Predicting gestational age improves newborn screening for congenital adrenal hyperplasia: a retrospective cohort study

Danny Jomaa[1,2]^, Steven Hawken[3]^, Sarah Lawrence[4], Pranesh Chakraborty[2], Matthew Henderson[2]

[1]School of Medicine, Faculty of Health Sciences, Queen's University, Kingston, Canada; [2]Division of Metabolics and Newborn Screening, Department of Pediatrics, University of Ottawa, Ottawa, Canada; [3]Clinical Epidemiology Program, Ottawa Hospital Research Institute, Ottawa, Canada; [4]Division of Pediatric Endocrinology and Metabolism, Department of Pediatrics, University of Ottawa, Ottawa, Canada

Contributions: (I) Conception and design: M Henderson, P Chakraborty, D Jomaa; (II) Administrative support: M Henderson, P Chakraborty, S Lawrence; (III) Provision of study materials or patients: S Hawken, M Henderson, P Chakraborty; (IV) Collection and assembly of data: S Hawken, D Jomaa, M Henderson; (V) Data analysis and interpretation: D Jomaa, S Hawken, M Henderson; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

Correspondence to: Matthew Henderson. Newborn Screening Ontario, Children's Hospital of Eastern Ontario, 415 Smyth Road, Ottawa, K1H 8M8, Canada. Email: mhenderson@cheo.on.ca.

Background: Newborn screening for congenital adrenal hyperplasia (CAH) is increasingly performed using a two-tiered approach; 17-hydroxyprogesterone by immunoassay followed by steroid panel by liquid chromatography mass spectrometry. The first tier uses gestational age (GA)-based 17-hydroxyprogesterone screening thresholds. GA is unreported in approximately 5% of births and, in these cases, birth weight (BW)-based screening thresholds are used. However, BW based thresholds have a lower specificity, resulting in more first tier false positives. By combining newborn demographics and the screening analytes measured in the newborn blood spot screen, a predictive model can be used to estimate GA. In this study, GA was predicted in newborns with an unreported GA. Newborns underwent subsequent GA-based screening to determine whether this method results in a higher positive predictive value (PPV) than current BW-based screening methods.

Methods: Screening results were obtained from Newborn Screening Ontario for 702,020 infants that were born in Ontario, Canada between 2011 and 2015. Predicted GA was calculated using a model composed of demographic and screening analyte factors. Newborns with an unreported GA underwent screening using BW and predicted GA, and the PPV for each method was calculated and compared. Descriptive statistics were determined for newborns that screened positive and negative with each algorithm.

Results: The PPV of first-tier GA-based and BW-based screening was 1.30% and 0.82%, respectively. GA was unreported for 3.61% of infants born in the study time period. The PPV of predicted GA-based screening was higher than BW-based screening for these newborns (1.64% *vs.* 1.52%). Predicted GA-based screening also correctly identified the 2 infants with CAH whose GA was unreported. Subsequently, a sequential screening algorithm was used for newborns with an unreported GA. In this algorithm, BW-based screening was conducted first and newborns that screened positive underwent predicted GA-based screening. Compared to BW-based screening, this approach further increased the PPV of first-tier screening for newborns with unreported GA (1.52% *vs.* 1.89%). Sequential screenings also lead to a lower false positive rate and correctly identified true positive cases of CAH.

Conclusions: Reducing the rate of false positive results in first-tier screening will prevent unnecessary second-tier testing and diagnostic referrals for CAH. Newborns with unreported GA are currently screened used BW-based screening—the only available approach. However, this approach lacks specificity and has a high false positive rate. This study is the first to demonstrate an alternative approach to first-tier CAH screening for newborns with an unreported GA that has a higher PPV than the current strategy.

^ORCID: Danny Jomaa, 0000-0001-6604-0740; Steven Hawken, 0000-0002-3341-9022.

Introduction

Congenital adrenal hyperplasia (CAH) encompasses a group of autosomal recessive disorders that disrupt adrenal steroidogenesis (1). CAH is diagnosed in 1 in 15,000-20,000 newborns and the majority of cases (>90%) are caused by mutations to the *CYP21A2* gene encoding 21-hydroxylase (21-OH) (2). 21-OH is required to convert 17-hydroxyprogesterone (17-OHP) to precursors of aldosterone and cortisol; impaired function of this enzyme causes a buildup of 17-OHP and its diversion to androgen synthesis. In Ontario, Canada, almost all children undergo newborn screening for inborn diseases to intervene in the disease course early and prevent long-term consequences. A blood sample is obtained from the heel of newborns 24–72 hours after birth and analyte levels are measured to detect underlying metabolic and endocrine disorders (3). These results are combined with key demographic information from the newborn, including gestational age (GA) and birth weight (BW), to screen for over 25 disorders. The primary target of CAH screening is to identify newborns with the classic salt-wasting and simple virilizing forms of the disease. The salt-wasting phenotype is a result of mutations that completely inactivate *CYP21A2*, leading to severe aldosterone deficiency. Newborns typically present with a life-threatening adrenal crisis in the first weeks of life, making early detection of this form critical (4). In simple virilizing CAH, patients retain 1–2% of 21-OH activity, allowing enough aldosterone production to prevent a salt-wasting crisis. However, excess androgen synthesis and exposure continues until the disease is detected (5).

At Newborn Screening Ontario, CAH screening follows a two-tier approach. In first-tier screening, 17-OHP levels are compared to a threshold value. Samples that exceed this threshold undergo second-tier screening via a steroid panel by liquid chromatography mass spectrometry, and the final screen positive population is referred for diagnostic evaluation. First and second tier screening results are available within 48 hours of sample receipt. The positive

predictive value (PPV) of first-tier screening is 5–6% and reduces the number of samples that require further screening. The PPV of second-tier screening is subsequently 6–7%. Despite a high false positive rate, in 2019–2020 the screening program had 0 false negative results resulting in a negative predictive value (NPV) of 100%.

GA-based 17-OHP thresholds have been shown to have a higher specificity than BW-based 17-OHP thresholds and are the preferred approach for newborn screening for CAH (6,7). However, GA is unreported for approximately 5% of births, often due to poor recall of menstrual histories, late prenatal presentation, and/or inability to access prenatal ultrasound (8,9). For these newborns, no screening approaches have been described apart from BW-based screening. As a result, a surplus of false positive samples undergo resource-intensive second-tier testing and contribute to the final screen positive population that has to undergo diagnostic testing.

Recently, a predictive model for GA was developed using a combination of newborn screening analytes and multiple birth status (10,11). This model was capable of predicting GA with an accuracy of ±2 weeks. Here, we sought to evaluate the performance of first-tier GA-based screening using a predicted gestational age (PGA). Specifically, we hypothesized that, in newborns with unreported GA, GA-based screening using PGA will have a higher PPV and lower false positive rate than BW-based screening. We present the following article in accordance with the STROBE reporting checklist (available at https://dx.doi.org/10.21037/jlpm-21-18).

Methods

Newborn screening for CAH

A blood sample is drawn from the heel of newborns within 72 hours of birth. By measuring the concentrations of over 40 analytes, Newborn Screening Ontario (NSO) screens newborns for 29 rare diseases, including metabolic disorders, endocrine disorders, immune system disorders,

and hemoglobinopathies. First-tier screening for CAH requires measurement of 17-OHP using fluorometric immunoassay (PerkinElmer GSP). Screening results are primarily determined using GA-based 17-OHP thresholds. If GA is unreported, BW-based 17-OHP thresholds are used instead (Table S1). Samples that receive a positive screening result undergo second-tier testing involving a steroid panel of 17-OHP, cortisol, and androstenedione. Newborns whose second-tier screening result is positive are referred to a pediatric endocrinologist for a diagnostic evaluation and the outcome of these investigations is reported to NSO. Pediatricians in Ontario are requested to inform NSO when CAH is diagnosed and was not previously indicated by a positive newborn screening result. This is recorded as a false negative screening result.

Study design

This study used a retrospective cohort design. The study cohort was screened using BW-based screening and, for newborns with a reported GA, GA-based screening. Newborns with an unreported GA were screened using BW-based screening and GA-based screening using PGA. Subsequently, the cohort of newborns with an unreported GA was screened using a sequential screening approach: newborns with an unreported GA were screened using BW-based screening and only newborns with a positive screening result underwent GA-based screening using PGA. This study was determined to be a quality assurance project by the Children's Hospital of Eastern Ontario Research Ethics Board (REB) and therefore did not require REB review as outlined in the Tri-Council Policy Statement 2. Individual consent for this retrospective analysis was waived. The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013).

Study population

Newborn screening data was obtained from NSO for a total of 702,020 children born in Ontario between January 1st, 2011 and December 31st, 2015 (99.17% of all infants born in that time period). Data was filtered to include only those blood samples obtained before 14 days of life for the purposes of newborn screening.

PGA

Demographic and biochemical data required to calculate PGA were obtained from NSO. PGA was calculated using a linear regression model based on 40 newborn screening analytes, including acylcarnitines, amino acids, hemoglobins, and TSH, and multiple birth status, as described previously (10). 17-OHP concentration and infant sex were included in the original model development but were removed for the purpose of this study.

Statistical analysis

Regression analyses were performed using simple linear modelling of reported GA versus PGA. Statistical significance between means was determined using unpaired *t*-tests. All analyses were conducted using R v3.6.0 (12).

Results

Population characteristics

The GA and BW for all newborns are presented in *Table 1* and organized by the GA- and BW-based 17-OHP thresholds used in first-tier screening (Table S1). GA was unreported for 25,336 newborns (3.61%). Low BW (<2,500 g) and preterm (<37 weeks) newborns composed 6.46% and 7.93% of the population, respectively.

In the study's time period, 34 infants were diagnosed with CAH. Thirty-two (94.10%) had the more severe salt-wasting form and the remaining 2 (5.90%) had the simple virilizing form. The incidence of CAH in the population under investigation was 1 in 20,648 live births. Fifteen newborns (44.10%) were assigned male, 9 (26.50%) were assigned female, and 10 (29.40%) did not have an assigned sex. GA was reported for 32/34 affected infants. The mean GA was 38.23 weeks [standard deviation (SD) =3.30 weeks], the mean BW was 3,319.5 g (SD =777.9 g), and the mean 17-OHP concentration was 364.15 nmol/L (SD =186.9 nmol/L). Three infants (8.80%) were low BW and 7 (20.60%) were born preterm.

BW-based and GA-based screening

The performance of first-tier screening was evaluated using the BW and GA algorithms. CAH screen positive newborns had a significantly lower mean BW and mean GA than screen negative newborns (*Figure 1*). Newborns with low BW were over-represented in the screen positive population compared to the overall study population (34.05% *vs*. 6.46%; P<0.01). BW-based screening

Table 1 Demographic information for all newborns included in the study population

| Variable	Gestational age		
	All newborns, n (%), N=702,020	Reported, n (%), N=676,684	Unreported, n (%), N=25,336
Gestational age (weeks)			
≥36^{+1}	639,192 (91.05)	639,192 (94.46)	–
35^{+1}–36^{+0}	13,898 (1.98)	13,898 (2.05)	–
33^{+1}–35^{+0}	12,781 (1.82)	12,781 (1.89)	–
≤33^{+0}	10,813 (1.54)	10,813 (1.60)	–
Mean (SD)	38.99 (2.3)	38.99 (2.3)	–
Birth weight (g)			
≥2,500	656,657 (93.54)	632,622 (93.49)	24,035 (94.87)
1,500–2,499	38,989 (5.55)	37,858 (4.49)	1,131 (4.46)
1,000–1,499	4,089 (0.58)	3,973 (0.59)	116 (0.46)
<1,000	2,285 (0.33)	2,231 (0.33)	54 (0.21)
Mean (SD)	3,342.03 (575.33)	3,340.53 (575.95)	3,382 (556.91)

SD, standard deviation.
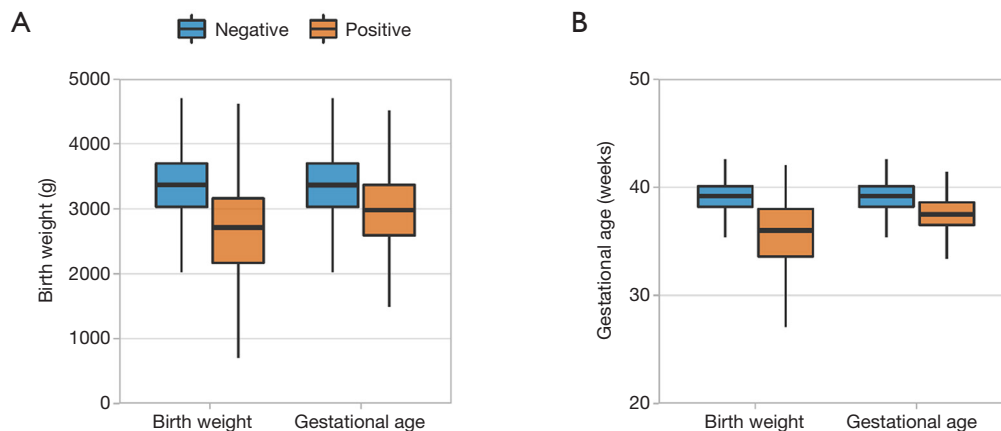
Figure 1 Characteristics of newborns who receive negative and positive screening results. (A & B) In both BW- and GA-based screening, screen positive newborns have a lower mean BW and lower mean GA than screen negative newborns. BW, birth weight; GA, gestational age.

identified 4,132 newborns as screen positive (*Table 2*). All infants with CAH received a positive screening result. The PPV of BW-based screening was 0.82% and its sensitivity and specificity were 100% and 99.42%, respectively.

For newborns with reported GA, screening was performed according to GA-based 17-OHP thresholds (*Table 2, Table S1*). 2,456 newborns received a positive screening result, including all true positive cases. First tier GA-based

screening had a PPV of 1.30%, indicating a lower false positive rate than BW-based screening, and had a sensitivity and specificity of 100% and 99.64%, respectively. Compared to the overall population, preterm newborns were over-represented in the group that screened positive by GA-based screening (7.64% *vs.* 29.68%; P<0.01). The marked increase in PPV of the GA algorithm suggests that screening newborns with an unreported GA using PGA and the GA

Table 2 Results of BW- and GA-based screening

Variable	Algorithm	
	BW, n (%)	GA, n (%)
Total	702,020 (100)	676,684 (100)
Screen positive	4,132 (0.59)	2,456 (0.36)
True positive[†]	34 (0.82)	32 (1.30)
Salt-wasting	32 (94.12)	30 (93.75)
Simple virilizing	2 (5.88)	2 (6.25)
False positive	4,098 (99.18)	2,424 (98.70)
Screen negative	697,888 (99.41)	674,228 (99.64)

[†]32/34 newborns with CAH screened positive by GA-based screening because 2 had unreported GA. BW, birth weight; GA, gestational age; CAH, congenital adrenal hyperplasia.

Table 3 Screening results for newborns with unreported GA using BW-based screening and GA-based screening with PGA

Variable	Algorithm	
	BW, n (%)	PGA, n (%)
Total	25,336 (100)	25,336 (100)
Screen positive	132 (0.52)	122 (0.47)
True positive	2 (1.52)	2 (1.64)
Salt-wasting	2 (100)	2 (100)
Simple virilizing	0 (0)	0 (0)
False positive	130 (98.48)	120 (98.36)
Screen negative	25,204 (99.48)	25,214 (99.52)

BW, birth weight; GA, gestational age; PGA, predicted gestational age.

algorithm may similarly improve screening.

Incorporating a PGA

Using PGA in place of reported GA for all newborns results in a lower PPV, thus reported GA is still the preferable parameter to use when available.

In the study population, GA was unreported for 25,336 newborns (3.61%). Using BW-based screening, 132 newborns screened positive, which included the 2 true positive cases with an unreported GA (PPV: 1.52%; *Table 3*). When PGA-based screening was performed in this sub-population, the PPV increased to 1.64% and continued to identify the 2 true positive cases. GA was reported for 32/34 infants with CAH and, as such, they were not included in this sub-population analysis.

However, when GA-based screening was performed using their PGA, all 32 infants screened positive.

Sequential screening using PGA

While PGA-based screening increased the study's PPV, it introduced a subset of false positive results that had not screened positive by BW-based screening (data not shown). These newborns all had BW <2,500 g. To reduce the false positives introduced by application of PGA, a sequential screening approach was examined. Newborns with unreported GA first underwent BW-based screening and 132 newborns (0.52%) screened positive (*Figure 2*). This subset then underwent PGA-based screening. Because BW-based screening has 100% sensitivity, this approach was expected to identify all true positive cases while avoiding
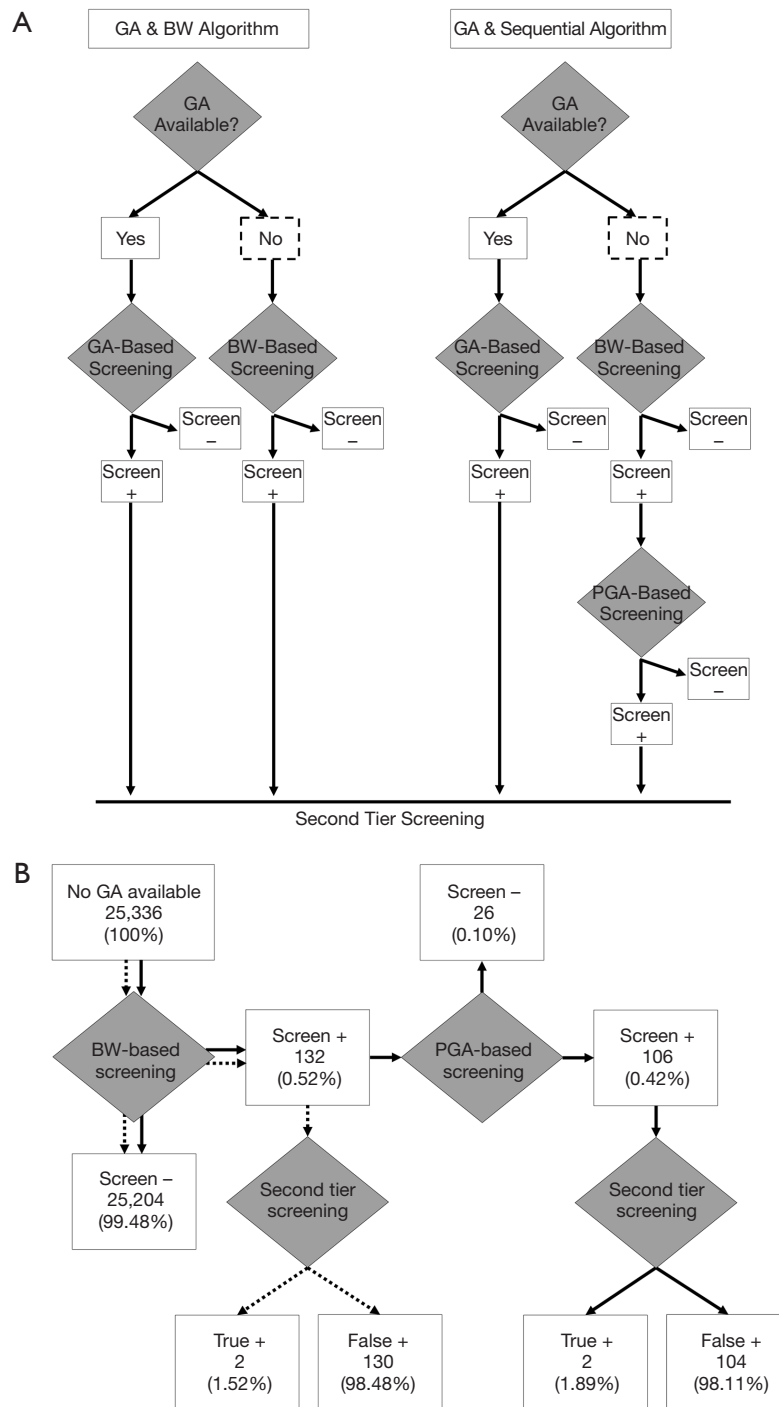
Figure 2 Flowchart of screening algorithms used. Grey diamonds represent screening decision nodes. White boxes represent screening results. (A) In the GA & BW Algorithm and the GA & Sequential Algorithm, newborns with a reported GA undergo GA-based screening. In the GA & BW Algorithm, newborns *without* a reported GA undergo BW-based screening. In the GA & Sequential Algorithm, newborns *without* a reported GA undergo BW-based screening and newborns that screen positive subsequently undergo PGA-based screening. Boxes with a dashed outline are further depicted in 2B. (B) Newborns without a reported GA. Dashed lines demonstrate the process of BW-based screening alone. Solid lines demonstrate the process of sequential screening; newborns that screen positive by BW-based screening undergo additional PGA-based screening. BW, birth weight; GA, gestational age; PGA, predicted gestational age.

Table 4 Screening results for newborns with unreported GA using sequential screening. BW-based screening identified newborns that screen positive and PGA-based screening was performed on this population

Variable	Algorithm	
	BW, n (%)	Sequential (BW & PGA), n (%)
Total	25,336 (100)	25,336 (100)
Screen positive	132 (0.52)	106 (0.42)
True positive	2 (1.52)	2 (1.89)
Salt-wasting	2 (100)	2 (100)
Simple virilizing	0 (0)	0 (0)
False positive	130 (98.48)	104 (98.11)
Screen negative	25,204 (99.48)	25,230 (99.58)

BW, birth weight; GA, gestational age; PGA, predicted gestational age.

the false positives introduced by PGA-based screening alone. Twenty-six newborns-one-fifth of the initial screen positive population-subsequently screened negative, thus reducing the screen positive population to 0.42% (*Figure 2*). Compared to BW-based screening and PGA-based screening performed in isolation, sequential screening had a higher PPV (1.52% *vs.* 1.64% *vs.* 1.89%, respectively) and continued to correctly identify the true positive cases (*Figure 2, Table 4*). Importantly, the 32 infants with CAH and reported GA also screened positive when this sequential screening approach was performed using their PGA.

Furthermore, of the 132 newborns that screened positive by BW-based screening, 25% had BW <2,500 g-approximately four-fold higher than the proportion of newborns with low BW in the overall study population (6.46%). After PGA-based screening, newborns with low BW represented only 16.04% of the screen positive population and the mean BW of newborns who screened positive and negative was 3,015.32 and 2,192.00 g, respectively (P<0.01). Thus, although newborns with low BW continued to be over-represented in the final screen positive population, the proportion was reduced by sequential screening. Additionally, after sequential screening, the mean PGA of the 26 newborns who screened negative was 33.53 weeks. The mean PGA of the final screen positive population was 37.73 weeks. Thus, in the final screen positive population, most newborns had PGA ≥37 weeks and BW ≥2,500 g.

To confirm if PGA-based screening is more specific in preterm newborns than full-term newborns, BW-based screening was performed on the 675,927 newborns with a reported GA. Of the newborns that screened

positive, 60.40% were preterm. After applying PGA-based screening, 54.24% of screen positive preterm newborns and 4.50% of screen positive full-term newborns became screen negative.

Whole population screening

The results of CAH screening for all newborns, with and without a reported GA, are shown in *Table 5*. Newborns with a reported GA were screened by GA-based screening; otherwise, BW-based screening was performed. With this approach, 2,588 newborns (0.37%) screened positive, including the infants with CAH. The population was screened again, but newborns with unreported GA were screened using the sequential algorithm. A total of 2,562 newborns (0.36%) screened positive, again with correct identification of all true positive cases. The elimination of 26 false positives resulted in a small increase in the test's PPV, from 1.31% to 1.33%. This change was due to a 24% increase in the PPV of screening for newborns with unreported GA.

Conclusions

Compared to BW, GA has been shown to more accurately predict 17-OHP levels because of its closer association with adrenal gland development (7). The implementation of GA-based 17-OHP thresholds has increased the PPV of first-tier CAH screening, however ours and other screening programs still report high rates of false positive screening results (13-16). In part, this occurs because 4–5% of newborns have unreported GA and undergo BW-based

Table 5 Screening results for the whole study population. As presented in *Figure 2*, the GA & BW Algorithm performs GA- and BW-based screening for newborns with and without a reported GA, respectively. In the GA & Sequential Algorithm, newborns with a reported GA undergo GA-based screening and newborns without a reported GA undergo BW-based screening and, for positive cases, PGA-based screening

Variable	GA & BW Algorithm, n (%)	GA & Sequential Algorithm, n (%)
Total	702,020 (100)	702,020 (100)
Screen positive	2,588 (0.37)	2,562 (0.36)
True positive	34 (1.31)	34 (1.33)
Salt-wasting	32 (94.12)	32 (94.12)
Simple virilizing	2 (5.88)	2 (5.88)
False positive	2,554 (98.69)	2,528 (98.67)
Screen negative	699,432 (99.63)	699,458 (99.64)

BW, birth weight; GA, predicted gestational age.

screening with lower specificity. Samples that screen positive undergo second-tier screening which has a higher PPV. However, improving the specificity of first-tier screening would reduce the number of samples sent to second-tier screening and the number of newborns ultimately referred for diagnostic work-up. In addition to GA-based screening, efforts to reduce the false positive rate have included adjusting 17-OHP thresholds to the age of blood sample collection and collecting a second blood sample from screen positive newborns and repeating screening (13,17). These examples have all increased the PPV of CAH screening but each adjustment requires rigorous evaluation to ensure that it does not unintentionally classify newborns with CAH as false negatives or delay diagnosis in newborns with salt-wasting CAH. Thus, further improvements to CAH screening must be incremental to preserve their high sensitivity and to continue to capture all true positive cases. This study presents a strategy for further reducing the false positive rate of first-tier CAH screening while continuing to identify all true positive cases.

This retrospective study evaluated newborn screening data from over 700,000 infants born in Ontario, Canada between 2011 and 2015. In the population described here, newborns who screened positive with BW- and GA-based screening were typically full-term, healthy BW, and had high 17-OHP levels. However, in both screen positive populations, preterm newborns were over-represented compared to the proportion they composed of the overall population, which is consistent with the results of previous studies (18,19). Although the reasons for this are not well understood, preterm and low BW newborns are known to have higher levels of 17-OHP and greater variability in these levels than full-term, healthy BW newborns (6,18). Thus, the specificity of a screening threshold-its ability to differentiate between healthy and pathologically elevated 17-OHP-is reduced in these populations. In addition to prematurity, several other conditions elevate 17-OHP levels and contribute to the overall high rate of false positives, including seizures, sepsis, respiratory conditions, and hydronephrosis (20,21).

Compared to BW-based screening, application of the sequential screening method described here resulted in a 24% improvement in PPV for newborns with unreported GA. In this approach, BW-based screening was performed first and newborns who screened positive were supplied with their PGA and underwent GA-based screening. Importantly, the two newborns with CAH and unreported GA were correctly identified by sequential screening. CAH screening that combines BW and GA has been reported previously, but this has used new thresholds based on both BW and GA, rather than sequential thresholds, as was done here (14,16). In our study, newborns whose positive BW-based screen became negative after PGA-based screening were predominantly pre-term and low BW. This can be attributed to how screening thresholds are structured. Using GA-based screening, newborns born at 36^{+1} weeks or later all have the same 17-OHP thresholds (Table S1); newborns with GA $<36^{+1}$ weeks are stratified into GA ranges with individualized 17-OHP thresholds. As a result, infants born prior to 36^{+1} weeks are more likely to be assessed using a 17-OHP threshold that is GA-specific, compared to infants born at 36^{+1} weeks or later. As described previously, prematurity is one of several factors contributing to the high rate of false positives so a strategy

to correctly identify false positives in this population is indeed necessary. Since most infants who remained screen positive in this study were full-term, future work should focus on developing GA-specific 17-OHP thresholds for infants born after 36^{+1} weeks.

This is the first study to propose a novel approach to CAH screening for newborns with unreported GA who would otherwise undergo BW-based screening. Here, the PPV of screening for this sub-population was increased from 1.52% (130 false positives) with BW-based screening to 1.89% (104 false positives) with sequential screening (*Figure 2*). A limitation of this study was the small number of newborns with CAH and unreported GA who could be included in the sub-population analysis. However, all newborns with CAH received a positive screening result when performing PGA-based screening, suggesting that this limitation is not a barrier to the study conclusions. Furthermore, the increase in PPV that we report reflects 26 newborns who were correctly identified as false positive cases and did not require second-tier or diagnostic testing. Thus, application of a PGA for newborns with unreported GA is one step that can be taken to improve the specificity of first-tier screening and reduce the need for subsequent tests.

Acknowledgments

Footnote

Reporting Checklist: The authors have completed the STROBE reporting checklist. Available at https://dx.doi.org/10.21037/jlpm-21-18

Data Sharing Statement: Available at https://dx.doi.org/10.21037/jlpm-21-18

Peer Review File: Available at https://dx.doi.org/10.21037/jlpm-21-18

Conflicts of Interest: All authors have completed the ICMJE uniform disclosure form (available at https://dx.doi.org/10.21037/jlpm-21-18). The authors have no conflicts of interest to declare.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013). This study was determined to be a quality assurance project by the Children's Hospital of Eastern Ontario Research Ethics Board (REB) and therefore did not require REB review as outlined in the Tri-Council Policy Statement 2. Individual consent for this retrospective analysis was waived.

References

1. El-Maouche D, Arlt W, Merke DP. Congenital adrenal hyperplasia. Lancet 2017;390:2194-210.
2. White PC. Neonatal screening for congenital adrenal hyperplasia. Nat Rev Endocrinol 2009;5:490-8.
3. Therrell BL, Adams J. Newborn screening in North America. J Inherit Metab Dis 2007;30:447-65.
4. Speiser PW, Azziz R, Baskin LS, et al. Congenital adrenal hyperplasia due to steroid 21-hydroxylase deficiency: an Endocrine Society clinical practice guideline. J Clin Endocrinol Metab 2010;95:4133-60.
5. Speiser PW, White PC. Congenital Adrenal Hyperplasia. N Engl J Med 2003;349:776-88.
6. Allen DB, Hoffman GL, Fitzpatrick P, et al. Improved precision of newborn screening for congenital adrenal hyperplasia using weight-adjusted criteria for 17-hydroxyprogesterone levels. J Pediatr 1997;130:128-33.
7. van der Kamp HJ, Oudshoorn CG, Elvers BH, et al. Cutoff levels of 17-alpha-hydroxyprogesterone in neonatal screening for congenital adrenal hyperplasia should be based on gestational age rather than on birth weight. J Clin Endocrinol Metab 2005;90:3904-7.
8. Martin JA. United States vital statistics and the measurement of gestational age. Paediatr Perinat Epidemiol 2007;21 Suppl 2:13-21.
9. Lee AC, Panchal P, Folger L, et al. Diagnostic Accuracy of Neonatal Assessment for Gestational Age Determination:

A Systematic Review. Pediatrics 2017;140:e20171423.

10. Wilson K, Hawken S, Murphy MSQ, et al. Postnatal Prediction of Gestational Age Using Newborn Fetal Hemoglobin Levels. EBioMedicine 2017;15:203-9.

11. Wilson K, Hawken S, Potter BK, et al. Accurate prediction of gestational age using newborn screening analyte data. Am J Obstet Gynecol 2016;214:513e1-e9.

12. R Core Team. R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing, 2018.

13. van der Linde AAA, Schonbeck Y, van der Kamp HJ, et al. Evaluation of the Dutch neonatal screening for congenital adrenal hyperplasia. Arch Dis Child 2019;104:653-7.

14. Pode-Shakked N, Blau A, Pode-Shakked B, et al. Combined Gestational Age- and Birth Weight-Adjusted Cutoffs for Newborn Screening of Congenital Adrenal Hyperplasia. J Clin Endocrinol Metab 2019;104:3172-80.

15. Gidlof S, Wedell A, Guthenberg C, et al. Nationwide neonatal screening for congenital adrenal hyperplasia in sweden: a 26-year longitudinal prospective population-based study. JAMA Pediatr 2014;168:567-74.

16. Guran T, Tezel B, Gurbuz F, et al. Neonatal screening for congenital adrenal hyperplasia in Turkey: a pilot study with 38,985 infants. J Clin Res Pediatr Endocrinol 2019;11:13-23.

17. Olgemoller B, Roscher AA, Liebl B, et al. Screening for congenital adrenal hyperplasia: adjustment of 17-hydroxyprogesterone cut-off values to both age and birth weight markedly improves the predictive value. J Clin Endocrinol Metab 2003;88:5790-4.

18. Nordenstrom A, Wedell A, Hagenfeldt L, et al. Neonatal screening for congenital adrenal hyperplasia: 17-hydroxyprogesterone levels and CYP21 genotypes in preterm infants. Pediatrics 2001;108:E68.

19. Torresani T, Gruters A, Scherz R, et al. Improving the efficacy of newborn screening for congenital adrenal hyperplasia by adjusting the cut-off level of 17-hydroxyprogesterone to gestational age. Screen 1994;3:77-84.

20. Anandi VS, Shaila B. Evaluation of factors associated with elevated newborn 17-hydroxyprogesterone levels. J Pediatr Endocrinol Metab 2017;30:677-81.

21. Pauwels G, Allegaert K, Regal L, et al. Risk factors for elevated levels of 17-hydroxyprogesterone during neonatal intensive care unit admission. Acta Clin Belg 2012;67:88-93.

Table S1 GA- and BW-based algorithms used in first-tier screening for CAH. Description of data: newborns with BW ≥2,500 g or post-natal age ≥21 days screen positive if 17-OHP ≥38 nmol/L. Newborns not meeting these criteria undergo GA-based screening. If GA is unavailable, BW-based screening is performed

Variable	Screen positive If 17-OHP
Criteria	
BW ≥ 2,500g	≥38 nmol/L
Age of sample collection ≥21 days	≥38 nmol/L
Otherwise GA Logic	
≥36^{+1} weeks	≥38 nmol/L
≥35^{+1} weeks & ≤36^{+0} weeks	≥55 nmol/L
≥33^{+1} weeks & ≤35^{+0} weeks	≥91 nmol/L
≤33^{+0} weeks	≥181 nmol/L
GA NULL	Use BW Logic
BW Logic	
≥1,500 g & ≤2,499 g	≥52 nmol/L
≥1,000 g & ≤1,499 g	≥78 nmol/L
<1,000 g	≥137 nmol/L
BW NULL	≥38 nmol/L