

Chapter 5

Newborn Screening for Congenital Disorders

CONTENTS

	<i>Page</i>
Introduction	93
Factors Affecting the Effectiveness of Newborn Screening	97
Percentage of Affected Infants Detected.....	98
Organization and Management of Newborn Screening Services	98
Timing and Number of Newborn Blood Specimens	100
Laboratory Performance: Quality Assurance and Proficiency Testing Programs +.....	102
Cost-Effectiveness of Seven Newborn Screening Strategies.	103
Components of OTA's Cost-Effectiveness Analysis	105
Findings of the Base Case Analysis	106
Components and Findings of the Sensitivity Analysis	108
Conclusions	109
Costs of Laboratory Testing for Four Additional Disorders	111
Biotinidase Deficiency	111
Sickle Cell Anemia.	112
Cystic Fibrosis.	112
Congenital Adrenal Hyperplasia,	112
Financing and Reimbursement for Newborn Screening and Treatment	112
Conclusions 4. ... <...	114

Box

<i>Box No.</i>	<i>Page</i>
s-A. Process of Newborn Screening and Treatment.	97

Figure

<i>Figure No.</i>	<i>Page</i>
s-I. Newborn Screening Strategies Compared in OTA's Cost-Effectiveness Analysis,.....	104

Tables

<i>Table No.</i>	<i>Page</i>
5-1. Nine Congenital Disorders Currently Detectable by Newborn Screening . . .	94
5-2. Effectiveness of Newborn Screening and Early Treatment for Nine Congenital Disorders +	95
5-3. Percentage of Affected Infants Detected by Newborn Screening for Five Congenital Disorders	99
5-4. Effectiveness and Health Care Savings of Newborn Screening Strategy I Compared to No Screening.	107
S-S. Incremental Effectiveness and Health Care Costs of Newborn Screening Strategies Compared to Strategy I.....	107
5-6. Sensitivity Analysis: Net Health Care Savings and Number of Cases Detected by Strategy I Compared to No Screening.	110
5-7. Sensitivity Analysis: Incremental Effectiveness and Health Care Costs of Newborn Screening Strategies Compared to Strategy I	110

Newborn Screening for Congenital Disorders

INTRODUCTION

The screening of large populations of newborns for congenital disorders began as a public health activity in 1961 with screening for phenylketonuria (PKU). PKU, an inherited disorder of metabolism of the amino acid phenylalanine, occurs in about 1 in 10,000 to 1 in 15,000 infants and is associated with severe mental retardation unless treated. The goal of newborn screening for this disorder is to detect the condition in the first week of life, confirm the diagnosis, and initiate treatment before 2 to 4 weeks of age. If the treatment is begun by this time and maintained at least throughout the child's development, the irreversible mental retardation that is the natural consequence of untreated PKU can be avoided.

Newborn screening has generally been limited to diseases such as PKU that are not clinically recognizable in time to treat before severe and irreversible consequences have occurred (57,442). These are diseases that are present throughout the life of an affected individual, do not get better (and often worsen) with time, and can result in severe mental retardation, physical disabilities, and even sudden death if untreated in the first days or weeks after birth. Although the number of conditions that fall into this category is small, and each of these conditions is relatively rare, newborn screening followed by early and sustained treatment can make the crucial difference in affected infants. Information about the incidence and natural course of PKU, congenital hypothyroidism, galactosemia, maple syrup urine disease (MSUD), and several other disorders that can be detected through tests on a blood specimen taken in the neonatal period are summarized in table s-1. I

¹ More detailed information about these and other disorders is presented in app. H.



Photo credit: March of Dimes Birth Defects Foundation

Screening newborns for certain congenital disorders can serve a critical function — alerting physicians to the need for treatment of infants who may not yet show signs or symptoms of a disorder.

In most States, newborn screening is mandated by law, except in the case of parental refusal on religious or other grounds. In a few States (North

Table 5-1.—Nine Congenital Disorders Detectable by Newborn Screening

Disorder	Approximate Incidence per 100,000 births	Problem	Natural course without adequate treatment
Phenylketonuria (PKU)	7	Impaired metabolism of the amino acid phenylalanine	Severe mental retardation, shortened lifespan
Congenital hypothyroidism (CH)	29	Deficiency of the hormone thyroxine needed for brain development and physical growth	Mental retardation, physical abnormalities, premature death, abnormal growth
Galactosemia (GA)	2	Deficiency of the enzyme needed to metabolize galactose, a type of sugar in milk	Life-threatening septicemia and liver damage in infancy; mental retardation and cataracts in survivors; leads to death if untreated
Maple syrup urine disease (MSUD)	0.4	Deficiency of the enzyme needed to metabolize branched-chain amino acids	Life-threatening acidemia and neurologic dysfunction in infancy, mental retardation in few survivors; leads to death if untreated
Homocystinuria (HC)	0.5	Deficiency of the enzyme needed to metabolize the amino acid homocystine	Developmental retardation, dislocation of ocular lenses, life-threatening thrombosis, skeletal manifestations
Biotinidase deficiency (BD)	3	Deficiency of the enzyme needed to metabolize the B vitamin biotin	Life-threatening neurologic dysfunction, developmental delay and hearing loss in survivors (milder and asymptomatic cases may occur)
Sickle cell anemia (SCA)	32	Abnormality of the red blood cells that causes them to be sickle-shaped	Life-threatening infections especially in infancy, chronic hemolytic anemia and vaso-occlusive crises in childhood and adulthood, risk of premature death, shortened lifespan
Cystic fibrosis (CF)	40	Disorder of the exocrine glands whose cause is unknown	Poor growth, digestive problems, life-threatening chronic obstructive lung disease with recurrent pneumonia; risk of death in early adulthood
Congenital adrenal hyperplasia (CAH)	5	Inability to produce the hormones needed to manage stress and to control, salt content of tissues, combined with the excessive buildup of male hormones	Life-threatening salt-wasting crises for some in infancy, reproductive dysfunction, and abnormal physical development

SOURCE Office of Technology Assessment, 1988

Carolina, Delaware, Vermont, and the District of Columbia) newborn screening is voluntary (32). Newborn screening for PKU and another condition—congenital hypothyroidism—is now conducted in all 50 States and the District of Columbia. Congenital hypothyroidism, resulting in a deficiency of the hormone thyroxine, is much more common than PKU, affecting about 1 in 3,000 to 1 in 4,000 infants. Like PKU, this disorder can cause permanent brain damage if untreated or treated too late. In addition to offering the tests for PKU and congenital hypothyroidism, some States offer tests for other disorders: 35 States test for galactosemia, 24 States test for MSUD, and 22 States test for homocystinuria (640a).² Galactosemia, MSUD, and homocystinuria are less

common than PKU, but all of them have serious adverse consequences if left untreated.

Recently, the scope of routine newborn screening has expanded to include tests for several additional diseases, some of which are more common in certain populations than PKU, but not all of which are as treatable as PKU. An increasing number of programs have begun pilot screening projects for biotinidase deficiency and sickle cell anemia, and a small number of programs are beginning to screen for cystic fibrosis and congenital adrenal hyperplasia. OTA found that as of 1986, approximately 18 States were screening for biotinidase deficiency, 9 States were offering tests for sickle cell anemia, 3 States were offering tests for cystic fibrosis, and at least 2 States were screening for congenital adrenal hyperplasia. Information about the effectiveness of newborn screening and treatment for these and other congenital disorders is presented in table 5-2.

²A few programs also offer screening tests for a number of other rare metabolic disorders, such as tyrosinemia or histidinemia.

Table 5-2.—Effectiveness of Newborn Screening and Early Treatment for Nine Congenital Disorders

Disorder	Optimal screening time after birth	Treatment following screening	Effectiveness of screening and early treatment
Phenylketonuria (PKU)	3-5 days	Phenylalanine-restricted dietary products continued indefinitely	Normal mental and physical development
Congenital hypothyroidism (CH)	3-5 days	Thyroxine supplements indefinitely	Normal mental and physical development
Galactosemia (GA)	Before 5 days (or in cord blood taken at birth)	Elimination of galactose-containing foods indefinitely	Life saved in neonatal period, normal mental and physical development in majority of cases; coordination and speech problems; gonadal failure in some females
Maple syrup urine disease (MSUD)	1-5 days	Dietary restriction of branched-chain amino acids indefinitely	Life saved in neonatal period normal mental and physical development, risk of sudden death at a later age in some cases
Homocystinuria (HC)	3-4 weeks	Dietary restriction of methionine and supplementation of cystine and vitamin B ₆ indefinitely	Normal mental development, some physical problems may remain
Biotinidase deficiency (BD)	Before 5 days	Oral biotin supplements	Life saved in neonatal period for some or avoidance of neurologic damage
Sickle cell anemia (SCA)	1st week	Prophylactic penicillin and pneumococcal vaccine, ongoing supportive therapy	Reduce risk of death in infancy and early childhood from complications of infection
Cystic fibrosis (CF)	1st week	Prophylactic vitamin and salt supplements, pancreatic enzyme replacement therapy, antibiotics, supportive respiratory therapy	May improve growth and development in childhood, long-term effects under investigation
Congenital adrenal hyperplasia (CAH)	2-5 days	Intravenous salt solution, hormone therapy	Life saved in neonatal period for some, aid sex assignment in infant girls with CAH, normal sexual development

SOURCE: Office of Technology Assessment 1988

In some cases, infants with biotinidase deficiency and some forms of congenital adrenal hyperplasia are at risk of sudden death if not immediately treated, often before physicians are able to make clinical diagnoses of their conditions. In these situations, newborn screening can serve a critical function—alerting physicians to the need for treatment in infants who may not yet show specific signs of their conditions. In general, however, little is known about the natural history of either biotinidase deficiency or congenital adrenal hyperplasia. For example, it is not known how many infants with biotinidase deficiency are at risk of sudden death and how many have a less severe form of the disease that does not require treatment. Congenital adrenal hyperplasia takes at least two forms, one requiring immediate life-sustaining treatment and another form that does not require such treatment and may be diagnosed clinically. Although some infants with biotinidase deficiency and congenital adrenal hyperplasia can certainly be helped by early diagnosis through

newborn screening, there is still not enough information available to judge whether screening of all infants for either of these conditions is desirable, and how the benefits of such screening compare to the benefits from newborn screening for other treatable conditions.

For sickle cell anemia and cystic fibrosis, treatment that will entirely prevent the major long-term disabilities characteristic of these disorders does not exist. In the case of sickle cell anemia, however, newborn screening may have an important intermediate goal. A certain percentage of infants with sickle cell anemia are at risk of overwhelming infection and sudden death in the first few years of life. If their sickle cell disease is identified before infection occurs, affected infants can be given prophylactic antibiotics that significantly reduce the risk of infection and lower the overall mortality rate from the disease in early life (198). For patients who have passed this critical period, there is no hard evidence that screening and diag-

nosis in the first few weeks of life leads to improved long-term survival. Nevertheless, some observers think that the effectiveness of prophylactic antibiotics to prevent infection in infants with presymptomatic cases of sickle cell anemia is probably a sufficient reason to include testing for sickle cell anemia in routine newborn screening. The issues of the long-term value and cost-effectiveness of early diagnosis and treatment of sickle cell anemia, however, remain to be resolved.

A similar situation exists with respect to newborn screening for cystic fibrosis. In recent years, earlier and more intensive treatment following clinical diagnosis of patients with cystic fibrosis has contributed to a generally longer survival among these individuals, who now live to early adulthood rather than dying in early childhood as patients did several decades ago. A variety of clinical observations suggest that early awareness of cystic fibrosis allows improvement in certain aspects of patients' physical condition—notably, in their early nutritional status (472). The major factor in long-term survival of patients with cystic fibrosis, however, is chronic lung disease; whether newborn screening and even earlier treatment will improve survival above the current average survival of 20 to 25 years is not known, although a controlled clinical trial being conducted in Wisconsin may help resolve the issue.³

The next section of this chapter describes the factors that influence the effectiveness of newborn screening. There are two distinct aspects of the overall effectiveness of newborn screening in detecting and treating affected infants. One aspect is the effectiveness of the screening test and the

overall screening process in detecting affected infants in need of treatment. This aspect includes how well the screening program coordinates abnormal laboratory findings with confirmatory diagnosis and initiation of adequate treatment. Another aspect is the efficacy of available treatments. Without effective treatment available to alter the natural course of the disease, early screening would be ineffective at best, and possibly even harmful. The discussion below addresses mainly the first aspect, but several issues pertaining to the treatment of specific diseases are discussed in appendix H and elsewhere in this chapter.

Using the best information available on the effectiveness and costs of screening for specific disorders, OTA performed a cost-effectiveness analysis of a basic strategy for newborn screening compared to no screening and of six expanded strategies compared to the basic strategy. The basic strategy consists of a one-specimen testing process to identify cases of PKU and congenital hypothyroidism. The expanded strategies include testing for additional selected congenital diseases and more intensive testing for PKU and congenital hypothyroidism. All of the various screening strategies involve combinations of tests for two or more of the following five disorders: PKU, congenital hypothyroidism, homocystinuria, galactosemia, and MSUD.

Tests for biotinidase deficiency, sickle cell anemia, cystic fibrosis, and congenital adrenal hyperplasia are not included in the strategies considered in the cost-effectiveness analysis. An increasing number of newborn screening programs have begun pilot screening projects for biotinidase deficiency and sickle cell anemia, and a small number of such programs are beginning to screen for congenital adrenal hyperplasia and cystic fibrosis. However, there is insufficient information on the long-term costs and effects of screening for these conditions to analyze the cost-effectiveness of strategies involving tests for these disorders. Nevertheless, this chapter does provide information on the costs of detecting these disorders in ongoing screening programs.

The chapter also discusses the financing of newborn screening and treatment programs. It emphasizes, in particular, the recent changes in the overall level of Federal support for such programs.

³The controlled clinical trial being conducted in the Wisconsin Cystic Fibrosis Centers is seeking to evaluate the potential pulmonary benefits, as well as the potential psychological risks, of newborn screening for cystic fibrosis (166). One-half the newborn population of the State will be screened at random, and newborns identified as having cystic fibrosis will be enrolled in a comprehensive evaluation and treatment protocol. At the end of 3.5 years, infants identified as having cystic fibrosis by the newborn screening test will be compared with age-matched patients diagnosed through conventional medical channels, and their health status will be monitored for at least another 3.5 years. Any differences in pulmonary status between the two groups at the end of this period may then be assumed to have resulted from the age at which the children were first diagnosed. Where possible, a controlled clinical trial such as this can be valuable in assessing the benefits and risks of a particular screening test before it is adopted on a routine basis in newborn screening programs.

FACTORS AFFECTING THE EFFECTIVENESS OF NEWBORN SCREENING

As described in box 5-A, newborn screening seeks to identify biochemical abnormalities that suggest the presence of disease in affected but as yet asymptomatic infants. Infants who test positive in the initial screening test can be evaluated further to diagnose the specific disorder and to determine the best mode of therapy. Since for most of the disorders targeted by newborn screening, treatment must begin in the first 2 to 3 weeks of life (as in PKU and congenital hypothyroidism) or even the first few days of life (as in galactosemia, MSUD, and congenital adrenal hyperplasia), there is a premium on identifying affected cases early, rapidly, and unambiguously.

The effectiveness of newborn screening in identifying affected infants depends in part on the abil-

ity of the screening program to collect blood specimens from all infants and to perform the tests properly and in time to initiate treatment. Potential errors in the overall screening process include not collecting blood specimens as needed; losing specimens in transit; collecting specimens too early or too late; reporting errors; and lack of adequate followup testing. Thus, the organization and management of newborn screening services, the timing and number of newborn blood specimens, and laboratory performance have major bearing on the effectiveness of newborn screening. Even without errors such as those just mentioned, however, there would still be upper limits on the technical ability of a newborn screening program to detect all affected infants, since none

Box 5-A.—Process of Newborn Screening and Treatment

The process of newborn screening involves collecting a few drops of blood from an infant's heel onto specially designed filter paper with accompanying information including the infant's name, date of birth, and other relevant data. The specimen is usually collected just prior to the infant's discharge from the hospital nursery or, for nonhospital births, during the first week of life.

Once the blood spots have dried, the filter paper is sent to a screening laboratory where specimens are analyzed using procedures and equipment designed to handle many samples simultaneously. (A medium-sized newborn screening program may screen 40,000 to 50,000 infants annually; larger programs, particularly those regional in scope, screen 100,000 to 150,000 specimens annually.) The distinction between affected and normal infants is either qualitative (presence or absence of a protein, for example) or quantitative (greater or lesser concentration of a chemical in the blood), although the distinction is made within certain bounds of sensitivity and specificity.¹ The test procedure used in newborn screening has to be sufficiently reliable and valid for use on a large-scale basis (280,281).

If the screening test result reveals an abnormality, the program staff notifies the family's physician or refers the family to a qualified specialist. The screening program may or may not be responsible for coordinating diagnostic services or treatment for infants with abnormal screening results. Followup services typically involve long-term treatment and evaluation and are guided by various health professionals, including pediatricians, geneticists, nutritionists, nurses, and social workers. The screening program may develop specific protocols for the treatment of unusual or rare diseases (378). It is usually the responsibility of the screening program to keep long-term records of the screening results and actions following identification of affected infants.

The initiation of timely treatment and provision of adequate followup requires close coordination among the various components of the screening program: the hospital of birth, the screening laboratory, the infant's family and physician, and specialists at an appropriate diagnostic and treatment center who oversee long-term treatment and monitoring. Therapy for diseases detected through newborn screening usually continues throughout an individual's lifetime.

¹Sensitivity is a measure of the accuracy of a screening or diagnostic test in correctly identifying those who have the disease in question; specificity is a measure of the accuracy of such a test in correctly identifying those who do not have the disease.

of the tests have complete sensitivity (i.e., the ability to classify correctly all affected infants as affected). Therefore, this section focuses not only on technical sensitivity of the tests used in newborn screening, but also on the organization and delivery of screening services that affect the effectiveness of newborn screening in practice.

Percentage of Affected Infants Detected

Data on the overall sensitivity and specificity of the tests for PKU and congenital hypothyroidism (two tests which have been in use for over 10 years) are not collected or evaluated on a national basis. The accuracy of these newborn tests in identifying blood samples with abnormal levels of phenylalanine or thyroxine, however, is generally considered to be very good. Estimates of the sensitivity and specificity of the tests for PKU and congenital hypothyroidism in practical use have been reported by various individual newborn screening programs, but the estimates vary, depending in part on the size of the program and how the tests are conducted. Furthermore, biological variation among individuals with a given condition makes it difficult to distinguish precisely between normal and abnormal findings, and the same result can be interpreted in different ways in different laboratories. In the case of tests to detect PKU, for example, the sensitivity depends on factors such as the cutoff point above which a sample is considered to have an abnormal amount of phenylalanine and also on the age of the infant when the blood sample was taken.

If there is limited information on the sensitivity of newborn screening tests for PKU and congenital hypothyroidism, there is even less information on the sensitivity of the newer tests for other conditions such as biotinidase deficiency, cystic fibrosis, and congenital adrenal hyperplasia. Like missed cases of PKU and congenital hypothyroidism, missed cases of these other conditions do not necessarily come to the attention of the appropriate State agency or to a central national office such as the Centers for Disease Control (CDC).

Lacking accurate data on sensitivity in a technical context as well as a broader context, OTA consulted with experts in newborn screening pro-

grams and in academic genetics and pediatrics departments to develop estimates and ranges for the percentage of affected infants detected by five newborn screening tests. Estimated incidence rates for PKU, congenital hypothyroidism, galactosemia, MSUD, and homocystinuria, along with estimated percentages of affected infants found by these tests on first and second specimens are shown in table 5-3. The estimates in the table reflect these experts' consensus regarding the technical sensitivity of the tests plus practical considerations in applying the tests to large populations. In MSUD testing, for example, the sensitivity of the test may generally be high, but MSUD is so rapidly fatal that, in practice, the test may not always be performed and reported in time to initiate effective treatment.

Organization and Management of Newborn Screening Services

Canada and the United States are the only developed countries offering newborn screening that do not have a *national* screening program.⁴ In the absence of a national newborn screening program or national set of minimum standards, each State has taken a slightly different approach to providing screening services. A few States have joined with neighboring States to form regional programs that together account for the screening of about 20 percent of births each year (279). Currently, there are three regional programs in the United States:

1. the New England regional program (covering Massachusetts, Rhode Island, Connecticut, Maine, and New Hampshire);
2. the Rocky Mountain States regional program (covering Colorado, Arizona, and Wyoming); and
3. the Pacific Northwest regional program (covering Oregon, Idaho, Nevada, and Alaska).

Although each of these regional programs has a central screening laboratory and coordinating center clinical followup services are provided at local medical centers.

⁴Approximately 30 countries outside the United States also offer newborn screening services for various diseases (640a).

Table 5-3.—Percentage of Affected Infants Detected by Newborn Screening for Five Congenital Disorders^a

Disorder	Incidence of the disorder		Percentage of affected cases found (sensitivity)			
			On first screening test		On second screening test	
	Median	Range	Median	Range	Median	Range
Phenylketonuria (PKU)	1 in 12,000	1 in 10,000 to 1 in 15,000	96/0	94-97%	99 %/0	98-1 00°/o
Congenital hypothyroidism (CH)	1 in 3,500	1 in 3,000 to 1 in 4,000	93 %/0	90-95 %/0	99 %/0	95-1 00 %/0
Galactosemia (GA)	1 in 62,000	1 in 40,000 to 1 in 80,000	85/0	80-90/0	—	—
Maple syrup urine disease (MSUD)	1 in 227,000	1 in 200,000 to 1 in 300,000	85/0	80-90/0	—	—
Homocystinuria (HC)	1 in 150,000	1 in 100,000 to 1 in 300,000	—	.	80 0/0	75-85/0

^aOTA solicited the estimates presented in this table for use in the cost-effectiveness analysis of seven newborn screening strategies, which is presented in this chapter. See text for discussion of sources.

SOURCE: Office of Technology Assessment, 1988.

The majority of infants—about 71 percent—are covered by newborn screening programs in individual States. Most State newborn screening programs do have a centralized screening laboratory, but only some (e.g., California) have an organized program of services linking the laboratory with followup, treatment, and monitoring. A few States, together accounting for about 9 percent of all births, operate without a central laboratory or a centrally organized program. These States (e.g., Nebraska, South Dakota, Oklahoma, and Hawaii) rely on an informal network of individual families, physicians, and a combination of public and private laboratories to provide screening and followup.

Because of the particular requirements and difficulties inherent in newborn screening, it seems reasonable that a coordinated system of services would provide the optimal organization for ensuring that all infants are satisfactorily screened and that affected infants are followed up and treated. In a 1975 publication, the National Academy of Sciences recommended the development in the United States of regional screening programs based on an area's birth rate rather than its State boundaries, particularly in areas where low population densities and low budgets would restrict access to high-quality screening services (442). Other sources have also recommended centralizing laboratories and coordinating various components of screening programs (e.g., 31,53). There is no ongoing system in place, however,

to assist States in implementing these recommendations and in developing coordinated screening programs.

In some areas, the lack of a coordinated network of services may be reducing the overall effectiveness of newborn screening by putting infants at risk for not being screened or for not receiving appropriate treatment. There are no national data on the number of infants at risk, however, because there is no central system for collecting comprehensive data with which to monitor and compare the outcomes of newborn screening in the State and regional programs. If a system were to be established, several specific indicators of effectiveness could be used to evaluate the performance of screening programs—e.g., the percentage of infants screened in a given area, the time that elapses between the completion of the screening tests and the initiation of treatment, and the frequency of “errors” in the process (from collecting unsatisfactory specimens to actually failing to identify an affected child).

A recent study of errors in the process and analysis of newborn screening specimens in Oregon's program suggests that collecting and monitoring such data on a national basis may be important (647). Using data from a computer-based surveillance system designed to track individual infants through the screening process and to monitor screening practices of individual hospitals, birth centers, and home deliveries, Tuerck and colleagues found that over one-half (58 percent) of

the 23,717 specimens collected in Oregon over a recent 4-month period were submitted with one or more screening practice “errors” (647). Any one of these errors, if uncorrected, could have allowed an affected child to go undetected and untreated or could have caused a serious delay in diagnosis and treatment. The five categories of errors, in order of frequency were:

1. 28.2 percent of the specimens had omissions of demographic information on the screening card (e. g., no name or unreadable information);
2. 27.7 percent of the specimens were taken at suboptimal times (e. g., the first one taken before 24 hours of age and the second taken after 14 days of age);
3. 22.9 percent were not retested as required by the State of Oregon;
4. 16.2 percent of the specimens took longer than 5 days to go from the birthing facility to the screening laboratory; and
5. 0.6 percent were unsatisfactory for testing, usually caused by poor techniques for blood collection.

A retrospective study that surveyed screening laboratory directors in 49 States identified 76 missed cases of PKU and congenital hypothyroidism that occurred during the history of the programs (279). The primary causes identified were the following: laboratory procedures (45 percent); errors in followup (16 percent); specimen collection errors, mostly the lack of an initial specimen (14 percent); false negative results⁵ due to biologic variation in disease expression (11 percent); and unidentified causes (14 percent). Although the investigators reported that 76 missed cases was probably an underestimate of the actual number of missed cases, these data do give a rough indication of the types of errors (mostly “human” errors) that allow affected infants to escape detection and treatment.

⁵ false negative is an affected person lease incorrectly identified by a test as not having the condition.

Timing and Number of Newborn Blood Specimens

Optimal Timing of Blood Specimen Collection

The optimal time for collecting a blood specimen for newborn screening during the newborn hospitalization⁶ depends on the characteristics of the disorders for which infants will be screened. The optimal time for collection is the earliest point at which the biochemical markers for the targeted disorders are present in the blood in high enough or low enough amounts (depending on the disorder) to identify the disorders accurately. As affected infants progress, these markers become more and more unmistakable, but irreversible symptoms may begin to occur. In PKU screening, for example, testing too early (e.g., in the first day of life) could miss an infant with the disease for biological reasons, and testing too late (e.g., in the second or third week of life) could lead to starting treatment too late to avert the severe long-term consequences of the disease.

Another consideration is that when tests for several disorders are performed on the same sample, the optimal times for the different assays may not overlap sufficiently to permit equally reliable results for each test. While testing at 3 to 5 days of life is usually considered optimal for PKU screening, testing for homocystinuria, for example, using the same blood sample would detect only about one-half of infants with homocystinuria; testing at 3 to 4 weeks, rather than in the first week, is considered optimal for detecting homocystinuria.

Number of Specimens

A single blood specimen collected from infants before discharge from the hospital has generally been considered sufficient in screening for PKU

⁶ The vast majority of U.S. infants are born in hospitals (712), and the collection of blood specimens for newborn screening is a routine procedure before infants are discharged. Births that occur outside of hospitals are difficult to monitor, and although birth attendants are instructed to collect a specimen, enforcement is difficult, even if testing is mandatory.

and congenital hypothyroidism.⁷ Two recent developments have cast doubt, however, on the adequacy of a single specimen to test for PKU and congenital hypothyroidism. These developments, discussed below, have led some people to advocate the collection of a second blood specimen during the second or third weeks of an infant's life. A second specimen can be collected from an infant either during a well-child visit to a physician or nurse practitioner or during an outpatient visit to a hospital. In Oregon, the collection of a second blood specimen is mandatory at the time of the first well-child visit between 2 and 6 weeks of age (325) and is achieved in about 85 to 90 percent of infants (81).

PKU Testing.—One development that casts doubt on the adequacy of a single specimen for newborn screening for PKU testing is the trend toward discharging an increasing percentage of newborns from the hospital before the optimal time for testing for PKU at about 3 to 5 days of age. In 1985, an estimated 41.4 percent of infants born in U.S. hospitals were discharged from the hospital nursery before the third day of life—an increase from 30.9 percent of infants in 1980 (707). The increasing percentage of blood specimens being taken in the first 1 or 2 days of life may increase the probability that tests on infants with PKU will be falsely negative, possibly causing some infants with PKU to escape detection and treatment.⁸ Since serum phenylalanine levels in PKU cases rise steadily over the first few days of

life and the difference between phenylalanine levels in normal and PKU infants increases with each day of life, some analysts suggested that up to 16 percent of infants with PKU could be missed for biologic reasons if tests were performed on blood specimens taken before the infant was 24 hours old (282,283,409).

So far, empirical data have not borne out these analysts' predictions (377,423,576). Although the available data are from clinics and institutions that care for children with mental disabilities and may underestimate the number of missed cases of PKU (because once a child is mentally retarded, the underlying cause may not be determined and reported as PKU), in practice, phenylalanine levels in PKU infants are usually higher than in normal infants even on the first day of life, and various technical adjustments can be made to raise the sensitivity of the assays used to identify PKU in first day samples (409,574). Thus, it does not appear likely that the collection of blood samples around 24 hours of age would result in missed cases of classic PKU as a result of low phenylalanine levels in the blood on the first day. Even so, however, an infant with PKU might be missed by the screening system for other reasons (e.g., because the infant was not being screened at all, or because the infant's first blood specimen was lost in transit).

One way of gaining assurance that infants with PKU are not being missed on the first test is to obtain a second blood specimen for retesting all infants. Experience in the Texas and Oregon screening programs, which perform a second screen on a majority of infants, suggests that the probability of missing PKU infants on the first test is low, because no additional cases of PKU have been found via tests on second blood specimens (81,640).

Congenital Hypothyroidism Testing.—A development that casts doubt on the adequacy of a single specimen in newborn screening for hypothyroidism is preliminary evidence that an additional 5 to 10 percent of infants with congenital hypothyroidism can be detected by second testing at 3 to 4 weeks of age—these are affected infants with no biochemical signs of congenital hypothyroidism on the first specimen taken during the first week of life (359,376); the severity of the hypothy-

⁷In some instances, second specimens have to be collected because of problems with the first specimen, but such instances are fairly uncommon. Reasons to collect a second specimen include failure of a lab to receive a first specimen, not enough blood present in the first specimen to complete the tests, incomplete demographic information on the specimen card (missing age of the infant when the sample was taken, illegible name, etc.), or a filter paper that contains anything other than the infant's blood.

⁸OTA is unaware of any cases of PKU that have been documented as being missed for this reason, but the theoretical probability that cases of PKU could be missed because of early hospital discharge of newborns is an important issue for newborn screening programs. Screening programs may be legally responsible for detecting all affected infants in the region covered—a fact that provides additional incentive to do as much as possible to identify all affected cases. Although routinely collecting a second specimen on all infants may be one of the most comprehensive approaches to guard against missed cases, this approach is certain to increase costs associated with newborn screening. It would have the advantage, however, of detecting diseases such as homocystinuria, which are not normally detectable in the first screening test.

roidism in the additional infants compared to the infants identified in the first test is still unknown. Several screening programs (e.g., Texas) routinely use a second specimen for hypothyroidism testing.

Laboratory Performance: Quality Assurance and Proficiency Testing Programs

The consequences of failing to identify an abnormal blood specimen, leading to a failure to diagnose and treat, can be catastrophic for infants with diseases such as PKU, congenital hypothyroidism, galactosemia, MSUD, biotinidase deficiency, and congenital adrenal hyperplasia. Furthermore, missed cases of PKU and congenital hypothyroidism have led to lawsuits against the State in which the infant was screened, the attending physician, the hospital of birth, or the Federal Government (in the case of a military birth) (279).⁹ The reasons for missed cases are diverse and involve errors in many stages of the process, from specimen collection, to laboratory analysis, to followup and treatment (279).

One step newborn screening laboratories can take to improve the reliability of their laboratory results and to maintain high-quality technical performance is to participate in proficiency testing or broader external quality assurance programs. Proficiency testing provides an opportunity for a laboratory to have an external check on its ability to identify abnormal specimens and recognize normal specimens. Such a check is especially important in low-volume laboratories in testing their laboratory ability to identify particularly rare conditions. By participating in proficiency testing on a regular basis, a laboratory can judge and compare its methods and test kits against others for precision and accuracy. In a broader sense, quality assurance programs, which include proficiency testing, facilitate the identification of laboratories

that are encountering technical problems and, through the combined experience of many participating laboratories, help to improve laboratory performance.

The major effort in external quality assurance for newborn screening programs in the United States has been that undertaken by CDC. To assist States in developing and maintaining high levels of accuracy and precision in their newborn screening programs, CDC operates two services:

1. a quality assurance program designed for newborn screening programs (called the "Standardization Program to Improve Laboratory Screening for Hypothyroidism, Phenylketonuria, and Other Inborn Metabolic Disorders"), and as part of that,
2. a proficiency testing program for newborn screening laboratories.¹⁰

The quality assurance program, which is jointly funded by CDC and the Health Resources and Services Administration, seeks to promote standards of good laboratory practice. Its objectives are to ensure that newborn screening laboratories accurately identify all cases of metabolic disorders in time to initiate treatment. Laboratory participation in CDC's quality assurance program for newborn screening is voluntary.

One of the central features of the quality assurance program operated by CDC is the production and distribution of quality control materials to laboratories for their internal use in maintaining the accuracy and precision of their screening tests for PKU and congenital hypothyroidism. This service is provided free of charge to State laboratories. Quality control materials are also provided to manufacturers to assist them in standardizing their equipment or testing kits (324).

CDC's proficiency testing service for newborn screening laboratories is one feature of an extensive proficiency testing program operated by CDC pursuant to the Clinical Laboratory Improvement Act of 1967 (Public Law 90-174, 42 U.S.C.). Newborn screening laboratories that qualify under the act's provisions as "interstate laboratories" (e.g.,

⁹Thus far, 20 to 25 lawsuits involving missed cases of PKU or congenital hypothyroidism have been brought to trial in the United States. Some of these cases have resulted in financial settlements as high as \$4 million (601). Each of these cases involved brain damage that could have been avoided by treatment if the children had been identified accurately in newborn screening. Settlements in such cases typically cover the cost of caring for a mentally retarded child, the economic value of the child's lost potential earnings throughout his lifetime, and may also include financial compensation for conscious pain and suffering (601).

¹⁰CDC discontinued its proficiency testing service for clinical laboratories in general in 1986, but it decided to retain these services for newborn screening for PKU and congenital hypothyroidism testing (592).

the Pacific Northwest and the New England regional programs) are required to participate in the program. Newborn screening laboratories that do not accept specimens across State boundaries may voluntarily participate in the CDC proficiency testing service for newborn screening, but they are not required to do so. Currently, almost all State newborn screening programs use CDC's proficiency testing service.

Although many observers believe that CDC's quality assurance program is an essential component of the process of newborn screening (324), no empirical studies have been done to measure the impact of the program on the effectiveness of

newborn screening in detecting and treating affected infants. According to interim data from a CDC analysis, however, the overall precision of newborn screening as measured by proficiency test results has improved in the past 5 years (64). Thirty-three percent of all PKU proficiency testing results in the first quarter of 1981 were more than 25 percent away from the consensus target value; by the end of 1985, the figure had dropped to 7.6 percent. This decrease in the range of values obtained from the participating laboratories indicates a general improvement in reliability among laboratories using different procedures (592).

COST= EFFECTIVENESS OF SEVEN NEWBORN SCREENING STRATEGIES

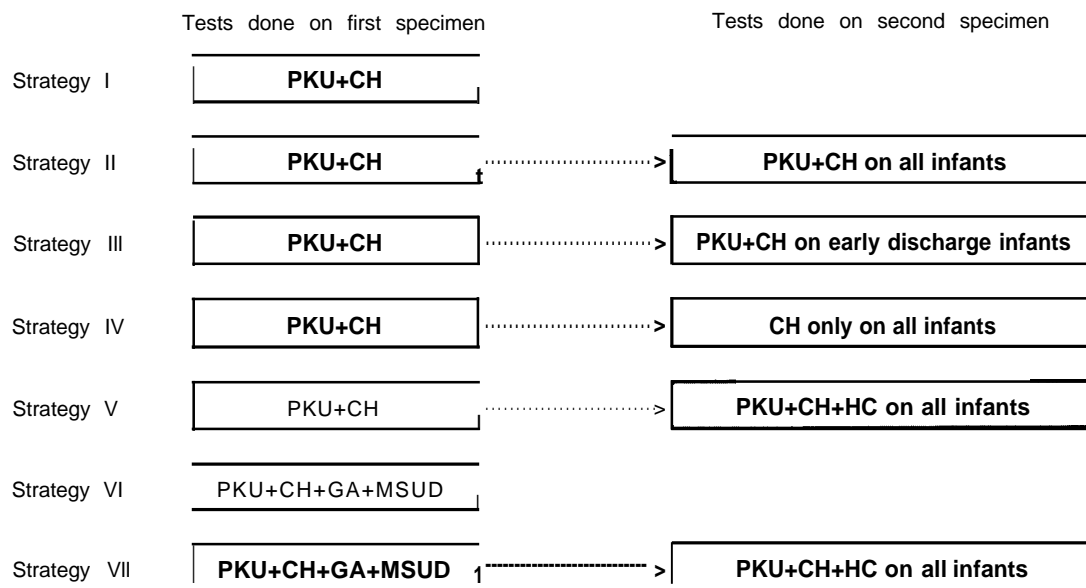
Available studies of the cost-effectiveness of various strategies for newborn screening have compared the cost of laboratory detection of PKU (47,85,471,621,729), congenital hypothyroidism (46,368), or several metabolic disorders together (123,650,727) with the averted costs to society associated with caring for a mentally retarded child. Most of the studies have omitted the costs of specimen collection and long term followup, and few of them applied a discount rate to future costs of treatment or institutional care. None of the studies has considered the costs and effectiveness of collecting a second specimen for additional screening. All of these studies have suggested that newborn screening for PKU, congenital hypothyroidism, and the rarer disorders result in a large cost savings to society in general.

OTA performed a cost-effectiveness analysis of the basic strategy for newborn screening—a one-specimen screening protocol for PKU and congenital hypothyroidism—compared to no screening and of six alternative newborn screening strategies compared to the basic strategy for screening. For the basic strategy, OTA ascertained the net costs or savings to the U.S. health care system and net effectiveness in identifying cases of PKU and congenital hypothyroidism requiring treatment. For the expanded strategies, OTA ascertained the incremental costs associated with detecting addi-

tional cases of selected congenital diseases beyond those detected in the basic strategy.

The seven newborn screening strategies considered in OTA's cost-effectiveness analysis are depicted in figure 5-1. These strategies vary with respect to two key features—the number of specimens tested and the types of congenital disorders tested for—but all of them involve various combinations of tests for two or more of the following five disorders: PKU, congenital hypothyroidism, homocystinuria, galactosemia, and MSUD. The costs and effectiveness of tests for biotinidase deficiency, sickle cell anemia, cystic fibrosis, and congenital adrenal hyperplasia are not included in the cost-effectiveness analysis, although OTA obtained preliminary information on the costs of adding tests for these four conditions to an ongoing screening program.

The basic screening strategy, Strategy I, involves the collection of a single blood specimen to test for PKU and congenital hypothyroidism and is compared in OTA's analysis to no screening at all. Since all 50 States and the District of Columbia offer newborn screening for PKU *and* congenital hypothyroidism, Strategy I (screening for PKU and congenital hypothyroidism using one specimen) reflects the minimum situation common to all U.S. newborn screening programs, and the

Figure 5-1 .—Newborn Screening Strategies Compared in OTA's Cost-Effectiveness Analysis

Abbreviations PKU - phenylketonuria; CH - congenital hypothyroidism, GA - galactosemia; MSUD - maple syrup urine disease; HC = homocystinuria
 SOURCE Office of Technology Assessment. 1988

other strategies reflect additional screening options some programs have taken. Accordingly, choices by screening programs are likely to be made not between some expanded screening strategy and no screening at all, but between a one-specimen strategy for PKU and congenital hypothyroidism (Strategy I) and expanded one- or two-test strategies (e. g., Strategies II through VII). Therefore, each of the other six screening strategies is compared to Strategy I rather than to no screening.

Strategies II, III, IV, and V follow an initial blood specimen to test for PKU and congenital hypothyroidism with a second specimen (the disorders tested for in the second specimen vary). Strategy VI uses a single specimen to test for four disorders (PKU, congenital hypothyroidism, galactosemia, and MSUD). Strategy VII, which is the most comprehensive screening strategy of all, uses a first specimen to test for four disorders (PKU, congenital hypothyroidism, galactosemia, and MSUD), then a second specimen to test for three disorders (PKU, congenital hypothyroidism, and homocystinuria).

For the basic strategy, OTA calculated net health care costs or savings to society per 100,000

infants screened. For comparisons of expanded screening strategies with Strategy I, OTA calculated the net incremental cost per extra case detected. All costs in OTA's analysis were expressed in 1986 dollars. In calculating net health care costs, OTA considered the costs of blood specimen collection, and laboratory detection, and medical treatment as well as costs of foster care, institutional care, and special education that would be averted by early treatment of affected infants. The value of avoiding premature death from the targeted conditions was not quantified in dollar terms. As a measure of effectiveness, OTA used the number of cases of targeted disorders detected and treated per 100,000 infants screened, and for the expanded strategies, the number of extra cases detected by these approaches. In all the analyses, this measure of effectiveness is used as a reasonable, though imperfect, proxy for the number of children whose lives would have been greatly diminished in quality or whose deaths would have occurred in childhood if their disorders had not been detected by newborn screening.

OTA's cost-effectiveness analysis included both a base case analysis and a sensitivity analysis. The

base case used what OTA deemed to be most likely values for major variables. The sensitivity analysis tested the sensitivity of the results to changes in assumptions about these variables. OTA estimated ranges for the major variables and then grouped together “worst case” and “best case” assumptions in the two parts of the sensitivity analysis.

Components of OTA's Cost-Effectiveness Analysis

Basic information on the data used in OTA's analysis of seven newborn screening strategies is presented below. For more detailed information on the data and methods used in the analysis, see appendix I.

Effectiveness Estimates

Estimates of the effectiveness of newborn screening for the five disorders were presented in the previous section. In the base case, OTA used the median incidence estimate and the median of the estimated percentage of affected infants detected shown in table 5-3. As noted earlier, these estimates represent a consensus of experts in the field and are the best estimates available.

Cost Estimates

OTA's estimate of the average costs of *collecting* a blood specimen from a newborn in the hospital was based on data from a time study at three Wisconsin hospitals (47). OTA estimated the average cost of collecting a first specimen in the hospital to be \$6.07. Unlike first specimens, the majority of second specimens are likely to be collected outside a hospital, possibly during a well-child visit at the physician's office or clinic. In the absence of data on the costs of specimen collection performed in a physician's office, OTA assumed in the base case that the cost of collecting a second blood specimen would be equal to the cost of collecting the first.

The resource costs of *detecting the five disorders* (PKU, congenital hypothyroidism, galactosemia, MSUD, and homocystinuria) through newborn screening were estimated for OTA's analysis on the basis of data provided by three State newborn screening programs: Washington (609), Wisconsin

(259), and Iowa (256). Since single-specimen tests for PKU and congenital hypothyroidism are available in all States, OTA combined costs for PKU and congenital hypothyroidism tests into a single estimate. In the base case analysis, OTA used the mean of the range of the combined cost of detecting these two disorders. The range was from \$3.88 to \$8.16 per specimen, giving a mean cost of \$5.65. The mean of the range of the cost of detecting galactosemia and MSUD was also used in the base case. With a range from \$1.25 to \$1.60 per specimen, the mean detection cost for galactosemia was \$1.43; the range for MSUD was \$0.98 to \$1.84, giving a mean detection cost for MSUD of \$1.41. Only one estimate for homocystinuria testing was available from the data provided by the three State programs, and that figure, \$0.93 per specimen, was used in the base case.

OTA's estimate of the costs of *treatment* for PKU and congenital hypothyroidism was derived from a study of PKU by Barden and colleagues (47) and from a study of congenital hypothyroidism by Barden and Kessel (46) and inflated to 1986 values: \$107,712 undiscounted total treatment costs for PKU and \$14,837 undiscounted total treatment costs for congenital hypothyroidism. Barden and colleagues discounted PKU and congenital hypothyroidism treatment costs to present value using a 7- and a 10-percent discount rate: for PKU treatment, \$53,855 at a 7-percent discount rate and \$42,670 at a 10-percent rate; and for congenital hypothyroidism treatment, \$4,260 at a 7-percent discount rate and \$3,588 at a 10-percent rate. The discount rate applied to future costs in OTA's base case was 7 percent. The sensitivity analysis used both a 7- and a 10-percent discount rate.

Data on treatment costs for galactosemia, MSUD, and homocystinuria comparable to data on treatment costs for PKU and congenital hypothyroidism are not available in the literature. Children with galactosemia need no special supplemental diet—just avoidance of foods containing galactose. In OTA's analysis, the costs of treatment for galactosemia were assumed to be close to the costs of treatment for congenital hypothyroidism mentioned above (46), because both these disorders include minor costs for medication and long-term costs of clinical care and monitoring. Treatment

costs for MSUD and homocystinuria in OTA's analysis were assumed to be approximated by the costs of long-term PKU treatment (47), which includes costs for a special diet and also long-term clinical care and monitoring.

In estimating the *health care costs averted* by newborn screening and treatment for PKU, OTA's analysis focused on the averted costs of custodial care and institutionalization and the averted costs of special education. OTA's estimate of the average net costs of residential care and special education for PKU was derived from Barden and colleagues (47) and inflated to 1986 values (see app. I for more details).

In estimating the health care costs averted by screening and treatment of congenital hypothyroidism, OTA focused similarly on averted costs of custodial care and institutionalization associated with mental retardation in individuals with untreated congenital hypothyroidism, with data derived from Barden and colleagues (46). In OTA's analysis, the averted costs of custodial care and institutionalization for individuals with untreated congenital hypothyroidism were combined with the averted costs of special education for untreated individuals and discounted to present value.

The health care costs averted by screening and treatment of galactosemia, MSUD, or homocystinuria are more difficult to quantify than those averted by screening and treatment of PKU and congenital hypothyroidism. No data are currently available to estimate the cost of the progressive deterioration and almost certain death that occur in the majority of cases of galactosemia or MSUD, or the long-term disabilities and risk of premature death that occur in cases of homocystinuria. Consequently, although OTA's analysis does quantify the costs of screening and treatment for these three conditions, it does not quantify the costs averted by screening and treatment for these conditions.

Findings of the Base Case Analysis

OTA's base case analysis indicates that, in comparison to no screening, Strategy I results in a net *savings* to the U.S. health care system of over \$3.2 million per 100,000 infants screened. This strategy results in the detection of 34.6 cases of PKU

and congenital hypothyroidism per 100,000 infants screened. For each of the cases identified and treated, net health care savings to society are approximately \$93,000 (see table 5-4). The net health care savings associated with Strategy I in comparison to no screening result from the detection and treatment of infants with PKU or congenital hypothyroidism who would have required custodial care or special education had their disorders not been treated.

OTA's base case analysis shows that each of the expanded strategies for screening are both more effective in detecting affected infants and more costly than Strategy I.¹¹ The number of additional cases of congenital disorders detected and the incremental costs incurred (i. e., the reductions in societal health care savings achieved by Strategy I) by six expanded newborn screening strategies in comparison to Strategy I are shown in table 5-5.

Strategy II (a first specimen to test for PKU and congenital hypothyroidism and then a second specimen to test for PKU and congenital hypothyroidism on all infants) detects 36.6 affected cases per 100,000 infants screened—or 2 cases more than Strategy I. The net incremental health care cost (i. e., loss of savings) per extra case detected and treated via this approach compared to Strategy I is very high—about \$466,000.

Strategy III (a first specimen to test for PKU and congenital hypothyroidism and then a second specimen to test for these disorders only in infants discharged early from the hospital whose blood specimens were collected before 3 days of age—41 percent of infants in 1985) results in the detection of 1.3 more affected cases per 100,000 infants screened than Strategy I. The net incremental cost per extra case detected and treated via this approach compared to Strategy I is approximately \$253,000.

Strategy IV (a two-specimen strategy that involves a first specimen to test for PKU and congenital hypothyroidism and a second specimen to

¹¹It should be noted, however, that each of the expanded screening strategies would result in a net savings to the health care system if they were compared to no screening at all, but in each case the net savings would be lower than that obtained for the basic strategy.

Table 5-4.—Effectiveness and Health Care Savings of Newborn Screening Strategy I Compared to No Screening (1986 dollars)

	Number of cases detected per 100,000 infants screened	Net health care savings per 100,000 infants screened	Net health care savings per case detected and treated
Strategy (1st test for PKU and CH only) v. No screening	34.6	\$3,218,000	\$93,000

Abbreviations: PKU = phenylketonuria; CH = congenital hypothyroidism
SOURCE Office of Technology Assessment 1988

Table 5-5.—Incremental Effectiveness and Health Care Costs of Newborn Screening Strategies Compared to Strategy I (1986 dollars)

Strategies compared	Number of extra cases detected per 100,000 infants screened	Net incremental costs per extra case detected and treated
Strategy II (1st test for PKU and CH followed by 2nd test for PKU and CH on all infants) v. Strategy I	2.0	\$466,000
Strategy III (1st test for PKU and CH followed by 2nd test for PKU and CH on early discharge infants only) v. Strategy I	1.3	\$253,000
Strategy IV (1st test for PKU and CH followed by 2nd test for CH only on all infants) v. Strategy I	1.7	\$432,000
Strategy V (1st test for PKU and CH followed by 2nd test for PKU, CH, and HC on all infants) v. Strategy I	2.5	\$421,000
Strategy VI (1st test for PKU, CH, GA, and MSUD) v. Strategy I	18	\$173,000
Strategy VII (1st test for PKU, CH, GA, and MSUD followed by 2nd test for PKU, CH, and HC) v. Strategy I	4.3	\$317,000

Abbreviations: PKU = phenylketonuria; CH = congenital hypothyroidism; HC = homocystinuria; GA = galactosemia; MSUD = maple syrup urine disease.
SOURCE Office of Technology Assessment, 1988

test for congenital hypothyroidism only) results in the detection of 36.3 affected cases, or 1.7 extra cases over Strategy I. The incremental cost of each case detected and treated via this approach compared to Strategy I is quite high—about **\$432,000**.

Strategy V, another two-specimen strategy, follows a first specimen for PKU and congenital hypothyroidism with a second specimen for these two disorders plus homocystinuria. Homocystinuria is the one condition of the five disorders considered in this analysis that is not optimally detected during the first week of life, so if a second specimen is being collected for PKU and congenital hypothyroidism, it might be advantageous to test for homocystinuria on the second round. Strategy V detects **2.5 cases** more per **100,000** infants screened than Strategy I.

The net incremental cost of detecting and treating an extra case via Strategy V relative to Strategy I is fairly high—approximately **\$421,000**. It is important to note, however, that in calculat-

ing the incremental costs associated with Strategy V, OTA did not include the costs averted by detecting and treating infants afflicted with homocystinuria (due to lack of available data); the inclusion of data on these averted costs, were such data available, would probably reduce the incremental costs associated with this strategy.

Strategy VI (a single-specimen strategy that adds tests for galactosemia and MSUD to the first specimen used to test for PKU and congenital hypothyroidism) detects 1.8 more cases per 100,000 infants screened than Strategy I. The net incremental health care cost per extra case found and treated by Strategy VI compared to Strategy I is low compared to the other strategies—about **\$173,000**. *The net* incremental cost associated with detecting additional cases via Strategy VI, in fact, is lower than the incremental costs associated with detecting additional cases via Strategies II, III, IV, or V—an observation that suggests that detecting extra cases by adding tests to an initial specimen for PKU and congenital hypothyroidism is less costly than detecting extra cases via a second

specimen. The cost of collecting additional specimens adds significantly to the incremental costs of the two-specimen strategies OTA considered.

Strategy VII involves all of the newborn screening tests considered in this analysis: a first specimen to test for PKU, congenital hypothyroidism, galactosemia, and MSUD, and a second specimen to test all infants for PKU, congenital hypothyroidism, and homocystinuria. Strategy VII detects almost 39 affected infants per 100,000 screened, or 4.3 more cases than Strategy I. The net incremental cost of detecting and treating an extra case via Strategy VII is rather high—approximately **\$317,000**.

Components and Findings of the Sensitivity Analysis

To test the sensitivity of the results of OTA's base case analysis of newborn screening strategies to changes in estimates of major variables—i.e., specimen collection costs, laboratory testing costs, percentage of affected cases detected, and discount rates—OTA examined the application of possible ranges of estimates for these variables in a sensitivity analysis. The estimates that were most favorable to the overall cost-effectiveness of newborn screening were combined in a “best” case analysis, and the least favorable estimates were combined into a “worst” case analysis. OTA performed best case and worst case analyses only for the four screening strategies with the most differences among them: Strategies I, II, VI, and VII.

Components of the Sensitivity Analysis

To vary the effectiveness of newborn screening, OTA varied the estimated percentage of affected infants detected by newborn screening (which reflects ranges in reported incidence as well as practical limitations on detection). The lowest of the range of estimated percentage rates for particular disorders were used in the worst case; the highest of the range of estimated percentage rates were used in the best case (see table 5-3).

To vary the costs of specimen collection in the sensitivity analysis, OTA used one approach to vary costs for the two one-specimen screening

strategies (Strategies I and VI) and a different approach to vary them for the two two-specimen strategies (Strategies II and VII). For the one-specimen strategies (I and VI), the cost of specimen collection used in the base case was varied by 50 percent: 50-percent higher (worst case) and 50-percent lower (best case). For the two-specimen strategies (II and VII), the cost of the first specimen collection was retained from the base case analysis, but the cost of the second specimen collection was varied: in the worst case, the cost of second specimen collection was assumed to be the same as collecting the first specimen; in the best case, it was assumed to be 25-percent lower.

To vary screening and treatment costs and discount rates, OTA used the same approach for all four strategies being compared:

- *Newborn screening and treatment costs:* The lowest estimate derived from data provided by one of the three State newborn screening programs was used in the best case; the highest estimate of the three was used in the worst case (see app. I).
- *Discount rate applied to future costs:* A 7-percent discount rate was applied to future costs in the best case; a 10-percent discount rate was used in the worst case.¹²

Together, all these changes in assumptions alter the expected number of cases detected per 100,000 infants screened and the costs of detection, treatment, and untreated disease. For an example of the calculation OTA used to arrive at an estimate of the overall costs or savings achieved by screening, see appendix I.

Findings of the Sensitivity Analysis

In the base case analysis, the net savings to the health care system associated with Strategy I (a single specimen for PKU and congenital hypothyroidism) compared to no screening was about \$3.2 million per 100,000 infants screened. As shown in table 5-6, net savings from Strategy I compared to no screening remain positive over the entire

¹²In general, costs that occur in the future are reduced more when a higher discount rate is used. Since averted costs of untreated disease are much greater than costs of preventive treatment, the use of a higher discount rate is appropriate for the worst case analysis since it lowers the averted costs.

range of assumptions tested, though they are as high as \$4.5 million per 100,000 infants screened in the best case and as low as \$626,000 per 100,000 infants screened in the worst case. Whereas 34.6 cases were detected per 100,000 infants screened by Strategy I in the base case, 41.3 cases are detected in the best case and only 28.8 are detected in the worst case.

Table 5-7 shows the incremental cost per extra case detected by Strategies II, VI, and VII, each compared to Strategy I. For Strategy VI (a single test strategy for PKU, congenital hypothyroidism, galactosemia, and MSUD), the sensitivity analysis shows that the incremental cost per each extra case detected by Strategy VI varies from \$277,000 in the worst case and \$85,000 in the best case, compared to \$173,000 in the base case analysis.

Strategies II and VII are both two-specimen testing strategies. For Strategy II (a first test for PKU and congenital hypothyroidism followed by a second test for PKU and congenital hypothyroidism on all infants), the incremental costs per extra case detected compared to Strategy I are about \$620,000 in the worst case and \$453,000 in the best case (see table 5-7), compared to \$466,000 in the base case analysis.

In Strategy VII (a first test for PKU, congenital hypothyroidism, galactosemia, and MSUD, followed by a second test for PKU, congenital hypothyroidism, and homocystinuria), the sensitivity analysis shows incremental costs per extra case found by Strategy VII compared to Strategy I are about \$474,000 in the worst case and \$218,000 in the best case, whereas the base case analysis estimated incremental costs of \$317,000 per extra case detected and treated.

The results of the sensitivity analysis represent extremes in the range of possible results, and it is unlikely that all the worst factors (or best factors) would occur together in a single situation. The sensitivity analysis shows, however, that the incremental costs of detecting additional infants with congenital disease are still somewhat high even under the most favorable situations, and can become substantially higher under the worst situations. It is worth noting, however, that under the best case assumptions, the cost of Strategy VI,

the cost of detecting an extra case in an expanded one-specimen strategy to test for two additional disorders, is low relative to the costs of many therapies currently considered standard medical procedure. The \$85,000 needed to detect an extra case of galactosemia or MSUD in Strategy VI (and with the best case assumptions) would buy an entire lifetime for a child with one of these disorders, compared to, for example, expenditures (in 1986 dollars) of about \$28,000 (162) to \$40,000 (98) per life-year gained from heart transplantation for congestive heart failure, or \$36,500 (530) per life-year gained from hemodialysis for end-stage renal disease.

Conclusions

In OTA's calculations, the costs of specimen collection and screening are important components of cost. By reducing the numbers of laboratories and avoiding duplication of fixed costs and highly trained personnel, costs of laboratory testing would probably be reduced. It follows that centralization of laboratories could make a substantial difference in the overall cost-effectiveness of newborn screening. Of the three State programs that provided data, the highest unit screening cost used in OTA's analysis was derived from the State program that had the lowest specimen volume, despite total overall costs similar to those of the other programs that provided data.

OTA's base case analysis compared the costs of screening by various expanded strategies with the consequences of doing less screening. This analysis showed that collecting additional specimens from a large portion of infants, whether to detect some percentage of extra cases of PKU, congenital hypothyroidism, or homocystinuria, or as a precautionary measure to guard against missed cases, is undoubtedly a costly strategy.

Only about one-half of all States screen for homocystinuria and MSUD, and about one-third screen for galactosemia, even though screening tests and treatment for these conditions have been available for many years. The rarity of these conditions is probably the main reason for their comparative unpopularity among screening programs. That rarity translates into high net costs of detecting each additional case, as reflected in OTA's

Table 5-6.—Sensitivity Analysis: Net Health Care Savings and Number of Cases Detected by Strategy I Compared to No Screening (1986 dollars)

Strategy	Sensitivity analysis					
	Base case		Worst case		Best case	
	Net health care savings per 100,000 Infants screened	Number of cases detected per 100,000 Infants screened	Net health care savings per 100,000 infants screened	Number of cases detected per 100,000 Infants screened	Net health care savings per 100,000 infants screened	Number of cases detected per 100,000 infants screened
Strategy I (1st test for PKU and CH only) v No screening	\$3,218,000	3 4 6	\$626,000	2 8 8	\$4,562,000	41 3

Abbreviations PKU = phenylketonuria CH = congenital hypothyroidism

SOURCE Off Ice of Technology Assessment 1988

Table 5-7.—Sensitivity Analysis: Incremental Effectiveness and Health Care Costs of Newborn Screening Strategies Compared to Strategy 1 (1986 dollars)

Strategies compared	Sensitivity analysis					
	Base case		Worst case		Best case	
	Number of extra cases detected per 100,000 Infants screened	Net Incremental cost per extra case detected & treated	Number of extra cases detected per 100,000 infants screened	Net Incremental cost per extra case detected & treated	Number of extra cases detected per 100,000 infants screened	Net Incremental cost per extra case detected & treated
Strategy II (1st test for PKU and CH followed by 2d test for PKU and CH on all Infants) v Strategy I	20	\$466,000	1 6	\$620,000	20	\$453,000
Strategy VI (1st test for PKU, CH, GA, and MSUD) v Strategy I	1 8	\$173,000	13	\$277,000	29	\$85,000
Strategy VII (1st test for PKU, CH, GA, and MSUD followed by 2d test for PKU, CH and HC) v Strategy I	4.3	\$317,000	31	\$474,000	57	\$218,000

Abbreviations PKU = phenylketonuria CH = congenital hypothyroidism HC = homocystinuria GA = galactosemia MSUD = maple syrup urine disease

SOURCE Off Ice of Technology Assessment 1988

analysis. Whether it is worth \$100,000 or more to detect an additional case of one of the treatable congenital disorders is, of course, a question that can only be answered in the context of larger societal decisions.

OTA's analysis incorporates the strategies that have been in use for many years in newborn

screening programs. Estimated net costs might be quite different, however, if new tests for additional disorders were considered. Combinations of tests other than the ones considered in OTA's analysis could change both the net costs and savings resulting from newborn screening and the incremental costs of detecting extra cases of congenital disease.

COSTS OF LABORATORY TESTING FOR FOUR ADDITIONAL DISORDERS

Four additional disorders not examined in OTA's cost-effectiveness analysis—biotinidase deficiency, sickle cell anemia, cystic fibrosis, and congenital adrenal hyperplasia—are being considered for inclusion in an increasing number of newborn screening programs. Screening for sickle cell anemia, in particular, is gaining widespread support as a result of recent evidence linking early detection and treatment of the disease with reduced mortality among infants with the disease in the first few years of life.

Few evaluations of the sensitivity and specificity of the screening tests and of the long-term value of early detection and treatment of biotinidase deficiency, congenital adrenal hyperplasia, sickle cell anemia, and cystic fibrosis have been conducted, so data on the long-term effectiveness of screening for these four disorders are unavailable. In the absence of more data on effectiveness, estimates of the cost of screening and treatment, not to mention costs averted by screening, would be incomplete at best. For that reason, OTA did not evaluate tests for these disorders in its cost-effectiveness analysis of newborn screening strategies. Since many newborn screening programs are incorporating these tests, however, preliminary cost estimates of detection are presented in this chapter. Cost estimates presented below are limited by the lack of sources of data; since only a few programs are screening for the disorders, and fewer have available cost information, the data below may not reflect representative additional costs of detecting these disorders.

OTA asked four State screening programs (Maryland, Iowa, Washington, and Colorado) to

identify and value the resources needed to test for the four additional disorders using the same blood specimens collected for PKU and congenital hypothyroidism (216,252,255,609). On the assumption that the tests for the four disorders would be added to the specimens for PKU and congenital hypothyroidism, the costs reported show no costs of additional specimen collection. Despite efforts to standardize the estimates, there may be somewhat more variability among programs in the methods used to derive costs of testing for the four disorders discussed below than to derive costs of testing for PKU, congenital hypothyroidism, galactosemia, homocystinuria, and MSUD.

Biotinidase Deficiency

Data from Maryland's newborn screening program suggest that testing for biotinidase deficiency adds about \$0.11 to the unit costs of laboratory testing for PKU and congenital hypothyroidism (216). With an approximate incidence rate for this deficiency of 1 in 45,000 live births, screening for biotinidase deficiency would yield approximately 2 to 3 infants with the disease per 100,000 screened.

Infants with biotinidase deficiency cannot recycle the B vitamin biotin, and treatment involves the oral administration of biotin. Experience to date suggests that such treatment saves severely affected infants from sudden death and prevents various kinds of neurologic damage in infants with milder cases. How many infants with milder cases could have been treated as effectively on the basis of a later clinical diagnosis without screening is not known.

Sickle Cell Anemia

Data from Iowa's newborn screening program suggest that screening for sickle cell anemia would add about \$3.51 per infant screened to the cost of PKU and congenital hypothyroidism screening (255).¹³ In the general population, approximately 32 cases of sickle cell anemia would be expected in 100,000 infants screened (the incidence among black newborns is about 1 in 500; 1984 census data indicate that approximately 16 percent of the total number of live births in the United States are black) (712).¹⁴

Many of the infants with sickle cell anemia are at risk for overwhelming infection and sudden death in the first year or two of life. Newborn screening, followed by the use of prophylactic antibiotics, may allow for a significant reduction of this risk (198). Implications for later treatment of infants diagnosed by newborn screening are unknown.

Cystic Fibrosis

Screening for cystic fibrosis would yield approximately 40 affected infants per 100,000 infants

¹³This estimate may be high because of the size of and cost allocation methods employed by the Iowa program.

¹⁴As noted earlier, the incidence of sickle cell anemia is higher among blacks and in people of Mediterranean and Middle Eastern descent. Screening for the disease in high-risk populations with a higher incidence of the disease will obviously yield different costs and effectiveness results than those yielded by screening in the general population.

screened at an estimated additional cost of \$1.32 per specimen screened, according to data from the Rocky Mountain States Regional Program (252). The most immediate potential benefit to presymptomatic diagnosis of cystic fibrosis may be the treatment of nutritional deficiencies which place some affected infants at high risk for neonatal death and impair the growth and development of other infants with cystic fibrosis.

Screening for cystic fibrosis is currently being performed on a pilot basis in a few States. It is unknown whether early diagnosis and treatment for cystic fibrosis improves long-term outcomes.

Congenital Adrenal Hyperplasia

It would cost an additional \$1.50 per infant screened to include testing for congenital adrenal hyperplasia in an ongoing screening program, according to data derived from the Washington State program (609). Screening for congenital adrenal hyperplasia would detect approximately 4 infants with the disease in 100,000 infants screened. Approximately half of the afflicted infants would have been at high risk for sudden death due to salt-wasting crises. Screening and early treatment can prevent such neonatal deaths. Infants with congenital adrenal hyperplasia who are not at risk for salt-wasting can benefit from accurate early diagnosis and treatment by receiving hormone therapy, possibly averting abnormal gender orientation and reproductive problems later in life.

FINANCING AND REIMBURSEMENT FOR NEWBORN SCREENING AND TREATMENT

State newborn screening and treatment programs are funded by a combination of Federal, State, and private sources. Since the passage of the Omnibus Budget Reconciliation Act of 1981 (OBRA-81), (Public Law 97-35) establishing the Maternal and Child Health (MCH) block grant under Title V of the Social Security Act, Federal funds for newborn screening have been included in the MCH block grants.¹⁵ For every \$0.75 of

¹⁵For more information on the MCH block grant program, see chs. 3 and 6.

State funds spent on maternal and child health services, \$1.00 of Federal funds is contributed. The total authorization for the MCH block grants was set at \$373 million under OBRA-81. MCH block grant funds were appropriated in two parts: 85 percent to be transferred directly to the States, and 15 percent to be set aside for special projects (in a Federal program entitled "Special Projects of Regional and National Significance"). Funding for certain genetic services was specified under the special projects portion of the budget, but the

States could also use the general block grant funds to support their newborn screening programs (182),

The Deficit Reduction Act of 1984 (Public Law 98-369) raised the authorization for MCH block grants to \$478 million. For fiscal year 1987, the Omnibus Reconciliation Act of 1986 (OBRA-86), (Public Law 99-509) increased the authorization for the MCH block grant to \$553 million. That act designated a percentage of the additional funds authorized for the MCH block grants (\$75 million in fiscal year 1987) to be used for newborn screening for "sickle cell anemia and other genetic disorders." The percentage was set at 7 percent in fiscal year 1987 (or \$5.25 million), 8 percent in fiscal year 1988, and 9 percent in fiscal year 1989. If these funds for screening are appropriated as outlined, newborn screening, particularly for sickle cell anemia, could be expanded in many State screening programs under the Division of Maternal and Child Health's general oversight provisions.

An increasing number of State newborn screening programs are charging user fees for the tests they perform. Currently, 12 States specify in their enabling legislation that a charge may be levied, and others specify the exact amount that may be charged (32). Reported charges, which are not necessarily related to actual costs, range from about \$3 to \$24¹⁶ per infant screened in State and private laboratories (32,562,592).

The costs to a family of newborn screening services are usually reimbursable under private insurance plans as well as various public programs, depending on the individual plan operating in the State. Under Medicaid, reimbursement for newborn screening is generally included in reimbursement for overall perinatal care. Newborn screening programs or hospitals may have to absorb the cost of screening if the family is unable to pay or if funds cannot be collected from the third party.

In contrast to the relatively small, one-time fee for the screening test (if one is charged at all), the costs of treatment are incurred over a long period of time in each case. Treatment for PKU, for example, consists of an essential dietary product that is initiated in the first weeks of the child's life and continues to be used through adolescence and, in some cases, indefinitely. Total cost of the special PKU formula over the first 20 years of a patient's life has been estimated at \$58,270 (47), or about \$3,000 per year. Similar dietary regimens are used to treat individuals with homocystinuria and MSUD.

Such dietary treatment is not normally reimbursed as medication or medical treatment, so standard health insurance may not cover the costs. Various sources of third-party payment aside from standard health insurance are available in general, but the availability of these sources varies widely among States and even within States. Some families may have access to programs that provide the diet for all patients regardless of ability to pay. In the long run, most families probably rely on a combination of different sources and may bear the entire cost of the diet during periods in which no reimbursement is available. Some families may discontinue the diet. In young children, however, going off the diet is highly likely to cause mental retardation. In older children, even brief gaps in maintenance of the diet therapy may diminish performance and alter the child's demeanor (12). For women with PKU, there is an additional problem. If women with PKU have gone off the diet before or during their childbearing years, it may be more difficult to ensure that they resume it before conception, especially if they are not aware of the significant risks to their fetus if they are not on the diet.

A survey of sources of reimbursement for the PKU diet showed that the State health agency was the sole or major source of third-party funds for PKU treatment in 54 out of 98 newborn screening programs surveyed, but that a variety of other sources were also used, each contributing a minor portion of the reimbursement: the Crippled Children's Program; Medicaid; the Women, Infants, and Children (WIC) program; military sources; private health insurance; the patient's family; the treatment clinic; and the formula manufacturer

¹⁶The State with the \$24 charge also includes the cost of followup and initiation of treatment in this charge (345).



Photo credit: Association for Retarded Citizens of the United States

Kammy and Sheila McGrath, poster children for 1961-62, are sisters who were born with PKU. Sheila is severely retarded. Mental retardation was avoided in Kammy, her younger sister, because newborn screening detected her disorder and allowed the prompt initiation of a special preventive diet.

CONCLUSIONS

Since newborn screening is organized at the State level, there are in practice about as many different strategies for screening as there are programs. OTA's analysis examined the net health care costs and effectiveness of the basic strategy common to all newborn screening programs and then calculated the incremental costs and effectiveness of six expanded strategies for screening compared to that basic strategy. The basic strategy—screening all infants using one specimen for PKU and congenital hypothyroidism—resulted in net health care savings of about \$3.2 million per

(570). In some cases, the WIC program covers the cost of the diet for PKU women during their pregnancy, but not before or after pregnancy. PKU formula is covered under Medicaid in some States (e.g., California, Colorado, and Washington), and is variably classified as a medical supply, a special medical item, a drug, or as a component of enteral or parenteral nutrition programs. In Wisconsin (where Medicaid does not cover the PKU formula), a \$3 "surcharge" is added to newborn screening charges for all infants to cover the cost of the formula and diet for those who need them, which places the cost on users of the program (i.e., the childbearing population) rather than general taxpayers (Wise. Assembly Bill 550, Act 157, 1983).

In some of the smaller States, or in some States without coordinated screening programs, the individual physician caring for the PKU child sometimes can become a key figure in obtaining reimbursement for the PKU diet. Moreover, differences in medical opinion about the number of years that a patient should remain on the diet can lead to different levels of advocacy on the part of physicians in securing long-term and consistent funding for the diet. In some cases, then, access to these sources of funding may be determined largely by the degree to which the individual physician or screening director becomes involved in the process.

100,000 infants screened and in the identification and treatment of about 34 infants with PKU and congenital hypothyroidism per 100,000 infants screened. These figures represent a net savings of about \$93,000 per case detected and treated. Under the best case and worst assumptions used in OTA's sensitivity analysis, the net savings per case detected and treated would amount to at best, \$110,000, and at worst, \$22,000.

While most State screening programs probably started with PKU testing and later added congen-

ital hypothyroidism testing, many programs have gradually expanded their activities over the last 10 years to include collection of second specimens or testing for galactosemia, homocystinuria, and/or MSUD. Each of the strategies involving a single-specimen with additional tests or involving two specimens detects more cases of congenital disorders than does the single specimen with tests for PKU and congenital hypothyroidism alone. In choosing among the various expanded strategies for detecting extra cases of congenital disorders, however, newborn screening programs have to take into account the incremental costs that are incurred through these approaches. OTA's analysis found that the incremental costs of any of the two-specimen strategies compared to the basic one-specimen strategy were quite high, ranging from \$253,000 to \$466,000 per extra case found and treated. The least costly strategy for detecting extra cases beyond those found in a single specimen to test for PKU and congenital hypothyroidism involved the addition of extra tests (i.e., for galactosemia and MSUD) onto the first specimen.

Overall, the data presented in OTA's analysis suggest that the basic newborn screening strategy—i.e., screening for PKU and congenital hypothyroidism in a single specimen—is less costly than no screening at all and that the six expanded screening strategies OTA considered in its analysis incur substantial additional costs. Since all State screening programs have at least adopted the basic strategy for newborn screening, the potential net savings nationwide have been large: testing for PKU and congenital hypothyroidism using a single specimen from each infant results in net savings to the U.S. health care sector of about \$120 million per year assuming that 3.7 million infants are screened in a given year. Many States have expanded screening beyond this basic strategy, however, and since these expanded strategies are costly, the actual savings to the health care sector are somewhat diminished.

Regardless of the strategy chosen for a particular newborn screening program, aspects of the organization and delivery of the services can have a major impact on the costs and effectiveness of screening. Foremost among them is laboratory costs, which are directly influenced by the effi-

ciency of the testing procedures. It is plausible that centralization of laboratory facilities for all States or multi-State regions could reduce the duplication of fixed costs, such as capital equipment, and could make better use of highly trained personnel. In addition, decentralized laboratories are less likely to encounter abnormal test results—a significant disadvantage in developing the expertise necessary to correctly identify cases of very rare diseases. At present, some States do not have centralized laboratories, although the trend in recent years has been toward the consolidation of facilities for newborn screening into one laboratory per State. For some States with small populations, however, even one centralized laboratory may not be the best use of resources and expertise; such States should be encouraged to form regional laboratories with neighboring States, while maintaining their own followup and treatment facilities on a local level.

Critical to the effectiveness of newborn screening are such tasks as ensuring the completeness of screening and the followup of positive screening tests. The accomplishment of both tasks is probably more difficult and less reliable in a screening program that lacks a tightly coordinated system of screening, followup, and treatment services. Data on the frequency and types of errors that can occur are emerging in a few programs. It would be particularly useful, however, to collect and compare data on the completeness and accuracy of screening across all the States' programs on an ongoing basis.

Newborn screening is currently expanding in scope to include tests for diseases such as sickle cell anemia, congenital adrenal hyperplasia, cystic fibrosis, and biotinidase deficiency, and tests for these conditions are rapidly being incorporated into routine screening practices. These diseases have little in common, except that in some cases, they can be life-threatening before appropriate medical care can be obtained. In these cases, newborn screening can provide essential early warning and treatment. In other cases, the benefit of earlier diagnosis and access to specialized treatment may be a demonstrable improvement in prognosis and course of the disease. At present, such tests appear promising, but reliable information to determine the overall value of newborn

screening for these four conditions compared to the effectiveness of their diagnosis and treatment through standard medical channels is still lacking. As data do become available, it would be useful to compare the incremental effectiveness and costs of incorporating these tests into various newborn screening strategies, such as the ones considered in this chapter.

For the established screening tests, the OTA analysis has shown that more screening, while leading to additional cases detected, can be quite costly, although some of this cost can be reduced

by modifying the screening protocols (e.g., performing the additional tests on the first specimen, without obtaining a second specimen from each infant). State screening programs should proceed cautiously with screening for biotinidase deficiency, sickle cell anemia, cystic fibrosis, and congenital adrenal hyperplasia. Concurrently, the Federal Government might put as a priority the collection and evaluation of data that would allow careful analysis in future of costs as well as effectiveness of widespread screening for these disorders.