NEWBORN SCREENING

Expanded newborn screening: reducing harm, assessing benefit

Bridget Wilcken

Abstract Achieving the goals of newborn screening is, as for any screening, a balancing act: getting the maximum benefit from screening while producing the minimum harm. The advent of "expanded" newborn screening, with a large number of disorders detectable using a single test, has also posed problems, not new, but now more obvious. One is the finding of many more cases by screening, the extra cases being largely patients who have attenuated phenotypes and may remain asymptomatic for many years, even throughout life. These may or may not require active management in the short term, but do need lifelong awareness. Additionally, disorders have been included that are now thought benign or largely so. Babies risk being unnecessarily medicalized. Assessing outcome has also proved difficult because of the rarity of some disorders and the impracticality of randomized controlled trials. The requirements for valid studies include the need for case definitions, comparable comparison groups and probably assessment on a whole-population basis. An Australia-wide study of tandem mass spectrometry newborn screening involving 2 million screened and unscreened babies has demonstrated benefits overall to screened patients at age 6 years. The study was too small to provide conclusions for individual disorders other than for medium-chain acyl-CoA dehydrogenase deficiency.

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

Achieving the goals of newborn screening is, as for any screening, a balancing act: getting the maximum benefit from screening while producing the minimum harm. The obvious potential harms from screening, and those most discussed, arise from the occurrence of false positive and false negative results, and the costs of the programme. But other potential harms are in reality well known, and over the years have been common to newborn screening programmes for individual disorders. Nevertheless, they appear to come as a surprise when they occur in the course of beginning a new programme. These common problems should be borne in mind so that they can be anticipated for each new screening adventure, and then perhaps more easily dealt with. Table 1 shows examples drawn from existing programmes.

The advent some 12 years ago of "expanded" newborn screening, with a large number of disorders detectable using a single test, was a great step forward but it has also posed problems, not new, but now much more obvious, largely because the disorders are on the whole so rare that pieces of the necessary knowledge base have been missing. This paper explores two of the currently most pressing and linked issues: the therapeutic challenges posed by expanded screening and, briefly, the current situation with assessment of outcomes.

Therapeutic challenges: whom to treat?

It has long been realised that screening for a disorder often reveals a higher frequency of cases than does clinical detection (Wilcken et al 2003; Marsden 2009). Although this may be due to inefficient diagnosis of some symptomatic patients (including those who may have died undiagnosed),

Table 1 Some of the problems common to newborn screening programmes

Problem	Possible effect	Example(s)	References
Over-diagnosis of mild cases	Unnecessary treatment; worry. Outcome seems more beneficial	? congenital hypothyroidism	La Franchi 2010
		Neuroblastoma	Woods et al. 2002
Newly discovered mild forms of disorders	Uncertainty about management	Citrullinaemia	Discussed in this article
Including disorders of little/no clinical significance	Unnecessary treatment; harmful treatment; worry	Histidinaemia SCADD	Popkin et al. 1974
Late effects of a disorder previously unrecognised	Effective management not in place	Maternal PKU	Lenke and Levy 1980
		MGC type I	Wortmann et al. 2010
New phenocopies not previously recognised	Wrong treatment instituted at first	Pterin disorders in PKU screening	Danks et al. 1978
Lack of evidence for treatment modalities	Guidelines based on practice (practice might be wrong)	VLCADD	Arnold et al. 2009
			Spiekerkoetter et al. 2009

SCADD Short-chain acylCoA dehydrogenase deficiency, *PKU* phenylketonuria, *MGC* methylglutaconyl CoA hydratase deficiency, *VLCADD* very-long-chain acylCoA dehydrogenase deficiency

the majority of these "extra" cases will be patients who have attenuated phenotypes and may remain asymptomatic for many years, even throughout life. As has been previously discussed (Wilcken 2008), newborns who develop early symptoms are clearly in need of treatment. For babies who are asymptomatic, immediate treatment is also indicated for those with disorders well known to lead to adverse outcomes without treatment, with damage slowly and silently accruing. Examples of this are many, and include phenylketonuria, (EC 1.14.16.1) and cystathionine β-synthase (EC 4.2.1.22) deficiency. There is no need in these cases to wait for symptoms. Also, in need of immediate and careful management plans, are newborns with disorders which pose a high risk of an adverse outcome of sudden onset, one obvious example being the fatty acid oxidation disorder medium-chain acyl CoA dehydrogenase deficiency (EC 1.3.99.3). Current estimates of mortality in this particular disorder after the first few days are about 5–7% over the first 6 years (Grosse et al. 2006; Wilcken 2010) unless both early diagnosis and appropriate management are in place.

Mild phenotypes: the example of citrullinaemia

The major therapeutic challenge comes either with disorders in which mild cases of usually severe disorders have come to light mainly through screening, or with those disorders that now seem to be largely benign. To introduce this discussion, it is helpful to consider argininosuccinate synthase (EC 6.3.4.5) deficiency (citrullinaemia type I) as an example of the former. Citrullinaemia type I is a disorder of the urea cycle. Without screening, many babies with this disorder will present with hyperammonaemia as newborns.

Others will come to clinical notice in the first year or so, typically when dietary protein increases with weaning. Now that this disorder is readily diagnosed by expanded newborn screening, a proportion of patients detected are asymptomatic, despite quite high levels of circulating citrulline, and many seem to remain asymptomatic. "It is currently unclear how to treat these patients, and how to counsel their parents" (Dimmock et al. 2008). This problem has been known for several years. Sander et al. (2003) described mild phenotypes detected by newborn screening. Among 610,000 babies screened, 4 babies had classical citrullinaemia type I and 8 had persistent asymptomatic hypercitrullinaemia. These mild cases were observed for up to 3 years, some being on no treatment at all. Haberle et al. (2003) reported on 21 mild patients, all but 2 being identified by newborn screening programmes including family studies. Even patients with plasma citrulline levels above 2,000 μmol/L remained asymptomatic on no treatment for up to 12 years. However, some of these mild cases have been symptomatic, and there is a suggestion that the early adoption of a low-protein diet may be protective (Dimmock et al. 2008). Much more worrying is the occurrence of patients with first decompensations during pregnancy or the postpartum period. At least 11 cases have been recorded, in which the diagnosis of citrullinaemia type I has come to light during a pregnancy or in the immediate post-partum period (Kurasawa et al. 1998; Gao et al. 2003; Ito et al. 2004; Dimmock et al. 2008; Berning et al. 2008; Häberle et al. 2010). One of these cases is of especial interest, as the patient, who died during the episode of decompensation, was homozygous for a mutation c.325G>A (p. Ala118Thr) which has been recorded in two other cases of pregnancy-related decom-

pensation, but is associated with only very slight reduction of enzyme activity, though altered kinetics (Berning et al. 2008). Mutations identified in citrullinaemia type I have recently been reviewed (Engel et al. 2009). These workers thought that in general "the effect of the genetic background on the phenotype remains a poorly defined phenomenon". For an asymptomatic baby with citrullinaemia type I, it must be concluded that mild protein restriction may be indicated in the early years but that otherwise specific treatment may not be needed except around pregnancy, which seems to create an especially risky situation. Lifelong awareness will be needed for all patients. These are difficult things to achieve.

Most other disorders have mild phenotypes being identified by newborn screening. Prominent here are fatty-acid oxidation disorders, such as medium-chain acyl-CoA dehydrogenase deficiency, where identified subjects may have genotypes not so far recorded in clearly symptomatic cases, and very-long-chain acyl-CoA dehydrogenase deficiency (EC 1.3.99.13). So far, it has not been possible to differentiate in the newborn period among cases that clearly need immediate management, and those that will remain largely asymptomatic (Spiekerkoetter et al. 2009).

Disorders with little or no clinical significance

Even more worrying is the possibility of the inclusion of benign metabolic variants in newborn screening panels. The early history of screening includes screening for histidinaemia (L-histidine ammonia lyase deficiency; EC 4.3.1.3). Children found to have this biochemical phenotype were in some instances subjected to a low-histidine diet (Lam et al. 1996), and sometimes invasive investigations such as liver biopsy (Barashnev et al. 1988). Yet even the first description of the clinical findings in this disorder suggested that they might not be a consequence of the enzyme deficiency (Ghadimi and Partington 1967) and, in a careful look at the literature in 1972, Neville et al. (1972) noted that patients were "discovered during the investigation of mental retardation .... and during subsequent investigation of members of the family. Even with this bias to selection the outstanding fact is that over half of those reported with histidinaemia are of normal intelligence. This proportion may perhaps rise now that this abnormality is being detected by screening in the newborn period (Neville et al. 1972)". In fact, the alleged phenotype was apparently due entirely to selection bias (Lam et al. 1996). Screening was carried out in some regions for over 17 years, and there was some evidence that the low histidine diet had been harmful (Widhalm and Virmani 1994). It seems that these sorts of lessons are not easily learned, and several disorders of uncertain significance are now included in screening programmes.

Short-chain acyl-CoA dehydrogenase deficiency as an example

One example of a disorder often included in routine screening, but without clear indications for any treatment at all, is short-chain acyl-CoA dehydrogenase (SCAD; EC 1.3.99.2) deficiency. The gene, ACADS, has two common variations present in homozygous or compound heterozygous state in about 7% of the population which mildly reduce function. This disorder is identified in newborn screening by elevated levels of butyrylcarnitine [which also identifies the apparently benign disorder, short/branched-chain acyl-CoA dehydrogenase deficiency (Sass et al. 2008), also included in some programmes as a secondary target]. Newborn babies identified with SCAD deficiency have remained asymptomatic, but commonly have inactivating mutations. Patients diagnosed clinically have had a broad range of symptoms, but mainly have common variations in the ACADS gene and not inactivating mutations (Pedersen et al. 2008). "Clinical symptoms .... are nonspecific, generally uncomplicated, often transient, and not correlated with specific ACADS genotypes (van Maldegem et al. 2006)". It is hotly debated whether or not deficiency of this enzyme has important consequences (Gregersen et al. 2008). Indeed, it must be supposed that there would be some consequences of a severe deficiency, perhaps being a risk factor for some as yet incompletely defined outcome, and this may prove a most interesting line of research. But this is not (yet) what newborn screening is about. It is true that newborn screening offers a route for research into these conditions of uncertain consequence (Jethva et al. 2008) which would be very valuable, but this should be made transparent. There is no place in *routine* screening for disorders widely believed to be benign. In Australasia, screening for SCAD deficiency has uniformly been abandoned.

Possibility of late effects, previously unrecognised

While it is most unfortunate to "medicalize" a subject whose disorder is a biochemical one with no clinical consequence, it would be perhaps worse to assume a disorder benign or at least in no need of treatment when there were as-yet unrecognised late complications. Such could well have been the case with citrullinaemia type II (citrin deficiency, caused by mutations in *SLC25A13*), had the situation not been elucidated earlier. Methylglutaconyl CoA hydratase (EC 4.2.1.18) is one such deficiency, which appeared likely to be benign (Ly et al. 2003). Recently, there is evidence to suggest that it may be associated with adult-onset leukoencephalopathy (Wortmann et al. 2010). It is not yet clear whether early treatment would be helpful.

How to treat patients?

Amidst this catalogue of uncertainty, it is true that for the majority of patients there is sufficient knowledge to put in place rational treatment. Even so, few treatments have a solid evidence base, and treatment guidelines are being developed (Arnold et al. 2008, 2009; Spiekerkoetter et al. 2009). A potential problem with this is that these are largely based on what is currently being done (expert opinion). The resulting guidance could be wrong, or not ideal, but would become accepted practice, negating possible trials. There is urgent need for randomized trials in those situations where this would be ethical. An example is the use of carnitine in any fatty acid oxidation disorders, except for the carnitine transporter defect, where use is clearly indicated for most patients. Usually, the issue of evidence for treatment modalities has no easy answer. Even where there is good evidence of treatment efficacy in some situations, the answers for newborn screening are not easy. Screening for lysosomal disorders is already here. Added to the obvious issues of distinguishing early-onset and late-onset disorders in a neonate, and the long-term efficacy of treatment is some of the disorders, there is the cost issue: will very expensive treatment started in infancy be able to be maintained lifelong?

Other harms: the issue of false positive results

The most discussed potential harms arising from screening are related to the effects of false positive results. Screening programmes nowadays strive to reduce these to a minimum, but for most analytes, the lower the false positive rate the greater is the risk of a significantly reduced sensitivity. Recent publications suggest that the adverse effects of false-positive results have been over-estimated. While a review in this journal (Hewlett and Waisbren 2006) showed that studies overall did show increased parental stress and anxiety following an abnormal result even after this was shown to be falsely positive, the authors also found that such stress was less in parents who were well informed about newborn screening. These workers later demonstrated that parents have a high tolerance for false-positive newborn screening results (Prosser et al. 2008), and that there was, contrary to their earlier study, no impact on early health-care utilization by babies and children who had received false-positive results (Lipstein et al. 2009). It may be that programmes are getting more skilled at ensuring good information for parents and health-care workers. Screening programmes should of course continue to strive to lower the false positive rate, but policy-makers may feel a little more comfortable about the tolerability of this, for parents.

Evaluating outcomes of newborn screening

The history of newborn screening has been replete with reports of optimization of testing procedures and assessments of birth prevalence and screening performance—sensitivity, specificity, positive predictive value, and so forth—but with rather few reports about the clinical outcome. While there is no longer much doubt about the overall benefit of the most important long-standing programmes, for phenylketonuria and congenital hypothyroidism, there are few conditions screened for at present for which the outcome has been rigorously assessed. This has occurred only for cystic fibrosis in two randomized controlled trials and several reasonably well-designed cohort studies. By contrast, for some disorders widely included in screening panels, there have been no apparent studies of outcome at all.

Assessing outcomes of expanded newborn screening does pose particular difficulties. We have recently published the results of our Australia-wide outcome study (Wilcken et al. 2009) and believe that this includes many features important for reliable assessment. We compared cohorts of children born from 1994 to 1998 (none screened) and from 1998 to 2002 (some screened, some unscreened, depending on where they were born). Children were assessed at 6 years of age. For those who presented clinically or died in the first 5 days of life and those with substantially benign disorders, we did not expect, nor find, any beneficial effect from screening. For the other patients, either presenting after 5 days of age or diagnosed by screening, we were able to show that screened patients had fewer deaths and fewer clinically significant disabilities on a whole population basis, taking into account the likely morbidity and mortality of never-diagnosed patients. This confirmed and extended the benefit we previously showed in screening for MCAD deficiency, and showed for the first time benefits for patients with all other disorders, taken as a whole.

The salient features of this study were that we had collaboration of all metabolic services and all laboratories conducting screening and metabolic testing in Australia, so that we had 100% ascertainment of *ever diagnosed* patients. All but 3 patients were cared for in a specialised metabolic service, and all but 3 could be followed up for 6 years. We were able to include two unscreened comparison groups (historical and contemporary). Even so, because of the increased diagnostic rates in the screened cohort, the cohorts were not comparable, and comparisons had to be made on a whole population basis. This reduces the power of a study. We were unable to give any separate assessment of benefit for individual disorders other than MCAD deficiency because of small numbers, although for some conditions benefit cannot really be doubted.

Clinical outcome studies are sorely needed. Usually randomized controlled trials will not be possible where there is a strong belief in the benefit of screening, or where the disorders are so rare as to make this approach not feasible. For valid cohort studies of clinical outcome, one will usually require:

- Case definitions (inclusion criteria)
- Comparison groups of patients from unscreened cohorts *or* a retrospectively screened population (see Nennstiel-Ratzel et al 2005 for an example of the latter)
- Full ascertainment of all *diagnosed* patients, screened and unscreened
- An assessment of the comparability of groups, including:

  Similar diagnostic rates for screened and unscreened patients
  Similar treatment regimens
  Similar populations—ethnicity, socioeconomic, etc.

- *or*: if groups are not comparable, outcomes must be assessed on a whole-population basis
- Other desirable features are described in the STROBE guidelines (von Elm et al. 2007).

Newborn screening has well and truly emerged from its backwater, and is in an interesting, and some might say risky, phase. Technology is advancing at an ever faster pace and almost anything seems possible in the future. For the present, planning of valid collaborative studies of rare diseases, their treatment, and outcome assessment are wanted, and wanted urgently, to tip the balance in the right way—towards maximum benefit and minimum harm.

References

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 93:363–370

Arnold GL, Van Hove J, Freedenberg D et al (2009) Clinical practice protocol for the management of very long chain acyl-CoA dehydrogenase deficiency. Mol Genet Metab 96:85–90

Barashnev JI, Nikolayeva EA, Klembovsky AI (1988) Histidinaemia: screening, diagnosis, clinical picture, therapy. Acta Paediatr Hung 29:343–351

Berning C, Bieger I, Pauli S et al (2008) Investigation of citrullinemia type I variants by in vitro expression studies. Hum Mutat 29:1222–1227

Danks DM, Bartholomé K, Clayton B et al (1978) Malignant hyperphenylalaninaemia—the current status. J Inherit Metab Dis 1:49–53

Dimmock DP, Trapane P, Feigenbaum A et al (2008) The role of molecular testing and enzyme analysis in the management of hypomorphic citrullinemia. Am J Med Genet 146A:2885–2890

Engel K, Höhne W, Häberle J (2009) Mutations and polymorphisms in the human argininosuccinate synthetase (ASS1) gene. Hum Mutat 30:300–307

Gao HZ, Kobayashi K, Tabata A et al (2003) Identification of 16 novel mutations in the argininosuccinate synthetase gene and genotype-phenotype correlation in 38 classical citrullinemia patients. Hum Mutat 22:24–34

Ghadimi H, Partington MW (1967) Salient features of histidinemia. Am J Dis Child 113:83–87

Gregersen N, Andresen BS, Pedersen CB, Olsen RK, Corydon TJ, Bross P (2008) Mitochondrial fatty acid oxidation defects-remaining challenges. J Inherit Metab Dis 31:643–657

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

Häberle J, Pauli S, Schmidt E, Schulze-Eilfing B, Berning C, Koch HG (2003) Mild citrullinemia in Caucasians is an allelic variant of argininosuccinate synthetase deficiency (citrullinemia type 1). Mol Genet Metab 80:302–306

Häberle J, Vilaseca MA, Meli C et al (2010) First manifestation of citrullinemia type I as differential diagnosis to postpartum psychosis in the puerperal period. Eur J Obstet Gynecol Reprod Biol, in press

Hewlett J, Waisbren SE (2006) A review of the psychosocial effects of false-positive results on parents and current communication practices in newborn screening. J Inherit Metab Dis 29:677–682

Ito S, Kurasawa G, Yamamoto K et al (2004) A pregnant patient with fulminant hepatic failure was found to carry a novel missense mutation in the argininosuccinate synthetase gene. J Gastroenterol 39:1115–1117

Jethva R, Bennett MJ, Vockley J (2008) Short-chain acyl-coenzyme A dehydrogenase deficiency. Mol Genet Metab 95:195–200

Kurasawa G, Shiotsuka S, Nakajo J, Nagao M, Ohisi M, Aida T, Tanaka T (1998) A case report of citrullinemia induced by pregnancy. Nissanpu Kantoren Kaihou 35:3–7

La Franchi S (2010) Newborn screening strategies for congenital hypothyroidism: an update. J Inherit Metab Dis, [this issue]

Lam WK, Cleary MA, Wraith JE, Walter JH (1996) Histidinaemia: a benign metabolic disorder. Arch Dis Child 74:343–346

Lenke RR, Levy HL (1980) Maternal phenylketonuria and hyperphenylalaninemia. An international survey of the outcome of untreated and treated pregnancies. N Engl J Med 303:1202–1208

Lipstein EA, Perrin JM, Waisbren SE, Prosser LA (2009) Impact of false-positive newborn metabolic screening results on early health care utilization. Genet Med 11:716–721

Ly TB, Peters V, Gibson KM et al (2003) Mutations in the AUH gene cause 3-methylglutaconic aciduria type I. Hum Mutat 21:401–407

Marsden DL (2009) Commentary on a Delphi clinical practice protocol for the diagnosis and management of very long chain acyl-CoA dehydrogenase deficiency by Arnold et al. Mol Genet Metab 96:81–82

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

Neville BG, Bentovim A, Clayton BE, Shepherd J (1972) Histidinaemia. Study of relation between clinical and biological findings in 7 subjects. Arch Dis Child 47(252):190–200

Pedersen CB, Kølvraa S, Kølvraa A et al (2008) The ACADS gene variation spectrum in 114 patients with short-chain acyl-CoA dehydrogenase (SCAD) deficiency is dominated by missense variations leading to protein misfolding at the cellular level. Hum Genet 124:43–56

Popkin JS, Clow CL, Scriver CR, Grove J (1974) Is hereditary histidinaemia harmful? Lancet 1(7860):721–722

Prosser LA, Ladapo JA, Rusinak D, Waisbren SE (2008) Parental tolerance of false-positive newborn screening results. Arch Pediatr Adolesc Med 162:870–876

Sander J, Janzen N, Sander S et al (2003) Neonatal screening for citrullinaemia. Eur J Pediatr 162:417–420

Sass JO, Ensenauer R, Röschinger W et al (2008) 2-Methylbutyryl-coenzyme A dehydrogenase deficiency: functional and molecular studies on a defect in isoleucine catabolism. Mol Genet Metab 93:30–35

Spiekerkoetter U, Lindner M, Santer R et al (2009) Treatment recommendations in long-chain fatty acid oxidation defects: consensus from a workshop. J Inherit Metab Dis 32:498– 505

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

von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP (2007) STROBE initiative. Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. BMJ 335(7624):806–808

Widhalm K, Virmani K (1994) Long-term follow-up of 58 patients with histidinemia treated with a histidine-restricted diet: no effect of therapy. Pediatrics 94:861–866

Wilcken B (2008) The consequences of extended newborn screening programmes: do we know who needs treatment? J Inherit Metab Dis 31:173–177

Wilcken B (2010) Fatty acid oxidation disorders: outcome and long-term prognosis. J Inherit Metab Dis, in press

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 (2009) Expanded newborn screening: outcome in screened and unscreened patients at age 6 years. Pediatrics 124:e241–e248

Woods WG, Gao RN, Shuster JJ et al (2002) Screening of infants and mortality due to neuroblastoma. N Engl J Med 346:1041– 1046

Wortmann SB, Kremer BPH, Graham A et al (2010) 3-Methylglutaconic aciduria type I redefined; a syndrome with late onset leukoence-phalopathy. Neurology, in press