Newborn screening has provided a model of a successful public health screening program for the past 40 years. However, the history of newborn screening is not without controversy. Many of these controversies have been rekindled with the introduction of tandem mass spectrometry, a technology that has greatly increased our ability to detect potential disease in asymptomatic newborns. This review highlights the challenges raised by this and future technological advances as we strive to maintain the success of newborn screening in the 21st century.

HISTORY OF NEWBORN SCREENING

The early history of newborn screening echoes many of the same controversies we currently face. Prior to Robert Guthrie's development of the bacterial inhibition assay for PKU in 1961, no technology could reliably identify asymptomatic infants with PKU within the first week of birth, when treatment impact is greatest. In contrast, Guthrie's test was cheap, easy, and reliable. These were critical features for any test used in a mass screening program. The movement from Guthrie's discovery to implementation of mass screening was swift, in part, because of previous case reports that suggested that a low phenylalanine diet could improve developmental outcomes in children with PKU.

Society at that time was desperate for medical advances in the area of mental retardation. In 1961, the National Association for Retarded Children, whose members included health professionals and parents of children with mental retardation, began a public campaign in support of PKU screening. Meanwhile, President John F. Kennedy promised to double funding of the National Institute of Health for mental retardation research and created a presidential advisory commission on mental retardation. Mental retardation was, and still is, an emotional issue that personally touched the lives of many involved in the PKU story. President Kennedy's sister, Rosemary, and Robert Guthrie's son and niece all lived with mental retardation, though only in the case of Guthrie's niece was it attributable to PKU.

In 1962, Guthrie began a field trial sponsored by the Child Health Bureau,
which enrolled 400,000 infants in 29 states. However, the move toward universal state newborn screening and mandatory screening, in particular, did not proceed unopposed. The medical community had strong reservations about mandatory mass screening and declared widespread institution of screening to be premature because it had been initiated prior to Guthrie’s publication of a peer-reviewed article in 1963. Critics voiced concern that understanding of the disease pathology and treatment was incomplete. Fundamental questions about the test’s validity, extent of treatment, and what realistic outcomes could be expected remained unanswered.9,12 Consequently, the American Academy of Pediatrics and the American Medical Association did not support mandatory screening in the 1960s.12,13

Despite the calls for cautious and careful implementation, political and advocacy groups, such as the National Association for Retarded Children, successfully lobbied for mandatory newborn screening.10 Troubled by poor enrollment in state programs without mandatory screening, the National Association for Retarded Children and the Maternal and Child Health Bureau, formerly the Children’s Health Bureau, saw legislation of mandatory testing as a means to ensure that all newborns would be screened.10 By 1965, 27 states mandated newborn screening, while a handful of others gave their public health departments the power to decide.

However, implementation of mandatory screening in the 1960s did not solve the controversies and issues raised in the PKU story. While screening was mandatory in many states, treatment was not. Many children were left without access to treatment for the disorders with which they had been diagnosed. Furthermore, while mandatory testing increased the number of infants screened, it did not ensure that parents were adequately informed about screening.14 Finally, the value of long-term follow-up in screening programs cannot be overlooked. Long-term cohort studies completed after implementation have informed decisions about the initiation,15,16 extent, and length of therapy17,18 and the reproductive effects of hyperphenylalanemia.19

This review will demonstrate that while tandem mass spectrometry—a technology recently introduced to newborn screening—may be new, the controversies surrounding it are not. Tandem mass spectrometry has simply served to rekindle these embers of controversy.

ADVENT OF TANDEM MASS SPECTROMETRY

In the early 1990s, scientists proposed that tandem mass spectrometry, a technology used in biochemistry laboratories, might prove useful in newborn screening.20 Until that point, adding more tests to newborn screening panels would have required an increase in the amount of blood collected. Tandem mass spectrometry circumvents this problem by measuring levels and patterns of numerous metabolites in a single drop of blood, which are then used to identify potential disease.21 Using this same drop of blood, tandem mass spectrometry enables the detection of at least 4 times the number of disorders than was possible with previous techniques.22

Much like the original assay by Guthrie, tandem mass spectrometry offers a relatively cheap and rapid method of testing in newborn screening.23 While the tandem mass spectrometer and associated equipment costs about $500,000, numerous disorders can be screened in minutes for an incremental marginal cost of about $10 more per birth.23,24 However, test performance represents only a portion of the cost of the screening program. Each infant with a disease-positive screen must be evaluated, and those identified will likely require lifelong treatment. Much, if not all, of this cost is absorbed by private and public health care payers.25

Although tandem mass spectrometry has many technical advantages over previously used tests, its implementation has not been without challenges. For example, a recent study revealed that, in states offering screening with tandem mass spectrometry, determination of appropriate cutoff values for positive screens, a lack of accepted protocols and guidelines for workup, and availability of adequate educational materials describing disorders have posed major challenges for implementation.26 Yet, despite these issues, the uptake of tandem mass spectrometry in newborn screening has been rapid, with a dramatic, albeit variable, increase in the number of disorders screened in different states across the country.27 The number of disorders screened across states ranged from 0 to 8 disorders in 199528 and from 7 to 52 in 2005.22 These differential screening practices prompted the Maternal and Child Health Bureau to commission the American College of Medical Genetics to examine available evidence and make recommendations for a uniform newborn screening panel. An American College of Medical Genetics report, released in March 2005, recommended that states test for 29 specific disorders, most of which would be screened for using tandem mass spectrometry.29

In contrast to the early days of PKU, the medical community’s response to the American College of Medical Genetics’ recommendation to expand newborn screening has been divided. While some groups, including the American Academy of Pediatrics,30 have endorsed the expansion, others have offered a more cautious response. Botkin and colleagues31 have expressed concern with the report’s reliance on expert opinion, its survey methodology, and its lack of attention to the impact of false-positive results as well as to other ethical, social, and legal issues. While recognizing the potential value of new technology for newborn screening, Botkin and colleagues31 worry about the adverse consequences of rapidly expanding newborn screening without establishing the appropriate research and resource infrastructure needed to provide high-quality medical services to children and their families.

Another concern of both critics and proponents of expanded newborn screening is the potential for “diagnostic odysseys.” An allusion to Homer’s Odyssey, this phenomenon describes a situation in which the patient embarks on a circuitous and extensive workup in the search for a diagnosis.32 Proponents of expanded screening fear that without screening, children with an unrecognized disease will undergo such odysseys.33 In contrast, critics worry that increased screening identi-
ties false-positive results that will trigger a cascade of additional unnecessary and potentially harmful testing. There is also a concern that some infants with positive screening results will receive inconclusive confirmatory testing, leaving the family unable to know whether or not the infant will ever develop disease. Unfortunately, little empirical evidence is available to help us assess the scope of diagnostic odysseys from the perspective of either critics or proponents.

While the medical community debates a uniform newborn screening panel, some parents have found other means to obtain testing that may be unavailable in their state. Parents who want additional testing—and can afford the fees ranging from $25 to $89—may send their child’s blood to a number of for-profit private or university-affiliated laboratories. However, for the most part, these laboratories provide only testing results. The collection of the specimen, follow-up, and any additional testing must be coordinated through the primary care provider.

REDEFINING BENEFIT

In discussing the recent technological advances in newborn screening, it is helpful to step back and consider its larger goal. Shortly after the implementation of newborn screening, the World Health Organization commissioned a report that set forth the principles and practice of screening for diseases in industrialized countries (Table). The resultant criteria focused heavily on the impact of screening and subsequent diagnosis, treatment, and evaluation on the individual being screened. The requirement for a clear and direct benefit of testing to the infant in newborn screening has been reiterated by the National Academy of Sciences and Institute of Medicine.

Technological advances enabling detection of many more disorders have led proponents of expanded screening to challenge the traditional justification for newborn screening as dogmatic and narrow and to advocate for broadening the view of benefit, which includes family and society. For example, even in the absence of proven treatments, Bailey and colleagues contend that access to early therapeutic and psychosocial support services for the child provides an important and overlooked benefit to the family. Additionally, Bailey et al maintains that parents consider the information about their child’s disease a benefit in and of itself. He argues that society benefits from the knowledge gained about incidence and the natural history of disease, even without evidence of immediate therapeutic benefit to the infant.

Clearly, this conceptualization of benefit represents a radical shift from the traditional goals of screening, which are to identify disease in asymptomatic individuals and to intervene to prevent or mitigate disease. Additional support for this reframing of benefit stems from the difficulty of providing conclusive evidence of therapeutic efficacy for some disorders. In the case of newborn screening, many of these disorders are rare, and there may not be enough children to demonstrate a definitive benefit using a randomized controlled trial. The heterogeneous spectrum of disease and the long-term follow-up needed to fully assess disease complications further compound this problem.

The issue of whether individuals other than newborns with disease should benefit from newborn screening is not new. For example, the potential use of newborn screening results for reproductive planning through identification of parental and newborn carriers is a longstanding debate that began with the advent of sickle cell screening, and has been revisited with regard to cystic fibrosis screening, and will likely continue with the forthcoming prospect of screening for fragile X syndrome. Focusing on identification of carriers has important legal and ethical implications, because it creates an intersection between state-mandated testing and reproductive rights. As a society, we must be mindful that this broader interpretation of benefit will redefine the very essence of newborn screening and have far-reaching economic, public health, and legal ramifications. If evidence of therapeutic benefit is not necessary, where will we draw the line for adding more tests to newborn screening panels? We must make clear and informed decisions now, given the rapid pace of technological advancements, as the sky may soon be the limit in terms of what genetic, molecular, and biochemical abnormalities we can detect.

Finally, the concept of benefit must be considered within the context of the mandatory nature of newborn screening. Parens patriae is the legal doctrine behind mandatory newborn screening. With the literal translation of “father of the people,” it provides the state with the right to usurp parental rights to protect a child when parents are unable or unwilling to do so. By mandating the disorders on a newborn screening panel, the state deems that parents who refuse testing are endangering their child’s health and so the state must intervene to protect the child’s welfare. Often these mandatory statutes allow parents rare justifications, usually only religious reasons, for opting out of testing. When there is clear evidence that early case detection and treatment improves medical outcomes, then mandatory testing seems justified. However, when evidence on treatment efficacy is lacking or when physicians are unable to predict the se-

Table. Principles of Early Disease Detection

<table>
<thead>
<tr>
<th>Principle</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>The condition sought should be an important health problem.</td>
<td></td>
</tr>
<tr>
<td>There should be an accepted treatment for patients with recognized disease.</td>
<td></td>
</tr>
<tr>
<td>Facilities for diagnosis and treatment should be available.</td>
<td></td>
</tr>
<tr>
<td>There should be a recognizable latent or early symptomatic stage.</td>
<td></td>
</tr>
<tr>
<td>There should be a suitable test or examination.</td>
<td></td>
</tr>
<tr>
<td>The test should be acceptable to the population.</td>
<td></td>
</tr>
<tr>
<td>The natural history of the condition, including development from latent to declared disease, should be adequately understood.</td>
<td></td>
</tr>
<tr>
<td>There should be an agreed policy on whom to treat as patients.</td>
<td></td>
</tr>
<tr>
<td>The cost of case-finding (including diagnosis and treatment of patients diagnosed) should be economically balanced in relation to possible expenditure on medical care as a whole.</td>
<td></td>
</tr>
<tr>
<td>Case-finding should be a continuing process and not a “once and for all” project.</td>
<td></td>
</tr>
</tbody>
</table>

*aFrom Principles and Practice of Screening for Disease.*
verity of disease, then mandatory screening does not seem as reasonable. For example, short-chain acyl-CoA dehydrogenase deficiency and 3-methylcrotonyl-CoA carboxylase deficiency are currently screened disorders for which recent large-scale population-based studies report difficulty in determining who will develop severe vs asymptomatic disease. Currently, 30 states have mandatory screening for short-chain acyl-CoA dehydrogenase deficiency, and 38 states have mandatory testing for 3-methylcrotonyl-CoA carboxylase deficiency.

**FIRST, DO NO HARM**

A discussion of the benefits of any screening program is incomplete without addressing the potential harms involved. Harm can develop in a number of ways in a screening program, the most fundamental being misidentification of individuals as having a disease when they do not (false-positive result) or as not having a disease when they in fact do (false-negative result). While it would be ideal to use a test that avoids false positives or false negatives, the reality is that maximizing a test’s ability to detect disease when it is present (sensitivity) necessitates a trade-off with the test’s ability to provide a negative result when a condition is absent (specificity). Because the goal of screening programs is to detect early disease, a screening test is chosen to maximize the identification of individuals with a disease (minimizing false negatives). However, this leads some people without a disease to be preliminarily misidentified as at high risk for having the disease (false positive) until further confirmatory testing excludes disease. Importantly, when the number of multiple independent screening tests performed for each individual increases, the number of false-positive results inevitably rises, even when a test is highly specific.

The potential harm generated from false-positive results is a familiar controversy within newborn screening. In the early years of PKU screening, medical professionals observed that mothers of newborns with false-positive PKU results manifested persistent anxiety about the general health of their child despite reassurance that the initial test result was a false positive. This behavior was termed the PKU anxiety syndrome, and physicians feared the problem would persist and perhaps worsen with increased mass screening. Their foresight was prophetic. During the past 40 years of newborn screening, many studies have demonstrated similar parental anxiety associated with false positives from newborn screening for multiple disorders, including congenital hypothyroidism and cystic fibrosis. This problem persists even in the current era of screening with tandem mass spectrometry. A recent study revealed that infants with false-positive newborn screening results had twice as many hospitalizations for unrelated illness compared with infants with normal results (21% vs 10%; P = .006). Mothers of infants with false-positive results scored higher on measures of stress and parental dysfunction, even after the result was deemed a false positive.

Some believe that this parental anxiety is a manifestation of the vulnerable child syndrome. First described in 1964 by Green and Solnit, this phenomenon arises when a parent experiences a real or imagined threat to his or her child’s health. The threat causes parents to misperceive their child as vulnerable and, as a result, to overestimate their child’s risk of illness, both at baseline and in the context of a benign illness. This parental anxiety and misperception of risk is not without significant short- and long-term consequences for both parent and child. Children perceived as vulnerable by their parents are inappropriately restricted from normal activity, have more school absences and behavioral disorders, suffer from anxiety, and use more health care resources.

It should be noted that the identification of potential harms associated with newborn screening does not mean that the system should be dismantled. Rather, we must recognize and understand these harms so that we can develop interventions to mitigate them, while preserving the benefits derived from newborn screening.

**COMMUNICATING WITH PARENTS**

While the debate over risks and benefits continues, health care providers face the important task of communicating with parents about the benefits and risks of newborn screening as well as the implications of test results. This undertaking is complicated by the logistics of screening and the increasing number of disorders that providers must discuss with parents. Because newborn screening is performed before hospital discharge, the infant’s primary care physician often does not have a chance to discuss screening with a parent until after an abnormal test result has been received, at which time the parent may be understandably distressed.

As a result, public health officials have advocated placing some communication responsibility in the hands of obstetricians. While a plausible alternative, it will require that we provide obstetricians with the educational resources necessary to provide optimal counseling to parents. In a recent survey of prenatal providers, only 33% of respondents (11% response rate) discussed newborn screening with their patients. Lack of knowledge and the belief that pediatricians would discuss these issues were the most commonly cited reasons for not discussing newborn screening and tandem mass spectrometry. Of note, even pediatric providers have reported feeling poorly prepared to discuss the results of newborn screening tests with parents. This discomfort will likely grow with the addition of rarer disorders with which the child’s provider is unfamiliar.

In an effort to assist the primary care physician, the American Academy of Pediatrics has provided fact sheets, which contain information regarding the natural history, etiology, and management for many of the disorders being screened. Though, for some disorders, such as short-chain acyl-CoA dehydrogenase deficiency, a variable prognosis makes it difficult to predict whether a child will have transient and mild symptoms or severe disease. These situations only add to the challenges physicians face in counseling parents about the implications of test results and the uncertainty that a family must confront. Such scenarios are likely to become more common as our ability to detect abnormalities outpaces our understanding of a disease’s natural history.
Given these issues, how well are we doing with our current communication practices? Surprisingly, we have little systematic data regarding what prospective and new parents understand about newborn screening. We do not know what proportion of new parents in this country are even aware that newborn screening testing has been done at birth, though we have reason to suspect the percentage is low. As of 1999, 10 states did not notify parents before screening infants. A more recent evaluation of educational newborn screening materials for parents revealed that none of the literature included all of the information recommended by the American Academy of Pediatrics Task Force on Newborn Screening, and only 13% explicitly mentioned the possibility of receiving a false-positive result. Clearly, parents seem less than ideally prepared to receive information about their child’s positive result, whether it turns out to be true or false. This is troubling, because recent studies have suggested that parental understanding of false-positive results and physician communication are closely linked to lingering parental anxiety about false-positive results.

WHERE DO WE GO FROM HERE?

The increase in newborn screening brought on by tandem mass spectrometry is already stressing the current newborn screening system. Meanwhile, the next technological advance, DNA microarrays, is on the horizon and fast approaching. It will permit screening for hundreds of potentially disease-causing genetic mutations via newborn screening. The challenge of distinguishing disease-causing from benign mutations looms large as we seek to avoid labeling children as potentially diseased and causing undue distress and “medicalization.”

Society has spent considerable resources refining the technology used in expanded newborn screening. Yet, it should be noted that technology is only one facet of a well-functioning newborn screening program, which must have both excellent detection and follow-up services. Important questions about this process remain unanswered. For an infant with a positive screening result, what barriers does a primary care physician face in coordinating a medical evaluation and communicating with the family? What obstacles do families confront in the time after a newborn screening result returns positive? When a positive newborn screening result is confirmed, how can coordination of follow-up care be optimized? These fundamental questions must be addressed to optimize collaboration between primary care and specialty care physicians and to ensure the continued success of newborn screening in the 21st century.

Accepted for Publication: January 25, 2007.
Correspondence: Beth A. Tarini, MD, MS, Division of General Pediatrics, Child Health Evaluation and Research Unit, University of Michigan, 300 N Ingalls St, Room 6C11, Ann Arbor, MI 48109-0456 (btarini@umich.edu).
Financial Disclosure: None reported.
Disclaimer: The views expressed herein do not necessarily represent the views of the University of Michigan.
Additional Contributions: Wylie Burke, MD, PhD, Gary Freed, MD, MPH, and Sarah Clarke, MPH, provided helpful comments during the preparation of the manuscript. Kristin Poole, BA, provided editing assistance.

REFERENCES


18. Diamond A. Phenylalanine levels of 6-10 mg/dL may not be as benign as once thought. Acta Paediatr Suppl. 1994;407:89-91.


©2007 American Medical Association. All rights reserved.