Newborn screening protocols and positive predictive value for congenital adrenal hyperplasia vary across the United States

P. W. Speiser  
*Zucker School of Medicine at Hofstra/Northwell*, pspeiser@northwell.edu

R. Chawla

M. Chen

A. Diaz-Thomas

C. Finlayson

*See next page for additional authors*

Follow this and additional works at: [https://academicworks.medicine.hofstra.edu/articles](https://academicworks.medicine.hofstra.edu/articles)

**Recommended Citation**

Newborn Screening Protocols and Positive Predictive Value for Congenital Adrenal Hyperplasia Vary across the United States

Phyllis W. Speiser 1,* 1, Reeti Chawla 2, Ming Chen 3, Alicia Diaz-Thomas 4, Courtney Finlayson 5, Meilan M. Rutter 6, David E. Sandberg 7, Kim Shimy 8, Rashida Talib 1, Jane Cerise 9, Eric Vilain 10, Emmanuelle C. Délot 10 and on behalf of the Disorders/Differences of Sex Development-Translational Research Network (DSD-TRN)

1 Division of Endocrinology, Cohen Children’s Medical Ctr of New York, Feinstein Institute for Medical Research, Zucker School of Medicine at Hofstra University, New Hyde Park, NY 11040, USA; rtalib@northwell.edu
2 Division of Endocrinology, Phoenix Children’s Hospital, Phoenix, AZ 85016, USA; rchawla@phoenixchildrens.com
3 Division of Endocrinology, CS Mott Children’s Hospital, University of Michigan, Ann Arbor, MI 48109, USA; chenming@med.umich.edu
4 Division of Endocrinology, LeBonheur Children’s Hospital, University of Tennessee Health Science Center, Memphis, TN 18103, USA; adiaztho@uthsc.edu
5 Division of Endocrinology, Ann & Robert H. Lurie Children’s Hospital of Chicago, Northwestern University Feinberg School of Medicine, Chicago, IL 60611, USA; cfinlayson@luriechildrens.org
6 Division of Endocrinology, Cincinnati Children’s Hospital Medical Center, University of Cincinnati, Cincinnati, OH 45229, USA; mleilan.rutter@cchmc.org
7 Susan B. Meister Child Health Evaluation and Research Center, University of Michigan, Ann Arbor, MI 48109, USA; dsandber@med.umich.edu
8 Division of Endocrinology, Children’s National Medical Center, Washington, DC 20010, USA; km.shimy@childrensnational.org
9 Feinstein Institute for Medical Research, Northwell Health, Manhasset, NY 11030, USA; jcerise@northwell.edu
10 Children’s National Hospital, Children’s Research Institute and George Washington University, Washington, DC 20010, USA; evilain@gwu.edu (E.V.); edelot@cnmc.org (E.C.D.)
* Correspondence: pspeiser@northwell.edu

Abstract: Newborn screening for congenital adrenal hyperplasia (CAH) caused by 21-hydroxylase deficiency is mandated throughout the US. Filter paper blood specimens are assayed for 17-hydroxyprogesterone (17OHP). Prematurity, low birth weight, or critical illness cause falsely elevated results. The purpose of this report is to highlight differences in protocols among US state laboratories. We circulated a survey to state laboratory directors requesting qualitative and quantitative information about individual screening programs. Qualitative and quantitative information provided by 17 state programs were available for analysis. Disease prevalence ranged from 1:9941 to 1:28,661 live births. Four state laboratories mandated a second screen regardless of the initial screening results; most others did so for infants in intensive care units. All but one program utilized birthweight cut-points, but cutoffs varied widely: 17OHP values of 25 to 75 ng/mL for birthweights >2250–2500 g. The positive predictive values for normal birthweight infants varied from 0.7% to 50%, with the highest predictive values based in two of the states with a mandatory second screen. Data were unavailable for negative predictive values. These data imply differences in sensitivity and specificity in CAH screening in the US. Standardization of newborn screening protocols could improve the positive predictive value.

Keywords: adrenal hyperplasia; congenital; newborn screening; standardization

1. Introduction

Newborn screening for congenital adrenal hyperplasia caused by steroid 21-hydroxylase deficiency (CAH-21) is mandated in the United States, 35 other countries, and in portions of 17 additional countries. Most of the US states report participation of 99.9% or higher [1]. The rationale for screening is to recognize and promptly treat the potentially life-threatening severe salt-wasting classic form of CAH. The overall prevalence of classic CAH-21 is approximately 1:14,000 to 1:18,000, of whom 75% have the salt-wasting form and 25% have non-salt-wasting or so-called simple virilizing CAH (summarized in [2]). Case fatality rates (lethality) of ~4% have been estimated for the era prior to newborn screening, presumably due to salt-wasting adrenal crisis, less frequent in the era of screening [3].

Affected 46,XY males cannot readily be identified without screening, as external genital virilization will flag the diagnosis only in 46,XX individuals. Lethality estimates were based on a rising disease prevalence, increased male to female case ratios, and increased salt wasting to simple virilizing case ratios observed since the advent of screening in the 1980s [4–6]. Other benefits of screening, aside from reduced infant mortality, have included less severe hyponatremia, shorter hospitalizations for infants identified with CAH [7–10], fewer learning disabilities as sequelae of salt-wasting crises [11,12], fewer reports of regret in the assigned gender of rearing [13,14], and potentially better growth outcomes [15].

2. Materials and Methods

Pediatric endocrinologists and other members of the Disorders/Differences of Sex Development Translational Research Network (DSD-TRN) circulated a survey (Supplementary Data) to state laboratory directors of their own states and neighboring state programs requesting qualitative and quantitative information about CAH screening programs. Questions included basic epidemiologic statistics for the number of infants screened, number of positive screens, and number of cases confirmed as CAH. Details of the assay methods, cut-points, and factors considered for stratifying abnormal values were queried. We asked about the personnel responsible for reporting and receiving abnormal results for hospitalized and discharged infants, and criteria for follow-up testing. Data sets were returned from 17 states, and descriptive statistics were analyzed for the calendar year 2017. Results shown refer to those obtained in normal birthweight infants. Confirmed CAH cases (true positives) were defined as those reported to the state screening programs as meeting standard confirmatory diagnostic criteria for classic CAH, including a baseline serum 17OHP >10,000 ng/dL (or 300 nmol/L), cosyntropin testing resulting in markedly elevated stimulated serum 17OHP [16], and/or CYP21A2 genotyping indicating known classic disease-causing mutations. False positives included those infants with 17OHP results above the cut-off point for birthweight but ultimately confirmed as not having CAH. Some of these data were presented in spring 2019 at The Endocrine Society Meeting, New Orleans, LA, USA and at the National Newborn Screening Board Meeting, Bethesda, MD, USA.

Statistical Methods

The reported number of screens, number of positive screens (excluding low birth weight infants), and number of endocrinologist-confirmed CAH cases from 2017 were used to estimate the number of true negative cases, prevalence, and positive predictive value (PPV). We arbitrarily assumed no false negative results for all programs, as negative screens were not referred for definitive diagnosis, and their final status was not reported to the state laboratories. Therefore, the number of confirmed
cases was estimated for each state as the number of positive screens who were confirmed as having CAH, and the number of true negative cases was estimated for each state as the total number screened minus the sum of true positive and false positive cases. Under the assumption of no false negatives, the estimate of the number of confirmed cases served as a lower bound on the number of confirmed cases, and the estimate of the number of true negative cases served as an upper bound on the number of true negative cases. In other words, if there were any false negatives, this would result in a higher number of confirmed cases and a lower number of true negative cases. Similarly, the prevalence, estimated for each state as the number of confirmed positives divided by the total number screened, was a lower bound estimate. PPV for each state was estimated as 100 times the number of confirmed cases divided by the number of positive screens. The weighted mean prevalence was calculated as the sum of the prevalence for each state weighted by the proportion of patients screened for that state out of all 17 states. Under the assumption of no false negatives, because all newborns are screened, this is equal to the sum of all states’ confirmed cases divided by the sum of infants screened in all states.

Spearman correlation coefficients (shown as $r_s$ below) were calculated between PPV and the 17OHP cut-off point, and between PPV and prevalence.

A $p$-value < 0.05 was considered statistically significant. All analyses were conducted using SAS version 9.4 (SAS Institute Inc., Cary, NC, USA). Figure 1 was created using the package ggplot2 in R version 3.6.1.

Figure 1. The positive predictive values (PPVs) calculated for each state are shown, with their respective 95% confidence intervals. States are ordered from the smallest to largest number of screens performed (from bottom up on y axis). Dot size is related to the number of screens performed (numerical data used are shown in Table 1).
<table>
<thead>
<tr>
<th>Screened (#Total = 1,564,756)</th>
<th>Positive (#Total = 3217)</th>
<th>Confirmed CAH (#Total = 93)</th>
<th>Prevalence (Weighted Mean = 1:16,825)</th>
<th>PPV (95% CI) IQR = [2.02 to 8.33]</th>
<th>Birth Weight Cut-Off Point (Grams)</th>
<th>17OHP Cut-Off Point (ng/mL) (Mean = 41.2)</th>
<th>Two-Screens Mandate</th>
</tr>
</thead>
<tbody>
<tr>
<td>230,431</td>
<td>146</td>
<td>11</td>
<td>1:20,948</td>
<td>7.53 (3.82–13.08)</td>
<td>&gt;2251</td>
<td>35</td>
<td>N</td>
</tr>
<tr>
<td>171,964</td>
<td>506</td>
<td>6</td>
<td>1:28,661</td>
<td>1.19 (0.44–2.56)</td>
<td>&gt;2500</td>
<td>55</td>
<td>N</td>
</tr>
<tr>
<td>138,226</td>
<td>608</td>
<td>9</td>
<td>1:15,358</td>
<td>1.48 (0.68–2.79)</td>
<td>&gt;2500</td>
<td>70</td>
<td>N</td>
</tr>
<tr>
<td>135,590</td>
<td>66</td>
<td>6</td>
<td>1:22,598</td>
<td>9.09 (3.41–18.74)</td>
<td>&gt;2500</td>
<td>30</td>
<td>N</td>
</tr>
<tr>
<td>109,740</td>
<td>94</td>
<td>9</td>
<td>1:12,193</td>
<td>9.57 (4.47–17.4)</td>
<td>&gt;2500</td>
<td>65</td>
<td>N</td>
</tr>
<tr>
<td>104,000</td>
<td>387</td>
<td>7</td>
<td>1:14,857</td>
<td>1.81 (0.73–3.69)</td>
<td>&gt;2250</td>
<td>25</td>
<td>N</td>
</tr>
<tr>
<td>84,000</td>
<td>247</td>
<td>5</td>
<td>1:16,800</td>
<td>2.02 (0.66–4.66)</td>
<td>&gt;2500</td>
<td>30</td>
<td>N</td>
</tr>
<tr>
<td>81,117</td>
<td>112</td>
<td>6</td>
<td>1:13,520</td>
<td>5.36 (1.99–11.30)</td>
<td>&gt;2500</td>
<td>50</td>
<td>Y</td>
</tr>
<tr>
<td>79,948</td>
<td>144</td>
<td>5</td>
<td>1:15,990</td>
<td>3.47 (1.14–7.92)</td>
<td>&gt;2249</td>
<td>37</td>
<td>N</td>
</tr>
<tr>
<td>79,000</td>
<td>135</td>
<td>4</td>
<td>1:19,750</td>
<td>2.96 (0.81–7.41)</td>
<td>≥2500</td>
<td>35</td>
<td>Y</td>
</tr>
<tr>
<td>72,440</td>
<td>36</td>
<td>3</td>
<td>1:24,147</td>
<td>8.33 (1.75–22.47)</td>
<td>&gt;2251</td>
<td>75</td>
<td>N</td>
</tr>
<tr>
<td>61,500</td>
<td>8</td>
<td>4</td>
<td>1:15,375</td>
<td>50.00 (15.70–84.30)</td>
<td>&gt;2500</td>
<td>25</td>
<td>Y</td>
</tr>
<tr>
<td>59,643</td>
<td>113</td>
<td>6</td>
<td>1:9941</td>
<td>5.31 (1.97–11.2)</td>
<td>No values given</td>
<td>N</td>
<td></td>
</tr>
<tr>
<td>55,935</td>
<td>121</td>
<td>5</td>
<td>1:11,187</td>
<td>4.13 (1.36–9.38)</td>
<td>&gt;2500</td>
<td>25</td>
<td>N</td>
</tr>
<tr>
<td>53,361</td>
<td>434</td>
<td>3</td>
<td>1:17,767</td>
<td>0.69 (0.14–2.01)</td>
<td>&gt;2500</td>
<td>30</td>
<td>N</td>
</tr>
<tr>
<td>36,361</td>
<td>55</td>
<td>3</td>
<td>1:12,120</td>
<td>5.45 (1.14–15.12)</td>
<td>≥2500</td>
<td>38.3</td>
<td>N</td>
</tr>
<tr>
<td>11,500</td>
<td>5</td>
<td>1</td>
<td>1:11,500</td>
<td>20.00 (0.51–71.64)</td>
<td>≥2300</td>
<td>35</td>
<td>Y</td>
</tr>
</tbody>
</table>
3. Results

3.1. Epidemiology

Completed questionnaires were provided by 17 states and were analyzed (Table 1). Partial responses from 3 additional states were not included. The single largest state screening population in our survey included >230,000 infants in a single year, and the smallest only 11,500. More cases were referred for follow-up and confirmatory testing, and then were verified as having CAH. States mandating a second screening sample referred only those second positive screens. In this survey, there were a total of 93 confirmed cases. The prevalence of CAH ranged from 1:9941 to 1:28,661 live births, with a weighted mean prevalence of 1:16,825, similar to published estimates (Table 1). We were unable to obtain mortality or morbidity statistics specific to CAH.

3.2. Assays and Cut-Off Points

All laboratories used fluoroimmunoassay for 17OHP measurement in filter paper blood samples (DELFIA time-resolved fluorescence assay, Perkin Elmer, Waltham, MA, USA) obtained after 24 h of life. Most states performed the assay in their own laboratories, while some sent the samples to Perkin Elmer’s facilities using the same assay kits. All but one of the programs utilized birthweight cut-points, but cutoffs varied widely, from 17OHP values of 25 to 75 ng/mL (mean 41.2 ng/mL) for normal birthweights >2250–2500 g (Table 1). Cut-off points for lower birth weights were generally higher. Four state laboratories in our survey mandated a second screen regardless of initial screening results, and results of the second screen were used in this analysis where available. Most states utilized a later or second screen for infants in intensive care units, as early screening of sick, premature, or low birth weight infants gives many false positive results, contributing to lower positive predictive values [2].

The estimated positive predictive values for normal birthweight infants varied from 0.7% to 50% (mean 8.1%; median 5.3%). No statistically significant correlation was observed between PPV and the 17OHP cut-off point ($r_s = -0.0059, p = 0.98, n = 16$), nor between PPV and prevalence ($r_s = 0.21, p = 0.42, n = 17$). The highest predictive values were found in two of the four states with a mandatory second screen, with estimated PPV values greater than 1.5 times the interquartile range of 6.3. (Table 1, Figure 1).

3.3. Post-Analytical Procedures

There were no data concerning the cross-validation of assays carried out in different state laboratories. Each state reported positive results differently. Methods included phone calls, faxed documents, and/or mailed letters. Reporting went to a variety of reportees, including administrative aides, nurses, genetic counselors, or pediatric endocrinologists. Reports on infants with mildly or moderately abnormal screens who had been discharged from the birth hospital typically were sent to the primary care pediatrician listed on the hospital discharge. Information was not available regarding the content of counseling to the families. Infants with borderline or mildly elevated results usually underwent repeat screening. When an abnormal screen was obtained for a hospitalized infant, the neonatologist was informed, and a repeat sample obtained shortly thereafter. Those with more markedly abnormal results were most often referred directly to the pediatric endocrinologist at the state-designated center and underwent further confirmatory tests, such as a cosyntropin stimulation test if indicated, at the clinician’s discretion. We were unable to obtain data on the exact time between the reporting of results to the clinician and the start of treatment, as such data were not collected by the state laboratories. There was no consistent means of reporting infants missed by screening who later proved to be affected with classic CAH, and thus data were unavailable for false negative results or negative predictive values. Moreover, limited or no phenotypic data were recorded by state laboratories about the infants’ genital examination or other clinical features; no long-term outcomes
were collected. Thus, we could not ascertain whether screening contributed to the prevention of salt wasting, nor were we able to distinguish salt-wasting from non-salt-wasting types of CAH.

4. Discussion

The primary aim of this report was to document differences in CAH newborn screening protocols. Our survey of about one-third of the US states confirms that screening for CAH using a single sample primary fluoroimmunoassay for 17OHP is associated with low positive predictive values [17]. In this report, we only included data related to the positive predictive value for normal birthweight infants, as prematurity, low birthweight, and sampling within the first 24 h of life are major causes of false positive results (reviewed in [2]). Although all laboratories used the same fluoroimmunoassay kit, conditions in the individual laboratories may have contributed to the lack of correlation between the cut-off point and positive predictive value. Cross-validation of assays might reduce some of the variability. Alternatively, these observations may be due to the relatively small number of confirmed positive cases in each state sample.

Protocols for CAH newborn screening have been adapted and implemented separately by each state laboratory in the US over the past several decades. Publications have generally reported data unique to an individual state screening program with a few exceptions [18,19]. For instance, at present, 13 states mandate a second screen at about two weeks of life in part due to the relatively low positive predictive value of the initial screen. These programs tend to show a higher disease prevalence compared with those only performing a single screen [18]. However, those later diagnosed infants are often non-salt-wasting, which some argue is not the objective of screening. In this survey, we could not ascertain whether screening contributed to the prevention of salt wasting, nor were we able to distinguish salt-wasting from non-salt-wasting types of CAH. According to the current Endocrine Society CAH Guidelines [2], all infants detected with CAH in the newborn period are to be treated with a combination of glucocorticoids, mineralocorticoid, and sodium chloride supplements. Thus, most pediatric endocrinologists no longer wait until evidence of salt wasting is observed before starting the patient on these standard treatments.

Alternative diagnostic methods and screening protocols have previously been evaluated. Suggestions for quality improvement include gestational age-based cut-off points [5,20] or a combination of birthweight and gestational age [21]. Implementation of a second screen at two weeks of life improves the positive predictive value as noted but adds cost and may be impractical or logistically difficult in some regions. A second tier test, specifically liquid chromatography-tandem mass spectrometry (LC-MS/MS), has been cited as a useful adjunctive test [22,23] that may be adapted to measure an extended steroid profile [24], although it is not currently suitable for high-volume screening at the first tier due to time constraints. At present, genotyping CYP21A2 is not utilized in most US programs for either primary or secondary screening due to the complexity of test interpretation, time constraints, and costs [2], although this may change in the future.

Quality improvement efforts should be directed at enhanced communication between clinicians and state laboratories and transparent reporting of states’ efforts on a national platform. This would include not only the creation of common disease screening panels but also the standardization of protocols, with the aim of eventually achieving a uniform set of best practices [25,26]. With respect to the negative predictive value of screening, a centralized registry of CAH cases diagnosed after the newborn period would aid in capturing false negative tests from the newborn period. Colorado published data citing a 28% false negative rate in the primary screen once a second screen was implemented [27]. New York reported only three false negative cases from their single screen over a span of 7 years, but this is highly dependent on complete follow-up and physician reporting [19]. Surprisingly, Minnesota found a 32% false negative rate, even in the context of two screens [28]. The causes of such disparities are unclear and should be investigated. Finally, there are negative psychological implications to conveying inaccurate screening results to parents, and improving these protocols could have important mental health ramifications. Nationwide (NewSteps [29] and the Disorders/Differences of
Sex Development-Translational Research Network (DSD-TRN). [30,31]) and international (Clinical and Laboratory Standards Institute [32]) collaboratives offer opportunities to explore and compare screening outcomes and develop best practices.

Supplementary Materials: Supplementary materials can be found at http://www.mdpi.com/2409-515X/6/2/37/s1.


Funding: This research was funded by NIH-NICHD 1R01 HD093450, DSD-Translational Research Network.

Acknowledgments: We thank state newborn screening laboratory personnel for their cooperation.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>CAH</td>
<td>Congenital adrenal hyperplasia</td>
</tr>
<tr>
<td>CAH-21</td>
<td>CAH due to steroid 21-hydroxylase deficiency</td>
</tr>
<tr>
<td>17OHP</td>
<td>17-hydroxyprogesterone</td>
</tr>
<tr>
<td>DSD-TRN</td>
<td>Disorders/Differences of Sex Development Translational Research Network</td>
</tr>
<tr>
<td>PPV</td>
<td>Positive predictive value</td>
</tr>
<tr>
<td>IQR</td>
<td>Interquartile range</td>
</tr>
<tr>
<td>LC-MS/MS</td>
<td>Liquid chromatography-tandem mass spectrometry</td>
</tr>
</tbody>
</table>

References


17. White, P.C. Optimizing newborn screening for congenital adrenal hyperplasia. *J. Pediatr.* 2013, 163, 10–12. [CrossRef]


© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).