a child's life or total admissions and the Total Stress score or Parental Distress factor from the PSI. Similarly, there were no significant differences between the mean scores (displayed in Table 5) of the three admission groups on these PSI measures.

DISCUSSION

Cognitive Functioning

This study examined a group of children with MCADD diagnosed in one of two ways; either clinically or by tandem mass spectrometry as part of newborn screening. Overall the condition did not appear to have an impact on general cognitive functioning with the cohort showing average intellectual ability. There were no differences between the screened and the unscreened group at the level of basic intellectual skills. The literature, whilst scarce in terms of formal neuropsychological testing, clearly indicates that this is a population at certain risk for developmental delay. However, the incidence of intellectual impairment we found was no higher than that found in normal populations. This could indicate more rapid and appropriate triage and treatment in Australia now, compared to some years ago.

Despite overall good intellectual outcome we did find evidence of differences in specific abilities between screened and unscreened participants. We found the verbal ability of the screened group to be superior to that of the unscreened group (with a moderate effect size). However the trend for higher/above average performance in the younger screened group, in addition to the finding of only average performance in the older unscreened group, may be attributable to the increasing task complexity at different ages. Indeed, among the four tasks that comprise the verbal measure in the WJIII, the scores of the younger participants tend to be more heavily weighted on their performance on picture vocabulary and naming tasks, while the tasks undertaken by older children require more verbal-reasoning skills. However, due to the way these tests are scored it is not possible to distinguish between individual task scores. While the study protocol did not allow us to assess memory functions in detail, there was no evidence of differences between the groups on two tasks (Visual Auditory Learning within the Thinking Ability Factor and Retrieval Fluency) that measured long-term retrieval of

information. This would suggest that retention and retrieval of information is comparable in both groups at least in the early years. However, memory testing at this age is still limited and these results do not exclude the presence of specific memory impairments. As these children are at increased risk for episodes of hypoglycemia, there is a potential for selective memory impairments and this will need to be explored in more detail at an older age.

We also found that the screened group performed better on the Planning, Visual Matching, and Rapid Picture Naming subtests. The latter two of these tests are performed under time pressure tapping into processing speed and all of these tests can be characterized as part of the repertoire of executive abilities an individual develops through early childhood to late adolescence. Research such as that by Welsh, Pennington, and Groisser (1991) identifies distinct developmental stages beginning to develop at age 6 such as the ability to resist distraction, and continuing development through to the age of 10 with the emergence of organized search, hypothesis testing, and impulse control. Some executive functions, for example, verbal fluency, motor sequencing, and planning, have also been found to only just begin to reach adult levels at age 12 (Welsh et al., 1991). Although there is no evidence of deficits in the unscreened group, the reduced performance relative to the screened group may suggest differences in executive functioning. As these functions are still developing throughout childhood and adolescence it may be that difficulties in more specific components are unable to be fully characterized until late childhood to early adolescence. Of direct relevance to our results is the implication that we cannot expect to see mastery of these skills until early adolescence. If indeed upon further follow-up we found differences between our groups, it would be consistent with the research indicating that executive functioning and processing speed may be impacted upon by episodes of hypoglycemia (Deary, Sommerfield, McAulay, & Frier, 2003; Hannonen et al., 2003). This would be an important finding as the majority of our unscreened group were diagnosed upon clinical presentation usually characterized by hypoglycemia leading to lethargy and even coma.

Adaptive Functioning

Adaptive functioning was relatively intact and there were no significant differences between groups. However, the daily living skills of the whole group were reduced overall (low average range) compared to the other domains measured and 23.7% of our participants returned scores that were more than one standard deviation below the population mean (compared to a normal sample prediction of 16%). This relative discrepancy between daily living skills and other areas adaptive functions was unexpected and at odds with the standardized data and difficult to interpret. It may be that these children have less practice at performing self-help skills (e.g., dressing, toileting, eating) if having a child with a disorder leads parents to provide a higher level of support for these activities. Alternatively, this may reflect a change in societal expectations and perhaps also be an indirect impact of the increase in the number of households with two working parents. Anecdotally many parents in this study reported that their children probably possess the skills required to attain normal levels of functioning but that leading busy lifestyles led to them doing many things for their children to save time.

Diagnosis and Complications

Another surprising finding was that children with more hospital admissions in the first 2 years of life had more advanced socialization and communication skills than those

with fewer than 2 or no admissions. In the same manner as previously discussed, these results highlight the potential importance of environmental factors and possibly reflect the effect of high levels of adult contact on the child's development.

Those who were diagnosed by screening, by broad definition were picked up very early in life before there was any chance for them to suffer any illness and subsequent neurological damage. However, it is also the case that some unscreened patients can present very early in life and are diagnosed and treated early. Our results indicate that those diagnosed within the first month of life exhibited higher mean Verbal Ability scores than those diagnosed later in life regardless of their screening status. This highlights the critical importance of early diagnosis and management in MCADD. This result also provides strong support for continued screening as the only way to increase an infant's chance of being diagnosed and managed early in life is via a program that screens for MCADD.

In conclusion, this study has served the purpose of beginning to describe the neuropsychological functioning of a group of children with MCADD beyond the broad classifications of being impaired versus unimpaired using formal neuropsychological testing, as opposed to clinical observations. Very little difference was seen in overall cognitive functioning in our sample; however, as noted, participant numbers for some of our measures were necessarily small due to the age range of children being tested. As such results should be interpreted with caution. We found little evidence of impairment overall but some suggestion of the presence of differences in specific skills. Of particular interest are the group's emerging executive skills and some differences that may be indicative of problems that may occur further along in development. These are preliminary findings, however, and long-term follow-up is needed to track outcomes and to adequately assess specific skills and executive functioning over time. Early diagnosis and management for children with MCADD appear to lead to better outcomes that will only become more apparent as development and research continues.

Original manuscript received August 23, 2007
Revised manuscript accepted December 13, 2007
First published online Month day, year

REFERENCES

Abidin, R. R. (1995). *Parenting Stress Index* (3rd ed.). Lutz, FL: Psychological Assessment Resources, Inc.

Anderson, V., Northam, E., Wrennall, J., & Hendy, J. (2001). *Pediatric neuropsychology: A clinical approach.* London: Erlbaum.

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society, Series B, 57,* 289–300.

Carpenter, K., Wiley, V., Sim, K. G., Heath, D., & Wilcken, B. (2001). Evaluation of newborn screening for medium chain acyl-CoA dehydrogenase deficiency in 275 000 babies. *Archives of Disease in Childhood Fetal and Neonatal Edition, 85,* 105–109.

Chalmers, J., Risk, M. T., Kean, D. M., Grant, R., Ashworth, B., & Campbell, I. W. (1991). Severe amnesia after hypoglycemia. Clinical, psychometric, and magnetic resonance imaging correlations. *Diabetes Care, 14,* 922–925.

Cohen, J. (1988). *Statistical power analysis for the behavioural sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.

Deary, I. J., Sommerfield, A. J., McAulay, V., & Frier, B. M. (2003). Moderate hypoglycaemia obliterates working memory in humans with and without insulin treated diabetes. *Journal of Neurology, Neurosurgery & Psychiatry, 74,* 278–279.

Hannonen, R., Tupola, S., Ahonen, T., & Riikonen, R. (2003). Neurocognitive functioning in children with type-1 diabetes with and without episodes of severe hypoglycaemia. *Developmental Medicine & Child Neurology, 45,* 262–268.

Hawdon, J. M. (1999). Hypoglycaemia and the neonatal brain. *European Journal of Pediatrics, 158,* S9– S12.

Iafolla, A. K., Thompson, R. J., & Roe, C. R. (1994). Medium-chain acyl-coenzyme A dehydrogenase deficiency: Clinical course in 120 affected children. *The Journal of Pediatrics, 124,* 409–415.

Kovacs, M., Ryan, C., & Obrosky, D. S. (1994). Verbal intellectual and verbal memory performance of youths with childhood-onset insulin-dependent diabetes mellitus. *Journal of Pediatric Psychology, 19,* 475–483.

Leonard, J. V., & Dezateux, C. (2002). Screening for inherited metabolic disease in newborn infants using tandem mass spectrometry. *BMJ (British Medical Journal), 324,* 4–5.

Pollitt, R. J., & Leonard, J. V. (1998). Prospective surveillance study of medium chain acyl-CoA dehydrogenase deficiency in the UK. *Archives of Disease in Childhood, 79,* 116–119.

Roe, C., & Coates, P. (1997). Mitochondrial fatty acid oxidation disorders. In C. Scriver, A. Beaudet, W. Sly, & D. Valle (Eds.), *The metabolic and molecular basis of inherited disease.* (pp. 1501–1534). New York: McGraw-Hill.

Sparrow, S. S., Balla, D. A., & Cicchetti, D. V. (1984). *Vineland Adaptive Behaviour Scales: Interview Edition, Survey Form.* Circle Pines, MN: AGS.

Taylor, H. G., & Alden, J. (1997). Age-related differences in outcomes following childhood brain insults: An introduction and overview. *Journal of the International Neuropsychological Society, 3,* 555– 567.

Touma, E. H., & Charpentier, C. (1992). Medium chain acyl-CoA dehydrogenase deficiency. *Archives of Disease in Childhood, 67,* 142–145.

Welsh, M. C., Pennington, B. F., & Groisser, D. B. (1991). A normative-developmental study of executive function: A window on prefrontal function in children. *Developmental Neuropsychology, 7,* 131–149.

Wilcken, B., Carpenter, K. C., & Hammond, J. (1993). Neonatal symptoms in medium chain acyl coenzyme A dehydrogenase deficiency. *Archives of Disease in Childhood, 69,* 292–294.

Wilcken, B., Haas, M., Joy, P., Wiley, V., Chaplin, M., Black, C., et al. (2007). Outcome of neonatal screening for medium-chain acyl-CoA dehydrogenase deficiency in Australia: A cohort study. *The Lancet, 369,* 37–42.

Wilcken, B., Hammond, J., & Silink, M. (1994). Morbidity and mortality in medium chain acyl coenzyme A dehydrogenase deficiency. *Archives of Disease in Childhood, 70,* 410–412.

Wilson, C. J., Champion, M. P., Collins, J. E., Clayton, P. T., & Leonard, J. V. (1999). Outcome of medium chain acyl-CoA dehydrogenase deficiency after diagnosis. *Archives of Disease in Childhood, 80,* 459–462.

Woodcock, R. W., McGrew, K. S., & Mather, N. (2001). *Woodcock-Johnson III.* Itasca, IL: Riverside Publishing.

J Inherit Metab Dis
DOI 10.1007/s10545-008-0912-z

ORIGINAL ARTICLE

Glutaric aciduria type I: outcome following detection by newborn screening

S. Bijarnia · V. Wiley · K. Carpenter · J. Christodoulou · C. J. Ellaway · B. Wilcken

Received: 13 March 2008 / Submitted in revised form: 23 April 2008 / Accepted: 29 April 2008 / Published online: 9 August 2008

Summary Glutaric aciduria type I (GA I), a cerebral organic acidaemia with the potential for severe neurological consequences, can now be detected by tandem mass spectrometry newborn screening. Early detection with implementation of careful management strategies appears to lessen the likelihood of neurological damage. We assessed the outcome in all 10 GA I patients detected in New South Wales during the last decade. Three patients were detected clinically and 7 by newborn screening. Diagnosis was confirmed by detection of significantly elevated urinary 3-hydroxybutyrate and glutarate in urine, isolated elevation of glutarylcarnitine in plasma, typical clinical and MRI findings in several, and mutation analysis or enzyme analysis on cultured skin fibroblasts in 4 cases. The birth frequency was 1:90 000. Following diagnosis, treatment was initiated in all children with oral carnitine (100 mg/kg per day) and a low-protein diet supplemented with a lysine-free, low-tryptophan amino acid formula. Disability was assessed in fields of motor, cognitive and speech development and scored according to Kyllerman. Clinically diagnosed patients were all symptomatic, with severity scores (out of 9) of 3, 5 and 9. Six of seven patients detected by newborn screening are asymptomatic, 4 being aged 2–6 years. One patient had a severe decompensation at 7 months, despite full management advice and treatment, and later died. Our data support previous findings that early diagnosis reduces neurological complications, but show that even with early diagnosis and careful management severe complications may ensue in some.

Communicating editor: Georg Hoffman

Competing interests: None declared

References to electronic databases: Glutaric aciduria type I (synonym glutaryl-CoA dehydrogenase deficiency): OMIM # 231670. Glutaryl-CoA dehydrogenase: EC 1.3.99.7.

S. Bijarnia · J. Christodoulou · C. J. Ellaway · B. Wilcken
Genetic Metabolic Disorders Service,
The Children's Hospital at Westmead,
Sydney, NSW, Australia

V. Wiley · B. Wilcken
NSW Newborn Screening Programme,
The Children's Hospital at Westmead,
Sydney, NSW, Australia

K. Carpenter · B. Wilcken
NSW Biochemical Genetics Service,
The Children's Hospital at Westmead,
Sydney, NSW, Australia

S. Bijarnia · V. Wiley · K. Carpenter · J. Christodoulou · C. J. Ellaway · B. Wilcken
Disciplines of Genetic Medicine and Paediatrics & Child Health, University of Sydney,
Sydney, NSW, Australia

B. Wilcken (✉)
NSW Biochemical Genetics Service and
NSW Newborn Screening Programme,
The Children's Hospital at Westmead,
Locked Bag 4001, Westmead, NSW 2145, Australia
e-mail: bridgetw@chw.edu.au

Abbreviations

CT	computed tomography
GA I	glutaric aciduria type I
GCDH	glutaryl Co-A dehydrogenase deficiency
GC-MS	gas chromatography–mass spectrometry
IQ	intelligence quotient
MRI	magnetic resonance imaging
NBS	newborn screening
NSW	New South Wales

Introduction

Glutaric Aciduria type I (glutaryl co-enzyme A dehydrogenase deficiency, GA I, OMIM # 231670), first described in 1975, is a cerebral organic aciduria resulting from an enzymatic block in the final common degradation pathway for L-lysine, L-hydroxylysine and L-tryptophan (Goodman et al 1975), due to a deficiency of glutaryl-CoA dehydrogenase (GCDH, EC 1.3.99.7) (Christensen 1983). GA I shows autosomal recessive inheritance. More than 150 pathogenic mutations have been identified in the *GCDH* gene on chromosome 19p13.2 (Kölker et al 2006).

GA I is characterized clinically by dystonia and dyskinesia appearing during the first years of life, chemically by excretion of glutaric and 3-hydroxyglutaric acids in urine, and pathologically by neuronal degeneration of the caudate and putamen (Goodman and Freeman 2001). CT and MRI scans often show frontotemporal atrophy and/or arachnoid cysts before the onset of symptoms (Goodman and Freeman 2001). The striatal damage usually occurs suddenly, during an intercurrent illness, although there can be an insidious onset (Strauss et al 2003). Presymptomatic treatment appears to be effective to some extent in preventing such an irreversible brain damage (Kölker et al 2007b; Naughten et al 2004; Strauss et al 2003), so timely diagnosis, with early initiation of therapy before the onset of neurological disease, is vital. Newborn screening provides an apt solution, if it is followed by intervention with carnitine supplementation and dietary restriction of lysine and tryptophan (Kölker et al 2006, 2007b). In addition, vigorous treatment of intercurrent illnesses with fluids and adequate calories is an essential part of management (Kölker et al 2007a).

In New South Wales (NSW), Australia, newborn screening for GA I started in 1999 using tandem mass spectrometry. We present the follow-up data on all children with GA I born in the past decade.

Methods

All children with GA I born in NSW since 1998 were studied. All children diagnosed both symptomatically and by newborn screening were included in the study. Newborn screening (NBS) is performed in NSW by electrospray tandem mass spectrometry. Blood spot glutarylcarnitine levels of 0.3 μmol/L or greater suggested the possibility of GA I, and immediate confirmatory testing was achieved by urinary organic acid analysis by gas chromatography–mass spectrometry (GC-MS) on a urine sample, based on increased 3-hydroxyglutaric and glutaric acids, as well as plasma acylcarnitine analysis. Ratios of other analytes to glutarylcarnitine were not used as a second-tier strategy, but a low carnitine level in any sample triggered automatic review of the profile, with possible request for a plasma sample if the profile indicated this. Glutaryl-CoA dehydrogenase enzyme assay was performed in two patients and mutation analysis in the *GCDH* gene was performed in 5 patients.

Following diagnosis, treatment was initiated in all children, including those diagnosed by screening, with oral carnitine (100 mg/kg per day) and a low-protein diet with supplementation with special amino acid formula free of lysine and low in tryptophan (XLys LowTry Amino acid mixtures, SHS International Ltd, Liverpool, UK). Very recently, the formula has been free of tryptophan. All children were provided with an individually tailored unwell emergency treatment regime which included a diet high in calories and free of natural protein, and instructions to contact the metabolic doctor on call at our hospital.

The children were followed up regularly in a metabolic clinic. Developmental assessment was performed in all in the last six months. Disability was assessed in fields of motor, cognitive and speech development and scored according to Kyllerman and colleagues (2004):

A Motor disability was scored as 1 (mild) in the case of no or mild motor dysfunction and no disability in daily life; 2 (moderate) in case of mild to moderate motor dysfunction and some, but not limiting, disability in daily life; and as 3 (severe) in the case of wheelchair dependency and severe disability in daily life.

B Cognitive function was judged from data derived from clinical records of development, normal or special school education and formal IQ tests. A score of 1 was given when considered normal, 2 when intermediate, and 3 when IQ level was below 70.

C Speech and language were assessed and scored as 1 when fluent or nearly fluent, as 2 when only single words could be expressed, and as 3 when there was no speech.

The sum of scores for all the above was considered a measure of outcome, with the minimum score being 3 (normal or near normal outcome) and maximum of 9 (poor outcome).

Results

Ten children were diagnosed with GA I in the last decade in nine families. The birth prevalence was

Table 1 Summary of individual cases and outcome

Subject	DOB/Sex	Age at diagnosis	Acute crises	Motor disability score	Cognitive disability score	Speech disability score	Maximum head circumference as centile, with corresponding age	Mutation analysis	MRI findings (age performed)
Case 1	13/04/98 Male	8 months	Head injury at 8 months (no neurological signs, MRI +ve) Acute ataxia with gastrointestinal illness at 6 years of age, full recovery	1 Very mild choreoathetosis, clumsy gait, hypotonia	1 Year 3 in mainstream school	1 Normal	95th centile 9 months	–	Cerebral atrophy. Subtle high signal putamen (8 months)
Case 2	10/02/00 Male	10 months	Nil. Insidious onset	3 Choreoathetosis, wheelchair-dependent, disability unit at school	1 Year 1, intellectual abilities normal	1 Dysarthria, can express well	25th centile 12 months	p.R227P, c.1244-2A>C (3% enzyme activity)	–
Case 3	27/09/00 Female	2.5 years	Two episodes: first at 8 months severe; second at 2.5 years	3 Very hypotonic, wheelchair dependent	2 Responds with appropriate facial expressions	3 Absent speech	50th–75th centile 8 months	p.R128Q and 852+5G>A (10% enzyme activity)	Diffuse abnormal signal bilateral basal ganglia (8 months)
Case 4	14/04/02 Female	NBS	Nil. Hospital admissions 10 times (6 in first 2 years)	1 Normal	1 Normal, in pre-school	1 Normal	90th–95th centile 5 months	–	Cerebral atrophy, increase signal intensity globus pallidus (3 months)
Case 5	22/03/04 Female	NBS	Nil, hospital admissions 4	1 Normal	1 Normal	1 Normal	90th–95th centile 5 months	p.G425V, p.R227P	Dilated extra axial spaces (2 months)
Case 6	11/12/05 Male	NBS	Severe neurological crises 7 months; never recovered—died 13 months	3 Severe dystonia, seizures	3 No recognition	3 No speech	95th centile 4 months	Homozygous p.R383C	Severe changes—atrophy basal ganglia, diffuse white-matter changes (7 months, 12 months)
Case 7	16/01/06 Male	NBS	Mild, 16 months–improved. Hospital admissions 5	1 Minimal dyskinesia; mild motor delay, with recovery	1	1	50th centile 6 months	–	–
Case 8	13/04/06 Female	NBS	Nil	1 Normal development	1	1	97th centile 7 months	–	–
Case 9	12/03/07 Male	NBS	Nil. Hospital admissions 2	1	1	1	90th centile 7 months	–	–
Case 10	7/07/07 Female	NBS	Nil	1 Very mild upper limb dystonia	1	1	>97th centile 7 months	Three mutations p.R257W, p.L245F, p.R243G	–

approximately 1:90 000. The male-to-female ratio was equal (5:5) and there was a history of consanguinity in one family. Seven babies were detected by newborn screening (NBS diagnosis). Three children were diagnosed after they presented in infancy with neurological symptoms (symptomatic diagnosis). The results are shown in Table 1. Our laboratory is the only one in New South Wales able to measure the relevant analytes, and we have close collaboration with the other five Australian laboratories, so that we are sure we know of all diagnosed cases of GA I.

Mean follow up of all surviving children was 55 months (range 9 months to 9.9 years, with 6 children followed for more than 2.2 years).

Three children were diagnosed clinically. The first child, a boy was born just before screening started in New South Wales. He was diagnosed at 8 months of age following the finding of an abnormal brain MRI scan following a minor head injury. Subsequent urine GC-MS analysis confirmed the diagnosis of GA I. There was no history suggestive of a previous neurological crisis. The second child had elevated glutarylcarnitine on NBS, but a resample level was below the established cut-off and he was not followed further. At 10 months, however, he presented with gradual onset of dyskinesia and hypotonia. Urine GC-MS analysis showed elevated glutaric and 3-hydroxyglutaric acids at the time. This child's results triggered changes in the NBS resampling protocols and cut-off levels, with a lower cut-off point, and immediate follow-up including urine GC-MS with no further resampling (Wilcken et al 2003). The third child was born outside NSW, and was not screened for GA I at birth. She was diagnosed at 2.5 years of age, after two acute illnesses with neurological crises at 8 months and 2.5 years of age. Urine GC-MS showed a trace increase of 3-hydroxyglutarate during a period of ketosis, but two subsequent tests gave normal results. At 2.5 years of age, raised urinary glutaric and 3-hydroxyglutaric acid levels in the urine and typical MRI findings confirmed the diagnosis, which was subsequently confirmed by enzymatic assay and mutation analysis (Table 1). Despite appearing alert and interactive, she has severe intellectual disability.

Among the 7 children diagnosed by newborn screening, several have had intercurrent illnesses, some requiring hospital admission (see Table 1). One died at 13 months from sequelae of an acute severe neurological injury (striatal necrosis), sustained at 7 months of age, following delayed presentation to hospital during an intercurrent illness, resulting in severe dystonia and seizures. Six children have remained well. However the youngest, now aged only 9 months, has developed mild dystonia affecting the arms, without any obvious episode of illness, and her routine immunizations were prospectively covered with a brief period of the emergency treatment regimen. An additional child had influenza A infection and developed mild dyskinesias, but appears to have made a full recovery.

Discussion

Newborn screening, started over four decades ago, is carried out to prevent significant morbidity, mortality and mental handicap in otherwise healthy children by early intervention and preventive therapy. Glutaric aciduria type I is one of the disorders that fit the criteria for newborn screening, as early diagnosis, an action plan for management, diet, and supplemental carnitine, appears likely to prevent, largely, the devastating outcome of a neurological crisis sustained early in life (Kölker et al 2007b; Naughten et al 2004).

The poor outcome in clinically diagnosed GA I is clear from many studies. Kyllerman and colleagues reported long-term follow up in 28 Nordic patients with GA I (Kyllerman et al 2004), mostly diagnosed symptomatically between 1975 and 2001. Three were siblings diagnosed presymptomatically. The overall outcome was described in terms of survival, and morbidity, with scores for motor, speech and cognitive function. Survival was 89% at 10 years, and 44% at 35 years of age. Only three patients of 28 had a total disability score of 3, indicating a normal or near-normal neurological outcome. Twenty-four patients had either dystonia [20] or dyskinesia [4] as a dominant feature. A strong correlation was found between disability in motor function and speech, whereas no correlation was detected between motor and cognitive function. Kölker and colleagues investigated the natural history, outcome and treatment efficacy in patients with GA I (Kölker et al 2006). The study enrolled 279 patients from 35 metabolic centres in Europe and North and South America. Two hundred and eighteen patients were diagnosed clinically, and compared with 23, 24 and 14 patients diagnosed by neonatal screening, high-risk screening and macrocephaly, respectively. Most patients (185 of 218) in the symptomatically diagnosed group had neurological complications following an encephalopathic crisis, including death in 49 children. A survival rate of 50% was reported at the age of 25 years in the study. The study demonstrated that the best predictor of a good outcome was a timely diagnosis and appropriate management with diet and L-carnitine. Treatment efficacy was low in patients diagnosed after the onset of neurological disease.

Only early diagnosis and presymptomatic treatment seem to be effective in preventing the acute and irreversible striatal damage. Naughten and colleagues reported the outcome of 21 GA I patients, 11 diagnosed clinically and 10 following high-risk sibling screening (Naughten et al 2004). Ten of the 11 clinically diagnosed patients had neurological sequelae (dystonic or spastic cerebral palsy), whereas none of the 10 siblings diagnosed presymptomatically and managed appropriately had any permanent neurological damage, although MRI was abnormal in 6 of these children. Seven of 11 clinically diagnosed patients died. One sibling with no neurological abnormality died following pneumonia. Newborn screening for GA I has been carried out in Germany using tandem mass spectrometry since 1999. Kölker and colleagues reported on 100 German patients, showing the effectiveness of newborn screening in prevention of severe neurological morbidity and mortality with the treatment strategies mentioned above (Kölker et al 2007b). The study population consisted of two patient groups: those diagnosed after NBS ($n = 38$, no patient known to be missed) and those diagnosed symptomatically or otherwise prior to NBS (historical controls). The birth prevalence of GA I was calculated in three groups: from 1979 to 1990 (1:611 400), from 1991 to 1998 (1:187 200) and from 1999 to 2005 (1:100 200). It is clear that there were many 'missing' patients from the unscreened groups (1979 to 1998). The exact number of patients cannot be calculated from the data available, but a rough estimate would be about 30. These patients could be asymptomatic, symptomatic but undiagnosed, or dead. Encephalopathic crises were observed in 41 of 53 symptomatic patients from the historical cohort (77%), 4 of 38 from the NBS cohort (10%); and 1 of 9 from the presymptomatically diagnosed historical cohort (11%). There were 11 deaths in the historical control group and none in the newborn screened group. The results show a significant benefit favouring NBS despite the fact that the existence of 'missing' patients means that the actual number of encephalopathic crises averted in the historical control cohort is not known.

Neuroradiology of GAI was performed in half of our patients, all of whom showed abnormal findings. This is comparable with data published by Twomey and colleagues from Ireland. Twomey and colleagues reported neuroradiological findings from 20 children (8 diagnosed neonatally) in the same cohort from Ireland as was reported by Naughten and colleagues in 2004. Multiple cranial ultrasound, CT scan and MRI of brain were performed on 9, 13 and 14 patients, respectively. Widening of the sylvian fissure (13 of 14) and basal ganglia abnormalities (9 of 14) were the most common findings using MRI (Twomey et al 2003). Our finding of relative macrocephaly in the majority of our patients is in agreement with other studies (Kyllerman et al 2004) but does not carry any significant correlation with clinical outcome.

Results in our small cohort support the overall findings in the larger German study (Kölker et al 2007b) and the Irish data (Naughten et al 2004) showing the benefit of early diagnosis by newborn screening in the prevention of adverse neurological outcomes in most patients with GA I. Despite the generally favourable outcome, it is clear that not all severe decompensations can be prevented; additionally, there are as yet no data on the long-term prognosis.

References

Christensen E (1983) Improved assay of glutaryl CoA dehydrogenase in cultured cells and liver. *Clin Chim Acta* **129**: 91–97.

Goodman SI, Freeman FE (2001) Organic acidemias due to defects in lysine oxidation: 2-ketoadipic acidemia and glutaric acidemia. In: Scriver CR, Beaudet AL, Sly WS, Valle D, eds; Childs B, Kinzler KW, Vogelstein B, assoc. eds. *The Metabolic and Molecular Bases of Inherited Disease*, 8th edn. New York: McGraw-Hill, 2195–2204.

Goodman SI, Markey SP, Moe PG, Miles BS, Teng CC (1975) Glutaric aciduria: a 'new' disorder of amino acid metabolism. *Biochem Med* **12**: 12–21.

Kölker S, Garbade SF, Greenberg CR, et al (2006) Natural history, outcome and therapeutic efficacy in children and adults with glutaryl-CoA dehydrogenase deficiency. *Pediatr Res* **59**: 840–847.

Kölker S, Christensen E, Leonard JV, et al (2007a) Guideline for the diagnosis and management. *J Inherit Metab Dis* **30**: 5–22.

Kölker S, Garbade SF, Boy N, et al (2007b) Decline of acute encephalopathic crises in children with glutaryl -CoA dehydrogenase deficiency identified by newborn screening in Germany. *Pediatr Res* **62**: 357–363.

Kyllerman M, Skjeldal O, Christensen E, et al (2004) Long-term follow up, neurological outcome and survival rate in 28 Nordic patients with glutaric acidemia type 1. *Eur J Paediatr Neurol* **8**: 121–129.

Naughten ER, Mayne PD, Monavari AA, Goodman SI, Sulaiman G, Croke DT (2004) Glutaric aciduria type I: Outcome in the Republic of Ireland. *J Inherit Metab Dis* **27**: 917–920.

Strauss KA, Puffenberger EG, Robinson DL, Morton DH (2003) Type 1 glutaric aciduria, part I: natural history of 77 patients. *Am J Med Gen semin Med Genet* **121**: 38–52.

Twomey EL, Naughten ER, Donoghue VB, Ryan S. (2003). Neuroimaging findings in glutaric aciduria type 1. *Pediatr Radiol* **33**: 823–830.

Wilcken B, Wiley V, Hammond J, Carpenter K (2003) Screening newborns for inborn errors of metabolism by tandem mass spectrometry. *N Engl J Med* **348**: 2304–2312.

PAPERS ADDRESSING GENRAL ASPECTS OF NEWBORN SCREENING

Eur J Pediatr (2003) 162: S62–S66
DOI 10.1007/s00431-003-1355-z

ARTICLE

Bridget Wilcken

Ethical issues in newborn screening and the impact of new technologies

Published online: 14 November 2003

Abstract Medical ethics is an integral part of medical practice. The general principles are well known: autonomy (the right to choose), beneficence (do good), nonmaleficence (do no harm), and justice (be fair and equitable). In newborn screening these principles must be especially carefully applied, as the intervention, screening, has not been sought by the patient, but is a form of preventive medicine.In proposing a screening programme questions to be asked are: should we do it? (is there enough benefit, not too much harm?); can we do it? (do we have the technology and skill to find the cases sought); can we afford it? The first question, with its ethical implications, is often ignored. New issues have arisen with new technology, but underlying ethical themes are the same. Tandem mass spectrometry can be used to detect about 30 very rare disorders in a single test. Proving the benefit of this (and other screening tests) is difficult because randomised controlled trials seem impractical, because of power considerations, long follow-up time, and because there is already a perceived benefit. Best possible evidence of a lower order must be sought. In future, DNA microarray technology is likely to become sufficiently inexpensive to apply to newborn screening. It is difficult to predict all the future possibilities of DNA technology in this fast-moving field. Major ethical problems are likely with the ability to detect adult-onset disorders or susceptibility to these in babies. Under what circumstances would this be ethical? We need to start debating these issues.

Keywords DNA microarray · Medical ethics · Newborn screening · Tandem mass spectrometry

B. Wilcken
The New South Wales Newborn Screening and Biochemical Genetics Departments, and the University of Sydney, The Children's Hospital at Westmead,
Locked Bag 400, NSW 2145 Westmead, Australia
E-mail: bridgetw@chw.edu.au
Tel.: +61-2-98453654
Fax: +61-2-98453121

Abbreviations *GA2* glutaric aciduria type 2 · *MCAD* medium-chain acyl-CoA dehydrogenase deficiency · *MS/MS* tandem mass spectrometry · *PKU* phenylketonuria

Introduction

Ethics, the Shorter Oxford Dictionary tells us, is "the branch of knowledge that deals with the principles of human duty" and "the rules of conduct recognised in a particular profession". In the medical profession one of the sets of rules recognised for centuries has been that expounded by Hippocrates in the 4th century BC. The Hippocratic oath has been sworn to by young doctors on their graduation until quite recently, but few could say exactly what it encompasses. Rather loosely translated the oath has some general themes – to do no harm, to keep patient's confidences and it has some fairly specific clauses, such as not to assist suicide or to cause abortion, not to practice beyond one's competence, and not to abuse professional relationships. These "rules" were widely adopted by Christian Western Europe. More recently, ethical principles in medicine were enunciated in more general ways [3]. The four principles governing medical ethics are recognised as autonomy (the right to choose), beneficence (doing good), nonmaleficence (doing no harm), and justice (being fair and equitable). These are the principles that must be applied in a consideration of how newborn screening should be conducted. In this essay the consideration of these principles is intertwined with the consideration of criteria for conducting screening programmes. There is no way of untangling these. Scientific validity of a screening programme is an essential component of the ethical considerations.

The questions to be explored in the consideration of setting up a screening programme can be simplified to three: (1) should we do it? (2) can we do it? and (3) can we afford to do it? There are ethical aspects to all these three questions. While the first seems likely to be the

most important one from an ethical perspective, it is the most often ignored, for reasons that will be explored. These three questions are a distillation of the classic criteria for screening, set out in the seminal paper of Wilson and Jungner in their World Health Organisation publication of 1968 [19]. In the light of new technology, however, several of these classical criteria need to be rethought, and another question added: how should we do it?

One of the new technologies that should be explained before consideration of the ethical aspects is tandem mass spectrometry (MS/MS). MS/MS can be used to detect disorders of amino acid, organic acid, and fatty acid metabolism by measuring amino acids and acyl-carnitines in dried blood spots in a single rapid assay. About 30 extremely rare disorders can be diagnosed simultaneously. It is possible to select which compounds to analyse, and so specify which disorders are tested for. Detection of markers for disorders known to be benign need not occur. Most of the disorders detectable have a birth prevalence of much less than 1:50,000.

Rethinking Wilson and Jungner

"Should we do it?" asks the question of whether there is a benefit in making an early diagnosis, and whether the possible harm in screening is much more than offset by the potential benefit. Wilson and Jungner address the aspect of benefit in six of their ten points. "The condition sought should be an important health problem" seems incontrovertible. "There should be an accepted treatment for patients with recognised disease" and "facilities for diagnosis and treatment should be available" also seem obvious, but all of these require reinterpretation when we consider the impact of one of the new technologies, MS/MS. MS/MS allows the identification of a number of extremely rare disorders simultaneously in the one test. The disorders, so individually rare, might not be considered to constitute "an important health problem" [19] or to occur frequently enough to justify screening the whole population [2]. However, another way to look at frequency and seriousness would be to consider newborn screening, up to a point, as a package. If a single sample can be used for a number of tests, then the rarity of some of the individual disorders will not matter if the whole aggregate of disorders occurs frequently enough to justify a mass screening programme. Currently, in many newborn screening programmes, around one baby per thousand tested has one of the disorders sought overall.

A current controversy about MS/MS screening is related to the possibility of detecting disorders which do not consistently respond to treatment. Examples are the severe, classical organic acidaemias, and neonatal onset glutaric aciduria type 2 (multiple acyl-CoA dehydrogenase deficiency, GA 2, McKusick 231680). Many babies with these conditions will die even with the implementation of early and appropriate treatment, and in addition, most will become sick before the availability of any screening test result. However, not all of such babies have in the past received the correct diagnosis, and there is great benefit from the early and correct diagnosis of a baby with a largely untreatable disorder if this might otherwise be mis-diagnosed as sepsis, birth trauma or other neonatal problem. Correct and definite diagnosis gives the parents comfort and makes a sort of closure possible, as well as offering future reproductive choices. But the implications are quite broad. The screening process is supposed to confer a benefit, and the implications are, benefit to the baby. Is benefit to the family enough, when there is no benefit to the baby tested? That might be so if the baby will go on to an impaired childhood, before dying, as benefit to the family will also certainly benefit the child during his (her) life. This could occur, for example, with Duchenne muscular dystrophy, a presently untreatable disorder for which there are a number of optional screening programmes offered [4]. But if the disorder is quickly fatal, then there will be no obvious benefit to the baby. But nor should there be deemed to be any harm, since the screening sample collected would be used for disorders which more clearly fit the criteria for screening. So if there is clear benefit to the family, this seems like a non-issue, but does cast considerable doubt on whether effective treatment is a sine qua non for disorders included in a screening programme. A benefit to the family when there is no benefit to the baby however would not seem contraindicated if there is no harm to the baby, and as the MS/MS newborn screening tests have the possibility to diagnose a number of disorders simultaneously, to exclude one because there might be no benefit to the baby, although parents could benefit, flies in the face of common sense.

Wilson and Jungner go on to make more points about benefit which are less easy to accept. It should be remembered that their excellent paper was written not specifically about newborn screening, but addressed cancer, infectious disease, chronic problems etc. Thus "there should be a recognised latent or early symptomatic stage" and "the natural history of the condition should be understood" may not seem applicable, as only with screening may the natural history of a disease begin to be understood. Recent papers about medium-chain acyl-CoA dehydrogenase (MCAD McKusick 201450) deficiency illustrate the difficulty there has been in unravelling the probable natural history without the aid of some population screening [5,11]. Surprisingly, Wilson and Junger do not really address sources of harm from screening, except perhaps in their 8th point, "there should be an agreed policy on whom to treat as patients". As in other forms of screening, newborn screening detects many mild variations of disease the need for treatment of which has only been established in severe cases. This is a very important issue, present in screening programmes already established, such as those for phenylketonuria (PKU, McKusick 261600), and likely to become more apparent with the emerging technologies of MS/MS and DNA approaches (see below).

Can we do it? asks the question of whether the technology and skill is available to find the cases. Wilson and Jungner say "there should be a suitable test or examination" (attention to sensitivity, specificity, positive predictive value, reliability and so forth), and "the test should be acceptable to the public". Acceptability has until now posed little problem in newborn screening. A heel-prick blood sample taken from a baby is well accepted. But a better informed public has become aware of the potential for DNA testing on dried blood samples, and a few parents have raised concerns about what might happen to the sample, a somewhat different issue. Newborn screening programmes have an obligation to be informative about sample retention, and to have policies about use which prohibit any use of an identified sample after the completion of the screening tests without written permission form parents.

Can we afford it? The WHO paper does address cost in one of the ten points: "the cost of case-finding should be economically balanced in relation to the cost of medical care as a whole". This also relates to benefit, and to one of the key points of more modern ethics, that of justice. Will the cost of providing benefit to a few at the expense of another programme that might provide benefit to more people. A balancing act indeed.

Beneficence: finding the evidence of benefit

None of this addresses the core question of benefit, and that is really an ethical as well as scientific question. It is certainly not ethical to conduct routine testing of babies if the benefit is uncertain. It is important to evaluate the benefit from early diagnosis by newborn screening, compared with that from later clinical diagnosis, and also what harm might come from screening. As we previously noted, worldwide, formal evidence of clinical effectiveness of newborn screening has rarely been sought [6]. The only randomised, controlled trials carried out have been of cystic fibrosis screening, one in the United States of America [7] and one in the United Kingdom [7]. Clinical effectiveness of screening for PKU and hypothyroidism is considered clear-cut [2] although no formal trials were ever conducted, and there is also much evidence, for example, of the usefulness of screening for sickle-cell disease in regions where this is prevalent [9]. Part of the problem is the difficulty of mounting randomised, controlled trials of screening for very rare disorders, due to power considerations [17]. For example, even in a disorder with a birth prevalence of 1:50,000, if an adverse outcome could be halved by early detection, a randomised controlled trial would require 2,500,000 babies in each arm. Another feature militating against randomised controlled trials for some disorders has been the strong conviction that early diagnosis would be beneficial, based on clinical experience. This situation has led some to voice uncertainty about the scientific and ethical validity for instituting many screening programmes. Nowhere is this more apparent than in the case of MS/MS screening in general, and MCAD deficiency in particular. MCAD deficiency is a potentially lethal fatty acid oxidation defect, the commonest of the disorders detected by MS/MS (apart from PKU). Management of MCAD deficiency is simple, and after diagnosis, a fatal outcome is vanishingly rare. But 25% of cases diagnosed clinically have been fatal. There is also considerable morbidity among diagnosed cases. Screening by MS/MS reliably diagnoses cases of MCAD deficiency [12] but probably diagnoses in addition some cases who are not at risk of metabolic decompensation [5]. While the latter is not yet certain, some commentators have felt that screening should not proceed [10] and indeed, screening by MS/MS in the United Kingdom has not been allowed except for PKU. None of this argument should be seen to be denigrating randomised controlled trials. If we do not perform such trials, we are using an innovation to experiment on 100% of a population, rather than only 50%. But as there is a genuine difficulty in mounting randomised trials of screening for very rare diseases which often cannot be overcome, newborn screening practitioners should see themselves as both ethically and scientifically bound to employ the next best possible means of evaluation.

Another aspect of benefit that needs to be addressed is reaching agreement on the goals of screening for different disorders. Increase in life expectancy may not be a necessary component of benefit from an ethical point of view. Improving the span of healthy trouble-free life may be enough. This can be seen in cystic fibrosis, where reduction of early morbidity and improvement of nutrition have been clearly shown [8, 18] but as yet there is no evidence of an ultimately increased life expectancy. This has led some to conclude that there is not yet evidence of benefit [16].

Non-maleficence: finding the harm

Potential harm in screening programmes needs to be considered for all categories – those families where there is a false-positive screening result (the most obvious category), families of babies "missed" by the test (false-negative cases), and to some extent to babies and families successfully detected, especially if treatment is not, or not fully, effective, and the screening has provided a longer time-frame for worry and disturbance to family life. There may be harm also to the whole screened population, from the processes of the testing, and this includes the sampling and the storage of samples. In fact there has been little systematic study of harm from screening programmes. The problem of false-positive results has been investigated only a few times [14, 15]. In one study, about 20% of the parents remained anxious after such a result, whether they believed the repeat test was needed because of an initial positive result, or because of a technical problem [14, 15]. However, studies of patients correctly diagnosed with a disorder have not shown early diagnosis to affect adversely the parent-child interaction. Clearly more studies are needed in

relation to harm from screening, but so far the evidence is somewhat encouraging. Unwanted carrier detection has been a feature of recent protocols for newborn screening for cystic fibrosis [19]. Here, the overall numbers of cases are small, and can be dealt with by sensitive counselling, although there are many implications. This unwanted effect of screening could be justified on the principle of "greater good" since protocols which do not involve any carrier detection are more burdensome on more of the unaffected population [19]. When we come to consider new technologies involving DNA testing however, this is a major issue.

Autonomy: the right to choose, and the protection of those with diminished autonomy

Consent to newborn screening is emerging as an important issue of autonomy. On the one hand, there is a general expectation of the right to informed choice about any medical intervention. On the other hand, there is the view expressed in the World Health Organisation's Draft Guidelines on Ethical Issues in Medical Genetics that "newborn screening should be mandatory and free of charge if early diagnosis and treatment will benefit the newborn"[1]. That is, parents should not be able to refuse interventions that are beneficial to their babies. The usual thinking in the western world is to accord full autonomy to the parents, and this is inclined to occur in newborn screening. Even where such testing is mandated, there is generally accommodation for parents who have some religious or other strong objections. Refusal is very uncommon where good information is made available to parents, and this is certainly an obligation for newborn screening programmes. As tests become more complex and more wide-ranging, provision of high quality educational material is ever more necessary.

Justice: distributive justice

This has already been covered briefly under costs of screening. Not only does this general principle of ethics in medicine imply that all should be able to benefit from newborn screening (not easy in cash-strapped societies where there is a charge for the screening test), but that the benefits and burdens of healthcare should be distributed as fairly as possible in a society not yet used to the idea of rationing healthcare. It would not be enough to show that newborn screening with a very expensive technique was of overall benefit, if it could be shown that the money could be much more beneficially spent elsewhere.

Application to new technologies

Two new technologies are likely to impact massively on newborn screening. One is here now, MS/MS. The full potential for its use in screening is not yet established. The other technology whose time is surely close is some DNA screening method, probably involving microarray.

As explained earlier, the key feature about MS/MS is that one inexpensive test can detect up to at least 30 disorders, and the disorders to be included can be almost completely determined in advance. The incremental cost of including an extra disorder (among the suite of classes of disorders) is extremely small, and relates only to follow-up, and not to the testing process. A number of potential ethical problems have been identified. Most have already been dealt with – detecting a disorder for which treatment is often not successful; detecting disorders whose natural history is not well known; detecting maternal variance and carriers; and proving benefit. Proving benefit has been a sticking point for many in the consideration of whether this new form of screening should be adopted. This was recently addressed in the United Kingdom, where two Technology Assessments were commissioned [11, 13]. They considered the evidence for newborn screening for different disorders, with particular reference to MS/MS, and came to somewhat different conclusions. One group found a strong case for introducing MS/MS screening for a large number of disorders, as a pilot study, [11] while the other recommended a pilot study of screening for only two disorders by MS/MS, MCAD deficiency and GA 2 (plus PKU) and an embargo on the inclusion of any other disorders [13]. Although the latter may seem a very well-based ethical approach, the special features of MS/MS make it seem unnecessarily restrictive. As we have seen, adding other tests is almost cost-free, and many disorders are well known to benefit from early treatment. For example, it was considered that there was no evidence of benefit in the early diagnosis of several amino acid disorders, including homocystinuria (cystathionine beta-synthase deficiency), whereas there is plenty of compelling evidence, simply no randomised controlled trials e.g. [21]. In fact, by 2003 no such pilot studies had been funded in the UK, despite there being several laboratories equipped to carry them out. There are serious questions of whether this strict stance is itself ethical, particularly for MCAD deficiency, as already discussed above. A simple test can detect a disorder and almost certainly prevent some mortality and morbidity. Should it be denied to children? Studies less rigorous than randomised controlled trials are under way, and will no doubt give some answers, but some years down the line. In the meantime a very real dilemma exists for policy makers.

DNA tests and newborn screening

Microarray technology will produce much worse dilemmas. Newborns are an attractive population for screening initiatives, total population screening being easily achieved, and economical sample collection possible. Filter paper blood samples are useful for many

analytes, and not least for DNA testing. What might be added to newborn screening tests depends on new-found benefit from early diagnosis (new understanding of pathophysiology, new treatments), and new technology. DNA tests, now relatively expensive, will certainly become much cheaper. For single gene disorders a microarray strategy might be used for primary screening. It is quite possible that the predictive value of the test would be much less good than for functional (biochemical) tests, would detect probable mild disorders not requiring intervention, and could also not readily encompass private mutations. Worse still, for recessively inherited disorders, there would be the certainty of identifying all carriers of the mutations represented on the array. Also, there is likely to be pressure to screen for adult onset disorders. Screening for susceptibility genes in newborns will of course also be possible. Here the uncertainty of predicting disease is much greater, and the pious hope that knowledge might lead to beneficial lifestyle alterations seems likely to remain just that – a pious hope. We have not yet learned how to ensure that people take advice about harmful lifestyles. There is a great need to consider all the ethical issues that are sure to be raised in future DNA screening, particularly in newborns, and to address these early and thoughtfully.

Summary

Ethical issues in newborn screening are like those in all of medicine. Special issues centre round the problems of proving benefit without, in the main, the conventional means of randomised controlled trials, and there is a need to educate policy forming committees about the appropriate considerations and the use of lower order evidence. New DNA technologies will pose new and very difficult problems which, although they cannot be foreseen in full, should be explored now, well before the technology could be applied to large population screening.

References

1. Anon (2000) Proposed international guidelines on ethical issues in medical genetics and genetics services. WHO/HGN/GL/ETH/98 1
2. Anon (2000) Serving the family from birth to the medical home. A report from the Newborn Screening Task Force convened in Washington DC, May 10–11, 1999. Pediatrics 106: 383–427
3. Beauchamp T, Childress J (1994) Principles of biomedical ethics, 4th edn. Oxford University Press, New York
4. Bradley DM, Parsons EP, Clarke AJ (1993) Experience with screening newborns for Duchenne muscular dystrophy in Wales. BMJ 306: 357–360
5. Carpenter K, Wiley V, Sim KG, Heath D, Wilcken B (2001) Evaluation of newborn screening for medium chain acyl-CoA dehydrogenase deficiency in 275,000 babies. Arch Dis Child Fetal Neonat Ed 85: F105–F109
6. Wilcken B, Wiley V, Hammond J, Carpenter K (2003) Screening newborns for inborn errors of metabolism by tandem mass spectrometry. N Engl J Med 348: 2300–2308
7. Chatfield S, Owen G, Ryley HC, Williams J, Alfaham M, Goodchild MC, Weller P (1991) Neonatal screening for cystic fibrosis in Wales and the West Midlands: clinical assessment after five years of screening. Arch Dis Child 66: 29–33
8. Farrell PM, Kosorok MR, Rock MJ, Laxova A, Zeng L, Lai HC, Hoffman G, Laessig RH, Splaingard ML (2001) Early diagnosis of cystic fibrosis through neonatal screening prevents severe malnutrition and improves long-term growth. Wisconsin Cystic Fibrosis Neonatal Screening Study Group. Pediatrics 107: 1–13
9. Lees CM, Davies S, Dezateux C (2000) Neonatal screening for sickle cell disease. Cochrane Database Syst Rev CD001913
10. Leonard JV, Dezateux C (2002) Screening for inherited metabolic disease in newborn infants using tandem mass spectrometry. BMJ 324: 4–5
11. Pollitt RJ, Green A, McCabe CJ, Booth A, Cooper NJ, Leonard JV, Nicholl J, Nicholson P, Tunaley JR, Virdi NK (2001) Neonatal screening for inborn errors of metabolism: cost, yield and outcome. Health Technol Assess 1: 1–202
12. Pourfarzam M, Morris A, Appleton M, Craft A, Bartlett K (2001) Neonatal screening for medium-chain acyl-CoA dehydrogenase deficiency. Lancet 358: 1063–1064
13. Seymour CA, Thomason MJ, Chalmers RA, Addison GM, Bain MD, Cockburn F, Littlejohns P, Lord J, Wilcox AH (2001) Newborn screening for inborn errors of metabolism: a systematic review. Health Technol Assess 1: 1–95
14. Sorenson JR, Levy HL, Mangione TW, Sepe SJ (1984) Parental response to repeat testing of infants with 'false-positive' results in a newborn screening program. Pediatrics 73: 183–187
15. Sveger T, Thelin T (1981) Four-year-old children with alpha 1-antitrypsin deficiency. Clinical follow-up and parental attitudes towards neonatal screening. Acta Paediatr Scand 70: 171–177
16. Wald NJ, Morris JK (1998) Neonatal screening for cystic fibrosis. BMJ 316: 404–405
17. Wilcken B (2001) Rare diseases and the assessment of intervention: what sorts of clinical trials can we use? J Inherit Metab Dis 24: 291–298
18. Wilcken B, Chalmers G (1986) Reduced morbidity in cystic fibrosis. Lancet 1: 439
19. Wilcken B, Wiley V, Sherry G, Bayliss U (1995) Neonatal screening for cystic fibrosis: a comparison of two strategies for case detection in 1.2 million babies. J Pediatr 127: 965–970
20. Wilson JMG, Jungner G (1968) Principles and practice of screening for disease. World Health Organization, Geneva
21. Yap S, Naughten E (1998) Homocystinuria due to cystathionine beta-synthase deficiency in Ireland: 25 years' experience of a newborn screened and treated population with reference to clinical outcome and biochemical control. J Inherit Metab Dis 21: 738–747

J Inherit Metab Dis (2007) 30:129–133
DOI 10.1007/s10545-007-0538-6

ICIEM 2006

Recent advances in newborn screening

B. Wilcken

Received: 13 December 2006 / Submitted in revised form: 10 January 2007 / Accepted: 10 January 2007 / Published online: 6 March 2007

Summary The introduction of tandem mass spectrometry has unquestionably been the most striking recent advance in newborn screening. A single test is applied for the simultaneous diagnosis of a number of disorders, making it possible to screen for some disorders that might otherwise have seemed too rare. Current screening is for disorders of metabolism of amino acids, organic acids and fatty acids. Assay performance for detection of disorders appears very good, but rarity of disorders, varied definitions and systems for follow-up and lack of databases for inborn errors of metabolism diagnosed clinically means that there is as yet insufficient information about most disorders. The technology can be applied to a much wider range of compounds, and the field looks set to expand. A key feature of newborn screening programmes must be the assessment of outcomes, and a major reason for the lack of uniformity in the approach adopted in different countries is the paucity of information on this. The available evidence points to overall advantages flowing from early diagnosis by screening, with reduction in mortality and morbidity. More studies are clearly needed and some are under way. The next new group of disorders already proposed for newborn screening is the lysosomal storage disorders. Attitudes may be changing about what it is desirable to include in a newborn screening programme, and this will indeed pose new ethical dilemmas.

Communicating editor: Verena Peters

Competing interests: None declared

B. Wilcken
Western Sydney Genetics Program, The Children's Hospital at Westmead, Westmead, and the University of Sydney, Sydney, New South Wales, Australia
e-mail bridgetw@chw.edu.au

Abbreviations

MCAD	medium-chain acyl-CoA dehydrogenase deficiency
MSUD	maple syrup urine disease
PPV	positive predictive value
PKU	phenylketonuria

Introduction

The history of newborn screening stretches back to the seminal work of Guthrie in the early 1960s. For the following 30 years or more screening expanded slowly, came into routine neonatal practice in developed countries, and was a successful public health enterprise. In 2006 newborn screening has become more widespread, and it has been estimated that about 25% of the world's newborns are tested at least for hypothyroidism.

Unquestionably the most striking recent advance in newborn screening has been the introduction of tandem mass spectrometry. This has triggered a profound revolution in the way newborn screening is regarded and its place within paediatrics, and, in some parts of the world, has greatly enhanced the very practice of newborn screening. The success in its introduction has spawned a new interest in further expansion, with already a large literature about current achievements.

Tandem mass spectrometry

The potential use of tandem mass spectrometry for newborn screening was first suggested in 1990 by Millington, Roe and colleagues (Millington et al 1990) and early studies soon demonstrated its practicality (Chace et al 1993, 1995; Rashed et al 1995; Ziadeh et al 1995). It is not too dramatic

to call this the beginning of a completely new era in newborn screening. Tandem mass spectrometry has seemed so new and revolutionary because for the first time a single test could be applied for the simultaneous diagnosis of a number of disorders, making it possible to screen for some disorders that might otherwise have seemed too rare. This very fact has blurred the standard criteria for screening. At present, the new 'expanded screening' is used to detect disorders of amino acid, organic acid and fatty acid metabolism, since these can be detected simultaneously in a single assay. However, the technology can be applied to a much wider range of compounds, and the field looks set to expand. The use of tandem mass spectrometry has become widespread since the late 1990s, and it is now possible to assess what has been learnt so far and what remains to be more fully investigated about the current screening.

Sensitivity, specificity and the positive predictive value

Reports from several newborn screening laboratories have given data on detection rates (Frazier et al 2006; Naylor and Chace 1999; Schulze et al 2003; Wilcken et al 2003; Zytkovicz et al 2001), but it has been difficult to get an accurate idea of sensitivity and specificity. What seems certain is that sensitivity is virtually 100% for phenylketonuria (PKU) (Hardy et al 2002; Schulze et al 2002) and for medium-chain acyl-CoA dehydrogenase (MCAD) deficiency (Wilcken et al 2007) with a high positive predictive value (PPV). Very high sensitivity (close to 100%) with acceptable PPV is also likely for a number of other disorders—early-onset organic acidurias, several amino acidurias including classical maple syrup urine disease (MSUD), citrullinaemia, argininosuccinic aciduria and tyrosinaemia type II—and for several of the fatty acid oxidation disorders. However, rarity of disorders, varied systems for follow-up and lack of databases for inborn errors of metabolism diagnosed clinically means that there is as yet insufficient information for most disorders. Germany (Hoffmann et al 2004) and Australia (Wilcken et al 2003) have been able to compare the rates of clinical and screening diagnosis for a range of disorders. It may be too early to say what the diagnostic potential is, as diagnostic algorithms are still being developed. In particular, Rinaldo has initiated a data collection that will enable more accurate use of this technology (P. Rinaldo, personal communication, 2006). The notion of an 'acceptable' PPV is necessarily vague: it will vary depending on what the particular disorder is, and on its frequency. It might be acceptable to some programmes to risk missing a case of 3-ketothiolase deficiency, say, rather than have an extremely high recall rate (Wilcken et al 2003), since an episode of decompensation is easily recognized and is treatable, but not acceptable to risk missing a case of classical PKU.

Some disorders cannot be detected with high sensitivity using current protocols: one example is homocystinuria–cystathionine β-synthase deficiency. Screening using methionine as a marker has been carried out for many years with bacterial inhibition assays, before the adoption of tandem mass spectrometry. No case of pyridoxine-responsive homocystinuria has been reported to have been detected. Use of an assay for homocystine might be successful, but this cannot be done simultaneously with assays in current use, and no mass screening has been reported. Tyrosinaemia type I cannot be reliably detected without using a separate assay for succinylacetone. There have been several reports of methods for succinylacetone, using dried blood samples, but all need a separate run and would therefore be used as a second-tier assay (e.g. Allard et al 2004).

There have been few systematic reports of false-negative results. The two screening programmes to report having missed cases (Frazier et al 2006; Wilcken et al 2003) showed a remarkable concordance: both reported missed cases of glutaric aciduria type I, cobalamin C, tyrosinaemia type I, 3-ketothiolase deficiency and mild MSUD. The NSW Australia programme also reported a missed case of very-long-chain acyl-CoA dehydrogenase deficiency. Many of these misses were early in the life of the screening programme. The missed mild cases of MSUD provide an illustration of the impossibility, with current screening, of detecting all mild cases of this condition (see Fig. 1) although clinical outcome is often poor (Bhattacharya et al 2006).

Greater rate of detection by screening

It is unsurprising, and has been well established, that for some disorders the detection rate is much higher by screening than by clinical diagnosis. This is particularly true of the fatty acid oxidation disorders, and especially MCAD deficiency, where the detection rate is more than twice as great by screening (Grosse et al 2006; Rhead 2006; Wilcken et al 2007). Other disorders with much higher rates of detection include the fatty acid oxidation defects very-long-chain acyl-CoA dehydrogenase deficiency and short-chain acyl-CoA dehydrogenase deficiency (Boneh et al 2006; van Maldegem et al 2006) and some organic acid disorders, especially 3-methylcrotonyl-CoA dehydrogenase deficiency, but others also (Koeberl et al 2003; Pedersen et al 2006). The spectrum of mutations may also vary in patients diagnosed by screening, compared with what was previously found by clinical diagnosis (Andresen et al 2001). Both of these features make it very difficult to compare outcomes of screening versus clinical diagnosis, as it is unclear how comparable the two groups of patients are (see below).

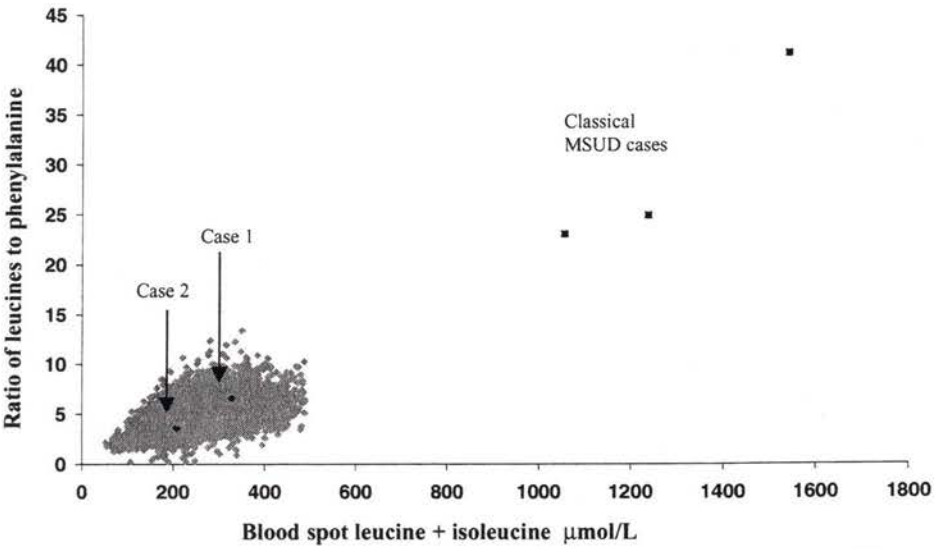


Fig. 1 Combined leucine + isoleucine to phenylalanine ratio versus leucine + isoleucine in 5 MSUD patients, 2 intermittent and 3 classical, within 30 000 newborn screening samples. The results for ratios to alanine were similar. The cut-off value for leucine + isoleucine for further action at the time was 500 μmol/L. No ratios were actually used for recall purposes

Which disorders to include?

Using tandem mass spectrometry, it is possible to choose which analytes to include and therefore, in the main, but not exclusively (because some analytes are markers for several disorders), to choose which disorders will be included in the screening suite (Garg and Dasouki 2006). The detection of disorders that may well be mild is a major problem for this new expanded screening. Screening puts a responsibility on the physician to provide some benefit to the person being screened (Frankenburg et al 1988), and there is also clearly a responsibility to avoid the harm that may come from making a diagnosis in a patient who has a benign condition. As mentioned above, some of the disorders currently included in screening programmes do seem to be either entirely or almost entirely benign (Rhead et al 2002; van Maldegem et al 2006). The American College of Medical Genetics has made use of a composite approach involving literature review and soliciting expert opinion to recommend a core panel of diseases for inclusion in a screening programme (American College of Medical Genetics Newborn Screening panel 2006), as well as 'secondary targets', disorders that are included in the differential diagnosis of a core disorder. This approach has been fairly successful overall but might need updating as more experience is gained, since disorders of unproven clinical significance are included. It seems a somewhat better approach than has been adopted in some other regions (Pollitt 2006). A major reason for the lack of uniformity in the approach adopted in different countries is the paucity of information on outcomes of screening.

Assessing outcomes

A key feature of newborn screening programmes must be the assessment of outcomes—is the introduction of screening actually doing any good? There are starting to be such assessments for tandem mass spectrometry. Reports of general outcomes (Schulze et al 2003; Waisbren et al 2003) and outcomes for specific disorders (Nennstiel-Ratzel et al 2005; Simon et al 2006) have been made. Most recently, in an Australia-wide study, we have demonstrated decreased morbidity and mortality up to the age of 4 years in patients with MCAD deficiency diagnosed by screening, as compared with those diagnosed clinically, but no differences in overall intellectual assessment. This study, unlike others, takes into account the greater rate of detection among screened patients (Wilcken et al 2007). To be able to make meaningful assessments, a clear definition of the disorder in question, a knowledge of the completeness of ascertainment, comparability of management and a uniform period of follow-up for screen-diagnosed and clinically diagnosed patients is necessary (Wilcken 2006). These have not always been available and, in particular, definitions need to be developed so that proper comparisons can be made. Nevertheless, the evidence available points to advantages flowing from early diagnosis by screening, with reduction in mortality and morbidity. More studies are clearly needed and some are under way.

Assessment of benefit is one thing, but possible harms should also be assessed. Reference has already been made to the probability of diagnosing entirely benign conditions, with the disadvantages to the patient and family that are likely to be a consequence. The positive predictive value of the testing is also important, as false-positive test results certainly have short-term implications for families, and perhaps long-term ones also. The false-positive rate of expanded newborn screening seems to be very low in the main, with reports of 0.2–0.33% of babies screened needing further testing (including for PKU) (Wiley et al 1999; Zytkovicz et al 2001), and PPVs seem to be quite acceptable. The burden imposed on laboratories and clinical services is not yet clear, but a

recent report shows only a small increase in work in New South Wales, Australia (Carpenter et al 2006). The incremental costs of screening by tandem mass spectrometry have been assessed theoretically in several studies. In Australia we have assessed the actual costs associated with the screening. One interesting finding is that the mean incremental cost of performing the screening test was on average only A$1.67 per baby screened in Australian laboratories (about US$1.25) and that this was the greatest component of the total cost of screening, diagnostic testing and costs of follow-up (M. Haas, personal communication, 2007).

Defining follow-up procedures and management

One result of the expansion of newborn screening has been its incorporation into the mainstream of metabolic medicine, and the consequent development of readily available targeted information on further investigation and treatment. The organizational model of care to some extent dictates what is needed in different regions: where clinical care is in specialized medical centres, the needs are different from those where care is largely by private physicians who may not be experts in metabolic disease. The fact sheets produced by the American Academy of Pediatrics Committee on Genetics have been recently revised in part to take account of the expanded newborn screening (Kaye et al 2006). They cover testing, follow-up of abnormal results, diagnostic tests, and disease management and are a particularly valuable resource where much treatment is away from specialized clinics.

Other advances involving tandem mass spectrometry

While expanded newborn screening by tandem mass spectrometry is not the only recent advance, it is by far the most spectacular. This technology will be important in the next new group of disorders to be proposed for newborn screening—lysosomal storage disorders. Tests have been developed for a range of disorders that have some treatment available, using either multiplexed direct enzyme assay (Gelb et al 2006) or multiplexed immune-quantification of lysosomal proteins (Meikle et al 2006). Already one disorder, Krabbe disease, or globoid-cell leukodystrophy, is included in routine newborn screening in New York State (http://genes-r-us.uthsca.edu/nbsdisorders.pdf).

What can be added to a newborn screening programme depends on the development of new ways to diagnose a disorder (new technologies, new biochemical or molecular markers), or the discovery of new treatments or new preventive measures that make a disorder worthy of consideration for inclusion in newborn screening. However, we may also be seeing the beginning of new attitudes to what is desirable: debates on whether we should embrace susceptibility testing for a range of disorders, such as type 1 diabetes, or whether it is proper to test newborns for adult-onset diseases. DNA testing, already commonplace in some newborn screening programmes, but only in an extremely limited way, may be the next new big expansion. This will indeed pose new ethical dilemmas.

References

Allard P, Grenier A, Korson MS, Zytkovicz TH (2004) Newborn screening for hepatorenal tyrosinemia by tandem mass spectrometry: analysis of succinylacetone extracted from dried blood spots. *Clin Biochem* **37**: 1010–1015.

American College of Medical Genetics Newborn Screening Panel (2006) Newborn screening: toward a uniform screening panel and system—executive summary. *Pediatrics* **117**: S296–S307.

Andresen BS, Dobrowolski SF, O'Reilly L, et al (2001) Medium-chain acyl-CoA dehydrogenase (MCAD) mutations identified by MS/MS-based prospective screening of newborns differ from those observed in patients with clinical symptoms: identification and characterization of a new, prevalent mutation that results in mild MCAD deficiency. *Am J Hum Genet* **68**: 1408–1418.

Bhattacharya K, Khalili V, Wiley V, Wilcken B (2006) Newborn screening may fail to identify intermediate forms of maple syrup urine disease. *J Inherit Metab Dis* **29**: 586.

Boneh A, Andresen BS, Gregersen N, et al (2006) VLCAD deficiency: pitfalls in newborn screening and confirmation of diagnosis by mutation analysis. *Mol Genet Metab* **88**: 166–170.

Carpenter K, Wiley V, Wilcken B (2006) The impact of expanded screening on diagnostic and clinical biochemical genetics services in New South Wales, Australia. *J Inherit Metab Dis* **29**: 17.

Chace DH, Millington DS, Terada N, Kahler SG, Roe CR, Hofman LF (1993) Rapid diagnosis of phenylketonuria by quantitative analysis for phenylalanine and tyrosine in neonatal blood spots by tandem mass spectrometry. *Clin Chem* **39**: 66–71.

Chace DH, Hillman SL, Millington DS, Kahler SG, Roe CR, Naylor EW (1995) Rapid diagnosis of maple syrup urine disease in blood spots from newborns by tandem mass spectrometry. *Clin Chem* **41**: 62–68.

Frankenburg WK, Chen JH, Thornton SM (1988) Common pitfalls in the evaluation of developmental screening tests. *J Pediatr* **113**: 1110–1113.

Frazier DM, Millington DS, McCandless SE, et al (2006) The tandem mass spectrometry newborn screening experience in North Carolina: 1997–2005. *J Inherit Metab Dis* **29**: 76–85.

Garg U, Dasouki M (2006) Expanded newborn screening of inherited metabolic disorders by tandem mass spectrometry: clinical and laboratory aspects. *Clin Biochem* **39**: 315–332.

Gelb MH, Turecek F, Scott CR, Chamoles NA (2006) Direct multiplex assay of enzymes in dried blood spots by tandem mass spectrometry for newborn screening of lysosomal disorders. *J Inherit Metab Dis* **29**: 397–404.

Grosse SD, Khoury MJ, Greene CL, Crider KS, Pollitt RJ (2006) The epidemiology of medium chain acyl-CoA dehydrogenase deficiency: an update. *Genet Med* **8**: 205–212.

Hardy DT, Hall SK, Preece MA, Green A (2002) Quantitative determination of plasma phenylalanine and tyrosine by electrospray ionization tandem mass spectrometry. *Ann Clin Biochem* **39**: 73–75.

Hoffmann GF, von Kries R, Klose D, et al (2004) Frequencies of inherited organic acidurias and disorders of mitochondrial fatty acid transport and oxidation in Germany. *Eur J Pediatr* **163**: 76–80.

Kaye CI, Accurso F, La FS, et al (2006) Introduction to the newborn screening fact sheets. *Pediatrics* **118**: 1304–1312.

Koeberl DD, Millington DS, Smith WE, et al (2003) Evaluation of 3-methylcrotonyl-CoA carboxylase deficiency detected by tandem mass spectrometry newborn screening. *J Inherit Metab Dis* **26**: 25–35.

Meikle PJ, Grasby DJ, Dean CJ, et al (2006) Newborn screening for lysosomal storage disorders. *Mol Genet Metab* **88**: 307–314.

Millington DS, Kodo N, Norwood DL, Roe CR (1990) Tandem mass spectrometry: a new method for acylcarnitine profiling with potential for neonatal screening for inborn errors of metabolism. *J Inherit Metab Dis* **13**: 321–324.

Naylor EW, Chace DH (1999) Automated tandem mass spectrometry for mass newborn screening for disorders in fatty acid, organic acid, and amino acid metabolism. *J Child Neurol* **14**(Supplement 1): S4–S8.

Nennstiel-Ratzel U, Arenz S, Maier EM, et al (2005) Reduced incidence of severe metabolic crisis or death in children with medium chain acyl-CoA dehydrogenase deficiency homozygous for c.985A>G identified by neonatal screening. *Mol Genet Metab* **85**: 157–159.

Pedersen CB, Bischoff C, Christensen E, et al (2006) Variations in IBD (ACAD8) in children with elevated C4-carnitine detected by tandem mass spectrometry newborn screening. *Pediatr Res* **60**: 315–320.

Pollitt R (2006) International perspectives on newborn screening. *J Inherit Metab Dis* **29**: 390–396.

Rashed MS, Ozand PT, Bucknall MP, Little D (1995) Diagnosis of inborn errors of metabolism from blood spots by acylcarnitines and amino acids profiling using automated electrospray tandem mass spectrometry. *Pediatr Res* **38**: 324–331.

Rhead WJ (2006) Newborn screening for medium-chain acyl-CoA dehydrogenase deficiency: a global perspective. *J Inherit Metab Dis* **29**: 370–377.

Rhead WJ, Allain D, Van Calcar S, et al (2002) Short chain acyl-CoA dehydrogenase (SCAD) and 3-methylcrotonyl-CoA caboxylase (MCC) deficiencies: tandem mass spectrometry newborn screening detects many clinically benign cases. *J Inherit Metab Dis* **25**: 4.

Schulze A, Mayatepek E, Hoffmann GF (2002) Evaluation of 6-year application of the enzymatic colorimetric phenylalanine assay in the setting of neonatal screening for phenylketonuria. *Clin Chim Acta* **317**: 27–37.

Schulze A, Lindner M, Kohlmuller D, Olgemoller K, Mayatepek E, Hoffmann GF (2003) Expanded newborn screening for inborn errors of metabolism by electrospray ionization-tandem mass spectrometry: results, outcome, and implications. *Pediatrics* **111**: 1399–1406.

Simon E, Fingerhut R, Baumkotter J, Konstantopoulou V, Ratschmann R, Wendel U (2006) Maple syrup urine disease: favourable effect of early diagnosis by newborn screening on the neonatal course of the disease. *J Inherit Metab Dis* **29**: 532–537.

van Maldegem BT, Duran M, Wanders RJ, et al (2006) Clinical, biochemical, and genetic heterogeneity in short-chain acyl-coenzyme A dehydrogenase deficiency. *JAMA* **296**: 943–952.

Waisbren S, Albers S, Amato S, et al (2003) Effect of expanded newborn screening for biochemical genetic disorders on child outcomes and parental stress. *JAMA* **290**: 2564–2572.

Wilcken B (2006) Mini-Symposium: Newborn screening for inborn errors of metabolism—clinical effectiveness. *J Inherit Metab Dis* **29**: 366–369.

Wilcken B, Wiley V, Hammond J, Carpenter K (2003) Screening newborns for inborn errors of metabolism by tandem mass spectrometry. *N Engl J Med* **348**: 2304–2312.

Wilcken B, Haas M, Joy P, et al (2007) Outcome of neonatal screening for medium-chain acyl-CoA dehydrogenase deficiency: a cohort study. *Lancet* **369**: 37–42.

Wiley V, Carpenter K, Wilcken B (1999) Newborn screening with tandem mass spectrometry: 12 months' experience in NSW Australia. *Acta Paediatrica Supplement* **88**: 48–51.

Ziadeh R, Hoffman EP, Finegold DN, et al (1995) Medium chain acyl-CoA dehydrogenase deficiency in Pennsylvania: neonatal screening shows high incidence and unexpected mutation frequencies. *Pediatr Res* **37**: 675–678.

Zytkovicz TH, Fitzgerald EF, Marsden D, et al (2001) Tandem mass spectrometric analysis for amino, organic, and fatty acid disorders in newborn dried blood spots: a two-year summary from the New England Newborn Screening Program. *Clin Chem* **47**: 1945–1955.

J Inherit Metab Dis (2008) 31:173–177
DOI 10.1007/s10545-008-0843-8

SSIEM SYMPOSIUM 2007

The consequences of extended newborn screening programmes: Do we know who needs treatment?

B. Wilcken

Received: 12 December 2007 / Submitted in revised form: 12 January 2008 / Accepted: 16 January 2008 /
Published online: 22 February 2008

Summary The development of an evidence base for newborn screening is especially difficult because of the rarity of disorders now detectable. One consequence of expanded newborn screening is that physicians are being called upon to manage asymptomatic babies with persistent biochemical disturbances that indicate likely enzyme deficiencies. Some of these may be very mild. There is not always agreement as to who should be treated. Particular problems are seen with disorders that were previously thought very rare but are now found frequently by newborn screening. Some of these disorders appear benign or nearly so, and in the present state of knowledge should clearly not be included in routine newborn screening panels.

Abbreviations

MCAD	medium-chain acyl-CoA dehydrogenase
3MCCC	3-methylcrotonyl-CoA carboxylase
PKU	phenylketonuria

Communicating editor: Rodney Pollitt

Competing interests: None declared

Presented at the Annual Symposium of the SSIEM, Hamburg, 4–7 September 2007.

B. Wilcken (✉)
Biochemical Genetics and Newborn Screening,
The Children's Hospital at Westmead,
Hawkesbury Road, Westmead,
NSW 2145, Australia
e-mail: bridgetw@chw.edu.au

B. Wilcken
The University of Sydney, Sydney,
Australia

Introduction

Since the late 1990s newborn screening has been transformed by the introduction of tandem mass spectrometry, which enables multiplex testing. It is now possible to detect over 50 disorders simultaneously from a 3 mm dried blood spot, in around 2 minutes per sample. As a consequence, for the first time very rare disorders can be included in a newborn screening programme, at virtually no extra laboratory cost, something that would not be possible if each needed a stand-alone test. Different jurisdictions have reacted very differently to the new capability, but most have thought there were good reasons to include a fairly wide range of disorders in the screening panel (Pollitt 2007). The possibilities for avoidance of unnecessary morbidity and mortality are great, and are now being well documented for a number of the diseases. However, another consequence is that physicians are being called upon to manage asymptomatic babies with persistent biochemical disturbances that indicate likely enzyme deficiencies—some of which have previously been described only a handful of times. This produces problems, and it is hard for health professionals and parents to know where to turn for fully informed guidance. This essay, which is not a comprehensive review, explores the problems that have arisen when it is not entirely clear which babies detected by newborn screening actually need treatment.

The most-quoted principles for screening are those in the 1968 WHO-sponsored paper of Wilson and Jungner (Wilson and Jungner 1968). While this was written largely in the pre-newborn screening era, screening for phenylketonuria is considered, although the main part of the paper takes into account screening

for infectious disease and cancer, as well as hypertension and other chronic problems. This document has stood the test of time but it is not now completely applicable throughout newborn screening. Nevertheless it does address all of the important aspects for consideration—features about the disorder under consideration, and the test itself, and the availability of facilities to make a final diagnosis and to treat identified subjects, and that there needs to be, 'an agreement on whom to treat as a patient'. This is where some programmes have found most difficulty. Who should be 'treated as a patient'? For most disorders there is no watertight answer; heterogeneity is the norm for natural history and clinical presentation of most inborn errors of metabolism. Because of the rarity of most disorders included in screening programmes, we still often need to rely on low orders of evidence—mainly 'expert opinion'—to form a judgement of what to do. It is possible to examine this problem by considering the disorders detectable by expanded newborn screening in just a few categories.

Patients who need active management

Firstly, those for whom there is a surety of the need for an active treatment or management programme: this group comprises patients with symptoms already present at diagnosis (for example, often the classic organic acidurias) and patients with disorders for which there is strong evidence that they will lead to preventable pathology. Phenylketonuria (PKU), glutaric aciduria type I (GA I), homocystinurias, maple syrup urine disease (MSUD), tyrosinaemia type I, and most fatty acid oxidation disorders fall into this category. There is, for example, ample evidence that early treatment for homocystinuria due to cystathionine synthase deficiency diagnosed by newborn screening is beneficial (Yap and Naughten 1998; Yap et al 2001), and a randomized controlled trial, even if feasible with the small numbers, would not be ethical. It was surprising that this disorder was not thought worthy of inclusion in one of the two technology assessments carried out in the United Kingdom (Seymour et al 2001).

In relation to those patients presenting with symptoms in the first days of life, we studied 2 million babies in Australia, born between 1994 and 2002, of whom 460 000 were screened by tandem mass spectrometry. The part of this study relating solely to medium-chain acyl-CoA dehydrogenase (MCAD) deficiency has already been published (Wilcken et al 2007). Overall, excluding PKU, 179 children, screened and unscreened, were found to have a disorder detectable by newborn screening, and 43 had symptoms by 5 days of age, of whom 17 had died by day 7 (see Table 1). There is no doubt that all such children will receive treatment, and that many do well despite an early presentation. Newborn screening per se is very unlikely to assist babies who die in the first week. However, it may provide a diagnosis in some cases where an inborn error had not been suspected in such a child who died. Screening may hasten the diagnosis for other early presenters, especially in regions where metabolic screening tests are not high on the list of priorities.

A second category comprises patients with disorders that are known to pose a risk for disease, but with reduced penetrance. Examples are MCAD deficiency and some organic acidurias such as 3-ketothiolase deficiency. Patients in this category need careful consideration, as inevitably some will have treatment imposed for which, in hindsight (never available to the treating physician), there will have been no need. After PKU, MCAD deficiency is the disorder most frequently detected by tandem mass spectrometry and exemplifies well the not unexpected problems of increased diagnosis by screening programmes. It has been well known for some time that far more patients with MCAD deficiency are diagnosed by screening

Table 1 Babies with inborn errors of metabolism detectable by tandem mass spectrometry newborn screening who developed symptoms in the first 5 days of life. The total cohort comprised 2 million babies born between 1998 and 2002, 640 000 screened

Class of disorder	Total diagnosed	Presented with symptoms by day 5	Died by day 7
Organic aciduria	51	11	3
Aminoacidurias	17	0	0
Urea cycle defect[a]	28 (14 OTC)	19	8 (6 OTC)
MCAD deficiency	59	8	3
Other FAOD	22	5	3
All disorders	177	43	17

[a] This includes ornithine transcarbamylase deficiency since 2 were detected by newborn screening, results coming after clinical presentation.

than by clinical presentation, and that the 'excess' patients may be dead, symptomatic but undiagnosed, or completely asymptomatic. Additionally, the finding of mutations never discovered in clinically presenting patients suggests that some with an undoubted degree of biochemical MCAD deficiency may never be at risk at all (Andresen et al 2001). In our first 50 MCAD patients detected by newborn screening 25 (50%) were homozygous for the common, severe mutation c.985A>G, whereas around 80% would have been expected from published data of clinically presenting cases (Tanaka et al 1992). Seven (14%) had the genotype c.985A>G/c.199T>C, discovered frequently in screening programmes (Andresen et al 2001) but not documented to have been found in clinically presenting cases. A further 5 of our 50 cases did not carry any copy of the common mutation. Some of this data has been published previously (Waddell et al 2005).

There is insufficient evidence at present to act on this, but at least for MCAD deficiency a special diet is not needed and the management is not unduly burdensome. Possibly enough evidence will emerge after several years of surveillance to indicate that babies found by newborn screening with certain genotypes such as c.985A>G/c.199T>C have a vanishingly small risk of decompensation, but as their newborn screening biochemical findings do overlap with those of homozygotes for the severe and common c.985A>G mutation they will probably continue to be reported by newborn screening programmes and physicians will feel bound to institute a management programme of some sort. It is not clear how soon this conundrum could be resolved.

Patients who may not need any management: Disorders that should possibly be eliminated from screening programmes

There are newborns who do have a biochemically recognizable disorder but who should probably not be treated and indeed probably not be detected. The disorders in this category are certainly controversial, so this poses a large problem for some screening programmes. Some disorders have been described in only one or a very few patients and the phenotype is really not known. Dienoyl-CoA reductase deficiency, with a single description in 1990, is a case in point. In the American College of Human Genetics Guidelines (American College of Medical Genetics 2006) it was recommended for screening as a secondary target, although the marker compound $C_{10:2}$ is not used to detect any of the core recommended disorders. The evidence level cited for the inclusion of this disorder in the screening panel was the highest (4/4). This disorder should probably not be sought, unless for a defined research project.

Other disorders are those that were considered extremely rare before newborn screening was undertaken but are now found in numbers of asymptomatic babies or sometimes their mothers. A prime example is 3-methylcrotonyl-CoA carboxylase (3MCCC) deficiency. This disorder of leucine metabolism, very rarely reported before expanded newborn screening, is the commonest organic aciduria found by screening, with a birth incidence of about. 1:50 000. Screening also detects asymptomatic affected mothers. We investigated the initial 8 mothers and 5 infants detected by our New South Wales newborn screening. All have been asymptomatic. Stadler considered that this disorder 'largely presents as non-disease' and that 'only a few individuals (<10%) develop symptoms, ...none reported so far can clearly be attributed to 3MCCC' (Stadler et al 2006). Recent reports do indeed show a wide range of possible symptoms including severe hypoglycaemia (Bartlett et al 1984; Ficicioglu and Payan 2006; Oude Luttikhuis et al 2005), metabolic stroke (Pinto et al 2006; Steen et al 1999) hypotonia (Elpeleg et al 1992), hypertonia and failure to thrive (Tuchman et al 2007), severe muscle pain in an adult (Boneh et al 2005), cardiomyopathy with asymptomatic affected family members (Visser et al 2000), necrotizing encephalopathy (Baykal et al 2005), leukodystrophy (de Kremer et al 2002), progressive respiratory failure (Wisemann et al 1998), and developmental delay (Yap et al 1998). There is also report of asymptomatic children, pre-dating newborn screening (Mourmans et al 1995; Pearson et al 1995), and asymptomatic adults detected by newborn screening of their children (Gibson et al 1998). Considering that this is now known to be a not uncommon disorder, particularly in some areas such as Taiwan, where it occurs in 1:20 000 newborns, and that it would have been nearly impossible to miss the diagnosis on urinary organic acid analysis, it is very likely that many of these reported patients had other unrelated problems.

It is therefore not clear how babies diagnosed with 3MCCC deficiency by newborn screening should be managed. Should they have no treatment at all? simple advice about intercurrent illness and what to do? just carnitine medication, perhaps plus mild protein restriction? or even this plus a leucine-free formula? We, in New South Wales, have opted for simple low-key advice about intercurrent illness.

Suggested treatment by The Screening, Technology and Research in Genetics (STAR-G) Project, supported

by the US Health Resources and Services Administration (HRSA) includes:

> Low-leucine diet... most of the food will be carbohydrate; ....Some children may be on a special food plan throughout life.

- In addition to a low-protein diet, some children are given a special leucine-free medical formula. Your metabolic doctor and dietician will decide whether your child needs this formula. Some states offer help with payment, or require private insurance to pay for the formula and other special medical foods.
- 'Even with treatment, some children still have repeated bouts of metabolic crisis. This can cause brain damage and may lead to life-long learning problems or mental retardation.

(http://www.newbornscreening.info/Parents/organicaciddisorders/3MCC.html; accessed 12/12/2007).

Introducing a very restrictive diet with little evidence for doing so has a serious impact on children's lives. It is questionable whether this disorder should be included in newborn screening panels. As several other disorders are detected by the marker analyte in question, C_5OH (3-hydroxyisovalerylcarnitine or its isomers), it is important for a close analysis to be carried out of the benefits and drawbacks for including this analyte at all. Germany has decided not to include it (Pollitt 2007).

Since the presentation on which this essay was based, and its first submission, an excellent paper has been published describing a Delphi-based consensus clinical practice guideline for the management of 3MCCC deficiency, and readers are advised also to refer to that (Arnold et al 2008).

The same problem has arisen with short-chain acyl-CoA dehydrogenase (SCAD) deficiency—also found much more commonly by screening programmes. However, there is considerable concern that this is largely a benign disorder (van Maldegem et al 2006). Exclusion of the marker butyrylcarnitine (C_4) would also exclude detection of isobutyryl-CoA dehydrogenase deficiency, a disorder that is of unknown significance but also apparently benign, and little known before screening was undertaken. These are conditions for which there is no evidence at present to warrant their inclusion in screening panels.

These problems encountered with screening programmes are not new. There are many lessons from phenylketonuria screening. Benign phenylketonuria is now well recognized. Babies and children with persistently but mildly raised levels of phenylalanine, never exceeding, say, 400 μmol/L, who need no dietary restriction and suffer no intellectual problems were indeed subjected to intrusive dietary restriction when screening started. On the other side of the coin, of course, were patients (some with only mildly elevated phenylalanine levels) who did badly despite treatment, developed severe neurological problems and died. These were found to have defects affecting the pterin co-factor pathway, defects unknown before PKU screening was undertaken.

Conclusions

There is much knowledge now about treatment of inborn errors of metabolism, and there are methods for monitoring the successes of treatment. But it is very important to ensure as far as possible that only those at risk will be treated, or good outcomes will be assumed—wrongly—in patients who never needed interfering treatment in the first place. The aim of newborn screening is primarily to produce a good clinical outcome for babies by early diagnosis of treatable disorders, and facilitation of appropriate treatment. Additionally, as discussed above, there must be avoidance of unnecessary 'medicalization' of subjects with no significant disease. Thus, the main responsibility for those who design and run screening programmes is to ensure a suitably conservative approach, close follow-up, careful data collection and, most importantly, the courage to stop any part of a programme if that becomes indicated. We have a way to go.

References

American College of Medical Genetics (2006) Newborn screening: towards a uniform screening panel and system. *Genet Med* **8**: 1S–252S.

Andresen BS, Dobrowolski SF, O'Reilly L, et al (2001) Medium-chain acyl-CoA dehydrogenase (MCAD) mutations identified by MS/MS-based prospective screening of newborns differ from those observed in patients with clinical symptoms: identification and characterization of a new, prevalent mutation that results in mild MCAD deficiency. *Am J Hum Genet* **68**: 1408–1418.

Arnold GL, Koeberl DD, Matern D, et al (2008) A Delphi-based consensus clinical practice protocol for the diagnosis and management of 3-methylcrotonyl CoA carboxylase deficiency. *Mol Genet Metab* [In press].

Bartlett K, Bennett MJ, Hill RP, Lashford LS, Pollitt RJ, Worth HG (1984) Isolated biotin-resistant 3-methylcrotonyl CoA carboxylase deficiency presenting with life-threatening hypoglycaemia. *J Inherit Metab Dis* **7**: 182.

Baykal T, Gokcay GH, Ince Z, et al (2005) Consanguineous 3-methycrotonyl-CoA carboxylase deficiency: early-onset

necrotizing encephalopathy with lethal outcome. *J Inherit Metab Dis* **28**: 229–233.

Boneh A, Baumgartner M, Hayman M, Peters H (2005) Methylcrotonyl-CoA carboxylase (MCC) deficiency associated with severe muscle pain and physical disability in an adult. *J Inherit Metab Dis* **28**: 1139–1140.

de Kremer RD, Latini A, Suormala T, et al (2002) Leukodystrophy and CSF purine abnormalities associated with isolated 3-methylcrotonyl-CoA carboxylase deficiency. *Metab Brain Dis* **17**: 13–18.

Elpeleg ON, Havkin S, Barash V, Jakobs C, Glick B, Shalev RS (1992) Familial hypotonia of childhood caused by isolated 3-methylcrotonyl-coenzyme A carboxylase deficiency. *J Pediatr* **121**: 407–410.

Ficicioglu C, Payan I (2006) 3-Methylcrotonyl-CoA carboxylade deficiency: metabolic decompensation in a noncompliant child detected through newborn screening. *Pediatrics* **118**: 2555–2556.

Gibson KM, Bennett MJ, Naylor EW, Morton DH (1998) 3-Methylcrotonyl-coenzyme A carboxylase deficiency in Amish/Mennonite adults identified by detection of increased acylcarnitines in blood spots of their children. *J Pediatr* **132**: 519–523.

Mourmans J, Bakkeren J, de Jong J, et al (1995) Isolated (biotin-resistant) 3-methylcrotonyl-CoA carboxylase deficiency: four sibs devoid of pathology. *J Inherit Metab Dis* **18**: 643–645.

Oude Luttikhuis HG, Touati G, Rabier D, Williams M, Jakobs C, Saudubray JM (2005) Severe hypoglycaemia in isolated 3-methylcrotonyl-CoA carboxylase deficiency; a rare, severe clinical presentation. *J Inherit Metab Dis* **28**: 1136–1138.

Pearson MA, Aleck KA, Heidenreich RA (1995) Benign clinical presentation of 3-methylcrotonylglycinuria. *J Inherit Metab Dis* **18**: 640–641.

Pinto L, Zen P, Rosa R, et al (2006) Isolated 3-methylcrotonyl-coenzyme A carboxylase deficiency in a child with metabolic stroke. *J Inherit Metab Dis* **29**: 205–206.

Pollitt RJ (2007) Introducing new screens: why are we all doing different things? *J Inherit Metab Dis* **30**: 423–429.

Seymour CA, Thomason MJ, Chalmers RA, et al (2001) Newborn screening for inborn errors of metabolism: a systematic review. *Health Technol Assess* **1**: 23–26.

Stadler SC, Polanetz R, Maier EM, et al (2006) Newborn screening for 3-methylcrotonyl-CoA carboxylase deficiency: population heterogeneity of MCCA and MCCB mutations and impact on risk assessment. *Hum Mutat* **27**: 748–759.

Steen C, Baumgartner ER, Duran M, et al (1999) Metabolic stroke in isolated 3-methycrotonyl-CoA carboxylase deficiency. *Eur J Pediatr* **158**: 730–733.

Tanaka K, Yokota I, Coates PM, et al (1992) Mutations in the medium chain acyl-CoA dehydrogenase (MCAD) gene. *Hum Mutat* **1**: 271–279.

Tuchman M, Berry SA, Thuy LP, Nyhan WL (2007) Partial methylcroyonyl-coenzyme A carboxylase deficiency in an infant with failure to thrive, gastrointestinal dysfunction, and hypertonia. *Pediatrics* **1993**: 664–666.

van Maldegem BT, Duran M, Wanders RJ, et al (2006) Clinical, biochemical, and genetic heterogeneity in short-chain acyl-coenzyme A dehydrogenase deficiency. *JAMA* **296**: 943–952.

Visser G, Suormala T, Smit GP, et al (2000) 3-methylcrotonyl-CoA carboxylase deficiency in an infant with cardiomyopathy, in her brother with developmental delay and in their asymptomatic father. *Eur J Pediatr* **159**: 901–904.

Waddell L, Wiley V, Carpenter KH, et al (2005) Medium-chain acyl-CoA dehydrogenase deficiency: genotype–biochemical phenotype correlations. *Mol Genet Metab* **87**: 32–39.

Wilcken B, Haas M, Joy P, et al (2007) Outcome of neonatal screening for medium-chain acyl-CoA dehydrogenase deficiency: a cohort study. *Lancet* **369**: 37–42.

Wilson JMG, Jungner G (1968) *Principles and Practice of Screening for Disease*, Geneva: World Health Organization.

Wisemann UN, Suormala T, Pfenninger J, Baumgartner ER (1998) Partial 3-methycrotonyl-CoA carboxylase deficiency in an infant with fatal outcome due to progressive respiratory failure. *Eur J Pediatr* **157**: 225–229.

Yap S, Monavari AA, Thornton P, Naughten E (1998) Late-infantile 3-methylcrotonyl-CoA carboxylase deficiency presenting as global developmental delay. *J Inherit Metab Dis* **21**: 175–176.

Yap S, Naughten E (1998) Homocystinuria due to cystathionine beta-synthase deficiency in Ireland: 25 years' experience of a newborn screened and treated population with reference to clinical outcome and biochemical control. *J Inherit Metab Dis* **21**: 738–747.

Yap S, Rushe H, Howard PM, Naughten ER (2001) The intellectual abilities of early-treated individuals with pyridoxine-nonresponsive homocystinuria due to cystathionine beta-synthase deficiency. *J Inherit Metab Dis* **24**: 437–447.

John Fernandes
Jean-Marie Saudubray
Georges van den Berghe
John H. Walter (Editors)

Inborn Metabolic Diseases

Diagnosis and Treatment

Fourth, Revised Edition

With 65 Figures and 63 Tables

Springer

2 Newborn Screening for Inborn Errors of Metabolism

Bridget Wilcken

2.1 Introduction

Newborn screening was first applied to the detection of phenylketonuria (PKU) by a bacterial inhibition assay pioneered in 1961 by Guthrie, who was also responsible for the introduction of the use of a dried blood sample [1]. This was followed by further bacterial inhibition assays to detect other aminoacidopathies (maple syrup urine disease, homocystinuria, urea cycle disorders and so on) but only screening for PKU was widely adopted. In 1975 Dussault described screening for congenital hypothyroidism (CH) [2], and since then other disorders covered in some screening programmes have included congenital adrenal hyperplasia, the galactosaemias, cystic fibrosis, biotinidase deficiency, glucose-6-phosphate dehydrogenase deficiency and many others. The application of tandem mass spectrometry to newborn screening was first described in 1990 [3]. This new technology has greatly changed both newborn screening and the diagnosis of many inborn errors of metabolism.

2.2 General Aspects of Newborn Screening

2.2.1 Aims and Criteria

The initial aim of newborn screening was to identify infants with serious but treatable disorders, so as to facilitate interventions to prevent or ameliorate the clinical consequences of the disease. In recent years, with the advent of tandem mass-spectrometry which can detect many disorders at one time, and hence the ability for early detection of currently untreatable disorders (► below), there has been discussion about how the aims of screening might be widened to encompass a benefit to families, rather than individual babies.

The classic criteria for screening are those of Wilson and Jungner [4]. More recently the World Health Organisation has published guidelines [5] as has the United States [6], and the United Kingdom National Screening Committee has extended the Wilson and Jungner criteria [7]. In reality the criteria can be simplified and reduced to two main considerations which would justify screening for any specific disorder: there should be a benefit from neonatal detection, and the overall benefit should be reasonably balanced by the costs of all kinds: the financial costs (opportunity costs) and the cost of the harm, if any, to individuals by early detection of the disorder, or false assignment of a positive or negative result. It is important to remember that newborn screening covers the whole process from sampling to the appropriate referral of an affected baby for the start of treatment, and assessment of overall outcome.

2.2.2 Sensitivity, Specificity, and Positive Predictive Value

In assessing screening tests and understanding the screening process, some definitions are important:
Sensitivity: The proportion of subjects with the disorder in question detected by the test.
Specificity: The proportion of subjects without the disorder that have a negative test result.
False negative rate: The percentage of affected subjects not detected by the test.
False positive rate: The percentage of healthy subjects with a positive test result.
Positive predictive value: The chance that a positive result actually indicates an affected individual. Similarly, a *negative predictive value* is the chance that a negative result excludes the disorder. These values depend not only on the specificity or sensitivity of the test, but also on the frequency of the disorder.

The sensitivity of a test depends to a large extent on chosen cut-off values, and is a balancing act: the higher the sensitivity the lower the specificity. The sensitivity and specificity of course will vary according to decisions about cut-off points and what range of false positive and negative results can be tolerated in a particular programme.

2.2.3 Technical Aspects of Newborn Screening Tests

Blood-collection-paper Samples

Newborn screening tests are mainly carried out on blood spots dried on specially manufactured filter paper, usually obtained by heel-stick. Although various forms of venous blood sampling may also be used, filter paper blood samples will be considered here.

Methods

A wide variety of technologies can be applied to filter paper samples, including bacterial inhibition assays, chromatographic techniques, enzyme-linked immunosorbent assays (ELISA), fluorescent immunoassays (FIA), radioimmunoassays (RIA), and most recently electrospray ionisation tandem mass spectrometry (MSMS). Methods for newborn screening prior to MSMS screening are well-described by Therrell [8]. DNA analysis is also performed as a part of some screening tests. The methodology to some extent varies with the analyte of interest (for example, hormone analyses are usually immunoassays), and cost, sensitivity and specificity vary according to the method used. While quantitative results are obtained, the precision of tests using a paper sample is less than for a plasma or serum sample because of the matrix, collection process, haematocrit variations and so forth.

Timing of the Test

The timing of the test also has an important influence on the results. A newborn during the first 72 h is catabolic to some extent, and this is very useful for detecting disorders of intermediary metabolism. There is a paucity of data, except for PKU and CH screening, about results of screening in the first 24 h of life, and this is generally not recommended. Cut-off values adopted to indicate a positive test result for the different analytes will vary with the age of the child at screening, and the sensitivity of a test for detecting certain disorders may also vary according to age.

Cut-off Values

Determination of the cut-off point for each analyte is always a compromise between the aim for perfect sensitivity (detecting all the cases) and keeping the false negative rate as low as possible. It is important for the laboratory to establish age-dependent cut-off values, as these can vary greatly. In general, the cut-points for amino acids and free carnitine may be lower for babies aged, say, 48–72 h than for babies aged 7–10 days. The converse is true for acylacarnitines, where levels decrease with increasing age, and use of a 2–3 day cut-point at 7–10 days could lead to a case being missed. Physicians must bear in mind that no screening test ever gives a perfect performance, although some may come close. If clinical presentation suggests a disorder which is included in newborn screening, a test should be done, even if the screening test was negative.

2.2.4 Range of Possibilities from Early Detection

Newborn screening has opened new perspectives in preventive medicine. Babies with disorders of amino acid, organic acid, and fatty acid metabolism are now often detected in the newborn screening laboratory, rather than by the clinical metabolic service. Early detection provides 3 possibilities:

1. The disorder may present in the first days of life, before any newborn screening result is likely. Disorders in this category include neonatal presentations of urea cycle defects, organic acidaemias such as methylmalonic acidaemia, and less commonly, almost any of the fatty acid oxidation defects. Detection by newborn screening is unlikely to benefit directly most cases in this category. However, it seems appropriate to include these early-presenting disorders in the screening suite, as some may have delayed diagnosis, and on occasion a diagnosis may never be made – the baby having been thought to have died from sepsis.
2. The disorder may be later presenting, and an effective treatment can beneficially alter the natural history. Cases in this category include the less severe urea cycle disorders, most aminoacidopathies, such as homocystinuria, some organic acidaemias, and most fatty acid oxidation disorder cases.
3. The disorder may be benign, or largely so, and most cases will have no benefit from early diagnosis. It is hard to know yet which cases will fit into this category. If that was clear, then the disorders could be removed from the screening suite, but newborn screening, if carefully and sensitively conducted, provides an excellent opportunity for elucidating the natural history of disorders which might or might not fall into this category. What is clear is that mild forms of several disorders will readily be detected by newborn screening, but may not need treatment. One example is mild citrullinaemia (argininosuccinate synthase deficiency) [9].

2.2.5 Tandem Mass Spectrometry

MSMS has so revolutionised newborn screening that a brief general description of the method is appropriate here.

Methods

MSMS is a method of measuring analytes by both mass and structure. After ionisation of the compounds, the first mass spectrometer selects ions of interest, sorted by weight. They pass into a collision cell, are dissociated to »signature« fragments, and then pass into a second mass spectrometer where ions are selected for detection. The two mass spectrometers are linked by computer. Only selected ions can pass through the assembly – the others are deflected. Two important points to note are, firstly, the ions to be analysed can be selected by the operator, using the multiple reaction monitoring (MRM) mode, and thus in most cases it is possible to select what disorders might be detected. Secondly, using both mass and structure gives high specificity, except in the case of isomers. The initial methods described for aminoacid and acylcarnitine analysis used derivatisation of samples to form either methyl or butyl esters [10]. Most newborn screening programmes use derivatisation but with modern instruments adequate sensitivity can be achieved without derivatising, simplifying sample preparation. For some analytes, elevations may indicate more than one disorder and the use of ratios of analytes improves sensitivity and specificity [11]. Cut-off values need to be determined for each laboratory, and actual values are not discussed here. Confirmation of a disorder always requires appropriate follow-up tests in a specialised biochemical genetics laboratory. These follow-up tests may include aminoacid analysis, organic acid analysis by gas chromatography/mass spectrometry, and plasma acylcarnitine profile by MSMS.

Range of Disorders Detectable

Current screening by MSMS involves the measurement of aminoacids and acylcarnitines, to detect selected disorders of aminoacid, organic acid, and fatty acid metabolism. It is

Table 2.1. Reliability of detection by tandem mass spectrometry (MSMS) of some inborn errors

High	Probably high	Uncertain	Unreliable
Phenylketonuria	Many other organic acidaemias	LCHAD	Homocystinuria (CBS): B6-responsive cases probably not detected
Tyrosinaemia type II	VLCAD	CPT I	Fumaryl acetoacetase deficiency: requires succinyl acetone assay
MSUD classical	CACT/neonatal CPT II	SCAD	Glutaric aciduria I: low excretors may be missed
ASS; ASL			Carnitine transporter
PA, MMA, IVA neonatal			MSUD variant
MCAD	MADD		OTC, CPS, NAGS

MSUD, maple syrup urine disease; *ASS*, argininosuccinate synthase deficiency; *ASL*, argininosuccinate lyase deficiency; *PA*, propionic acidaemia; *MMA*, methylmalonic acidaemia; *IVA*, isovaleric acidaemia; *SCAD, MCAD, VLCAD, LCHAD, MADD*, short-chain, medium-chain, very long-chain, long-chain 3-hydroxy and multiple acyl-CoA dehydrogenase deficiencies; *CACT*, carnitine acylcarnitine transporter defect; *CPT I* and *II*, carnitine palmitoyl transferase deficiencies; *CBS*, cystathionine beta-synthase deficiency; *B6*, pyridoxine; *OTC, CPS, NAGS*, ornithine transcarbamylase, carbamyl phosphate synthetase and N-acetyl glutamate synthetase deficiencies.

still too early to be sure of the reliability of MSMS for the detection of the rarer disorders. ◘ Table 2.1 gives an estimate, based on current knowledge, of the reliability of detection by MSMS (when the sample is appropriate). The selection of disorders that should be included is the subject of intense discussion and some disagreement in various screening jurisdictions. Other factors which affect performance and interpretation include the flux through the metabolic pathway, depending on the degree of catabolism present.

Performance and Outcome

There are as yet no comprehensive studies of outcomes for MSMS screening, but these are on the way. A false positive rate of approximately 0.2–0.3% has been achieved by several programmes that have also attained an apparently high sensitivity [11, 12]. While it is clear which disorders selected analytes can indicate, what is not yet clear is the screening performance for each disorder – i.e. the sensitivity and specificity that can be achieved. Moreover, it has emerged, not surprisingly, that screening detects more cases than clinical diagnosis. This is particularly so for medium-chain acyl-CoA dehydrogenase deficiency, to a lesser extent other fatty acid oxidation disorders, and additionally, for some organic acidopathies, previously thought very rare [13, 14]. Some maternal disorders can also be detected by testing of the baby. Reported results from several comprehensive programmes are available [11, 12, 14].

2.3 Screening for Individual Inborn Errors of Metabolism

Well over 40 inborn errors of metabolism can now be detected by newborn screening, with varying degrees of certainty. This section will concentrate on inborn errors that are the province of the metabolic physician.

2.3.1 Phenylketonuria

Test Methods

Screening for PKU has been usual in most developed countries since the late 1960s. The initial test was the »Guthrie test«, a bacterial inhibition assay, and this is still widely used. Alternative methods include those using fluorimetry and calorimetry. More recently, PKU screening has been by MSMS where this is available.

Timing

Screening in the United States may take place after 24 h, but elsewhere 48 h plus is usual. As with other aminoacids, levels of phenylalanine rise steadily during the first days of life, and there is a theoretical risk of missing PKU if the test is conducted too early. However, testing by MSMS and measuring the ratio of phenylalanine to tyrosine overcomes this problem [11, 15].

Reliability

The bacterial inhibition assay is a robust test, and if it is applied to samples taken at 48 h or more, any false negative tests in subjects with classical PKU have almost always been due to errors of process (sample not taken, clerical error etc.) rather than biological variation or problems with the actual test procedure [16]. MSMS screening may be even more reliable, and overall, the sensitivity of all PKU tests is very high indeed. With experience, the specificity is also high. There are no established benchmarks.

Outcome

Patients obtaining good control of phenylalanine levels by 3–4 weeks and maintaining good average control have a good neuropsychological outcome. There are still minor deficits, and *maternal PKU* remains a potential problem. These aspects are reviewed in Chap. 17.

2.3.2 Galactosaemias

Galactose-1-phosphate uridyl transferase (GALT) deficiency, galactokinase deficiency and galactose epimerase deficiency can all be detected by newborn screening.

Test Methods

Methods used in screening are measures of metabolites, galactose and galactose-1-phosphate, or measures of enzyme activity, confined to the Beutler test for GALT. Most commonly nowadays, a metabolite assay is followed by confirmatory testing using a GALT assay and quantitative determination of galactose and galactose-1-phosphate.

Reliability

A combination of these methods will provide a precise diagnosis of transferase deficiency (but ▶ below), and an indication of galactokinase deficiency, in which there is elevation of galactose alone and a normal GALT activity. However, the differentiation of red-cell epimerase deficiency (a benign condition) and systemic epimerase deficiency, clinically similar to transferase deficiency, may not be clear. Additionally, moderate metabolite elevations and severely reduced, but not absent, GALT activity is seen in combined heterozygosity for a severe mutation in the transferase *GALT* gene, and a common »Duarte« mutation. The differentiation is important, as a Duarte/galactosaemia double heterozygote needs no treatment.

Outcome

Despite early identification and treatment, the long-term outcome for transferase deficiency is not particularly good, with about half the children having early intellectual problems, and some evidence of ongoing deterioration in most (▶ Chap. 7). There is no evidence that pre-symptomatic treatment alters outcome, (although death may be avoided in a few) and because of this, not all developed countries screen for the galactosaemias. Treated galactokinase deficiency would be expected to have a good outcome, but is much rarer than transferase deficiency, and systemic epimerase deficiency is rarer still, and little is known of long-term effects of screening.

2.3.3 Aminoacidopathies

■ Table 2.2 shows the amino acids most commonly analysed by tandem mass-spectrometry, the secondary markers in use, and the follow-up tests indicated.

Disorders of the Urea Cycle

Citrullinaemia and *argininosuccinic aciduria*, either severe or later-presenting, can be diagnosed with apparently high sensitivity by measuring citrulline. There are problems because of the recent description of mild, asymptomatic citrullinaemia. Detection of severe, early-presenting *argininaemia*, by measuring arginine, has also been described. *Carbamyl phosphate synthase* and *ornithine transcarbamylase (OTC) deficiencies* cannot be so easily detected. Low citrulline is an indicator, but a low cut-off for citrulline

■ **Table 2.2** The main amino acids analysed in tandem mass-spectrometry for the detection of aminoacidopathies, also showing the possible diagnoses, the secondary markers used and the follow-up tests indicated

Amino acids	Possible diagnoses	Secondary markers	Initial follow-up tests
Arginine (ARG)	Arginase deficiency		PLAA, enzyme analysis
Citrulline (CIT)	ASS, ASL, Citrullinaemia type II	CIT/ARG ratio	PLAA, UAA, LFTs
Leucines (LEUs)	Maple syrup urine disease	LEUs/PHE	PLAA
Methionine (MET)	Cystathionine synthase deficiency, hypermethioninaemias	MET/PHE	PLAA, tHCY
Phenylalanine (PHE)	Phenylketonuria, hyperphenylalaninaemia, pterin disorders	PHE/TYR	PLAA, pterin load, urinary pterins, DHPR
Tyrosine (TYR)	Tyrosinaemias: fumaryl acetoacetase deficiency, tyrosine aminotransferase deficiency, tyrosinaemia III	Succinylacetone	UOAs, PLAA, LFTs etc.

ASS, argininosuccinate synthase deficiency (citrullinaemia type I); *ASL*, argininosuccinate lyase deficiency; *PLAA*, plasma amino acid analysis; *UAA*, urinary amino acid analysis; *UOAs*, urinary organic acid analysis; *LFTs*, liver function tests; *tHCY*, total homocysteine (plasma); *DHPR*, dihydropteridine reductase assay (dried blood spot).

overlaps with low citrulline seen in sick neonates in general. Diagnosis of OTC deficiency has been described by the detection of pyroglutamic acid, derived from glutamine, and a blood-spot method for detecting glutamine has been described [17]. Probably newborn screening is quite unreliable for these early disorders. A related disorder, *citrin deficiency* causing neonatal hepatitis, (citrullinaemia type II) could be detected by a disturbance of several amino acids, especially moderately elevated citrulline, and detection of one case by MSMS has been described. If ornithine is one of the analytes included, *hyperornithinaemia, hyperammonaemia, homocitrullinuria (HHH syndrome)* could theoretically be detected, but no cases have been reported and this analyte is often not included. Hyperornithinaemia due to *ornithine aminotransferase deficiency* probably does not cause elevated ornithine levels in early infancy, and would not be detected [18].

Other Aminoacidopathies

When screening for *phenylketonuria* by MSMS is the method used, some but not all programmes have found it useful to use a phenylalanine/tyrosine ratio for identifying positive results [15]. The *tyrosinaemias* present some problem. In *tyrosinaemia type 1 (fumaryl acetoacetase deficiency)* the blood tyrosine level in newborns is often not high, and there is considerable overlap with transient tyrosinaemia cases. Several cases of tyrosinaemia type 1 have been reported as missed by MSMS screening, and unless a separate assay of succinylacetone is performed as a back-up, it is likely that this disorder is usually not detectable by MSMS without an unacceptable false-positive rate [12]. *Tyrosinaemia type II* is readily detectable, and at least one case of tyrosinaemia type III has been found. *Maple syrup urine disease* is detected by assay of leucines and perhaps valine. As MSMS detection cannot distinguish among isotopes, the leucine peak encompasses isoleucine, alloisoleucine, and hydroxyproline in addition. Classical MSUD can readily be detected (although a positive result could also indicate the benign hyperhydroxyprolinaemia) but it is not clear that all variant cases can be distinguished, even by the use of the leucine/phenylalanine ratio. It goes without saying that a result indicating classic MSUD needs to be handled as an emergency. *Cystathionine synthase deficiency (homocystinuria)* is currently detected by an elevated methionine level. The likelihood is that this test, as with the bacterial inhibition assay for methionine formerly used, will miss most cases of pyridoxine-responsive homocystinuria. It is too early yet to be sure of the sensitivity for non-responsive cases, but it is likely to be good. A back-up assay of homocysteine, which cannot be done using the same method, would be ideal, but no experience of this approach has been reported.

Several other aminoacids can be measured simultaneously by MSMS without altering the method, including ornithine, serine, valine and glutamine, but most have not been reported as useful in newborn screening. We found that elevated alanine was not a useful discriminator for babies later diagnosed with mitochondrial respiratory chain disorders.

2.3.4 Organic Acid Disorders

Organic acids that form acylcarnitines can be detected by MSMS, and a large number of organic acid disorders have been so detected (■ Table 2.3). The classic organic acid disorders, *methylmalonic* (MMA), *propionic* (PA), and *isovaleric acidaemias* (IVA) can readily be detected, although the baby will probably be symptomatic before newborn screening results are available. An elevation of propionylcarnitine (C_3) might indicate either PA, MMA, vitamin B_{12} deficiency secondary to maternal deficiency, or *cobalamin C defect* (methylmalonic aciduria with homocystinuria), or possibly *cobalamin D or F defects*. While severe neonatal onset MMA will have elevated C_3 levels, other defects can be more reliably detected by using the ratio of C_3 to acetylcarnitine. Of special importance is *glutaric aciduria type I* (glutaryl CoA dehydrogenase deficiency). The marker compound, glutarylcarnitine (C_5DC), which is also one of several markers for *glutaric aciduria type II*, may be only marginally elevated, and may not be detected, especially if the infant is sampled after the first week. Again, a ratio, this time with palmitoylcarnitine, is more discriminatory [11].

Newborn screening has uncovered an unexpectedly high frequency of cases of *3-methylcrotonyl CoA carboxylase deficiency* (MCCC), previously thought to be exceptionally rare, and asymptomatic cases of maternal MCCC are also detected regularly by newborn screening [19]. This disorder is one of several that might be benign in most instances. Other maternal disorders uncovered by abnormal results on neonatal screening include mild holocarboxylase synthase deficiency and vitamin B_{12} deficiency.

2.3.5 Fatty Acid Oxidation Disorders

Disorders of carnitine uptake, the carnitine cycle, and mitochondrial beta-oxidation can be detected by MSMS testing of acylcarnitines (■ Table 2.3). For several disorders, newborn screening programmes have detected more cases than have historically presented clinically [12, 13]. While some of these subjects might never have experienced episodes of decompensation it is not possible at present to distinguish who is at most risk, and all have by definition a functional defect in oxidation rates. This is especially true of *medium-chain acyl-CoA dehydrogenase deficiency (MCAD)*, the most frequently occurring fatty acid oxidation disorder, in which the detection rate is nearly doubled. MCAD is reliably indicated by elevated octanoylcarnitine

Table 2.3. The major acylcarnitines measured in newborn screening by tandem mass spectrometry. Also shown are the possible diagnoses, the secondary markers and the initial follow-up tests indicated for confirmation. Other acylcarnitine species are often routinely measured simultaneously with those listed, and give further support to possible diagnoses

Abbreviation	Acyl-carnitine	Possible diagnoses	Secondary markers	Initial follow-up tests
C_0	Carnitine	Low: carnitine transporter defect High: CPT I		PLAC, Fibroblast assay, uptake enzyme
C_2	Acetylcarnitine	No inborn error, (high: ketoacidosis, low: low carnitine status)		
C_3	Propionylcarnitine	Methylmalonic acidaemias (incl CblC) propionic acidaemia, vitamin B_{12} deficiency	C_3/C_2	UOAs, tHCY, enzyme assays, serum B12, infant and mother
C_3DC	Malonylcarnitine	Malonic aciduria		UOAs
C_4	Butyrylcarnitine	SCAD, MADD, isobutyryl-CoA dehydrogenase		PLAC, UOAs
C_5	Isovalerylcarnitine, 3-methylbutyrylcarnitine	Isovaleric acidaemia, or 2-methylbutyryl CoA dehydrogenase deficiency		UOAs
C_6	Hexanoylcarnitine	MCAD, MADD		
C_5-OH	3-Hydroxyisovaleryl carnitine, or 2-methyl-3-hydroxy-butyrylcarnitine	Biotinidase, HMG-CoA lyase, β-ketothiolase 3-methylcrotonyl-CoA carboxylase, or multiple carboxylase deficiencies		UOAs, maternal UOAs
C_8	Octanoylcarnitine	MCAD, MADD	C_8/C_{10}	UOAs, DNA mutation
C_{10}	Decanoylcarnitine	MCAD, MADD		
C_5DC	Glutarylcarnitine	Glutaric acidaemia I		UOAs, enzyme assay
CH_3-C_5DC	3-Methyl-glutarylcarnitine	HMG-CoA lyase deficiency		
$C_{14:1}$	Tetradecenoylcarnitine	VLCAD, LCHAD		PLAC
C_{14}	Tetradecanoylcarnitine	LCHAD, VLCAD		PLAC
C_{16}	Hexadecanoylcarnitine	CPT II, CACT (low: CPT 1)	$C_{16}+C_{18}/C_2$	PLAC
C_{16}-OH	Hydroxyhexadecanoyl-carnitine	LCHAD		PLAC

CPT, carnitine palmitoyltransferase; *CblC*, cobalamin C defect; *DC*, dicarboxylic; *SCAD, MCAD, VLCAD, LCHAD, MADD*, short-chain, medium-chain, very long-chain, long-chain 3-hydroxy and multiple acyl-CoA dehydrogenase deficiencies; CACT, carnitine acylcarnitine transporter defect; *CPT I* and *II*, carnitine palmitoyl transferase deficiencies; *UOAs*, urinary organic acids; *tHCY*, plasma total homocysteine; B_{12}, vitamin B_{12}; *PLAC*, plasma acylcarnitine profile.

(C_8). The ratio of C_8 to C_{10} (decanoylcarnitine) is also used as a secondary marker. Often, mutational analysis to determine the presence of at least the common Northern European mutation, A985G, is performed as a secondary test. However, the allele frequency of this mutation is somewhat lower (around 75%) in screen-detected than in clinically presenting patients, and some other mutations found by screening have not been recorded in clinical presentations [20]. Elevated urinary acylglycines and an abnormal plasma acylcarnitine profile virtually confirm the diagnosis of MCAD. More definitive confirmation can be obtained by enzymatic or DNA analysis or more simply by acylcarnitine profiling in cultured skin fibroblasts. *Short-chain acyl-CoA dehydrogenase deficiency* is found more commonly in screened patients than would have been expected from previous experience, and all cases reported so far have been asymptomatic. Other fatty acid oxidation disorders found by newborn screening are shown in ◘ Table 2.3. Amongst these, some cannot be reliably detected, but with a multiplex test this does not seem to be a reason not to screen. An example is the carnitine transporter defect. It is certain that some cases can be detected, but not necessarily all. Treatment is simple and as the disorder is potentially fatal, detection of even some of the cases seems justified.

2.3.6 Other Neonatal Screening Programmes

Other disorders often tested for in various combinations in newborn screening programmes, but of more importance for general paediatricians or those with other specialties, include the following:

Congenital Hypothyroidism

Screening for primary CH is universal in countries with well-developed health systems, and widely practised in less well-developed countries. Usually the primary test is measurement of thyroid stimulating hormone (TSH). False negative results may occur in cases with well compensated ectopic thyroid glands, or sometimes with dyshormonogenesis. Trial withdrawal of treatment at 2–3 years will identify cases that were transient and not needing life-long treatment.

Cystic Fibrosis

Screening and early diagnosis has produced clear-cut nutritional benefit and probable improvement in lung function. The test involves measurement of immunoreactive trypsin (trypsinogen) and usually some mutational analysis on the original blood spot. Some screening programmes have been functioning for almost 25 years, and screening is widely recommended.

Congenital Adrenal Hyperplasia

Measurement of 17-hydroxyprogesterone is the primary test for congenital adrenal hyperplasia (CAH). The test readily identifies cases, but the false positive rate is high, especially among pre-term babies, and strategies to reduce this are necessary. In the future, a secondary MSMS test may be the best option.

Biotinidase Deficiency

A specific enzyme assay on dried blood spots is often used to detect biotinidase deficiency. This is more sensitive than MSMS testing. The utility of using a separate test for this disorder has not been substantiated.

Glucose-6-phosphate Dehydrogenase Deficiency

An NAD/NADH-based enzyme assay can detect this disorder. The test is widely used in Asian countries. The usefulness in avoiding kernicterus and haemolytic crises is likely but has not been well demonstrated.

Haemoglobinopathies

Sickle-cell disease is included in areas where this is prevalent. There is now the possibility of a wider screen for haemoglobinopathies using MSMS.

References

1. Guthrie R, Susi A (1963) A simple phenylalanine method for detecting phenylketonuria in large populations of newborn infants. Pediatrics 32:338–343
2. Dussault JH, Coulombe P, Laberge C et al. (1975). Preliminary report on a mass screening program for neonatal hypothyroidism. J Pediatr 6:670–674
3. Millington DS, Kodo N, Norwood DL, Roe CR (1990) Tandem mass spectrometry: a new method for acylcarnitine profiling with potential for neonatal screening for inborn errors of metabolism. J Inherit Metab Dis 13:321–324
4. Wilson JMG, Jungner G (1968) Principles and practice of screening for disease. World Health Organization, Geneva
5. Proposed international guidelines on ethical issues in medical genetics and genetics services. WHO/HGN/GL/ETH/98 1 2000
6. Serving the family from birth to the medical home. A report from the Newborn Screening Task Force convened in Washington DC, May 10-11, 1999 (2000). Pediatrics 106(2 Pt 2):383–427
7. Green A, Pollitt RJ (1999) Population newborn screening for inherited metabolic disease: current UK perspectives. J Inherit Metab Dis 22:572–579
8. Therrell BL Jr (1993) Laboratory methods for neonatal screening. American Public Health Association, Washington, DC
9. Sander J, Janzen N, Sander S et al. (2003) Neonatal screening for citrullinaemia. Eur J Pediatr 162:417–420
10. Rashed MS, Rahbeeni Z, Ozand PT (1999) Application of electrospray tandem mass spectrometry to neonatal screening. Semin Perinatol 23:183–193
11. Zytkovicz TH, Fitzgerald EF, Marsden D et al. (2001) Tandem mass spectrometric analysis for amino, organic, and fatty acid disorders in newborn dried blood spots: a two-year summary from the New England Newborn Screening Program. Clin Chem 47:1945-1955
12. Wilcken B, Wiley V, Hammond J, Carpenter K (2003) Screening newborns for inborn errors of metabolism by tandem mass spectrometry. N Engl J Med 348:2304–2312
13. Schulze A, Lindner M, Kohlmuller D et al. (2003) Expanded newborn screening for inborn errors of metabolism by electrospray ionization-tandem mass spectrometry: results, outcome, and implications. Pediatrics 111:1399–1406
14. Chace DH, Kalas TA, Naylor EW (2003) Use of tandem mass spectrometry for multianalyte screening of dried blood specimens from newborns. Clin Chem 49:1797–1817
15. Chace DH, Millington DS, Terada N et al. (1993) Rapid diagnosis of phenylketonuria by quantitative analysis for phenylalanine and tyrosine in neonatal blood spots by tandem mass spectrometry. Clin Chem 39:66–71
16. Holzman , Slazyk WE, Cordero JF, Hannon WH (1986) Descriptive epidemiology of missed cases of phenylketonuria and congenital hypothyroidism. Pediatrics 78:553–558
17. Trinh MU, Blake J, Harrison JR et al. (2003) Quantification of glutamine in dried blood spots and plasma by tandem mass spectrometry for the biochemical diagnosis and monitoring of ornithine transcarbamylase deficiency. Clin Chem 49:681–684
18. Wang T, Lawler AM, Steel G et al. (1995) Mice lacking ornithine-delta-aminotransferase have paradoxical neonatal hypoornithinaemia and retinal degeneration. Nat Genet 11:185–190
19. Gibson KM, Bennett MJ, Naylor EW, Morton DH (1998) 3-Methylcrotonyl-coenzyme A carboxylase deficiency in Amish/Mennonite adults identified by detection of increased acylcarnitines in blood spots of their children. J Pediatr 132(3 Pt 1):519–523
20. Andresen BS, Dobrowolski SF, O'Reilly L et al. (2001) Medium-chain acyl-CoA dehydrogenase (MCAD) mutations identified by MS/MS-based prospective screening of newborns differ from those observed in patients with clinical symptoms: identification and characterization of a new, prevalent mutation that results in mild MCAD deficiency. Am J Hum Genet 68:1408–1418

Pathology (February 2008) 40(2), pp. 104–115

CURRENT OPINIONS

Newborn screening

BRIDGET WILCKEN AND VERONICA WILEY

NSW Newborn Screening Programme, The Children's Hospital at Westmead and the University of Sydney, Sydney, Australia

Summary

The aim of newborn screening is to detect newborns with serious, treatable disorders so as to facilitate appropriate interventions to avoid or ameliorate adverse outcomes. Mass biochemical testing of newborn babies was pioneered in the 1960s with the introduction of screening for phenylketonuria, a rare inborn error of metabolism, tested by using a dried blood spot sample. The next disorder introduced into screening programs was congenital hypothyroidism and a few more much rarer disorders were gradually included.

Two recent advances have greatly changed the pace: modification of tandem mass spectrometry and DNA extraction and analysis from newborn screening dried blood spot. These two technologies make the future possibilities of newborn screening seem almost unlimited.

Newborn screening tests are usually carried out on a dried blood spot sample, for which there are special analytical considerations. Dried blood spot calibrators and controls, prepared on the same lot number of filter paper, are needed. Methods have a co-efficient of variation of about 10% due to the increased variability of a dried filter paper sample compared with other biochemical samples. The haematocrit is an additional variable not able to be measured. Also of importance is obtaining a balance between the sensitivity and specificity of each assay. Fixing cut-off points for action needs consideration of what is an acceptable percentage of the population to recall for further testing. Few assays are 100% discriminatory.

Programs in Australasia currently screen for at least 30 disorders. Detection of these requires not only the assay of a primary marker but often determination of a ratio of that marker with another, or possibly an alternative assay, for example a DNA mutation. The most important disorders screened for are described briefly: phenylketonuria, primary congenital hypothyroidism, cystic fibrosis, the galactosaemias, medium-chain acyl-CoA dehydrogenase deficiency, glutaryl-CoA dehydrogenase deficiency and congenital adrenal hyperplasia, together with several other disorders detectable by tandem mass spectrometry.

Newborn screening deals with rare disorders and benefit cannot be shown easily without very large pilot studies. There have been randomised controlled trials of screening for cystic fibrosis, and now several studies are beginning to establish the benefit of tandem mass spectrometry screening for disorders of fatty acid and amino acid metabolism.

Two things will influence the new directions for newborn screening: the development of effective treatments for hitherto untreatable disorders, and advancing technology, enabling new testing strategies to be developed. There are novel treatments on the horizon for many discrete disorders. Susceptibility testing has recently been considered for newborn screening application, but is more controversial. Newborn screening has entered a new and exciting phase, with an explosion of new treatments, new technologies, and, possibly in the future, new preventive strategies.

Key words: Newborn screening, tandem mass spectrometry, dried blood spot, congenital hypothyroidism, cystic fibrosis, medium chain acyl-CoA dehydrogenase deficiency.

Abbreviations: CAH, congenital adrenal hyperplasia; CH, congenital hypothyroidism; IRT, immunoreactive trypsin; MSMS, tandem mass spectrometry; PKU, phenylketonuria; TSH, thyroid stimulating hormone.

Received 16 August, revised 12 October, accepted 18 October 2007

SCOPE OF THIS ARTICLE

In this article we discuss mass newborn screening to identify certain inborn errors of metabolism, hormonal defects and other disorders by means of blood testing performed at central laboratories. Thus, other important aspects of screening newborns, such as locally performed hearing screening or testing for hip dysplasia, etc., are not considered. We discuss the range of screening tests carried out in Australasia with a mention of some tests carried out elsewhere, the different methodologies available and their limitations, and exciting future possibilities.

BRIEF HISTORY OF NEWBORN SCREENING

Mass biochemical testing of newborn babies was pioneered in the 1960s by Guthrie, who introduced screening for phenylketonuria (PKU), a rare inborn error which had been first described some 25 years before.[1] Untreated PKU resulted in severe mental retardation in over 90% of sufferers. Treatment with a low-phenylalanine diet had been investigated in the 1950s[2,3] and it became apparent that great gains were possible if affected children were diagnosed and treated early. A urine test with ferric chloride was then available but was unreliable, especially in the neonatal period. Guthrie, a microbiologist who had a mentally retarded son and niece, the latter shown later to have PKU, in 1960 had devised a bacterial inhibition assay which would detect raised phenylalanine levels in blood. His seminal paper describing this was not published until 1963.[4] There were two major aspects to his test: firstly, it

Print ISSN 0031-3025/Online ISSN 1465-3931
DOI: 10.1080/00313020701813743

could be carried out rapidly on large numbers of samples and secondly, and most importantly, the matrix he used was a blood sample dried on filter-paper. The samples could thus be sent to central laboratories and processed cheaply. The possibility of mass screening was investigated by the Children's Bureau in the United States, and in 1963 Massachussetts became the first state to mandate testing of all infants for PKU.[5] Newborn screening was born. Testing for PKU became universal in the United States, Australasia, most of Western Europe, Japan, and some other countries, and hundreds of children have now been spared from severe mental retardation. PKU screening has served as a model against which other newborn screening programs are often measured. It was not all plain sailing and PKU screening uncovered unexpected findings (see below) but, overall, it was a signal success.

The dried blood samples could be used for other testing; Guthrie and colleagues devised assays for other amino acids but, because of the rarity of the diseases (and one was benign – histidinaemia[6]), these tests were not incoporated in screening programs to a great extent. In 1975 screening for hypothyroidism was described.[7] This is probably the most important disorder tested for at present. Initial assays were for thyroxine. Later, an assay for thyrotropin (TSH) became possible and was more widely used as a primary test.[8] Other important disorders for which tests were introduced include congenital adrenal hyperplasia[9,10] in 1979, cystic fibrosis,[11] sickle cell anaemia,[12] biotinidase deficiency,[13] and the galactosaemias.[14] These and other disorders that have been tested for are discussed fully below.

Two recent advances have greatly changed newborn screening. In the 1990s a new era came when tandem mass spectrometry could be modified to make it suitable for newborn screening. The possible application to newborn screening had been described in 1990[15] and by the mid-1990s this was able to be exploited; upwards of 30–40 disorders of aminoacid, organic acid and fatty acid metabolism could be detected simultaneously using a single assay. Prior to this, it had been shown that the newborn screening dried blood spot could readily be used for DNA extraction and analysis.[16] These two technologies make the future possibilities of newborn screening seem almost unlimited.

GENERAL AIMS OF NEWBORN SCREENING, AND CRITERIA FOR SCREENING

Simply put, the aim of newborn screening is to detect newborns with serious, treatable disorders, so as to facilitate appropriate interventions to avoid or ameliorate adverse outcomes. There has been recent discussion about the possible widening of this aim. Some screening programs in the past were put in place to detect disorders which were then untreatable. An example is Duchenne's muscular dystrophy screening in Wales.[17,18] Even when no treatment seemed to be on the horizon (treatment possibilities are being discussed now),[19,20] parents and the public in general were in favour of this screening program. Currently, tandem mass spectrometry can detect many disorders simultaneously and some of these detected patients may have an essentially untreatable disorder. There are arguments in favour of the detection of an untreatable disorder which rely on the clear benefit to the family of a diagnosis being made, rather than the possibility of a death from unknown cause.

The classic criteria for screening are those in the 1969 WHO-sponsored paper of Wilson and Jungner, written mainly to discuss screening for infectious disease, chronic disease, cancer and so forth, but PKU does get a consideration.[21] This beautifully written document has stood the test of time but is not now completely applicable throughout to newborn screening. Other recent important documents discussing screening criteria include those from the American College of Human Genetics[22] and the United Kingdom National Screening Committee.[23] The Human Genetics Society of Australasia has a policy on Newborn Screening which is regularly updated and simpler and yet encapsulates all the important points. Briefly, newborn screening is recommended if:

1. There is benefit for the baby from early diagnosis. (In some untreatable disorders, benefit to the family as a whole can benefit the baby.)
2. The benefit is reasonably balanced against financial and other costs.
3. There is a reliable test suitable for newborn screening.
4. There is a satisfactory system in operation to deal with diagnostic testing, counselling, treatment and follow-up of patients identified by the test.

More specific parts of the policy are available at: http://hgsa.com.au/ (accessed September 2007).

GENERAL LABORATORY ASPECTS OF NEWBORN SCREENING

Screening includes all processes to ensure that each baby in the target population is offered testing, all appropriate follow-up is performed, and the baby is referred for early treatment where indicated. Ideally screening should be conducted in large centralised laboratories to minimise costs and ensure expertise of detecting rare disorders.

The sample

Many developing or pilot programs that only screen for one or a few disorders, e.g., primary congenital hypothyroidism, use cord blood to ensure that a sample is collected from each baby, because of very early discharge. Most developed programs collect blood spot samples when the baby is older onto filter paper with predetermined characteristics for absorption and elution. In Australia samples are ideally collected at 48 to 72 hours of age. The sample cards also contain all relevant identifying and other demographic information and may include relevant clinical details. These samples have distinct advantages compared with cord blood samples: once dried they are easy to transport using the post to a centralised laboratory, and they can be used to determine many different analytes. For many analytes, cord blood reflects the mother's and not the baby's situation.

Consent

Consent is required for newborn screening tests. Before sample collection, the staff performing the collection should

explain to the parent(s) the screening test and its importance. A pamphlet detailing all aspects of the screening program is also provided to the parent(s) so that they can give informed consent. Unlike other pathology tests there is no request form associated with screening. The request is implied by the collection and sending of a sample. An example of guidelines for testing in one Australian state can be found at: http://www.health.nsw.gov.au/policies/pd/2006/pdf/PD2006_099.pdf (accessed September 2007).

Sample integrity

Samples indicating contamination or haemolysis or with insufficient volume collected are deemed unsuitable for screening. In most screening programs these unsuitable samples are tested, as an abnormal result may need urgent follow up; however, a numerical result is not released and a repeat sample is requested.

Computing

All samples are given a unique sample identification number and information is entered into the laboratory information management system. Currently there are two basic approaches to data storage for screening: short term storage of results with a long term record only for babies diagnosed with a disorder; or a comprehensive secure system with long-term storage for all patient demographics, sample details including date of collection, tests performed, test results and interpretations. A comprehensive system provides easy access for statistical analysis of results and allows for result stratification for any parameter stored; for example, feed status or age of baby.

Testing

Over the years, various analytical techniques became well established for use within a routine newborn screening laboratory. In general, screening protocols included determination of an analyte as an initial screen followed by testing of a subset of samples with a secondary test to optimise sensitivity and specificity. Techniques for primary screening included microbiological inhibition testing, colorimetric, fluorescence, enzymatic or chromatography assays and various immunoassays. Many of these techniques could also be used as secondary assays. However, DNA mutational analysis was useful as a secondary screening assay to further differentiate true positives from the normal population. This has been most commonly incorporated into the screening for cystic fibrosis. Until the 1990s, most screening tests determined one analyte as a marker for a particular disorder, e.g., measurement of phenylalanine for phenylketonuria or thyroid stimulating hormone for congenital hypothyroidism. However, in the late 1990s tandem mass spectrometry introduced into screening programs provided determination of many analytes and thus many disorders on the same sample at the same time. Tandem mass spectrometry is considered in more detail below. Commercially available kits can be obtained for any assay/disorder currently being screened. As with any pathology assay, kits can provide the advantage of reliability but the disadvantage can often be the cost of supply. Many programs therefore prepare in-house reagents for screening tests. Whether a kit or in-house assay is used, the characteristics in that particular laboratory need to be established.

Special analytical considerations in newborn screening

As for any pathology assay it is important to use matrix-matched calibrators and in newborn screening this generally means dried blood spot calibrators. Calibrators and controls should be prepared on the same type and lot number of filter paper to minimise confounders for interpretation of results. Similarly, the amount of sample, i.e., the size of punched circle, must be consistent for calibrators, controls and test samples. In general, good laboratory methods for newborn screening assays using matrix-matched calibrators with a set haematocrit of 55% have a co-efficient of variation of approximately 10% due to the increased variability of a dried filter paper sample compared with other biochemical samples. The haematocrit, however, is an additional variable not able to be measured. In Australasia it is recommended that samples are taken between 48 and 72 hours of life. This gives a timely receipt of the sample in the laboratory, while avoiding problems associated with early sampling, for example with thyroid stimulating hormone (TSH) levels. Normal ranges of most analytes vary considerably during the first hours and days. Fig. 1 and 2 show the distribution of blood spot levels of phenylalanine and TSH on days 1–10. This also means that separate cut-off levels for action are needed for samples which have had to be repeated, and so are taken later than the recommended time. Fig. 3 shows the increase in pre-treatment phenylalanine levels with time in babies affected with PKU.

Sensitivity, specificity, positive predictive value, and setting the action limits/cut-off points

To understand the screening process, and consider cut-off points, some definitions are important:

Sensitivity is the proportion of subjects with the disorder in question detected by the test.

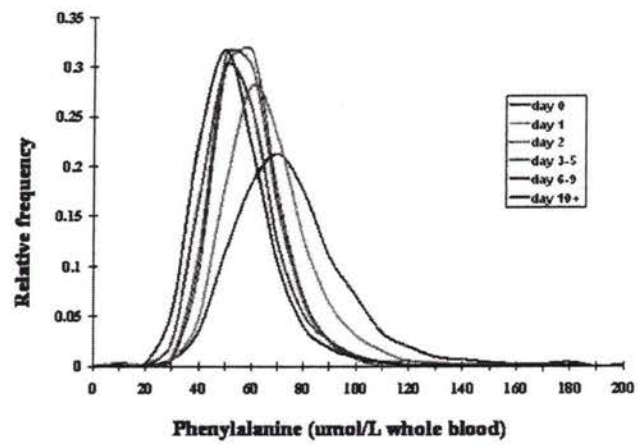


FIG. 1 The effect of age of the baby at time of dried blood spot sample collection on phenylalanine concentration. Each histogram indicates the relative frequency of results: day 0, $n=1500$; day 1, $n=3500$; day 2, $n=55000$; days 3–5, $n=510000$; days 6–9, $n=3500$; day 10+, $n=4000$. The population histogram is left shifted with age.

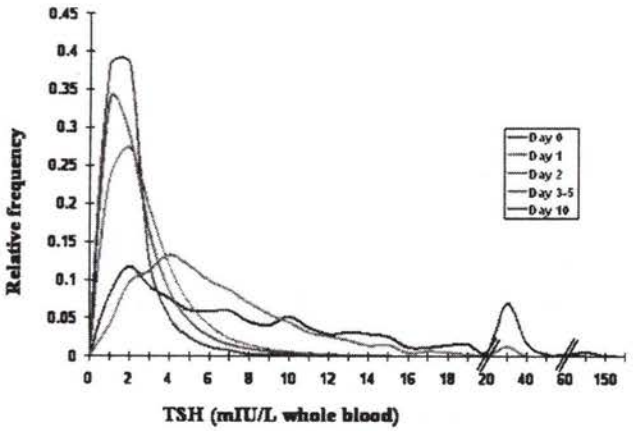


FIG. 2 The effect of age of the baby at time of dried blood spot sample collection on thyroid stimulating hormone (TSH) concentration. Each histogram indicates the relative frequency of results: day 0, $n = 1500$; day 1, $n = 3500$; day 2, $n = 55\,000$; days 3–5, $n = 510\,000$; days 6–9, $n = 3500$; day 10+, $n = 4000$. The population histogram is left shifted with age. Samples collected from babies on day 0 and day 1 demonstrate the need for age-specific reference ranges.

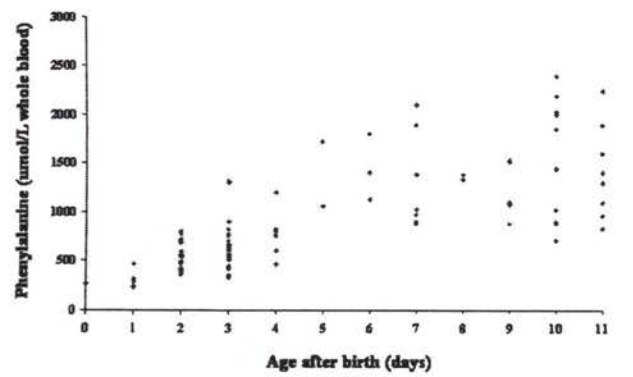


FIG. 3 The effect of age of the baby on phenylalanine concentration in untreated patients diagnosed with phenylketonuria (PKU). Each baby has two results: one from the initial screening sample and another from a confirmatory sample before treatment was commenced. As each baby with untreated PKU ages, the concentration of phenylalanine increases.

Specificity is the proportion of subjects without the disorder that have a negative test result.

False negative rate is the percentage of affected subjects not detected by the test.

False positive rate is the percentage of healthy subjects with a positive test result.

Positive predictive value is the chance that a positive result actually indicates an affected individual.

Negative predictive value is the chance that a negative result excludes the disorder.

These values depend not only on the specificity or sensitivity of the test, but also on the frequency of the disorder.

There is a natural trade-off between sensitivity and specificity: the higher the sensitivity for most disorders, the higher will be the false positive rate. Ideally there is no overlap between the normal distribution and that obtained from affected subjects. In practice there may be significant overlap. For many screening programs the action limit is determined by what is an acceptable percentage of the population to recall for further testing. Few assays are 100% discriminatory. To improve sensitivity and specificity for many disorders a proportion of the samples are analysed by a secondary screening assay. This may be the evaluation of the result compared with other results; for example, the ratio of phenylalanine to tyrosine concentrations aids discrimination between phenylketonuria and other causes of an increased phenylalanine level, such as total parenteral nutrition or liver dysfunction. Alternatively, the secondary test may be a separate assay; for example, DNA mutational analysis or quantitation of enzyme activity. Depending on the overlap of the 'normal' versus 'affected' population histograms, for most analytes there are three action limits. The lowest action limit takes into account the measurement of uncertainty for dried blood spot samples and is for in-house retesting of the analyte in duplicate to ensure that the reportable value is not above the recall/resample action limit. The next level, i.e., the lower resample action limit, is used when a baby might have a disorder but this is unlikely. Usually a repeat sample is required for another primary screen test which is then evaluated against the appropriate range for that analyte in an older baby. The upper limit is to indicate those with a high degree of suspicion of having the disorder; urgent further samples are then required for diagnostic testing.

Governance and organisational issues

Newborn screening in Australasia is considered a mature process as it has been established since the late 1960s. Each of the five Australian reference centres (located in Adelaide, Brisbane, Melbourne, Perth and Sydney) where screening is performed is funded by the relevant state government. Similarly, there is one government funded testing centre for New Zealand. The screening test is not mandatory and consent must be obtained from the parent before sampling, as discussed above. Although each program has its own local advisory committee, the recommended screening policy for the programs is developed by a joint subcommittee of the Human Genetics Society of Australasia and the Division of Paediatrics of the Royal Australasian College of Physicians. A sub-committee established by the Australian Health Ministers' Advisory Council is currently developing principles and guidelines for newborn blood spot screening in Australia.

Tandem mass spectrometry

Undoubtedly the most exciting development in newborn screening in recent years has been tandem mass spectrometry (MSMS). This has already completely revolutionised newborn screening and will soon expand the possibilities even further. It is important to understand some aspects of the methodology.

The instrumentation in use for newborn screening is an electrospray ionisation tandem mass spectrometer. A triple quadrupole tandem mass spectrometer consists of an ionisation source followed by three mass filters connected in tandem and a detection system. Once the source has produced ions of the various analytes in the sample, the ions are selected according to their mass by the first filter (MS1). These are precursor ions. In the second filter the selected ions are then fragmented by collision, most

commonly with the inert gas, argon (although nitrogen may also be used), to form smaller product ions before selection of the product ions occurs in the third filter (MS2). The computer software controlling the MSMS sequentially analyses compounds for detection. In general, positive identification of an analyte can be made on the basis of its fragmentation pattern; however, analytes with identical masses will not be distinguished without modified sample preparation, introduction of a separation technique or alternative scan functions.

Three scan modes can be used: product ion – where all fragments produced from a single precursor are detected (for example, used for galactose-1-phosphate); precursor ion – where all precursors producing a single product are detected (used for acylcarnitines); and neutral loss – where all precursors undergoing the loss of a common fragment are detected (used for amino acids). The acquisition of mass spectral data can be obtained in full scan mode, where every fragment is detected, or multiple reaction monitoring (MRM) where only specific fragmentations are detected. As MRM provides greater analytical sensitivity as well as an ability to select analytes, it is commonly used for newborn screening. Until recently most programs used a sample preparation method based on that of Millington *et al.*[15] In short, blood from a blood disc was eluted into methanol containing isotopically labelled internal standards; the supernatant was decanted and evaporated to dryness; analytes were derivatised with heating before again being evaporated to dryness. The sample was finally dissolved in an acetonitrile solution for injection into the MSMS. Derivatisation of analytes was used to improve analytical sensitivity. Instruments produced today are inherently more sensitive and many laboratories are therefore investigating the possibility of simplifying sample preparation by eliminating the derivatisation step. In the above method calibration was obtained using internal standard alone; however, this did not take into account variations in elution characteristics associated with dried blood spot samples. Therefore, many laboratories have incorporated matrix matched non-isotopically labelled standard curves for each analyte of interest.

This technology just described enables multiplex testing for the first time. At present, over 50 disorders can be detected simultaneously from a 3 mm dried blood spot in around 2 minutes per sample. This greatly expanded capability means that very rare disorders can be included in a newborn screening program for the first time, something that would not be possible if each needed a stand-alone test. Current screening is mainly confined to the measurement of amino acids and acylcarnitines, to detect selected disorders of amino acid, organic acid, and fatty acid metabolism. The selection of disorders that should be included is the subject of intense discussion and much disagreement in various screening jurisdictions. The most influential decisions have been those of the United States Maternal and Child Health Bureau and the American College of Medical Genetics which, after consultation, have endorsed a list of 29 'core disorders' for screening, plus 25 secondary targets – disorders which form part of a differential diagnosis of the core 29.[22] By contrast, the United Kingdom's National Screening Committee has at this time only permitted MSMS to be used for screening for PKU and medium-chain acyl-CoA dehydrogenase (MCAD) deficiency (described below) and no other disorders may be sought. In Australasia, where MSMS screening is now universal, no such restrictions are in place.

MSMS screening has meant that babies with disorders of amino acid, organic acid, and fatty acid metabolism are now often detected in the newborn screening laboratory, rather than by the clinical metabolic service and more cases are detected by screening than by clinical presentation, as with all screening. Because of this, detected babies fall into one of three categories:

1. The baby might be symptomatic before the newborn screening result is known. Disorders where this may occur include neonatal presentations of urea cycle defects, organic acidaemias such as methylmalonic acidaemia, galactosaemia (not usually diagnosed by MSMS), and less commonly, almost any of the fatty acid oxidation defects. Detection by newborn screening may be of benefit as some affected babies may have delayed diagnosis, or may die undiagnosed, for example, having been thought to have died from sepsis.
2. The baby may not yet have symptoms, and an effective treatment can be introduced before damage is done. Cases in this category include phenylketonuria, the less severe urea cycle disorders, most other aminoacidopathies, such as homocystinuria, some organic acidaemias, and most fatty acid oxidation disorder cases.
3. The baby may have a disorder which would have proved benign, or largely so, and would thus derive no benefit from early diagnosis. It is clear that mild forms of several disorders will readily be detected by newborn screening, but will not need treatment. It may not be possible, however, to distinguish all those with mild forms from patients at risk who do need treatment. But beyond this, there are disorders now being commonly detected which previously were thought extremely rare. The disorder itself may prove to be always or almost always benign. Examples to be discussed later include 3-methylcrotonyl CoA carboxylase (3MCCC) deficiency and short-chain acyl-CoA dehydrogenase (SCAD) deficiency.

SCREENING FOR SPECIFIC DISORDERS

The disorders screened for at present (2007) in Australasia are shown in Table 1, which also shows the year that screening started. Disorders tested for in other countries, but not currently being considered in Australasia, include sickle-cell disease, glucose-6-phosphate dehydrogenase deficiency, human immunodeficiency virus and toxoplasmosis infections, and Krabbé's disease. Descriptions of all these disorders are to be found in general paediatric texts, and more detailed descriptions of the rarer disorders can be accessed[24] and will not be further referenced here. Very useful fact sheets addressing screening for various disorders have been published.[25]

Phenylketonuria (PKU)

This disorder, the first to be diagnosed by newborn screening, leads to severe mental retardation in over 90%

TABLE 1 Disorders screened for in Australian states and New Zealand

Disorder/class	Screened in:	Started*	Birth prevalence	Screening test	Confirmation	Sensitivity	Comment
PKU	All	1967	1:14 000†	MSMS (BIA and colourimetric previously)		~100%	Robust tests, FU 100%, effective Rx
CH (primary)	All	1977	1:2750†	TSH (EIA, FIA)	TFTs, imaging	~97%	Does not detect central hypothyroidism
CF	All	1981	1:3000†	IRT (EIA, FIA), DNA mutation on upper 1%	Sweat test, extended mutation	94–98%	Variable mutation panel
Gal	All but Victoria	1983	1:50 000†	Manual NAD/NADH metabolite test	Enzymatic (Beutler; other)	~98%	Detects transferase, kinase, and epimerase defects
CAH	NZ only	1984	1:23 000	17OHP (RIA, FIA)	17OHP plasma, mutation	~100%	
Biotinidase	NZ only	1986	1:150 000	Enzyme (colourimetric)	Plasma enzyme assay	~100%	Only complete deficiency sought; partial deficiencies do not need treatment
MSMS	All	1998–2006	1:6000‡	MSMS amino acids, acylcarnitines	Plasma amino acids, acylcarnitine, urine organic acids, mutation, enzymatic	Variable (see text)	

*Screening started at different times in different states. The earliest date is given here.
†Birth prevalence Australia, 2000–2004.
‡New South Wales 1998–2006, excluding PKU, ornithine transcarbamylase, non-ketotic hyperglycinaemia, and maternal defects.
BIA, bacterial inhibition assay; CAH, congenital adrenal hyperplasia; CF, cystic fibrosis; CH, congenital hypothyroidism; EIA, enzyme immunoassay; FIA, fluorescence immunoassay; Gal, galactosaemias; IRT, immunoreactive trypsin; MSMS, tandem mass spectrometry; NZ, New Zealand; 17OHP, 17hydroxy progesterone; PKU, phenylketonuria; RIA, radioimmunoassay; TFT, thyroid function tests; TSH, thyroid stimulating hormone.

of sufferers if it is untreated. The newborn screening test now used throughout Australasia (and in most developed countries) is tandem mass spectrometry measurement of phenylalanine. Use of a secondary measurement, the phenylalanine/tyrosine ratio, may reduce the recall rate for some laboratories. The sensitivity is close to 100% for tests performed after 24 hours, and the false positive rate is very low indeed, and is associated with preterm infants, and use of total parenteral nutrition. Before the advent of MSMS, screening was carried out mainly by the original 'Guthrie' bacterial inhibition assay. This was robust and cheap, with good performance, but semi-quantitative. Also used were fluorimetric and colourimetric methods. To distinguish between classical PKU and several disorders of the pterin co-factor requires further testing, which involves the measurement of urinary pterins, a specialised test using high performance liquid chromatography (HPLC), and often a tetrahydrobiopterin loading test, carried out in a specialised metabolic ward setting. Early treatment, by about day 10, with a low phenylalanine (low protein) diet and phenylalanine-free amino acid supplements leads to an excellent outcome with good development and growth. Treatment is now thought to be needed for life for the best outcome. The discovery of the teratogenic nature of elevated phenylalanine, causing the maternal phenylketonuria syndrome, meant that dietary control of phenylalanine levels was especially important in women of childbearing age.[26]

Primary congenital hypothyroidism (CH)

CH occurs in around 1:3000 babies in all ethnic groups. Affected neonates generally appear normal, and are thus not readily diagnosed without a screening program. Delayed treatment leads to mental and growth retardation. In Australasia the analyte used for screening is TSH, currently detected by immunoassay, most commonly using DELFIA (dissociation-enhanced lanthanide fluorescent immunoassay). This approach of course does not detect central hypothyroidism due to defects in pituitary or hypothalamic function, and the optimal strategy for screening has been the subject of debate. Initial screening tests in the 1970s employed a primary thyroxine (T4) test, and this approach was almost universal in the United States for many years, with a TSH test used as a secondary test for samples with a low T4. Several different protocols have been proposed for hypothyroid screening, including using simultaneous assays for TSH and T4, adding a thyroid binding globulin assay, and using a routine second blood collection at 2–6 weeks of age. Screening strategies have recently been comprehensively reviewed.[27] Using TSH as the primary test for screening has advantages over other approaches, and in patients discharged after 24 hours of age it shows lower patient recall rates with very low false-negative test results.[27] A positive test result requires confirmation by full thyroid function tests. Additionally, thyroid imaging, and for prognosis a measurement of bone age, are usually recommended. Some patients may be missed by the test, especially a few with dyshormonogenesis. Very low birthweight babies require a follow-up test at 4 weeks, as they may have transient hypothalamic immaturity.[28] In babies with compensated hypothyroidism, especially seen in those

with ectopic glands, the TSH may be moderately elevated, but formal thyroid function tests may show a normal T4 level. These babies need careful follow up, and age related normal ranges must be used in interpretation. Early treatment of CH produces excellent outcome for both growth and development.[29]

Cystic fibrosis (CF)

Cystic fibrosis, with its severe pulmonary and gastrointestinal manifestations, is one of the most common life-shortening genetic disorders. Screening was made possible by the finding of increased immunoreactive trypsin (trypsinogen; IRT) in the blood of newborns. Systematic screening was started in New Zealand in 1980, and New South Wales in 1981, and now is carried out throughout Australasia. Screening was controversial at first, and has been adopted slowly in developed countries,[30] despite the early findings of nutritional and growth benefit,[31,32] and later demonstration of pulmonary benefit[33] and better survival in childhood.[34] Initial screening was by measurement of dried blood spot IRT by immunoassay. The IRT test had a sensitivity of about 98% but poor specificity, and a repeat sample was needed on about 0.7% of babies.[35] Those positive on the second test needed a sweat test for diagnosis. The second test at around 4 weeks of age was much more discriminatory, with a positive predictive value of at least 50%. The identification of the cystic fibrosis transmembrane conductance regulator (CFTR) gene in 1989 and the discovery of a common mutation, ΔF508, immediately made it possible to screen babies for CF using a single sample, carrying out mutation testing on samples with an elevated IRT level, without the immediate need for requesting a second sample.[36] Since that time, various strategies have been developed, mostly using a combination of IRT and DNA mutation on a dried blood spot; this has recently been reviewed.[37] In Australasia, all screening laboratories follow a similar protocol, with an initially raised level of IRT triggering a mutation analysis for either a single mutation, or a panel of mutations, with consideration for the mutation spectrum in the state CF population, and the costs and expected effectiveness. The mutations currently included are shown in Table 2. Overall the current sensitivity is 95–98%; some affected babies are missed by the test. Additionally, the clinical spectrum of CF is very broad, and patients with mild disease, destined not to become clinically evident until late adolescence or adulthood, are probably not detected.[38] At present most screening programs avoid including known mild mutations in their DNA screening panels.

The galactosaemias

The main target of galactosaemia screening is to detect galactose-1-phosphate uridyl transferase (gal-1-PUT) deficiency (classical galactosaemia), a disorder which can be fatal in early infancy. The birth prevalence is about 1:50 000 in Australasia. Other galactosaemias which are also detected are the extremely rare galactokinase deficiency, which does not cause systemic disease but does lead to cataract in infancy, and galactose epimerase deficiency. Most cases of the latter are benign partial deficiencies but a severe systemic form has been recorded a handful of times in the literature (not yet diagnosed in Australasia). Methods used in screening are measures of metabolites, galactose and galactose-1-phosphate, or measures of enzyme activity, confined to the Beutler test for gal-1-PUT.[14] Most commonly nowadays, a metabolite assay is followed by confirmatory testing using a gal-1-PUT assay and quantitative determination of galactose and galactose-1-phosphate, and this is what is used in Australasia.[39] A combination of these methods will provide a precise diagnosis of transferase deficiency (but see below), and an indication of galactokinase deficiency, in which there is elevation of galactose alone and a normal gal-1-PUT activity. However, the differentiation of red-cell epimerase deficiency (a benign condition) and systemic epimerase deficiency, clinically similar to transferase deficiency, may not be clear. Additionally, moderate metabolite elevations and severely reduced, but not absent, gal-1-PUT activity is seen in combined heterozygosity for a severe mutation in the transferase GALT gene, and a common 'Duarte' mutation. The differentiation (with trace gal-1-PUT activity, increasing to 25% during the first year) is important, as a Duarte/galactosaemia double heterozygote usually needs no treatment.[40] The outcome for early-treated galactosaemia patients is not uniformly good. Certainly timely screening largely avoids the danger of a potentially fatal early neonatal illness, but affected children on galactose-free diets are still at risk for learning and speech difficulties (in about 50%) and a very small percentage have severe intellectual disability, the reason for which is currently unknown. Treatment of galactokinase deficiency certainly avoids the development of cataract in childhood. Longer follow-up is not reported.

TABLE 2 Cystic fibrosis: mutations tested for in Australasian screening programs

Mutation	NSW	NZ	Qld	SA	Vic	WA
p.F508 del	+	+	+	+	+	+
p.I507 del	+	+	+	+	+	+
DF508C	–	–	–	–	–	+
c.489 + 1G > T	–	–	+	–	+	–
c.1585 – 1G > A	–	–	–	–	+	–
c.3718 – 2477C > T	–	–	–	–	+	–
p.W1282X	–	–	–	–	+	–
p.R553X	–	–	+	+	+	–
p.R560T	–	–	–	–	+	–
p.N1303K	–	–	+	–	+	–
p.G542X	–	+	+	+	+	+
p.G551D	–	+	+	+	+	+
p.520F	–	–	+	–	+	–
621 + 1	–	–	–	–	–	+
p.R117H	–	–	+	–	–	–

Nomenclature is based on the cDNA reference sequence GenBank L13923; nucleotide 1 is the A of the ATG translation initiation codon, 134 bases from the start of the cDNA sequence.
NSW, New South Wales; NZ, New Zealand; Qld, Queensland; SA, South Australia; Vic, Victoria; WA, Western Australia.

Congenital adrenal hyperplasia (CAH)

CAH is caused by disorders of adrenal steroid biosynthesis, and can lead to salt-loss, vomiting, dehydration and death in the most severe 'salt-wasting' forms which comprise

70% of cases. Symptoms include virilisation, which is most evident in girls. Simple virilising forms also occur, often with later presentation. Over 90% cases are due to 21-hydroxylase deficiency. The birth prevalence of newborn screening-detectable cases is around 1:18 000.[10] Newborn screening for 21-hydroxyase deficiency has been carried out for over 25 years,[10,41] relying on measurement of 17-hydroxyprogesterone in dried blood spots by immunoassay. The false positive rate is high in preterm babies (even using differential cut-off levels) but overall the performance is good, with high sensitivity, close to 100%. Methods using liquid chromatography–tandem mass spectrometry have recently been described which have better specificity and thus offer advantages over routine immunoassays due to the elimination of interferences, but the run-times at present make these only suitable for second-tier testing (on the same blood spot, after a positive immunoassay test result). These methods would require additional analytical time, and thus possibly additional MSMS equipment, as they cannot be run simultaneously with methods for detecting amino acids and acylcarnitines.[42,43] The advantages of newborn screening for CAH include avoidance of adrenal crises with related morbidity and mortality, and avoidance of wrong gender assignment in girls which, although rare, does still happen.

Disorders detected by tandem mass spectrometry

Over 40 disorders apart from PKU are detectable by MSMS. Table 3 shows the analytes usually included in MSMS newborn screening, and mentions disorders detectable. There are many recent reviews of MSMS screening[44–46] and reports of individual screening programs.[47–51] Here we describe screening for two of the most important disorders detected, and give brief notes on several others to illustrate different problems.

Medium chain acyl-CoA dehydrogenase (MCAD) deficiency

This recessively inherited disorder is the most common disorder of fatty-acid oxidation, with a birth prevalence of around 1:10 000 to 1:20 000 in populations derived from Europe. Affected children may have life-threatening and neurologically damaging episodes of hypoketotic hypoglycaemia during catabolic stress but are well between episodes and may indeed never become sick. MCAD deficiency is reliably indicated by elevated dried blood spot octanoylcarnitine (C8). The ratio of C8 to C10 (decanoylcarnitine) is also used as a secondary marker. Often, mutational analysis to determine the presence of at least the common Northern European mutation, c.985 A > G, is performed as an additional secondary test. Confirmation,

TABLE 3 Disorders detectable by tandem mass spectrometry and currently sought in Australasia

Class	Disorder	Primary analyte	Secondary marker
Amino acids	Argininaemia/arginase deficiency	Arginine (ARG)	
	Argininosuccinic aciduria (ASA lyase deficiency)	Citrulline (CIT)	CIT/ARG ratio
	Citrullinaemia (argininosuccinate synthase deficiency, citrin deficiency)	Citrulline (CIT)	CIT/ARG ratio
	Citrin deficiency	Citrulline (CIT)	
	Fumaryl acetoacetase deficiency (tyrosinaemia type I)*	Tyrosine (TYR)	Succinylacetone
	Homocystinuria (cystathionine β-synthase deficiency)	Methionine (MET)	MET/PHE
	Maple syrup urine disease	Leucines (LEUs)	LEUs/ALA, LEUs/PHE
	Phenylketonuria	Phenylalanine (PHE)	PHE/TYR
	Pterin defects	Phenylalanine (PHE)	PHE/TYR
	Tyrosine aminotransferase deficiency	Tyrosine (TYR)	
Organic acid disorders	Beta-ketothiolase deficiency (mitochondrial acetoacetyl-CoA thiolase deficiency)	C5OH	
	Cobalamin C defect (homocystinuria with methylmalonic aciduria)	C3/C2	
	Glutaryl CoA dehydrogenase deficiency (glutaric acidaemia type I)	C5DC	
	Holocarboxylase synthetase deficiency	C5OH	
	3-hydroxy-3-methylglutaryl-CoA lyase (HMGCoA lyase) deficiency	C5OH	
	Isobutyryl-CoA dehydrogenase deficiency†	C4	
	Isovaleric acidaemia	C5	
	Methylmalonic acidurias (mutase deficiency, CblA and CblB defects)	C3	
	Propionic acidaemia	C3	
	3-methylcrotonyl-CoA carboxylase deficiency	C5OH	
	2-methylbutyryl-CoA dehydrogenase deficiency	C5	
	3-methylglutaconyl-CoA hydratase deficiency	C5OH	
Fatty acid oxidation	Carnitine/acylcarnitine translocase deficiency	C0	
	Carnitine transporter defect	C0	
	CPTI (carnitine palmitoyl transferase deficiency type I)	C0	C0/C16 + C18
	CPTII (carnitine palmitoyl transferase deficiency type II)	C16	
	LCHAD (3-hydroxy long chain acyl-CoA dehydrogenase deficiency)	C16OH	C18OH
	MCAD (medium chain acyl-CoA dehydrogenase deficiency)	C8	C6, C10, C10:1, C8/C10
	MADD (multiple acyl-CoA dehydrogenase deficiency)	C4, C6, C8, C10, C12	
	SCAD (short chain acyl-CoA dehydrogenase deficiency)†	C4	
	SCHAD (short chain hydroxy acyl-CoA dehydrogenase deficiency)†	C4OH	
	TFP (trifunctional protein deficiency)	C16OH	C18OH
	VLCAD (very long chain acyl-CoA dehydrogenase deficiency)	C14:1	C12, C14, C16

C0, carnitine; C3, propionylcarnitine; C4, butyrylcarnitine; C5, isovalerylcarnitine/3-methylbutyrylcarnitine; C5OH, 3-hydroxyisovaleryl carnitine or 2-methyl-3-hydroxy-butyrylcarnitine; C5DC, glutarylcarnitine; C6, hexanoylcarnitine; C8, octanoylcarnitine; C10, decanoylcarnitine; C10:1, decanoylcarnitine; C14, tetradecanoylcarnitine; C14:1, tetradecenoylcarnitine; C16, hexadecanoylcarnitine; C16OH, 3-hydroxy hexadecanoylcarnitine.

*Fumarylacetoacetate is usually not detectable by methodology in current use.

†Some disorders in the table are not sought by several programs.

where required, is by urinary organic acid analysis for acylglycines (elevations of hexanoylglycine and subarylglycine), plasma acylcarnitine analysis, further mutation analysis and the acylcarnitine profile in cultured skin fibroblasts. Specific enzyme analysis is not normally needed. Screening has detected almost twice as many cases as were known by clinical detection and some mutations found by screening have not been recorded in clinical presentations.[52] It has been clearly shown that early diagnosis by screening with subsequent adoption of a management plan has resulted in a much lower incidence of severe episodes or death.[53,54]

Glutaryl CoA dehydrogenase deficiency (Glutaric aciduria type 1, GA1) GA1 is a cerebral organic aciduria. In the first few years of life metabolic decompensation triggered by catabolic stress usually results in damage to the basal ganglia, with the development of dystonia and severe choreoathetosis. The birth prevalence is about 1:100 000. Elevation of dried blood spot glutarylcarnitine indicates the likelihood of GA1, and sensitivity and specificity are very high at 24–72 hours. Confirmation is by measuring this analyte in plasma, and detecting elevated urinary glutaric acid and 3-hydroxyglutaric acid by gas chromatography mass spectrometry (GCMS). A small number of patients are 'low excretors' and a urine test may be unreliable if negative.[55] The ultimate confirmation involves enzyme analysis on cultured skin fibroblasts or mutation analysis. A recent study has unequivocally demonstrated the advantages of early detection by newborn screening with the institution of appropriate management, including emergency treatment of intercurrent illness, carnitine therapy, and a low lysine and tryptophan diet.[56]

Homocystinuria (cystathionine β-synthase deficiency) Present screening is by detection of elevated levels of blood methionine. This will detect the approximately 50% of patients who are pyridoxine non-responsive but not those who are responsive to pyridoxine.

Maple syrup urine disease Classical cases are detected using leucines (leucine + isoleucine + alloisoleucine), and sometimes valine in addition, as markers. Clear advantage of neonatal detection has been shown.[57] Not all mild cases, who nevertheless may develop neurological and intellectual problems, can be detected by current screening protocols.[58]

Very long chain acyl-CoA dehdrogenase deficiency MSMS screening detects this important fatty acid oxidation disorder by the elevation of unsaturated tetradecanoylcarnitine. While newborn screening appears sensitive, confirmatory testing is difficult as plasma acylcarnitines and urinary organic acids may be normal when the baby is well. Leukocyte acylcarnitine assays or acylcarnitine profiles in cultured skin fibroblasts may be required. Differentiation between affected and carriers is not always easy.[59]

Short chain acyl-CoA dehydrogenase deficiency This disorder has been detected surprisingly often by newborn screening programs and is now thought to be benign or almost invariably so.[60] Australasian programs no longer include this in their screening panel. Other detectable disorders thought to be usually benign include 3-methylcrotonyl CoA carboxylase deficiency (3-MCCC).

Detection of maternal disorders Several maternal disorders can be detected by early sampling of neonatal blood. Prominent has been 3-MCCC deficiency. Mothers have almost invariably been healthy despite very low levels of plasma total carnitine.[61] Other disorders detected include carnitine transporter defect, holocarboxylase synthase deficiency (in the NSW program), and vitamin B12 deficiency.

Biotinidase deficiency This is a disorder of biotin recycling. Untreated profound deficiency often results in seizures, developmental delay, seborrhoeic dermatitis, and sometimes hearing loss and optic atrophy. Partial deficiency is commonly asymptomatic. Screening tests using a semiquantitative colourimetric assay of biotinidase activity are used in many countries including New Zealand, but not in Australia.[13,55] Some cases may be detected by MSMS screening but this will not reliably detect all. Presymptomatic treatment with oral biotin is effective.

Sickle cell disease and other haemoglobinopathies Sickle-cell disease (SCD) represents a group of conditions causing chronic haemolysis and intermittent vascular occlusions. Screening for SCD is carried out throughout the USA and, recently, the United Kingdom and in other countries but not in Australasia, where the ethnic background of the population does not yet indicate this. Screening is done by electrophoretic or chromatographic means of separating haemoglobin species and can be performed by MSMS. Early treatment prevents death in the first year from pneumococcal septicaemia and from splenic sequestration.

Glucose-6-phosphatase deficiency Screening for this common X-linked disorder is not carried out routinely in Australasia but is commonly performed in the southeast Asian region. The benefits of early detection have not been clearly demonstrated.

EVALUATING THE COSTS AND BENEFITS OF NEWBORN SCREENING

We recently commented in *The Lancet* that 'it remains an enigma that newborn screening for cystic fibrosis has until now been so controversial, when there is so much evidence favouring it'.[30] Indeed there has been evidence from many cohort studies with historical controls and two randomised trials to support cystic fibrosis screening. This is in great contrast to the rapid uptake of MSMS newborn screening, for which initially there was no documentation of any benefit. MSMS, as one technology for multiplex testing, has led to a re-appraisal of the criteria for screening. Since testing for some additional disorders can be added with virtually no up-front cost for detection, it is tempting to include disorders without adequate consideration of the likely benefits and drawbacks. In fact, evaluation of the outcomes of newborn screening has been sadly lacking for most disorders overall.

Newborn screening deals with rare disorders and benefit cannot be shown easily without very large pilot studies; this fact has to be recognised. For the earliest two disorders

screened, PKU and hypothyroidism, benefit now seems obvious and no further studies are needed simply to establish this. For most disorders, randomised controlled trials are unlikely but this does not diminish the need for careful studies of a lesser degree of reliability. For MSMS screening these are beginning to be reported.[53–57,62,63] For some other disorders, there are long-standing state-wide or nation-wide programs for which there have been no studies reported but for which benefit is not obvious.[64] As newborn screening moves into new and more controversial areas, establishment of benefits and costs will be ever more important and adequate protocols need to be built into new experimental programs.

THE FUTURE: WHERE WILL NEWBORN SCREENING BE GOING?

Two things will influence the new directions for newborn screening: the development of effective treatments for hitherto untreatable disorders and advancing technology, enabling new testing strategies to be developed. We are already witnessing this in an interesting way. When the American College of Medical Genetics was developing recommendations for disorders to be included in US screening panels, widespread consultation was invoked. Among 88 disorders considered in 2002, Krabbé's disease scored the lowest marks; there was no available test and no treatment for this devastating disease. A subsequent report of successful treatment, at least in the short term, by umbilical cord blood transplant in the earliest weeks[65] and an MSMS-based test for the enzyme galactocerebrosidase,[66] as well as some public pressure, prompted newborn screening for this disorder in New York State.[57,67] The lysosomal storage disorders, of which Krabbé's disease is but one example, are becoming rapidly more treatable, with enzyme replacement therapies and haematopoetic cell transplantation well established for some. Methods for newborn screening have been developed and are being finalised, with trials likely in the next year or two. Two types of multiplex method have been proposed: protein profiling and enzyme analysis, both using tandem mass spectrometry.[33,68,69] This area is surely going to be the next new expansion, although there are many questions still to be answered. The treatments are extremely expensive, and so may not be available to all who could benefit, and the long-term outcome is not known. Treatment for some disorders could convert an unpleasant fatal disorder over the short term into a chronic degenerative disorder which is also ultimately fatal.

There are novel treatments on the horizon for many other disorders. One exciting development is the possibility of a drug to prompt ribosomes to read through premature stop codons. It has been appreciated for some time that gentamycin can do this but it is too toxic to be effective in clinical situations – the dose would need to be too high. Now a new compound, PTC124, is being trialled. PTC124 can detect premature stop (termination) codons (PTCs) that arise from non-sense mutations but does not interfere with 'genuine' stop codons. Drug-induced read-through could allow treatment of subsets of patients with many different genetic disorders. At present, trials with Duchenne's muscular dystrophy are producing promising results and, if ultimately successful, would re-kindle interest in newborn screening for this disorder.[70] The possible use of chemical chaperones is another field that may revolutionise treatment and thus influence newborn screening.[71]

Susceptibility testing has not been considered for a newborn screening application until recently. Particularly important in the area of new technology will be the development of cheaper DNA applications of microarrays, making mass screening an economic possibility. Here there will indeed be ethical issues, as it becomes possible to test for susceptibility for later-onset common diseases. One could readily envisage screening for susceptibility to cardiovascular disease with a 'heart disease' array, susceptibility to various cancers, diabetes and so forth. Newborn screening for type 1 diabetes susceptibility is already a possibility and is being promoted, principally using human leukocyte antigen (HLA) haplotypes as markers of susceptibility or protection;[72] an experimental program in Italy is under way.[73] However, there are particular problems for type 1 diabetes, since there is as yet no known preventive strategy. Health professionals appear to be not in favour of such programs at this time, and ethicists have many reservations and caveats.[74]

However, newborn screening has undoubtedly entered a new and exciting phase, with an explosion of new treatments to drive it, new technologies, and possibly in the future new preventive strategies. Close co-operation among newborn screening laboratories, biochemical genetics and molecular genetics laboratories, and the genetic metabolic and clinical genetics clinical services, as happily occurs throughout Australasia, will ensure a successful development. What was a quiet back-water looks to become a mainstream interest.

Address for correspondence: Professor B. Wilcken, The Children's Hospital at Westmead, Locked bag 4001, Westmead, NSW 2145, Australia. E-mail: Bridgetw@chw.edu.au

References

1. Folling A. Uber Ausscheidung von Phenylbrenztraubensaur in den Harn als Stoffwechselanomalie in Verbindung mit Imbezillitat. *Z Physiol Chem* 1934; 227: 169.
2. Woolf LI, Griffiths R, Moncrieff A. Treatment of phenylketonuria with a diet low in phenylalanine. *Br Med J* 1955; 1: 57–64.
3. Bickel H, Gerrard J, Hickmans EM. Influence of phenylalanine intake on phenyketonuria. *Lancet* 1953; 265: 812–3.
4. Guthrie R, Susi A. A simple phenylalanine method for detecting phenylketonuria in large populations of newborn infants. *Pediatrics* 1963; 32: 338–43.
5. MacCready RA, Hussey MG. Newborn phenylketonuria detection program in Massachusetts. *Am J Public Health Nations Health* 1964; 54: 2075–81.
6. Popkins JS, Clow CL, Scriver CR, Grove J. Is hereditary histidinaemia harmful? *Lancet* 1974; 1: 721–2.
7. Dussault JH, Coulombe P, Laberge C, Letarte J, Guyda H, Khoury K. Preliminary report on a mass screening program for neonatal hypothyroidism. *J Pediatr* 1975; 86: 670–4.
8. Illig R, Torresanti T, Sobradillo B. Early detection of neonatal hypothyroidism by serial TSH determination in dried blood. Six months experience with a reliable, efficient and inexpensive method. *Helv Paediatr Acta* 1977; 32: 289–97.
9. Pang S, Murphey W, Levine LS, *et al.* A pilot newborn screening for congenital adrenal hyperplasia in Alaska. *J Clin Endocrinol Metab* 1982; 55: 413–20.
10. Pang SY, Wallace MA, Hofman L, *et al.* Worldwide experience in newborn screening for classical congential adrenal hyperplasia due to 21-hydroxylase deficiency. *Pediatrics* 1988; 81: 866–74.
11. Crossley JR, Elliott RB, Smith PA. Dried-blood spot screening for cystic fibrosis in the newborn. *Lancet* 1979; 1: 472–4.
12. Lees CM, Davies S, Dezateux C. Neonatal screening for sickle cell disease. *Cochrane Database Syst Rev* 2000; CD001913.

13. Wolf B, Heard GS. Biotinidase deficiency. *Adv Pediatr* 1991; 38: 1–21.
14. Beutler E. Galactosemia: screening and diagnosis. *Clin Biochem* 1991; 24: 293–300.
15. Millington DS, Kodo N, Norwood DL, Roe CR. Tandem mass spectrometry: a new method for acylcarnitine profiling with potential for neonatal screening for inborn errors of metabolism. *J Inherit Metab Dis* 1990; 13: 321–4.
16. McCabe ER, Huang SZ, Seltzer WK, Law ML. DNA microextraction from dried blood spots on filter paper blotters: potential applications to newborn screening. *Hum Genet* 1987; 75: 213–6.
17. Bradley DM, Parsons EP, Clarke AJ. Experience with screening newborns for Duchenne muscular dystrophy in Wales. *BMJ* 1993; 306: 357–60.
18. Parsons EP, Bradley DM, Clarke AJ. Newborn screening for Duchenne muscular dystrophy. *Arch Dis Child* 2003; 88: 91–2.
19. Yokota T, Pistilli E, Duddy W, Nagaraju K. Potential of oligonucleotide-mediated exon-skipping therapy for Duchenne muscular dystrophy. *Expert Opin Biol Ther* 2007; 7: 831–42.
20. Aurino S, Nigro V. Readthrough strategies for stop codons in Duchenne muscular dystrophy. *Acta Myol* 2006; 25: 5–12.
21. Wilson JMG, Jungner G. *Principles and Practice of Screening for Disease*. Geneva: World Health Organization, 1968.
22. ACMG Newborn Screening Expert Group. Newborn screening panel and system. *Genet Med* 2006; 8: 1–252S.
23. National Screening Committee. Criteria for appraising the viability, effectiveness and appropriateness of a screening programme. London: Health Department of the United Kingdom, 2003. www.library.nhs.uk/screening (accessed October 2007).
24. Scriver CR, Beaudet AL, Valle D, *et al. The Metabolic and Molecular Bases of Inherited Disease*. 8th ed. New York: McGraw-Hill, 2001.
25. Kaye CI, Accurso F, La FS, *et al.* Introduction to the newborn screening fact sheets. *Pediatrics* 2006; 118: 1304–12.
26. Koch R, Hanley W, Levy H, *et al.* Maternal phenylketonuria: an international study. *Mol Genet Metab* 2000; 71: 233–9.
27. Rose SR, Brown RS, Foley T, *et al.* Update of newborn screening and therapy for congenital hypothyroidism. *Pediatrics* 2006; 117: 2290–303.
28. Australasian Paediatric Endocrine Group. *Guidelines for Management of Congenital Hypothyroidism*. http://www.apeg.org.au/Portals/0/documents/guidelines%20for%20management%20of%20congenital%20hypothyroidism.pdf (accessed October 2007).
29. Dubuis JM, Glorieux J, Richer F, Deal CL, Dussault JH, Van Vliet G. Outcome of severe congenital hypothyroidism: closing the developmental gap with early high dose levothyroxine treatment. *J Clin Endocrinol Metab* 1996; 81: 222–7.
30. Wilcken B, Gaskin K. More evidence to favour newborn screening for cystic fibrosis. *Lancet* 2007; 369: 1146–7.
31. Farrell PM, Kosorok MR, Laxova A, *et al.* Nutritional benefits of neonatal screening for cystic fibrosis. *N Engl J Med* 1997; 337: 963–9.
32. Waters DL, Wilcken B, Irwing L, *et al.* Clinical outcomes of newborn screening for cystic fibrosis. *Arch Dis Child Fetal Neonatal Ed* 1999; 80: F1–7.
33. McKay KO, Waters DL, Gaskin KJ. The influence of newborn screening for cystic fibrosis on pulmonary outcomes in New South Wales, Australia. *J Pediatr* 2005; 147: S47–50.
34. Grosse S, Boyle CA, Cordero JF. Newborn screening for cystic fibrosis: recommendations from the Centers for Disease Control and Prevention. *Am Fam Physician* 2006; 71: 1482–7.
35. Wilcken B, Wiley V, Sherry G, Bayliss U. Neonatal screening for cystic fibrosis: a comparison of two strategies for case detection in 1.2 million babies. *J Pediatr* 1995; 127: 965–70.
36. Seltzer WK, Accuros F, Fall MZ, *et al.* Screening for cystic fibrosis: feasibility of molecular genetic analysis of dried blood specimens. *Biochem Med Metab Biol* 1991; 46: 105–9.
37. Wilcken B. Newborn screening for cystic fibrosis: techniques and strategies. *J Inherit Metab Dis* 2007.
38. Padoan R, Genoni S, Moretti E, Seia M, Giunta A, Corbetta C. Genetic and clinical features of false-negative infants in a neonatal screening programme for cystic fibrosis. *Acta Paediatr* 2002; 91: 82–7.
39. Bowling FG, Brown ARD. Development of a protocol for newborn screening for disorders of the galactose metabolic pathway. *J Inherit Metab Dis* 1986; 9: 99–104.
40. Gitzelmann R, Bosshard NU. Partial deficiency of galactose-1-phosphate uridyltransferase. *Eur J Pediatr* 1995; 154: S40–4.
41. Cutfield WS, Webster D. Newborn screening for congenital adrenal hyperplasia in New Zealand. *J Pediatr* 1995; 126: 118–21.
42. Rauh M, Groschl M, Rascher W, Dorr HG. Automated, fast and sensitive quantification of 17 alpha-hydroxy-progesterone, androstenedione and testosterone by tandem mass spectrometry with on-line extraction. *Steroids* 2006; 71: 450–8.
43. Lacey JM, Minutti CZ, Magera MJ, *et al.* Improved specificity of newborn screening for congenital adrenal hyperplasia by second-tier steroid profiling using tandem mass spectrometry. *Clin Chem* 2004; 50: 621–5.
44. Rinaldo P, Tortorelli S, Matern D. Recent developments and new applications of tandem mass spectrometry in newborn screening. *Curr Opin Pediatr* 2004; 16: 427–33.
45. Chace DH, Kalas TA. A biochemical perspective on the use of tandem mass spectrometry for newborn screening and clinical testing. *Clin Biochem* 2005; 38: 296–309.
46. Chace DH, Kalas TA, Naylor EW. Use of tandem mass spectrometry for multianalyte screening of dried blood specimens from newborns. *Clin Chem* 2003; 49: 1797–817.
47. 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: 2304–12.
48. Zytkovicz TH, Fitzgerald EF, Marsden D, *et al.* Tandem mass spectrometric analysis for amino, organic, and fatty acid disorders in newborn dried blood spots: a two-year summary from the New England Newborn Screening Program. *Clin Chem* 2001; 47: 1945–55.
49. Wiley V, Carpenter K, Wilcken B. Newborn screening with tandem mass spectrometry: 12 months' experience in NSW Australia. *Acta Paediatr Suppl* 1999; 88: 48–51.
50. Frazier DM, Millington DS, McCandless SE, *et al.* The tandem mass spectrometry newborn screening experience in North Carolina: 1997–2005. *J Inherit Metab Dis* 2006; 29: 76–85.
51. Schulze A, Lindner M, Kohlmuller D, Olgemoller 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: 1399–406.
52. Andresen BS, Dobrowolski SF, O'Reilly L, *et al.* Medium-chain acyl-CoA dehydrogenase (MCAD) mutations identified by MS/MS-based prospective screening of newborns differ from those observed in patients with clinical symptoms: identification and characterization of a new, prevalent mutation that results in mild MCAD deficiency. *Am J Hum Genet* 2001; 68: 1408–18.
53. Wilcken B, Haas M, Joy P, *et al.* Outcome of neonatal screening for medium-chain acyl-CoA dehydrogenase deficiency: a cohort study. *Lancet* 2007; 369: 37–42.
54. Nennstiel-Ratzel U, Arenz S, Maier EM, *et al.* Reduced incidence of severe metabolic crisis or death in children with medium chain acyl-CoA dehydrogenase deficiency homozygous for c.985A > G identified by neonatal screening. *Mol Genet Metab* 2005; 85: 157–9.
55. Lindner M, Ho S, Fang-Hoffmann, Hoffman GF. Neonatal screening for glutaric aciduria type 1: strategies to procede. *J Inherit Metab Dis* 2006; 29: 378–82.
56. Kolker S, Garbade SF, Boy N, *et al.* Decline of acute encephalopathic crises in children with glutaryl-CoA dehydrogenase deficiency identified by newborn screening in Germany. *Pediatr Res* 2007; 62: 1–7.
57. Simon E, Fingerhut R, Baumkotter J, Konstantopoulou V, Ratschmann R, Wendel U. Maple syrup urine disease: favourable effect of early diagnosis by newborn screening on the neonatal course of the disease. *J Inherit Metab Dis* 2006; 29: 532–7.
58. 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: 586.
59. Liebig M, Schymik I, Mueller M, *et al.* Neonatal screening for very long chain acyl-CoA dehydrogenase deficiency: enzymatic and molecular evaluation of neonates with elevated C14:1-carnitine. *Pediatrics* 2006; 118: 1065–9.
60. van Maldegem BT, Duran M, Wanders RJ, *et al.* Clinical, biochemical, and genetic heterogeneity in short-chain acyl-coenzyme A dehydrogenase deficiency. *JAMA* 2006; 296: 943–52.
61. Koeberl DD, Millington DS, Smith WE, *et al.* Evaluation of 3-methylcrotonyl-CoA carboxylase deficiency detected by tandem mass spectrometry newborn screening. *J Inherit Metab Dis* 2003; 26: 25–35.
62. Scriver CR, Levy HL. Histidinaemia. Part I: Reconciling retrospective and prospective findings. *J Inherit Metab Dis* 1983; 6: 51–3.
63. Haas M, Chaplin M, Joy P, Wiley V, Black C, Wilcken B. Healthcare use and costs of medium-chain acyl-CoA dehydrogenase deficiency in Australia: screening versus no screening. *J Pediatr* 2007; 151: 121–6.
64. Gilbert R, Dezateux C. Newborn screening for congenital toxoplasmosis: feasible, but benefits are not established. *Arch Dis Child* 2006; 91: 629–31.
65. Escolar ML, Poe MD, Provenzale JM, *et al.* Transplantation of umbilical-cord blood in babies with infantile Krabbe's disease. *N Engl J Med* 2005; 352: 2069–81.
66. Li Y, Brockmann K, Turecek F, Scott CR, Gelb MH. Tandem mass spectrometry for the direct assay of enzymes in dried blood spots: application to newborn screening for Krabbe disease. *Clin Chem* 2004; 50: 638–40.

67. Pass KA, Green NS, Lorey F, Sherwin JE, Comeau AM. Pilot programs in Newborn Screening. *Ment Retard Dev Disabil Res Rev* 2006; 12: 293–300.
68. Meikle PJ, Grasby DJ, Dean CJ, *et al.* Newborn screening for lysosomal storage disorders. *Mol Genet Metab* 2006; 88: 307–14.
69. Gelb MH, Turecek F, Scott CR, Chamoles NA. Direct multiplex assay of enzymes in dried blood spots by tandem mass spectrometry for the newborn screening of lysosomal storage disorders. *J Inherit Metab Dis* 2006; 29: 397–404.
70. Welch EM, Barton ER, Zhuo J, *et al.* PTC124 targets genetic disorders caused by nonsense mutations. *Nature* 2007; 447: 87–91.
71. Singh LR, Chen X, Kozich V, Kruger WD. Chemical chaperone rescue of mutant human cystathionine beta-synthase. *Mol Genet Metab* 2007; 91: 335–42.
72. Dantonio P, Meredith N, Earley M, *et al.* A screening system for detecting genetic risk markers of type 1 diabetes in dried blood spots. *Diabetes Technol Ther* 2006; 8: 433–43.
73. Lorini R, Minicucci L, Napoli F, *et al.* Screening for type 1 diabetes genetic risk in newborns of continental Italy. *Acta Biomed* 2005; 76: 31–5.
74. Hiraki S, Ormond KE, Kim K, Ross LF. Attitudes of genetic counselors towards expanding newborn screening and offering predictive genetic testing to children. *Am J Med Genet* 2006; 140: 2312–9.