Effect of Newborn Screening for Congenital Adrenal Hyperplasia

Patrick G. Brosnan, MD; Christine A. Brosnan, DrPH; Stephen F. Kemp, MD, PhD; David B. Domek, MD; David H. Jelley, MD; Piers R. Blackett, MD; William J. Riley, MD

Objective: To compare the incidence of diagnosis and morbidity in newborns who were screened with newborns who were not screened for congenital adrenal hyperplasia (CAH).

Design: A retrospective cohort study.

Setting: Arkansas, Oklahoma, and Texas.

Patients: An unscreened population in Arkansas and Oklahoma (n = 400 118) was compared with a screened population in Texas (n = 1 613 378) during a 5-year period. Simultaneous data were collected on the incidence of diagnosis and associated morbidity in patients with CAH.

Main Outcome Measures: Diagnosis of CAH, age (in days) at diagnosis, and frequency and length of initial hospitalization.

Results: The incidence of diagnosis of classic CAH per 100 000 newborns in the unscreened cohort (5.75) and in the screened cohort (6.26) was similar (relative risk, 0.92; 95% confidence interval, 0.58-1.44). The unscreened group had 0.73 fewer male newborns with salt-wasting CAH diagnosed per 100 000 newborns (relative risk, 0.73; 95% confidence interval, 0.35-1.56). The median age at diagnosis was 26 days for male newborns with salt-wasting CAH in the unscreened cohort vs 12 days in the screened cohort (z = 2.49; P = .01). Male newborns with simple-virilizing CAH and newborns with nonclassic CAH were detected only in the screened cohort.

Conclusions: There was not a statistically significant (P = .73) increase in the diagnosis of salt-wasting CAH in the screened cohort. Male newborns benefited as a result of significantly (P = .01) earlier diagnosis, reduced morbidity, and shorter lengths of hospitalization. Large collaborative studies or meta-analyses are needed to determine the life-saving benefits of screening.


Editor’s Note: The authors are to be congratulated for using a naturally controlled environment to perform this study. The apparent benefit to male newborns is intriguing; now we need the big study to see if these or other effects result from early screening.

Catherine D. DeAngelis, MD

Many newborn screening programs have become possible and have been mandated for various US populations since phenylketonuria screening was introduced in 1961. In times of financial constraint, all cannot be universally adopted, and it is necessary to decide which screens are the most valuable. However, there is often little scientific evidence on which to base recommendations. A recent systematic literature review of screening for inborn errors of metabolism concluded that congenital adrenal hyperplasia (CAH) is 1 of 4 available screens deserving a widespread trial.

The worldwide incidence of classic CAH due to 21-hydroxylase deficiency is estimated to be 1 per 15 000 newborns (6.6 per 100 000). Classic CAH includes those forms evident in early childhood. The salt-wasting (SW) variant comprises 75% of the total and is known to cause hypovolemic shock and death in newborns. The simple-virilizing (SV) variant comprises 25% of the total and does not result in spontaneous hypovolemia. Both cause inappropriately virilization of female newborns and early overgrowth in both sexes, which compromises final height. Nonclassic (NC) CAH is a mild, late-presenting form of the disorder.

The newborn screening programs of 19 states include testing for CAH, and other states are considering doing the same. The justification for screening is pre-
PATIENTS AND METHODS

STUDY POPULATIONS

We gathered simultaneous data on the incidence of CAH diagnosis and associated morbidity in 400 118 newborns born between July 1, 1989, and June 30, 1994, in Arkansas and Oklahoma (the unscreened cohort) and in 1 613 378 newborns born during the same period in Texas (the screened cohort). Births by occurrence, rather than residence, were used. All Arkansas and Oklahoma residents born and screened in Texas were counted in the screened cohort; Texas residents born in Arkansas or Oklahoma and not screened were counted in the unscreened cohort.

The Texas cohort was a subset of newborns previously studied for whom screening, diagnostic, and follow-up data have been reported. In brief, newborns in Texas receive screening tests in the first days of life and again at 1 to 2 weeks of age. Dried blood collected onto filter paper cards is analyzed for 17α-hydroxyprogesterone concentration by radioimmunoassay. Newborns with elevated 17α-hydroxyprogesterone above specified levels are referred to pediatric endocrinologists for evaluation. Newborns confirmed to have CAH are regularly seen by those specialists who provide follow-up data to the Texas Department of Health using standardized forms. Data from the initial and first follow-up visits form the basis of the Texas component of this comparison.

The Arkansas and Oklahoma data were collected from medical record review of every patient with CAH from the birth cohort seen at least once by a pediatric endocrinologist. Every pediatric endocrinologist in Arkansas and Oklahoma participated in this study. Data on these children were collected during 1995 by one of us (C.A.B.), using a version of the form used in Texas.

DIAGNOSIS OF CAH

In Texas, classification into the variant forms was made according to commonly used criteria, as described previously. Since knowledge of a positive screen result or recognition of sex ambiguity can encourage overclassification, we reviewed all diagnostic assignments in the unscreened and screened cohorts for consistency. One of us, a pediatric endocrinologist in Texas (W.J.R.), used the Texas criteria to independently assign CAH variant status to cases in the unscreened cohort. A Spearman rank correlation coefficient found a significant association between the reviewer's diagnosis and the caretaker endocrinologist's diagnosis (r = 0.73, P < .001). The reviewer disagreed with the classification of 2 female newborns. The reviewer diagnosed one female newborn classified as having SV CAH and another female newborn classified as having SW CAH as having SV CAH. The prevalence of classic CAH is at least as great in Hispanics as in white non-Hispanics, and the percentage of African Americans is similar in both cohorts. Relative risk was estimated for each variant with Taylor series 95% confidence intervals. A corresponding χ² test was used to calculate if the differences in incidence were significant, and the Fisher exact test and Poisson probabilities were used when data were not appropriate for the χ² test. Differences in age at diagnosis and length of stay were assessed with the Mann-Whitney U test because the morbidity data were skewed. A P value of less than .05 was considered statistically significant.

As having SW CAH and another female newborn classified as having SW CAH as having SV CAH. The diagnosis of SW was based on a record of hyponatremia with hyperkalemia and high urine sodium or high renin level after cortisol replacement. Most unscreened male newborns had poor weight gain or frank hypovolemia. Clinical course after diagnosis was considered (eg, patients with SV CAH with subsequent crises were reclassified by their caretaker and by the reviewer as having SW CAH).

MORBIDITY

Early morbidity in newborns with SW CAH was calculated using 3 indicators. First, age (in days) at diagnosis estimates morbidity since SW symptoms worsen over time, increasing the risk of adrenal crisis. The age was calculated as the interval from birth to the day on which a diagnosis of CAH was suspected and confirmatory laboratory tests were ordered. Second, hospitalization at the time of diagnosis was defined as any peridiagnostic hospitalization, regardless of admitting diagnosis. Third, length of stay was defined as total days of initial hospitalization regardless of diagnosis or the number of hospital sites. Thus, if a newborn was first admitted to a community hospital and later transferred to a regional hospital, the days spent in each hospital were included. Although data on weight loss, vomiting, “shock,” “crisis,” and other manifestations of SW were available in the patient records, they had not been uniformly defined for the study; therefore, days to diagnosis, frequency of hospitalization, and length of stay were chosen as the most objective and economically meaningful measures of morbidity.

ANALYSES

The cumulative incidence of diagnosis per 100 000 newborns was calculated since we were interested in the detection of the condition rather than population prevalence. Adjustments for ethnicity were not necessary. Although Texas has a larger Hispanic population than Arkansas or Oklahoma, the prevalence of classic CAH is at least as great in Hispanics as in white non-Hispanics, and the percentage of African Americans is similar in both cohorts. Relative risk was estimated for each variant with Taylor series 95% confidence intervals. A corresponding χ² test was used to calculate if the differences in incidence were significant, and the Fisher exact test and Poisson probabilities were used when data were not appropriate for the χ² test. Differences in age at diagnosis and length of stay were assessed with the Mann-Whitney U test because the morbidity data were skewed. A P value of less than .05 was considered statistically significant.
tified clinically or by family history. All SW CAH cases were detected clinically or on the first screen, while newborns with SV CAH and NC CAH were more likely to be identified on the second screen. The program's incremental cost in 1994 for the screening detection of 8 newborns with classic CAH who were not yet recognized clinically was $147,093 per case.10

Table 1. The 5-Year Cumulative Incidence of Congenital Adrenal Hyperplasia (CAH) in an Unscreened and Screened Cohort by Type and Sex*

<table>
<thead>
<tr>
<th>CAH Type</th>
<th>Unscreened Cohort†</th>
<th>Screened Cohort‡</th>
<th>Worldwide Estimate§</th>
<th>Relative Risk (95% Confidence Interval)¶</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Cases</td>
<td>Incidence per</td>
<td>Cases</td>
<td>Incidence per</td>
</tr>
<tr>
<td></td>
<td>Diagnosed</td>
<td>100 000 Newborns</td>
<td>Diagnosed</td>
<td>100 000 Newborns</td>
</tr>
<tr>
<td>SW</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>8</td>
<td>1.99</td>
<td>44</td>
<td>2.72</td>
</tr>
<tr>
<td>Female</td>
<td>12</td>
<td>2.99</td>
<td>30</td>
<td>1.86</td>
</tr>
<tr>
<td>Total</td>
<td>20</td>
<td>4.99</td>
<td>74</td>
<td>4.58</td>
</tr>
<tr>
<td>SV</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>0</td>
<td>0</td>
<td>11</td>
<td>0.68</td>
</tr>
<tr>
<td>Female</td>
<td>3</td>
<td>0.75</td>
<td>16</td>
<td>0.99</td>
</tr>
<tr>
<td>Total</td>
<td>3</td>
<td>0.75</td>
<td>27</td>
<td>1.67</td>
</tr>
<tr>
<td>Total Classic</td>
<td>23</td>
<td>5.75</td>
<td>101</td>
<td>6.26</td>
</tr>
<tr>
<td>Nonclassic</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>0</td>
<td>0</td>
<td>27</td>
<td>1.67</td>
</tr>
<tr>
<td>Female</td>
<td>1</td>
<td>0.25</td>
<td>19</td>
<td>1.17</td>
</tr>
<tr>
<td>Total</td>
<td>1</td>
<td>0.25</td>
<td>46</td>
<td>2.85</td>
</tr>
<tr>
<td>Total Cases</td>
<td>24</td>
<td>5.99</td>
<td>147</td>
<td>9.11</td>
</tr>
</tbody>
</table>

*SW indicates salt wasting; SV, simple virilizing; and ellipses, data not applicable.
†Included 400,118 births from Arkansas and Oklahoma (sources: Planning and Evaluation Section, Oklahoma State Department of Health; and Center for Health Statistics, Arkansas Department of Health).
‡Included 1,613,378 births from Texas (source: Bureau of Vital Statistics, Texas Department of Health).
§Given as the incidence per 100,000 newborns in the unscreened cohort divided by the incidence per 100,000 newborns in the screened cohort.
¶Significant difference (P < .01).

The cumulative incidence of classic CAH diagnosis in the unscreened Arkansas and Oklahoma cohort was 1 per 17,396, while the incidence in the Texas cohort was 1 per 15,974. The incidence among newborns classified as white, Hispanic, or other was 1 per 15,277 in the unscreened vs 1 per 14,158 in the screened cohort. The incidence among newborns classified as African American was 1 per 64,018 in the unscreened and 1 per 75,291 in the screened population.

Comparisons of the cumulative incidence of CAH diagnosis, by type and sex, are shown in Table 1. The estimated worldwide incidence is also provided.4,5 The difference in incidence of the classic CAH diagnosis between the unscreened cohort and the screened cohort was not statistically significant (P = .71).

Newborns with SW CAH accounted for 87% (20/23) of all classic cases in the unscreened vs 73% (74/101) in the screened cohort. The overall incidence of SW CAH in the unscreened group was not significantly (P = .73) different from the incidence in the screened cohort. Although not significant (P = .42), there were 0.73 fewer male newborns per 100,000 newborns (1.46 per 100,000 male newborns) in whom CAH was diagnosed in the unscreened cohort. Using Poisson probability, the incidence of undiagnosed male newborns with SW CAH in the unscreened cohort ranged from 0 to 1.5 per 100,000 newborns (95% confidence interval). There were 1.13 per 100,000 more female newborns with SW CAH in the unscreened population. For SW CAH, the male-female ratio was 0.67:1.00 in the unscreened and 1.47:1.00 in the screened population.

No male newborns with SV CAH were found in the unscreened group, while 0.68 per 100,000 were found by screening. There was a significant association between screening and diagnosis of NC CAH among male newborns (Poisson probability = .001) and among male and female newborns combined (χ² = 9.29; P = .002). There were 2.85 per 100,000 newborns diagnosed as having NC CAH in Texas, while only 1 female child (who presented at 4½ years with premature adrenarche) was diagnosed as having NC CAH in Arkansas and Oklahoma.

We were unable to state with certainty that any newborn death attributable to CAH occurred in either group before diagnosis by the comparison of incidence technique used in the prior literature on this subject. (One “male” newborn with a positive screen result died in Texas before confirmation and was found at autopsy to be a fe-
male newborn with hypertrophic adrenal glands. To our knowledge, no child in whom CAH was diagnosed in either cohort has died as a result of the disease.

Results relating to early morbidity in newborns with SW CAH are shown in Table 2. Male newborns with SW CAH were identified significantly later in the unscreened cohort ($z = 2.49; P = .01$). Male newborns with SW CAH from both cohorts tended to have some clinical evidence of SW at presentation (weight loss, vomiting, or poor feeding). Of those newborns for whom data were available, 88% (7/8) in the unscreened cohort and 77% (30/39) in the screened group manifested at least 1 of these symptoms when recognized. Female newborns with SW CAH in both cohorts tended to be recognized early because of genital ambiguity, but 17% (2/12) in the unscreened and 44% (11/25) in the screened cohorts had some clinical evidence of SW during their initial workups. According to the caretaker endocrinologists in Arkansas and Oklahoma, no child in the unscreened cohort has neurodevelopmental disability ascribed to CAH. Two female newborns in the unscreened cohort were severely masculinized and were assigned a male sex until SW symptoms appeared at 22 and 35 days. Sex misassignments in Texas were limited to the notification time of the screen, which ranged from 9 to 13 days of age in 1994. In both cohorts, newborns with SW CAH were likely to be hospitalized at diagnosis, 79% (15/19) in the unscreened cohort and 82% (55/67) in the screened cohort. Female newborns with SW CAH or SV CAH from both cohorts were often hospitalized for diagnostic evaluation of genital ambiguity. Although not statistically significant ($P = .32$), the median length of hospitalization was 7.5 days longer for male newborns with SW in the unscreened cohort vs the screened cohort.

The incremental differences in cases detected between the unscreened population and the screened population after 1 and 2 screens are shown in the Figure. All screened newborns with SW CAH were detected either on the first screen (43% [32/74]) or clinically before screening results were reported (57% [42/74]). Among screened newborns with SV CAH, 48% (13/27) were detected only by the second screen. These second screen diagnoses account for almost the entire 0.92 per 100 000 difference in SV CAH diagnosis between the unscreened and screened populations.
screened and screened cohorts. There were 2.85 additional cases of NC CAH per 100,000 newborns detected in Texas, and 83% (38/46) of these were initially identified on the second screen.

COMMENT

Thirteen neonatal screens are conducted in at least 1 state in the United States. There is a great deal of interest in ascertaining how these screens rank for morbidity and mortality saved, so that scarce screening dollars can be put where they are most useful. However, there is a lack of evidence on which to base decisions. Our experience, and a critical review of the literature on which adoption of CAH screening is based, illustrates the difficulty of evaluation. There is general agreement that newborn screening can find newborns with CAH; however, every newborn found by screening is not necessarily a large benefit, if that patient was already recognized or was likely to be found clinically in a few days with little added morbidity.

Newborn screening for CAH by testing blood dried onto filter paper for 17α-hydroxyprogesterone was introduced in Alaska. The incidence of SW CAH among the Yupik population there was found to be 1 per 490. Timely clinical diagnosis was rare, and neonatal deaths from CAH occurred frequently. In a remote group of high prevalence such as the Yupik, there was little doubt that the missing males in the clinical series were SW cases who had died without diagnosis.

Suwa estimated that the incidence of clinically diagnosed CAH in Japan before 1981 was 1 per 43,764; in 1994, he reported an incidence of 1 per 18,877, ascertained by screening 4 million newborns. The difference in incidence was significant, but since the studies were separated by a decade, an increase in practitioners’ knowledge of CAH might have changed the diagnostic rate even without screening. Suwa’s study indicated that the 5-year prevalence of CAH with screening was 1 per 9,800.

Our findings are similar to these recent comparisons. The incidences of classic CAH in the unscreened (5.75 per 100,000) and screened (6.26 per 100,000) cohorts were not significantly different and were similar to that observed worldwide (6.60 per 100,000). The incidences of SW CAH in both cohorts were comparable. Although it was not a significant finding, the diagnosis was less likely to be made in SW male newborns in the unscreened group than in SW male newborns who were born in Texas. The high incidence of diagnosed SW CAH in female newborns and the 0.67:1.00 male-female ratio for SW CAH in the unscreened cohort does not suggest that not all asymptomatic screen-identified male newborns would have presented clinically.

An alternate reason for the failure to detect significant differences between unscreened and screened populations in current studies may be that better access to health care, increased frequency of electrolyte measurement in vomiting newborns, and increased awareness of CAH as a potential cause of hypovolemia and genital ambiguity have improved recognition of SW CAH in all newborns and of SV CAH in female newborns.

An alternate reason for the failure to detect significant differences between unscreened and screened cohorts in large regional studies, such as ours, the study by Balsamo et al, and the study by Thilen et al, may be low statistical power due to the rarity of the disease. The expected incidence of SW CAH in male newborns is 1 per 40,000, or 10 in the entire unscreened cohort of 400,118. If 2 male newborns with SW CAH were missed, they would represent almost 10% of newborns born with classic CAH in Arkansas and Oklahoma during the study period. Yet, the detection of these cases would only increase the incidence of classic CAH from 1 per 17,396 to 1 per 16,004 births, both similar to the incidence of 1 per 15,974 in the screened cohort. Large collaborative investigations are needed before the lifelong effects of screening can be measured. However, studies like ours estimate a range for the effect size of screening.

Although we did not demonstrate that neonatal screening averts mortality, we did find that indicators associated with morbidity were reduced for SW newborns in Texas. Screening significantly shortened the time to diagnosis for male newborns with SW CAH. Since age at diagnosis has been shown to correlate with severity of SW symptoms, screening might reduce the cost and risk of initial care. Lengths of stay were shorter for male and female newborns in the screened cohort,
possibly reflecting milder illness or decreased diagnostic uncertainty. Sex misassignments were limited to the notification time of the screen, while 2 misassigned female newborns in the unscreened cohort presented with severe SW symptoms at 25 and 33 days of age. However, continued education of physicians should improve the timeliness of clinical detection and reduce sex misassignment.

Screening remains the only reliable way to recognize SV CAH in male newborns. Children in whom SV CAH is diagnosed late often have severe epiphyseal advancement with poor prognosis for final height. The condition was not diagnosed clinically in male newborns during the newborn period, and these male newborns did not present in the unscreened cohort during the 5 years of the study. Therefore, it is possible that screening may prevent loss of final height.26 Therrell et al26 found that 1 screen did not effectively detect all male newborns with this variant and that a 2-screen approach was needed.

The worldwide incidence of NC 21-hydroxylase deficiency is debated, and estimates range as high as 1% of some populations.27 If this estimate is valid, screening programs detect only a small proportion of newborns with NC CAH. No one has proposed screening for the purpose of finding NC CAH, although some newborns are at risk for short stature and may be helped by early treatment.26 Since there are not uniformly accepted criteria for distinguishing this variant from SV CAH in male newborns, many identified newborns must be observed for evidence of overgrowth.27,28 Criteria for treatment of screen-detected newborns with NC CAH will only evolve as these children are observed over time. Risks of screening include the possibility of overtreating newborns with mild forms of CAH and the parental anxiety caused by prolonged follow-up of positive screen results and uncertainty about when to initiate therapy.7,20,30

Brosnan et al31 reported elsewhere the costs of CAH screening in Texas. Based on those data, the cost to all payers for the addition of a single screen for CAH to an existing program (including a physician examination, electrolyte profile, and rescreen for positive first screen results) would be $257 735 per 100 000 newborns (1994 US dollars). Adding 2 screens for CAH to an existing 2-screen program would cost $348 839 per 100 000 newborns. Setting up a second screen de novo would cost $918 839 per 100 000 newborns, mostly because of expenses related to specimen collection. Estimates included the cost of diagnostic evaluation for false-positive results. The false-positive rate was 0.65% for the first screen and 0.40% for 2 screens.

Our results are consistent with those of others21,22,25,33 in suggesting that screening for CAH has benefit. First, screening detects 0.73 additional male newborns affected with the severe SW variant per 100 000 newborns. For male newborns with the SW variant, the 95th percentile confidence interval on “missed diagnosis” in the unscreened population is 0 to 1.5 per 100 000. Second, screening once ensures that the condition is diagnosed in male newborns at risk for adrenal crisis at an early age when they are not so seriously ill. Third, screening may reduce the length of hospitalization. Further quantification of economic savings requires comparison studies using standardized measures of care acuity and is beyond the scope of this study. Finally, screening results in the diagnosis and treatment of male newborns with SV CAH and in the identification and surveillance of some newborns with NC CAH. However, a second screen is needed to detect most newborns with these milder variants.

In evaluating potential additions to newborn screening programs, the risks of adverse outcome preventable by screening should be considered in addition to the incidence of the disease.22,33 Phenylketonuria and CAH have a similar prevalence, but phenylketonuria is rarely recognized clinically before causing costly developmental delay, while CAH is often recognized clinically and is rarely associated with developmental disability.23 The primary value of CAH screening results from its ability to prevent death and to avert serious illness during the neonatal period before the diagnosis is established. Since the worldwide incidence of severe SW CAH in male newborns is 1 per 40 000 (2.5 per 100 000)12 and our data suggest that 30% may go undiagnosed, it is reasonable to compare CAH with diseases in which the incidence of preventable morbidity or mortality is in the order of 1 to 1.5 per 100 000. Congenital adrenal hyperplasia is associated with an estimated 1.5% risk of death after diagnosis,34 and so newborns who are detected through screening are likely to have productive lives. Convincing evidence that screening saves lives awaits a large collaborative study or, if the data are available, a carefully constructed meta-analysis.

Accepted for publication May 14, 1999.

This study was supported by the Genentech Foundation for Growth and Development, Charlottesville, Va.

From the Department of Pediatrics, School of Medicine (Dr P. G. Brosnan), and the Department of Systems and Technology, School of Nursing (Dr C. A. Brosnan), The University of Texas–Houston Health Science Center; the Department of Pediatrics, University of Arkansas for Medical Sciences, Little Rock (Dr Kemp); Department of Pediatrics, Integris Baptist Medical Center, Oklahoma City, Okla (Dr Domek); the Warren Clinic Diabetes Center and the University of Oklahoma Health Sciences Center, Tulsa (Dr Jelley); the Department of Pediatrics, University of Oklahoma Health Sciences Center, Oklahoma City (Dr Blackett); and the Department of Medical Education, Driscoll Children’s Hospital, Corpus Christi, Tex (Dr Riley).

Presented in abstract form at the Pediatric Endocrine Society of Texas, Oklahoma, Arkansas, and Louisiana Meeting, Dallas, Tex, April 19, 1996; the Pediatric Endocrine Society of Texas, Oklahoma, Arkansas, and Louisiana Meeting, Houston, Tex, April 16, 1999; and the Endocrine Advisory Council of the Texas Department of Health Meeting, Austin, Tex, October 24, 1997.

We thank John F. Annegers, PhD, School of Public Health, The University of Texas–Houston, who provided consultation in study design and epidemiological statistics; and Robert Franks, MD, who made valuable editorial suggestions.

Reprints: Patrick G. Brosnan, MD, Division of Pediatric Endocrinology, Department of Pediatrics, The University of Texas–Houston Health Science Center, 6431 Fann
REFERENCES


ARCH PEDIATR ADOLESC MED/VOL 153, DEC 1999

©1999 American Medical Association. All rights reserved.