

Comprehensive Cost-Utility Analysis of Newborn Screening Strategies

Aaron E. Carroll, MD, MS^{a,b}, Stephen M. Downs, MD, MS^{a,b}

^aChildren's Health Services Research, Indiana University School of Medicine, Indianapolis, Indiana; ^bRegenstrief Institute for Health Care, Indianapolis, Indiana

The authors have indicated they have no financial relationships relevant to this article to disclose.

ABSTRACT

BACKGROUND. Inborn errors of metabolism are a significant cause of morbidity and death among children. Inconsistencies in how individual states arrive at screening strategies, however, lead to marked variations in testing between states.

OBJECTIVE. To determine the cost-effectiveness of each component test of a multitest newborn screening program, including screening for phenylketonuria, congenital adrenal hyperplasia, congenital hypothyroidism, biotinidase deficiency, maple syrup urine disease, galactosemia, homocystinuria, and medium-chain acyl-CoA dehydrogenase deficiency.

METHODS. A decision model was used, with cohort studies, government reports, secondary analyses, and other sources. Discounted costs, quality-adjusted life-years (QALYs), and incremental cost-effectiveness ratios were measured.

RESULTS. All except 2 screening tests dominated the "no-test" strategy. The 2 exceptions were screening for congenital adrenal hyperplasia, which cost slightly more than \$20 000 per QALY gained, and screening for galactosemia, which cost \$94 000 per QALY gained. The screening test with the lowest expected cost was tandem mass spectrometry. The results found in our base-case analysis were stable across variations in nearly all variables. In instances in which changes in risks, sequelae, costs, or utilities did affect our results, the variation from base-case estimates was quite large.

CONCLUSIONS. Newborn screening seems to be one of the rare health care interventions that is beneficial to patients and, in many cases, cost saving. Over the long term, funding comprehensive newborn screening programs is likely to save money for society.

www.pediatrics.org/cgi/doi/10.1542/peds.2005-2633H

doi:10.1542/peds.2005-2633H

The views expressed in this article are those of the authors and do not necessarily represent the views of the Indiana University School of Medicine.

Key Words

newborn, screening, decision analysis, cost-effectiveness, cost utility

Abbreviations

PKU—phenylketonuria
MS/MS—tandem mass spectrometry
CAH—congenital adrenal hyperplasia
CH—congenital hypothyroidism
MSUD—maple syrup urine disease
QALY—quality-adjusted life-year
CP—cerebral palsy
MCAD—medium-chain acyl-CoA dehydrogenase

Accepted for publication Dec 27, 2005

Address correspondence to Aaron E. Carroll, MD, MS, Riley Research 330, 699 West Dr, Indianapolis, IN 46074. E-mail: aaecarro@iupui.edu

PEDIATRICS (ISSN Numbers: Print, 0031-4005; Online, 1098-4275); published in the public domain by the American Academy of Pediatrics

INBORN ERRORS OF metabolism are a significant cause of morbidity and death among children.¹⁻⁴ Technology has advanced remarkably since screening began for phenylketonuria (PKU) >40 years ago.⁵ Although we have technology today to screen for a host of diseases, which diseases are screened for differs according to state. Every year ~4 million dried blood spots are analyzed for metabolic and hematologic disorders and endocrinopathies.⁶ These diseases account for ~3000 cases of potentially disabling or fatal outcomes each year.⁶ The diseases for which we screen and the manner in which we screen vary significantly according to state. These decisions should be based on disease prevalence, costs of testing, treatment, and complications, and effectiveness of timely intervention. Inconsistencies regarding how states arrive at these decisions, however, lead to marked variations in testing among states.

Tandem mass spectrometry (MS/MS) offers the ability to increase drastically the number of diseases for which we screen.⁶⁻⁸ More than 20 disorders of fatty acid oxidation and organic acidemia are already screened for routinely with MS/MS methods.⁹ The number of diseases screened for and the number of states using the technology are increasing rapidly.¹⁰

The significant start-up costs of MS/MS screening, however, have prevented some states from adopting it as part of their standard screening process.¹¹ Many have pointed out that most of the diseases for which MS/MS methods screen have a relatively low prevalence.^{9,10} MS/MS screening can lead to false-positive or abnormal results for well children, yielding increased costs and unnecessary parental angst.^{12,13} In addition, some diseases are so severe that, even with an ideal screening program, many children would not be treated in time to prevent morbidity and death.

Other data argue for the merits of broader screening even in light of these issues. Although costs for wider screening can be significant, the costs saved through preventing hospitalizations for diagnosis or long-term disability are also significant.¹¹ Studies show that even diseases without a consistently effective treatment can benefit from early diagnosis.⁷ Evidence indicates that, overall, screening reduces stress while improving outcomes.¹²

A more objective method for determining the usefulness of any new technology is cost-effectiveness analysis. A number of these analyses were conducted previously for other aspects of newborn screening^{14,15} or the use of MS/MS screening for a small number of disease processes.^{1,9,10} Currently, programs do not screen for biotinidase deficiency and galactosemia with MS/MS methods, but there is clear evidence that galactosemia could be detected with this technology.¹⁶

To date, no cost-effectiveness study has examined the prospect of incorporating broad MS/MS testing into a comprehensive newborn screening program. The objec-

tive of this study was to determine the cost-effectiveness of a comprehensive newborn screening program, including screening for PKU, congenital adrenal hyperplasia (CAH), congenital hypothyroidism (CH), biotinidase deficiency, maple syrup urine disease (MSUD), galactosemia, homocystinuria, and medium-chain acyl-CoA dehydrogenase (MCAD) deficiency, alone or in combination with diseases detectable with MS/MS methods.

METHODS

Study Design

We developed a decision model with decision analysis software (DATA 4.0; TreeAge Software, Williamstown, MA), to compare various strategies for screening for newborn diseases. We adopted the recommendations of the Panel on Cost-Effectiveness in Health and Medicine for conducting and reporting a base-case analysis.¹⁷ We assumed a societal perspective to produce results, permitting comparisons across different interventions, and expressed results in terms of discounted costs, quality-adjusted life-years (QALYs), and incremental cost-effectiveness ratios.

Decision Model Structure and Assumptions

To frame the model for analysis, we made a number of simplifying assumptions. We assumed that any child could have, at most, 1 of the conditions for which the tests screen. We used a multiplicative utility model, that is, for outcomes with >1 complication, we multiplied the quality adjustments (utilities) for all relevant complications together. We assumed the existence of a comprehensive screening program that ensures timely testing, result notification, and appropriate treatment of all positive screens. Finally, we assumed that MS/MS methods would be used to screen for multiple conditions.

The Decision Model

Figure 1 shows the decision model structure. The decision tree consisted of a decision node whose branches represent 8 screening tests, that is, PKU, CAH, CH, biotinidase deficiency, MSUD, galactosemia, homocystinuria, and MS/MS tests. We assumed that MS/MS would be used to detect PKU, biotinidase deficiency, MSUD, galactosemia, and homocystinuria, as well as MCAD deficiency. In a second analysis, we compared MS/MS directly with a screening panel that included PKU, biotinidase deficiency, MSUD, galactosemia, and homocystinuria detection with conventional methods.

Following each test branch is a chance node whose branches represent the 8 conditions that the child may have, ie, PKU, CAH, CH, biotinidase deficiency, MSUD, galactosemia, homocystinuria, or MCAD deficiency, or "no disease." Each disease branch is followed by a subtree that represents potential sequelae of the disease. In Fig 1, final outcomes are shown only for homocystin-

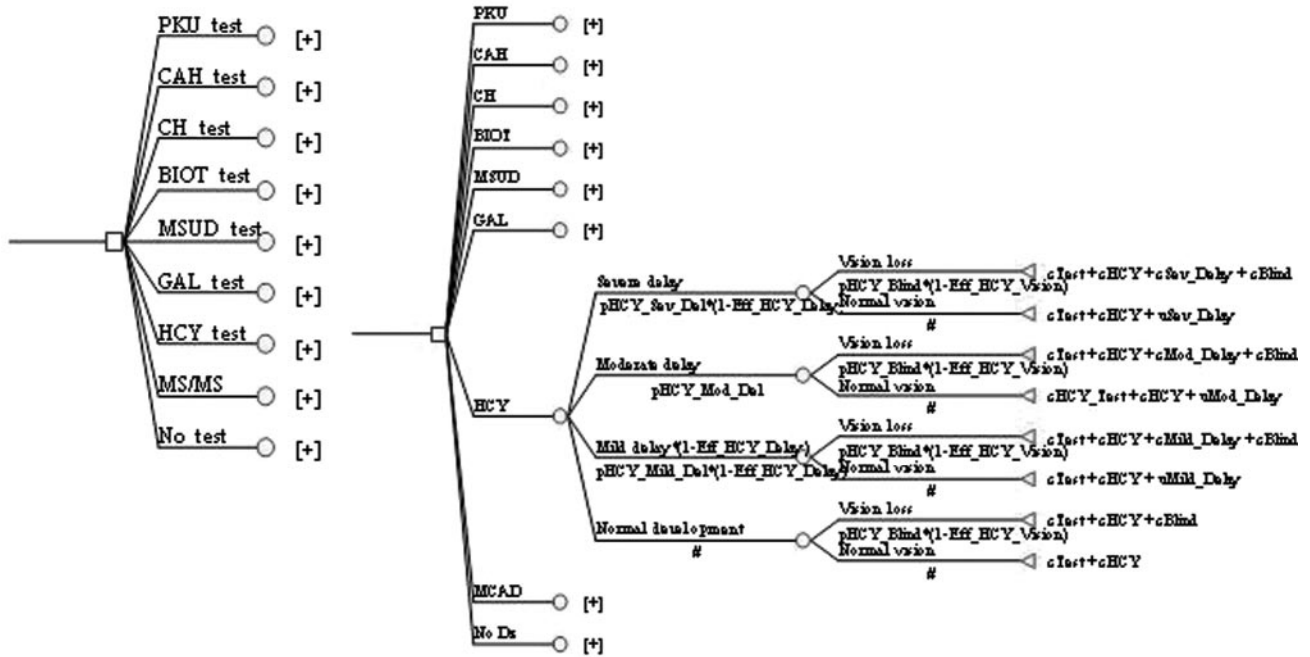


FIGURE 1

Decision model. Branches from the square decision node represent the various screening tests. Following each of these is a subtree representing each of the conditions that might be present. Only one of these subtrees is shown. Following each condition is a subtree depicting the possible outcomes of the conditions. Final outcomes are shown for homocystinuria (HCY) only. # indicates the probability of the lower branch is 1 minus the probabilities of the branches above. BIOT indicates biotinidase deficiency; GAL, galactosemia; Dz, disease.

uria. PKU was associated with a risk of mild, moderate, or severe developmental delay and/or seizure disorder. CAH could lead to early childhood death (average life expectancy: 1 year). CH could cause mild, moderate, or severe delay. Biotinidase deficiency could lead to any combination of seizure disorder, severe developmental delay, hearing loss, or blindness. MSUD could lead to any combination of severe or mild cerebral palsy (CP) or mild, moderate, or severe developmental delay. Galactosemia could lead to death in infancy or moderate or severe delay. Homocystinuria had a risk of mild, moderate, or severe delay and/or vision loss. MCAD deficiency could lead to early death (modeled as death at 1 year, on average) or CP and/or mild-to-severe developmental delay. Children with no disease face a risk of a false-positive newborn screening result and its attendant expenses.

The probability of each of the outcomes with screening, $P(\text{outcome}|\text{screen})$, was represented by the formula $P(\text{outcome}|\text{screen}) = P(\text{outcome}) \times [1 - (\text{efficacy} \times \text{sensitivity})]$, where $P(\text{outcome})$ is the probability of the outcome (eg, CP or delay), without screening or treatment, derived from the literature (see below), efficacy is the efficacy of early detection and treatment in preventing the outcome, and sensitivity is the sensitivity of the screening test for the disease.

Data and Assumptions

Probabilities of sequelae, quality-adjusted survival rates, and costs for treatment were derived from various clin-

ical and administrative data sources. We searched Medline (from January 1980 to November 2002) for studies with the terms “newborn screening” and any of the disease entities we were investigating. We also searched online Internet databases for the same terms. We supplemented this strategy by hand-searching citation lists from articles identified in the electronic search and recent reviews and cost-effectiveness analyses of other newborn screening strategies.

The prevalence of each disease was estimated, when possible, from the National Newborn Screening Report.¹⁸ Supplemental information was taken from the American Academy of Pediatrics newborn screening fact sheets and secondary analyses.^{8,19} We determined the sensitivities and specificities of the various screening tests through national reports and other secondary analyses.^{10,13,18,20} Probabilities of various sequelae and the effectiveness of early intervention in preventing these sequelae were also estimated from the newborn screening fact sheets, secondary data analyses, and cohort studies of patients with the diseases in question.^{19,21–28} Base-case probabilities for major complications are summarized in Table 1.

Estimation of Life Expectancy

We estimated the average life expectancy of normal patients from the *National Vital Statistics Reports* published by the Centers for Disease Control and Prevention.²⁹ We estimated the effect of disability on life expectancy through a cohort study on the subject.³⁰

TABLE 1 Data Used In Decision Tree

	Base Case	Reference	Threshold
Test characteristics			
Sensitivity of newborn screening tests (for all conditions)	1	13	Table 4
Specificity of BIOT screening test	0.9998	13	0.958
Specificity of CAH screening test	0.992	13	
Specificity of CH screening test	0.985	13	0.970
Specificity of GAL screening test	0.997	13	
Specificity of HCY screening test	0.9993	20	0.994
Specificity of MS/MS testing	0.9975	10	0.854
Specificity of MSUD screening test	0.9987	10	0.997
Specificity of PKU screening test	0.998	13	0.881
Disease prevalence, cases per 100 000 infants			
BIOT	1.1	18	0.2
CAH	4.9	18	3500
CH	39.6	18	28.0
GAL	1.5	18	
HCY	0.4	18	0.16
MCAD deficiency	9.0	19	
MSUD	0.6	18	0.48
PKU	6.6	18	0.8
Probability of disease causing sequelae			
BIOT, hearing loss	0.76	25	
BIOT, severe developmental delay	0.50	24	
BIOT, seizure disorder	0.70	24	
BIOT, vision loss	0.50	24	
Risk of death from CAH in first 5 y	0.10	Expert opinion	
CH, mild developmental delay	0.27	19	0.018
CH, moderate developmental delay	0.021	19	
CH, severe developmental delay	0.019	19	0.0082
GAL, neonatal death	0.14	27	>0.27
GAL, moderate developmental delay	0.90	26	
GAL, severe developmental delay	0.10	26	
HCY, blindness	1	22	0.4
HCY, mild developmental delay	0.11	21	
HCY, moderate developmental delay	0.33	21	
HCY, severe developmental delay	0.05	21	
MCAD deficiency, moderate CP	0.10	28	
MCAD deficiency, neonatal death	0.204	21	
MCAD deficiency, mild developmental delay	0.05	9	
MCAD deficiency, severe developmental delay	0.05	9	
MSUD, moderate CP	0.50	23	
MSUD, severe CP	0.25	23	0.14
MSUD, severe developmental delay	0.50	23	0.34
PKU, mild developmental delay	0.05	24	
PKU, moderate developmental delay	0.475	24	
PKU, severe developmental delay	0.475	24	
PKU, seizure disorder	0.25	24	
Effectiveness of early screening in preventing sequelae			
BIOT, developmental delay	1	19	
BIOT, hearing loss	1	19	
BIOT, seizure disorder	1	19	
BIOT, vision loss	1	19	
CAH, preventing death from crisis	0.8	Expert opinion	
CH, developmental delay	1	19	0.67
GAL, neonatal death	0.66	21	
GAL, developmental delay	0	21	
HCY, blindness	1	22	0.31
HCY, developmental delay	1	22	
MCAD deficiency, neonatal death	0.98	21	
MSUD, CP	0.93	23	0.59
MSUD, developmental delay	0.53	19, 24	0.36
PKU, developmental delay	1	24	0.01
PKU, seizure disorder	1	24	

TABLE 1 Continued

	Base Case	Reference	Threshold
Cost of screening tests			
BIOT	\$1.83	33	\$14
CAH	\$3.63	33	
CH	\$4.59	33	\$8.70
GAL	\$3.79	33	
HCY	\$0.84	33	\$2.45
MS/MS	\$16.02	9	\$58
MSUD	\$2.49	33	\$2.90
PKU	\$3.43	33	\$38.20
Follow-up evaluation of false-positive result (for any test)	\$300	Expert opinion	Table 4
Cost of caring for disease			
BIOT	\$6592	21	
CAH	\$10 000	Expert opinion	
CH	\$9439	34	
GAL	\$9439	34	
HCY	\$122 515	34	
MCAD deficiency	\$10 173	21	
MSUD	\$122 515	34	
PKU	\$122 515	34	
Cost of sequelae			
Blindness	\$581 688		\$180 000 (HCY)
Hospitalization before death	\$27 809	36	
MCAD deficiency crisis	\$308 315	21	
Deafness	\$445 255	14	
Mild developmental delay	\$44 192	35, expert opinion	\$3000
Moderate developmental delay	\$77 079	35, expert opinion	
Severe developmental delay	\$1 042 110	35	Table 4
Moderate CP	\$77 079	35, expert opinion	
Severe CP	\$946 528	35	\$540 000 (MSUD)
Seizure disorder	\$216 848	21	
Life expectancy, y			
Normal	77.2	29	
Mild developmental delay	74	30	
Moderate developmental delay	67.6	30	
Severe developmental delay	58.6	30	
Utilities of having sequelae			
Blindness	0.514	32	
Deafness	0.8611	31	
Mild developmental delay	0.7396	31	
Moderate developmental delay	0.5625	31	
Severe developmental delay	0.3909	31	
Moderate CP	0.5625	31	
Severe CP	0.3909	31	

Base-case estimates are presented with the reference from which those estimates were drawn. Also presented are the threshold values obtained from sensitivity analyses (see text). BIOT indicates biotinidase deficiency; GAL, galactosemia; HCY, homocystinuria.

Quality-of-Life Adjustments

We adjusted life expectancy for quality of life by using health state utilities.¹⁷ We calculated QALYs by multiplying remaining years of life by the utility value for each health state. Utilities for most disabilities were drawn from a study assessing parents' utilities by using a standard gamble, adjusting for different severities when necessary.³¹ The study assessed utilities for outcomes of occult bacteremia and meningitis, which resemble closely outcomes of the diseases for which newborns are screened. Specifically, parent utilities were assessed for deafness and various degrees of developmental delay and CP. Respondents in the study were parents of children between 2 and 24 months of age who visited an

urban emergency department. These were the most relevant utilities available for these outcomes. Utilities for blindness were taken from a study that developed an equation for estimating utility on the basis of visual acuity.³² We did not adjust utility if a disease state did not result in a disability that was thought to affect quality of life. Base-case estimates for utilities are summarized in Table 1.

Costs

The costs for the specific screening tests were derived from 2 sources, ie, a PriceWaterhouseCoopers analysis of newborn screening costs and a prior cost-effectiveness study of MS/MS screening.^{9,33} These costs incorporated

amortized start-up costs and operating costs. The costs of treating selected diseases over the course of a lifetime were estimated with data from the Office of Technology Assessment analysis of strategies for newborn screening, data from a Washington State report on the cost-effectiveness of some screening strategies, and, in 1 case, expert opinion.^{21,34} Costs of treating sequelae of different diseases were estimated from sources including a *Morbidity and Mortality Weekly Report* on economic costs associated with mental retardation, CP, hearing loss, and visual impairment; an epidemiologic study of end-of-life care costs; and prior secondary analyses.^{14,21,35,36} All costs were adjusted to 2004 US dollars with an inflation calculator on the Web site of the Department of Labor, Bureau of Labor Statistics.³⁷ For this study, costs were a complex mixture of medical and nonmedical costs, and we elected to adjust costs with the General Consumer Price Index, rather than the Medical Care Component Consumer Price Index. Base-case estimates for costs are summarized in Table 1.

Time Preference

To reflect patient preferences for health outcomes and having material goods sooner rather than later, we discounted all costs and health effects at an annual rate of 3%.

Calculation of Incremental Cost-Effectiveness Ratios

For each treatment strategy, we calculated the expected total costs by multiplying the probability of each unique outcome with its associated costs and then adding these values for all possible outcomes. Total QALYs for each treatment were calculated in a similar manner. We then calculated the incremental cost-effectiveness of each screening strategy by dividing the difference in costs by the difference in QALYs.

Sensitivity Analysis

We used 1-way sensitivity analysis to identify important model uncertainties. All probabilities were tested across the full range from 0 to 1. Costs were analyzed from \$0

to \$1 million. To test the impact of changing multiple assumptions, we developed a “pessimistic” case analysis that biased the model against newborn screening. In the pessimistic case, we eliminated indirect costs from the estimated costs of developmental delay and CP.³⁵ We eliminated the costs of developmental delay from any outcomes that included CP. We increased the cost of MS/MS to \$20, the high end of published estimates. We increased the cost of evaluating false-positive screening test results to \$1000. We used low-end estimates for the prevalence of MCAD deficiency (1 case per 18 000 infants) and PKU (1 case per 20 000 infants). We also used a low-end estimate of the risk of death resulting from CAH (3%).

RESULTS

Base-Case Analysis

The results of the base-case analysis are shown in Table 2. The first column of Table 2 shows the testing strategies in ascending order of average cost. The second column shows the average cost of each strategy. This includes the cost of the test and the average costs (per person screened) associated with all of the treatments and outcomes for the diseases represented in the model. The third column shows the difference in average costs between the testing strategy and not testing. A negative number means the strategy saves money, on average, over not screening. The fourth column shows the average effectiveness of each strategy in QALYs. The fifth column shows the difference in effectiveness between each strategy and not testing. A positive number means there is a gain in QALYs. The sixth column is the average cost-effectiveness ratio for each strategy, and the seventh column is the incremental cost-effectiveness ratio, the incremental cost per QALY gained. Tests marked as “Dominates” save money and improve outcomes relative to not testing.

All except 2 screening tests dominated the no-testing strategy. That is, these screening tests improved outcomes and reduced costs relative to not screening. The 2

TABLE 2 Results of the Base-Case Analysis

Strategy	Cost	Incremental Cost	Effectiveness	Incremental Effectiveness	C/E	Incremental C/E
MS/MS testing	\$55	−\$43	77.19358	0.004	0.72	Dominates
PKU test	\$63	−\$35	77.19264	0.00306	0.81	Dominates
BIOT test	\$85	−\$13	77.19008	0.0005	1.1	Dominates
CH test	\$93	−\$5	77.19272	0.00314	1.21	Dominates
HCY test	\$96	−\$2	77.18977	0.00019	1.24	Dominates
MSUD test	\$97	−\$1	77.18978	0.0002	1.26	Dominates
No test	\$98		77.18958		1.26	
GAL test	\$102	\$5	77.18963	0.00005	1.32	\$94 000
CAH test	\$103	\$6	77.18986	0.00028	1.34	\$20 357

Presented are overall cost and effectiveness, incremental cost and incremental effectiveness over no testing, and both cost-effectiveness and incremental cost-effectiveness ratios. All strategies were compared with not testing. C/E indicates cost-effectiveness; BIOT, biotinidase deficiency; HCY, homocystinuria; GAL, galactosemia.

TABLE 3 Comparison of Costs and Effectiveness

Strategy	Cost	Incremental Cost	Effectiveness	Incremental Effectiveness	C/E	Incremental C/E
MS/MS	\$49.00		77.1957		0.64	Dominates
Panel	\$52.80	\$3.70	77.19358	-0.00212	0.68	

MS/MS testing for PKU, biotinidase deficiency, MSUD, galactosemia, and homocystinuria was compared with a panel of conventional testing methods for the same diseases. MS/MS also tests for medium-chain acyl-CoA deficiency, for which there is no other screening test available. C/E indicates cost-effectiveness.

exceptions are screening for CAH, which costs slightly more than \$20 000 per QALY gained, and screening for galactosemia, which costs approximately \$94 000 per QALY. The screening test with the lowest expected cost is MS/MS.

In addition to comparing each test with a no-testing strategy, we compared a panel of conventional tests for PKU, biotinidase deficiency, MSUD, galactosemia, and homocystinuria with MS/MS methods to screen for the same conditions and MCAD deficiency. The cost and effectiveness comparison is shown in Table 3. When used in this multiplex manner, MS/MS was less expensive and more effective than the comparable panel of conventional screening tests.

Sensitivity Analysis

Because most screening tests were dominant over not screening, we conducted sensitivity analysis on all of the variables in the model, to determine at what levels the tests were no longer cost saving. The fourth column of Table 1 shows the threshold value for each variable, at which the corresponding screening strategy becomes more costly than not screening.

In the base-case analysis, we assumed that the sensitivity of all of the screening tests was 100%. When we varied test sensitivity from 0% to 100%, the dominant strategies remained dominant to sensitivities well below those in the literature, as follows: PKU, 10%; CH, 67%; biotinidase deficiency, 13%; MSUD, 83%; homocystinuria, 37%; MS/MS, 27% (Table 4).

Because of the costs of evaluating false-positive results, the dominant testing strategies became more costly

TABLE 4 Thresholds at Which Tests Are No Longer Cost Saving

Strategy	Sensitivity	Cost of False-Positive Result	Cost of Severe Delay	Cost of Severe CP
MS/MS testing	0.27	>\$10 000	a	a
PKU test	0.1	a	a	a
BIOT test	0.13	\$600	a	a
CH test	0.67	>\$10 000	\$450 000	a
HCY test	0.37	\$2600	a	a
MSUD test	0.83	>\$10 000	\$685 000	\$540 000
GAL test	b	b	b	b
CAH test	b	b	b	b

BIOT indicates biotinidase deficiency; HCY, homocystinuria; GAL, galactosemia.

a Screening strategy is cost saving at any value of the variable.

b Screening strategy is not cost saving at any value of the variable.

TABLE 5 Cost-Effectiveness Analysis With the Pessimistic Case Analysis Assumptions

Strategy	Cost	Incremental Cost	Effectiveness	Incremental Effectiveness	C/E	Incremental C/E
No test	\$39.5		77.192		0.51	
Panel	\$40.5	\$1.00	77.196	0.00322	0.52	310.56
MS/MS	\$43.5	\$3.00	77.196	0.00062	0.56	4838.71

Cost-effectiveness (C/E) of screening for PKU, biotinidase deficiency, MSUD, galactosemia, and homocystinuria with a panel of conventional testing methods was compared with that of MS/MS testing for the same conditions plus medium-chain acyl-CoA dehydrogenase deficiency and no testing.

as specificity decreased. When the specificity of a screening test was lowered, a larger proportion of unaffected children had false-positive test results, which resulted in increased costs to rule out disease. When the specificity of a test was below a threshold value, some tests were no longer cost saving (Table 1); however, threshold values were below those reported in the literature.

The prevalence of a disease also affected its incremental cost. Biotinidase deficiency, CH, homocystinuria, MSUD, and PKU all became costly when the prevalence was decreased to values between 3 and 10 times lower than baseline values. When the prevalence of CAH was very high (3.5%), screening for CAH became cost saving. The cost of screening for galactosemia was not sensitive to prevalence, and the cost of MS/MS screening was not sensitive to the prevalence of MCAD deficiency.

The results were relatively insensitive to the rates at which individual sequelae occurred. Screening for CH was not cost saving if the rate of mild developmental delay without treatment was less than one tenth of baseline values or if the risk of severe delay was decreased by approximately one half. If the risk of blindness resulting from biotinidase deficiency was <40%, then screening for biotinidase deficiency alone was no longer cost saving. Screening for MSUD alone was not cost saving if, without treatment, the risk of severe CP was <14% or that of severe developmental delay was <34%. Because MS/MS screens for many conditions, it remained cost saving despite changes in any one of these parameters. Testing for galactosemia had a positive net cost in all sensitivity analyses; however, when the risk of neonatal death was >27%, the cost of screening fell below \$50 000 per QALY saved.

The cost savings with the screening tests depended in some cases on the effectiveness of treatment in preventing specific sequelae. Screening for CH was not cost saving if it was <67% effective in preventing severe delay. Screening for homocystinuria was not cost saving if treatment was <31% effective in preventing blindness. MSUD screening was not cost saving if treatment was <59% effective in preventing CP or <36% effective in preventing delay.

Our results were not sensitive to the costs of screening tests. Screening test costs that exceeded savings from

disease costs were between 2 and 10 times reported costs of screening tests. Conversely, screening for galactosemia or CAH was not cost saving even if tests could be performed for free.

Lifetime costs of diseases and their sequelae are difficult to estimate. However, we found that the cost saving for most of these screening tests was either insensitive or sensitive at values significantly below reported rates.

Our pessimistic case analysis biased the model against screening. Under this set of assumptions, newborn screening with MS/MS methods was not cost saving relative to a conventional screening panel, and the screening panel was not cost saving relative to no screening (Table 5). The conventional screening panel cost \$310.56 per QALY saved over not screening. Compared with this panel, MS/MS screening cost \$4838.71 per QALY saved.

DISCUSSION

Newborn screening for PKU, CH, biotinidase deficiency, MSUD, and homocystinuria individually not only was cost-effective but actually was cost saving in our base-case analysis. Use of MS/MS to screen for PKU, biotinidase deficiency, MSUD, and homocystinuria, as well as MCAD deficiency, had even greater cost savings because of the multiplicity of conditions detected with a single test. This was true even when MS/MS was compared directly with a panel of available conventional tests for the same conditions. Screening for CAH had a net cost per QALY gained; however the cost was less than the \$50 000 per QALY used conventionally as a benchmark for cost-effectiveness.^{38,39}

At more than \$90 000 per QALY, screening for galactosemia generally would not be considered cost-effective. This is because identifying galactosemia early may prevent a neonatal death but doing so induces high lifetime costs because of sequelae. In our sensitivity analyses, we found that under no circumstances would screening for galactosemia be cost saving. However, if the risk of neonatal death resulting from galactosemia is as high as 27%, then screening for galactosemia is cost-effective according to conventional criteria. In addition, the marginal cost of screening for galactosemia with MS/MS or in a panel of newborn screens seems to be cost-effective.

Some results are sensitive to changes in the effectiveness of screening and therapy, which emphasizes the importance of an effective timely system for notification and follow-up evaluation of positive screening tests. Loss of effectiveness in many areas may nullify the cost savings and health benefits of the screen. Similarly, the quality of the screening tests is important, specifically in minimizing false-positive results. The sensitivity of our results to the specificity of tests and the cost of evaluating false-positive results emphasizes that false-positive re-

sults must be minimized to realize the cost savings of these screening strategies.

Our pessimistic case analysis was intended to eliminate biases favoring newborn screening. With these assumptions, newborn screening with conventional testing or MS/MS methods was not cost saving. However, the incremental costs per QALY were very low, compared with conventional health economic benchmarks.^{38,39}

As with all cost-effectiveness analyses, our study is limited by the validity of our assumptions about risks, costs, and quality of life. For example, we assumed that data from small cohort studies would predict accurately the occurrence of sequelae in the population at large. We were forced to make some broad assumptions about lifetime costs that could not take into account changes in technology or care. It is also inherently difficult to estimate quality of life across different disease states, especially with respect to children. However, we found that incremental cost-effectiveness was relatively unaffected by even large changes in nearly all of our variables.

Another serious weakness of this study is the cost estimates for most of the conventional screening tests. We used estimates from PriceWaterhouseCoopers.³³ This study merely compiled global estimates by various newborn screening programs. There was no effort to divide fixed and marginal costs, and programs were not consistent in which costs were included in their estimates. Nevertheless, these are the best estimates currently available. A more-careful microeconomic analysis of screening program costs would be needed to improve these estimates; however, considering our sensitivity analyses, it seems unlikely that more carefully derived costs would show that newborn screening was not cost-effective.

Newborn screening seems to be one of the rare health care interventions that is beneficial to patients and, with some assumptions, cost saving. Over the long term, funding comprehensive newborn screening programs may save money for society. Even with pessimistic assumptions, newborn screening seems to be highly cost-effective. Although additional refinements of cost estimates may be justified, they are unlikely to change these conclusions materially. Cost-effectiveness analysis will continue to be a useful way to consider which screening tests to include.

ACKNOWLEDGMENTS

This work was conducted under a subcontract from the American College of Medical Genetics (agreement 240-01-0038), with funding from the Maternal and Child Health Bureau, Health Resources and Services Administration. Dr Carroll is supported by a grant from the National Institutes of Health (K23 DK67879-01).

We thank Alex Kemper, MD, MPH, MS, Tracy Lieu, MD, MPH, and Scott Grosse, PhD, for their thoughtful comments on a draft of this manuscript.

REFERENCES

- Venditti LN, Venditti CP, Berry GT, et al. Newborn screening by tandem mass spectrometry for medium-chain acyl-CoA dehydrogenase deficiency: a cost-effectiveness analysis. *Pediatrics*. 2003;112:1005–1015
- American Academy of Pediatrics, Committee on Genetics. Newborn screening fact sheets. *Pediatrics*. 1989;83:449–464
- Levy HL, Albers S. Genetic screening of newborns. *Annu Rev Genomics Hum Genet*. 2000;1:139–177
- Waisbren SE, Levy HL. Expanded screening of newborns for genetic disorders. *JAMA*. 2004;291:820–821
- Guthrie R, Susi A. A simple phenylalanine method for detecting phenylketonuria in large populations of newborn infants. *Pediatrics*. 1963;32:338–343
- Centers for Disease Control and Prevention. Using tandem mass spectrometry for metabolic disease screening among newborns. *MMWR Morb Mortal Wkly Rep*. 2001;50:1–22. Available at: www.cdc.gov/mmwr/preview/mmwrhtml/rr5003a1.htm. Accessed May 13, 2004
- American College of Medical Genetics/American Society of Human Genetics Test and Technology Transfer Committee Working Group. Tandem mass spectrometry in newborn screening. *Genet Med*. 2000;2:267–269
- Naylor EW, Chace DH. Automated tandem mass spectrometry for mass newborn screening for disorders in fatty acid, organic acid, and amino acid metabolism. *J Child Neurol*. 1999;14(suppl 1):S4–S8
- Insinga RP, Laessig RH, Hoffman GL. Newborn screening with tandem mass spectrometry: examining its cost-effectiveness in the Wisconsin Newborn Screening Panel. *J Pediatr*. 2002;141:524–531
- Schoen EJ, Baker JC, Colby CJ, To TT. Cost-benefit analysis of universal tandem mass spectrometry for newborn screening. *Pediatrics*. 2002;110:781–786
- Wood JC, Magera MJ, Rinaldo P, Seashore MR, Strauss AW, Friedman A. Diagnosis of very long chain acyl-dehydrogenase deficiency from an infant's newborn screening card. *Pediatrics*. 2001;108(1). Available at: www.pediatrics.org/cgi/content/full/108/1/e19
- Waisbren SE, Albers S, Amato S, et al. Effect of expanded newborn screening for biochemical genetic disorders on child outcomes and parental stress. *JAMA*. 2003;290:2564–2572
- Kwon C, Farrell PM. The magnitude and challenge of false-positive newborn screening test results. *Arch Pediatr Adolesc Med*. 2000;154:714–718
- Keren R, Helfand M, Homer C, McPhillips H, Lieu TA. Projected cost-effectiveness of statewide universal newborn hearing screening. *Pediatrics*. 2002;110:855–864
- Panepinto JA, Magid D, Rewers MJ, Lane PA. Universal versus targeted screening of infants for sickle cell disease: a cost-effectiveness analysis. *J Pediatr*. 2000;136:201–208
- Jensen UG, Brandt NJ, Christensen E, Skovby F, Norgaard-Pedersen B, Simonsen H. Neonatal screening for galactosemia by quantitative analysis of hexose monophosphates using tandem mass spectrometry: a retrospective study. *Clin Chem*. 2001;47:1364–1372
- Weinstein MC, Siegel JE, Gold MR, Kamlet MS, Russell LB. Recommendations of the Panel on Cost-Effectiveness in Health and Medicine. *JAMA*. 1996;276:1253–1258
- National Newborn Screening and Genetics Resource Center. *National Newborn Screening Report—2000*. Austin, TX: National Newborn Screening and Genetics Resource Center; 2003. Available at: <http://genes-r-us.uthscsa.edu/resources/newborn/00chapters.html>. Accessed June 3, 2004
- American Academy of Pediatrics, Committee on Genetics. Newborn screening fact sheets. *Pediatrics*. 1996;98:473–501
- National Newborn Screening and Genetics Resource Center. Homocystinuria. In: *National Newborn Screening Report—2000*. Austin, TX: National Newborn Screening and Genetics Resource Center; 2003:1–11. Available at: http://genes-r-us.uthscsa.edu/resources/newborn/00/ch7_complete.pdf. Accessed June 23, 2004
- Washington State Department of Health. Least Burden and Cost Benefit Analysis, Newborn Screening for Metabolic Disorders, WAC 246-650; August 12, 2003. Available at: www.sboh.wa.gov/Meetings/Meetings_2003/2003-10_15/documents/Tab09-NBS_analysis.pdf. Accessed June 23, 2004
- Cruysberg JR, Boers GH, Trijbels JM, Deutman AF. Delay in diagnosis of homocystinuria: retrospective study of consecutive patients. *BMJ*. 1996;313:1037–1040
- Naughten ER, Jenkins J, Francis DE, Leonard JV. Outcome of maple syrup urine disease. *Arch Dis Child*. 1982;57:918–921
- Scriver CR. *The Metabolic and Molecular Bases of Inherited Disease*. 8th ed. New York, NY: McGraw-Hill; 2001
- Wolf B, Spencer R, Gleason T. Hearing loss is a common feature of symptomatic children with profound biotinidase deficiency. *J Pediatr*. 2002;140:242–246
- Shield JP, Wadsworth EJ, MacDonald A, et al. The relationship of genotype to cognitive outcome in galactosaemia. *Arch Dis Child*. 2000;83:248–250
- Badawi N, Cahalane SF, McDonald M, et al. Galactosaemia: a controversial disorder: screening and outcome: Ireland 1972–1992. *Ir Med J*. 1996;89:16–17
- Pollitt RJ, Leonard JV. Prospective surveillance study of medium chain acyl-CoA dehydrogenase deficiency in the UK. *Arch Dis Child*. 1998;79:116–119
- Arias E, Smith BL. Deaths: preliminary data for 2001. *Natl Vital Stat Rep*. 2003;51:1–44
- Bittles AH, Petterson BA, Sullivan SG, Hussain R, Glasson EJ, Montgomery PD. The influence of intellectual disability on life expectancy. *J Gerontol A Biol Sci Med Sci*. 2002;57:M470–M472
- Bennett JE, Sumner W II, Downs SM, Jaffe DM. Parents' utilities for outcomes of occult bacteremia. *Arch Pediatr Adolesc Med*. 2000;154:43–48
- Sharma S, Brown GC, Brown MM, et al. Converting visual acuity to utilities. *Can J Ophthalmol*. 2000;35:267–272
- PriceWaterhouseCoopers. *Newborn Screening Programs: An Overview of Cost and Financing*. New York, NY: Price WaterhouseCoopers; 2002. Available at: www.marchofdimes.com/files/FinalPWC_NBS_Report2.pdf. Accessed June 3, 2004
- US Congress, Office of Technology Assessment. Data and methods used in OTA's cost-effectiveness analysis of strategies for newborn screening. In: *Healthy Children: Investing in the Future*. Washington, DC: US Government Printing Office; 1988: 236–241. Publication OTA-H-345. Available at: www.wps.princeton.edu/cgi-bin/byteserv/prl/~ota/disk2/1988/8819/881919.PDF. Accessed June 3, 2004
- Centers for Disease Control and Prevention. Economic costs associated with mental retardation, cerebral palsy, hearing loss, and vision impairment: United States, 2003. *MMWR Morb Mortal Wkly Rep*. 2004;53:57–59
- Angus DC, Barnato AE, Linde-Zwirble WT, et al. Use of intensive care at the end of life in the United States: an epidemiologic study. *Crit Care Med*. 2004;32:638–643
- US Department of Labor, Bureau of Labor Statistics. Inflation calculator. Available at: www.bls.gov/cpi. Accessed July 14, 2004
- Goldman L, Gordon DJ, Rifkind BM, et al. Cost and health implications of cholesterol lowering. *Circulation*. 1992;85:1960–1968
- Hirth RA, Chernew ME, Miller E, Fendrick AM, Weisert WG. Willingness to pay for a quality-adjusted life year: in search of a standard. *Med Decis Making*. 2000;20:332–342

Comprehensive Cost-Utility Analysis of Newborn Screening Strategies

Aaron E. Carroll and Stephen M. Downs

Pediatrics 2006;117;S287

DOI: 10.1542/peds.2005-2633H

Updated Information & Services

including high resolution figures, can be found at:
http://pediatrics.aappublications.org/content/117/Supplement_3/S287

References

This article cites 31 articles, 12 of which you can access for free at:
http://pediatrics.aappublications.org/content/117/Supplement_3/S287#BIBL

Subspecialty Collections

This article, along with others on similar topics, appears in the following collection(s):
Administration/Practice Management
http://www.aappublications.org/cgi/collection/administration:practice_management_sub
Fetus/Newborn Infant
http://www.aappublications.org/cgi/collection/fetus:newborn_infant_sub

Permissions & Licensing

Information about reproducing this article in parts (figures, tables) or in its entirety can be found online at:
<http://www.aappublications.org/site/misc/Permissions.xhtml>

Reprints

Information about ordering reprints can be found online:
<http://www.aappublications.org/site/misc/reprints.xhtml>

American Academy of Pediatrics

DEDICATED TO THE HEALTH OF ALL CHILDREN®



PEDIATRICS®

OFFICIAL JOURNAL OF THE AMERICAN ACADEMY OF PEDIATRICS

Comprehensive Cost-Utility Analysis of Newborn Screening Strategies

Aaron E. Carroll and Stephen M. Downs

Pediatrics 2006;117;S287

DOI: 10.1542/peds.2005-2633H

The online version of this article, along with updated information and services, is located on the World Wide Web at:

http://pediatrics.aappublications.org/content/117/Supplement_3/S287

Pediatrics is the official journal of the American Academy of Pediatrics. A monthly publication, it has been published continuously since 1948. Pediatrics is owned, published, and trademarked by the American Academy of Pediatrics, 345 Park Avenue, Itasca, Illinois, 60143. Copyright © 2006 by the American Academy of Pediatrics. All rights reserved. Print ISSN: 1073-0397.

American Academy of Pediatrics

DEDICATED TO THE HEALTH OF ALL CHILDREN®

