Single newborn screen or routine second screening for primary congenital hypothyroidism

Stuart K. Shapira[a,*], **Cynthia F. Hinton**[a], **Patrice K. Held**[b], **Elizabeth Jones**[c], **W. Harry Hannon**[d], and **Jelili Ojodu**[c]

[a]National Center on Birth Defects and Developmental Disabilities, Centers for Disease Control and Prevention, Atlanta, GA, USA

[b]Wisconsin State Laboratory of Hygiene, University of Wisconsin, Madison, WI, USA

[c]Newborn Screening and Genetics Program, Association of Public Health Laboratories, Silver Spring, MD, USA

[d]Division of Laboratory Sciences, Centers for Disease Control and Prevention (retired), Atlanta, GA, USA

Abstract

Routine second screening of most newborns at 8–14 days of life for a panel of newborn conditions occurs in 12 U.S. states, while newborns in the other states typically undergo only a single routine newborn screen. The study objective was to evaluate screening consequences for primary congenital hypothyroidism (CH) in one- and two-screen states according to laboratory practices and medical or biochemical characteristics of screen-positive cases. Individual-level medical and biochemical data were retrospectively collected and analyzed for 2,251 primary CH cases in one-screen (CA, WI) and two-screen (AL, DE, MD, OR, TX) states. Aggregate data were collected and analyzed for medical and biochemical characteristics of all screened newborns in the states. Among the states evaluated in this study, the detection rate of primary CH was higher in the one-screen states. In the two-screen states, 11.5% of cases were detected on the second screen. In multivariate analyses, only race/ethnicity was a significant predictor of cases identified on the first versus second screen, which likely reflects a physiologic difference in primary CH presentation. Newborn screening programs must heed the potential for newborns with CH not being detected by a single screen, particularly newborns of certain races/ethnicities. If the two-screen states converted to a single screen using their current algorithms, newborns currently identified on the routine second screen would presumably not be detected, resulting in probable delayed diagnosis and treatment. However, based on the one-screen state experiences, with appropriate modifications in screening method and algorithm, the two-screen states might convert to single screen operation for CH without loss in performance.

*Corresponding author at: NCBDDD, CDC, 4770 Buford Hwy., Mailstop E-87, Atlanta, GA 30341, USA., sshapira@cdc.gov (S.K. Shapira).

1. Introduction

Newborn screening (NBS) has been an effective and successful public health program for the early detection and treatment of disorders that can cause intellectual disability, morbidity, and in some cases, mortality. When NBS began in the 1960s, it was typical that the heel stick blood specimens were obtained at 48–96 hours following birth. This practice lowered the chance of cases screening falsely negative from inadequate nutritional intake to detect metabolic abnormalities or from a physiologic delay in elevation of thyroid stimulating hormone (TSH) concentrations [1]. Changes in the healthcare system have subsequently resulted in early hospital discharge of most mothers and newborns before 48 hours of life. Although the majority of blood specimens are collected in the 24–48 hour period, early discharge has significantly impacted NBS, such that many blood specimens are now collected before 24 hours of life [2, 3]. When only a single newborn screen is collected, there is an increased chance of missing cases (false-negatives), resulting in delayed diagnosis [4, 5]. To minimize the chance of clinically significant disorders being missed by a single screen, 9 states (AZ, CO, DE, NV, NM, OR, TX, UT, WY) have mandated that a second blood specimen be routinely collected on all newborns, preferably at 8–14 days of age, regardless of the results of the first routine newborn screen. An additional 3 states (AL, MD, WA) recommend a second screen, which is obtained at 8–14 days from birth on 85% of newborns in the state, although routine second screening was initiated in Maryland in 1976, prior to the impact of early discharges [5]. Thus, routine second screening in these 12 states occurs for approximately 22.5% of all U.S. newborns. Newborns in the other states typically undergo only a single newborn screen, although some newborns might receive a second screen (targeted second screening) according to state-based screening algorithms for preterm newborns or early specimen collection.

The evidence that a routine second screen detects clinically significant cases that would otherwise be missed by a single screen comes primarily from published reports by states performing routine second screens for primary congenital hypothyroidism (CH) and congenital adrenal hyperplasia (CAH) [6–11]. In spite of these reports, there has been no clear consensus regarding the utility of routine second screening, or whether it is the most appropriate public health approach to detect cases that might otherwise be missed (delayed diagnosis). Therefore, to evaluate the justification for the routine second screen, the Health and Human Services Secretary's Advisory Committee on Heritable Disorders in Newborns and Children (SACHDNC) endorsed that the study described here be undertaken to investigate the effects of the routine second screen for primary CH [12].

2. Methods

2.1. Study design and participants

Under guidance from the SACHDNC, a 5-year retrospective study was developed, with representation on the planning workgroup from 14 state NBS laboratories, 9 endocrinologists representing states performing only a single screen or a routine second screen, and representatives from the Health Resources Services Administration, the National Newborn Screening and Genetics Resource Center, the Centers for Disease Control and Prevention, the Food and Drug Administration, the Association of Public Health Laboratories (APHL), the SACHDNC, Pediatrix Screening, and the Congenital Adrenal Hyperplasia Research Education & Support Foundation. The NBS programs that were invited to participate in the study included those in the 9 states that mandate collection of a routine second screen on all newborns, the 3 states that recommend a second screen, and 3 states that collect only a single screen. NBS programs in 7 states agreed to and followed through with participation.

Participating programs were defined as "one-screen states" (CA, WI) or "two-screen states" (AL, DE, MD, OR, TX). All participating states received Institutional Review Board approval for the study. Programs submitted data on all confirmed cases of CH identified by the state's NBS program during a 3–5 year period (Fig. 1), which included primary CH, secondary CH, hypothyroxinemia of prematurity, thyroid binding globulin deficiency, T4 resistance, transient hypothyroidism, and uncertain hypothyroidism. Data analyses in this report were restricted to cases of primary CH.

2.2. NBS methodologies and algorithms for CH

The NBS programs utilized different methodologies and algorithms for identifying potential cases of CH (Fig. 2). Three participating programs (CA, DE, WI) quantified TSH only as the analytic marker for CH, and all used a fixed cutoff with some minor variations based on the age of the newborn at the time of specimen collection. Three programs (MD, OR, TX), quantified thyroxine (T4) as the primary analytic marker and reflexed to analyze TSH if the T4 level was below the screening cutoff; 2 of these programs used fixed cutoffs for T4 and TSH while one program used a fixed cutoff for TSH, but for T4 used a floating cutoff that was determined daily based upon a percent from the mean T4 value obtained on the entire population of newborns screened that day. The other participating program (AL) quantified both T4 and TSH on all newborns, using fixed cutoffs for both analytic markers.

2.3. Data elements

NBS programs were asked to submit individual-level de-identified data on each confirmed cases of CH to the APHL, which served as the repository for submitted data. A secure web-based portal developed by APHL was used for entering data elements, including the newborn's demographics (e.g., race/ethnicity, sex); factors that might affect the screening result (e.g., age at specimen collection, birth weight (BW), neonatal intensive care unit (NICU) admission, blood transfusion); laboratory factors (e.g., time from screening to laboratory assay, screening methods, screening algorithm); and clinical characteristics pertaining to case diagnosis (e.g., confirmatory serum test results, thyroid imaging results,

treatment and when initiated, presence of birth defects, Down syndrome, or other medical conditions, mode of delivery, topical iodine, dopamine, or steroid exposures, family history of thyroid disease). Each newborn's race/ethnicity was recorded on each state's NBS blood collection card, and was designated at the time of specimen collection. NBS programs also reported the CH type diagnosed for each case, and whether each case was identified on the first screen, second screen, or was a case not detected by NBS. Due to this being a retrospective study, data from states were missing some data elements. Therefore, only the following variables could be included in analyses: state, race/ethnicity (non-Hispanic white (NHW), non-Hispanic black (NHB), Hispanic, Asian/Pacific Islander (A/PI), Other), sex (male, female), feeding status at the time of first screen (breastfeeding only, formula only, breastfeeding and formula, total parenteral nutrition (TPN), other), BW (<1000 g, 1000–2499 g, 2500–3999 g, 4000 g), NICU admission at time of first screen (no, yes), blood transfusion prior to first screen (no, yes), age of newborn at first screen specimen collection (0 days, 1 day, 2 days, 3 days, 4 days), time from first screen specimen collection to laboratory assay completion (0–3 days, 4–5 days, 6–7 days, 8 days), and initial abnormal screen identifying the potential case (first screen, second screen—for the two-screen states), targeted second screen (for the one-screen states), detected clinically (i.e., no abnormal screening result(s)). A metabolic specialist (S.K.S.) reviewed all submitted data to confirm the initial abnormal screen identifying the potential case and the diagnosis as primary CH.

In addition to the above data on confirmed primary CH cases, NBS programs submitted aggregate data on all newborns screened during the time periods shown in Fig. 1 for the following variables: race/ethnicity, sex, feeding status at the time of first screen, BW, NICU admission at time of first screen, blood transfusion prior to first screen, age of newborn at first screen specimen collection, and time from first screen specimen collection to laboratory assay completion. Additionally, the two-screen states submitted aggregate data for race/ethnicity for all newborns receiving a routine second screen.

2.4. Data analysis

The total number of newborns screened and the number screened with each characteristic were used to calculate estimated detection rates of cases in the one- and two-screen states. The numerator for each detection rate was the number of identified primary CH cases with the characteristic, and the denominator was the number of all screened newborns with the same characteristic. Detection rates overall and for each characteristic were compared between the one-screen and two-screen states by Z-test for 2 proportions.

The variables described in section *2.3. Data elements* were evaluated by univariate analyses for inclusion in logistic regression models as predictive of a case of primary CH being identified in the two-screen states on the initial abnormal screen (first screen versus the second screen). Predictive modeling was then conducted with the significant variables using multivariate logistic regression and assessing which factors remained significantly associated with the initial abnormal screen after adjusting for other covariates.

For confirmed cases of primary CH in both the one- and two-screen states, TSH screening concentrations were log-transformed and compared between race/ethnicity categories using least squares means, adjusting for state, sex, BW (as a continuous variable), blood

transfusion prior to first screen, and age of newborn at first screen specimen collection. Analyses for cases identified on first screens or second screens were conducted separately. Adjusted means for the transformed serum TSH concentrations for each race/ethnicity were then untransformed for data presentation. Cases from Maryland were excluded from the analyses because the program reported BW as a categorical variable, in contrast to the other states that reported the individual BW values. Logistic regression and least squares means analyses were conducted using SAS 9.3 (SAS Institute, Inc., Cary, NC).

3. Results

3.1. One-screen versus two-screen states

Data for a total of 1056 cases of primary CH were submitted by the one-screen states and 1195 cases by the two-screen states (Table 1). Cases were identified by an abnormal newborn screen result on either the first screen, routine second screen (two-screen states), or targeted second screen (one-screen states). Cases were classified as unknown if an initial specimen submitted for NBS was deemed unsatisfactory by the laboratory and the case was detected on a subsequent screening specimen, but because of the initial unsatisfactory specimen, it is unknown whether the case would have been detected on the first screen. A small proportion of cases in both the one- and two-screen states were not identified by NBS, but were identified clinically and were subsequently reported to the NBS programs.

In the one-screen states, 1027 cases of primary CH (97.2%) were identified on the first screen, 19 cases (1.8%) were detected on a targeted second screen, and the remaining 10 cases were either unknown or not detected by NBS (Table 1). Based on the total of 2,010,531 newborns screened on the first screen in one-screen states, the detection rate of cases detected on the first screen was 1 in 1,958 and the overall detection rate for all cases of primary CH in one-screen states was 1 in 1,904. In the two-screen states, 1041 cases of primary CH (87.1%) were identified on the first screen, 137 cases (11.5%) were detected on the second screen, and the remaining 17 cases were either unknown or not detected by NBS (Table 1). Based on the total of 2,629,627 newborns screened on the first screen in the two-screen states, the detection rate of cases detected on the first screen was 1 in 2,526, and the detection rate for those detected initially on the routine second screen was 1 in 19,194. The overall detection rate for primary CH in the two-screen states was 1 in 2,201.

The detection rate for primary CH overall was significantly higher in the one-screen than in the two-screen states evaluated in this study (p=0.001) (Table 2). This difference was primarily accounted for by a higher detection rate in the one-screen states among NHW newborns (p=0.038), female newborns (p<0.001), and among normal BW newborns (p=0.009); there was also a higher detection rate among Hispanic newborns, but it was borderline non-significant (p=0.056). Additionally, the primary CH detection rate was higher in the one-screen states for newborns whose NBS specimen was collected at <24 hours of life (p=0.010) or at 24–47 hours (p<0.001).

Further investigation of this difference in the detection rate of primary CH related to age of specimen collection showed that the mean age of first specimen collection for primary CH cases detected on the first screen was significantly lower in the one-screen than in the two-

screen states (Table 3); this difference was statistically significant for each racial/ethnic group and overall.

No difference in detection rate was found for NHB or male newborns, newborns in the NICU when the first screen was collected, and newborns that were very low BW (<1000 g) or low BW (1000–2499 g) (Table 2). Several characteristics had a lower primary CH detection rate among newborns in the one-screen states: A/PI race (p=0.001), blood transfusion prior to the first NBS specimen collection (p=0.023), and NBS specimen collected at 96 hours of life (p<0.001) (Table 2).

The detection rate of primary CH differed by racial/ethnic group (Table 2). The highest detection rate of primary CH was found among A/PI and Hispanic newborns, followed by newborns in the Other race/ethnicity category and NHW newborns. NHB newborns had the lowest detection rate of primary CH. The race/ethnicity makeup of the screened populations differed between the one- and two-screen states. Among all screened newborns, there was a higher proportion of Hispanic (44% vs. 35%) and A/PI (8% vs. 3%) newborns and a lower proportion of NHW (34% vs. 41%) and NHB (6% vs. 14%) newborns in the one-screen states, compared with the two-screen states.

3.2. First screen versus second screen: two-screen states

In unadjusted analyses, numerous characteristics differed between cases detected on the first versus the second screen in the two-screen states (Table 4). First-screen cases (compared to second-screen cases) were less likely to be detected in Maryland and Oregon than in Texas (odds ratio (OR)=0.33; 95% confidence interval (CI)=0.19–0.56 and OR=0.44; 95% CI=0.26–0.75, respectively); less likely to be NHB or A/PI newborns than NHW newborns (OR=0.40; 95% CI=0.22–0.73 and OR=0.21; 95% CI=0.11–0.37, respectively); less likely to be very low BW (<1000 g) than to be normal BW (2500–3999 g) (OR=0.42; 95% CI=0.20–0.88); less likely to have received a blood transfusion before the first newborn screen (OR=0.43; 95% CI=0.19–0.97); more likely to be female than male (OR=1.51; 95% CI=1.05–2.16); and more likely to have had the first NBS specimen collected at 24–47 hours of life than collected at <24 hours (OR=2.06; 95% CI=1.26–3.35).

In multivariate logistic regression only race/ethnicity was significant, with NHB and A/PI newborns less likely than NHW newborns to be identified on the first versus second screen (Table 4). There was no significant difference for Hispanic newborns or newborns in the Other race/ethnicity category, compared with NHW newborns (OR=1.33; 95% CI=0.84–2.13, OR=0.49; 95% CI=0.20–1.18, respectively).

To investigate these race/ethnicity differences further between cases identified on the first versus the second screen in the two-screen states, least square means analyses, stratified by race/ethnicity, were conducted for the TSH screening concentrations of the cases reported by the NBS programs (right side of Fig. 3).

Geometric mean TSH serum concentrations were adjusted for the variables reported in Table 4 that were significant predictors for cases being identified on the first versus the second screen in univariate analyses. The race/ethnicity groups were NHW, NHB, Hispanic,

and Other, which included A/PI, Middle Eastern, Native American, and Other, non-specified. Among the cases in the Other category, 66% of first-screen cases and 85% of second-screen cases were A/PI. For first screen cases, the adjusted geometric mean TSH concentrations were significantly lower for NHB newborns, compared with NHW newborns (p<0.001) and Hispanic newborns (p<0.001); there was no significant difference for NHB compared with Other (p=0.062). For the second screen, the adjusted geometric mean TSH concentrations were still significantly lower for NHB newborns compared with NHW newborns (p=0.002) and Hispanic newborns (p<0.001), and there was no significant difference for NHB compared with Other (p=0.080). In order to confirm that the pattern of geometric means by race/ethnicity was not unique to the two-screen states, a similar least square means analysis was performed on cases identified on the first screen in one-screen states (left side of Fig. 3). For the cases in the Other category, 77% were A/PI. Similar to the two-screen states, the adjusted geometric mean TSH concentrations in the one-screen states were significantly lower for NHB newborns, compared with NHW newborns (p=0.006) and Hispanic newborns (p<0.001), and in this case also significantly lower when compared with Other newborns (p=0.002).

4. Discussion

Among the two-screen states evaluated in this study, 11.5% of primary CH cases were detected on the routine second screen, which is comparable to rates of second-screen identified cases in previous single state studies: 10.4% in the Northwest Regional Screening Program [9], 5.1% in Texas [13], 7.7% in Washington state [7], and 18.5% in Colorado [10]. All of the cases detected on the routine second screen in the current study appear to have been clinically significant since they were treated with thyroid replacement therapy. If the two-screen states performed only a single screen according to their current screening algorithms, these cases would not have been detected by the NBS programs since the first screen results for these cases were normal.

The only statistically significant predictor in multivariate analysis for identifying cases on the first screen compared to the second screen in the two-screen states was race/ethnicity. Compared to cases identified on the second screen, the first screen cases were less likely to be NHB or A/PI, than to be NHW cases. These race/ethnicity differences in detection on first versus second screen appear to be due to differences in the physiologic responses of the hypothalamic-pituitary axis to the hypothyroid state. Among cases in the two-screen states, NHB newborns, as well as newborns in the Other race/ethnicity category (primarily A/PI), had significantly lower serum TSH concentrations on NBS assays compared with NHW and Hispanic newborns. These differences were present for both first- and second-screen identified cases. For cases in the one-screen states, NHB newborns also had significantly lower serum TSH concentrations on NBS assays compared with newborns in all other racial/ethnic groups. These findings suggest that the TSH rise in response to low T4 in NHB and probably A/PI newborns is perhaps not as early or as elevated as among NHW and Hispanic newborns, which might account for the higher proportion of NHB and A/PI newborns being identified on the routine second screen in the two-screen states. Given these physiologic differences in biochemical presentation of primary CH, it appears that the routine second

screen is an effective way to detect these cases that might otherwise be missed by a single screen.

Among the states evaluated in this study, the overall detection rate of primary CH was significantly higher in the one-screen states compared with the two-screen states. Newborns with certain characteristics had a significantly higher detection rate in the one-screen states: NHW, female, normal BW, or being screened at <48 hours of life. It is unlikely that a single explanation accounts for these differences, although the different detection rates of primary CH for the race/ethnicity groups could be a significant contributing factor to the overall detection rate of primary CH being higher in the one-screen states. Previous studies have reported that compared to NHW newborns, primary CH is more common among Hispanic and A/PI newborns, and less common among NHB newborns [14–19]. The analyses performed here showed similar relationships between detection rates by race/ethnicity (Table 2). The populations screened by the one- and two-screen states differed by race/ethnicity, with higher proportions of the screened populations being Hispanic and A/PI in the one-screen, compared to the two-screen states. Since newborns in these 2 racial/ethnic groups have a higher detection rate of primary CH, compared with NHW and NHB newborns, the higher Hispanic and A/PI makeup of the screened population in the one-screen states likely contributes in part to the higher primary CH detection rate in the one-screen states. However, that cannot be the entire explanation because if race/ethnicity were the only factor predicting the primary CH detection rate, then the detection rate among each racial/ethnic group should be the same in the one- and two-screen states; contrary to this, the observed primary CH detection rate was significantly higher among NHW newborns, borderline higher among Hispanic newborns, and significantly lower among A/PI newborns in the one-screen states compared with the two-screen states. Therefore, other factors are also impacting the primary CH detection rate differences between the one- and two-screen states, Unfortunately it is not possible to examine these factors independently since individual-level data on all screened newborns was not available.

The significantly different primary CH detection rates between the one-and two-screen states could also be related in part to screening methodologies, which are notably different between the states. Both of the one-screen states only assay TSH as the screening analytic marker, while among the two-screen states, DE assays only TSH, AL assays both T4 and TSH, and the other two-screen states assay T4 as the primary analytic marker, with a reflex to TSH if the T4 concentration is below a specific screening cutoff. If the differences in screening methods caused the lower identified detection rate in the two-screen states, this would suggest that even with a routine second screen on the majority of newborns, a large number of cases were not detected in the two-screen states. However, there is no evidence for a large number of cases in these states not being detected by screening. Therefore, although screening methodologies differed between the one- and two-screen states, it is unlikely that screening methods in the two-screen states account in a substantial way for the lower detection rate.

Another potential cause for detection rate differences might be misclassification. For example, if substantially more cases of transient hypothyroidism were misdiagnosed as primary CH in the one-screen states, this might inflate the observed detection rate. Since the

female-to-male ratio for primary CH is approximately 2:1, and for transient hypothyroidism it is on the order of 0.5:1 to 1:1 [20–22], if a significant number of newborns with transient hypothyroidism were misclassified as primary CH, then the observed ratio would fall below 2:1. However, among the cases in the one-screen states, the sex ratio is 2.17:1, indicating that there is unlikely to have been substantial misclassification of this type to inflate the detection rate. Another potential source of misclassification involves the diagnoses assigned to 2 groups of newborns in the two-screen states: 209 newborns with hypothyroxinemia of prematurity and 93 newborns with uncertain hypothyroidism (compared to only 1 and 28 newborns, respectively reported as being in these 2 groups in the one-screen states). The former group consists of preterm newborns treated for hypothyroidism because of low T4, but generally normal TSH concentrations [23–25]. The latter group consists of newborns without a specific confirmatory diagnosis and some may not actually have hypothyroidism. These newborns did not have primary CH according to the data provided by the state laboratories, so they were excluded from the study. However, if enough newborns in these 2 groups actually had primary CH, but were misclassified as hypothyroxinemia of prematurity or uncertain hypothyroidism because of insufficient follow-up testing, that misclassification could have erroneously lowered the observed detection rate for primary CH in the two-screen states. One final source of misclassification might be related to the age of the newborn at the time of the initial specimen collection. Compared to the two-screen states, cases identified on the first screen in the one-screen states had (1) higher geometric mean TSH concentrations (Fig. 3), (2) specimens collected on average at a younger age (1.25 days vs. 2.34 days; Table 3), and (3) a higher overall rate of detected cases among those newborns with specimens collected at <48 hours of life (Table 2). These observations suggest that the time of specimen collection might impact the case detection rate since newborns undergoing earlier collection of specimens, closer to the TSH surge that occurs 30 minutes following birth, would more likely be screen-positive. Although screening closer to birth has the potential to increase the false-positive detection rate, the CH cases reported by each NBS program were cases subsequently confirmed by serum diagnostic testing. Since it is unlikely that the one-screen states have misclassified false-positive cases as actual cases, this potential source of misclassification does not seem related to the higher detected rate in the one-screen states.

Finally, less than 15% of primary CH cases have a known genetic basis [26–30], so the underlying causes for most cases remain unknown. If the observed primary CH detection rates reflect a true difference between the one- and two-screen states, then these rates could mirror the causes of primary CH from yet undetermined genetic factors and environmental exposures that differ between the populations of the one- and two-screen states evaluated in this study.

A limitation of this study is the retrospective data collection; data were incomplete for certain variables since the NBS laboratories and follow-up programs could only report the data that were available. Additionally, results could be biased by the states that contributed the largest number of cases. Furthermore, results and conclusions are not generalizable to all one-screen and two-screen states, but are limited to those that participated in this study. Another limitation is that the diagnosis of primary CH was not necessarily determined after adequate follow-up to differentiate between permanent and transient hypothyroidism.

Finally, the CH screening protocols, including the screening methods, screening analytic marker cutoffs, and screening algorithms differed between the laboratories, with some even instituting protocol changes during the course of the study; therefore, direct comparison of the effects of laboratory screening protocols on primary CH detection rates was not possible. The strengths of this study are that it is the only comparative study between one-screen and two-screen states; the sample included 2,251 cases of primary CH from among 4.64 million births, which is much larger than the previously reported within-state studies [7, 9, 10, 13]; and individual-level data for cases improved the ability to tease out specific associations.

5. Conclusion

Due to significant differences in screening algorithms between the one- and two-screen states evaluated in this study, it is not possible to make conclusions about the comparative utility of a routine second screen in contrast to a single newborn screen. However, it is notable that in the two-screen states, 11.5% of all primary CH cases were detected on the routine second screen and race/ethnicity of the newborn played a significant role in predicting whether a case would be identified on the first versus the second screen, with NHB and A/PI newborns more likely identified on the routine second screen compared with NHW and Hispanic newborns. Therefore, NBS programs that perform only a single newborn screen for primary CH must be aware of the potential for affected newborns not being detected by the single screen, particularly newborns of certain races/ethnicities. If the two-screen states converted to a single screen using their current NBS algorithms, the second screen-positive primary CH cases that they currently identify would presumably not be detected, resulting in probable delays in diagnosis and treatment. On the other hand, there is no evidence that the one-screen states, by not performing a routine second screen, are missing a substantial number of cases, since the overall primary CH detection rate in the one-screen states was found to be higher than in the two-screen states, and the one-screen states are aware of only a tiny proportion (<1%) of the primary CH cases not being detected by NBS. Therefore, with appropriate modifications in screening method and algorithm to align more closely with the processes in the one-screen states, the two-screen states might convert to single screen operation for CH with improved cost efficiency and without loss in performance.

Abbreviations

NBS	newborn screening
TSH	thyroid stimulating hormone
CH	congenital hypothyroidism
SACHDNC	Secretary's Advisory Committee on Heritable Disorders in Newborns and Children
APHL	Association of Public Health Laboratories
T4	thyroxine
BW	birth weight
NICU	neonatal intensive care unit
NHW	non-Hispanic white
NHB	non-Hispanic black
A/PI	Asian/Pacific Islander
TPN	total parenteral nutrition
OR	odds ratio
CI	confidence interval

References

1. Fisher DA, Dussault JH, Foley TP Jr, Klein AH, LaFranchi S, Larsen PR, Mitchell ML, Murphey WH, Walfish PG. Screening for congenital hypothyroidism: results of screening one million North American infants. J Pediatr. 1979; 94(5):700–705. [PubMed: 87512]

2. Lee KS, Perlman M. The impact of early obstetric discharge on newborn health care. Curr Opin Pediatr. 1996; 8(2):96–101. [PubMed: 8723801]

3. Therrell, BL., Jr; Pass, KA.; Levy, HL. In: Pass, KA.; Levy, HL., editors. Early hospital discharge: impact on newborn screening; Early Hospital Discharge: Impact on Newborn Screening: Proceedings of a Conference Held; Washington, D.C. March 31–April 1, 1995; Atlanta, GA: Council of Regional Networks for Genetics Services, Emory University School of Medicine; 1996. p. 75-86.

4. Hanley WB, Demshar H, Preston MA, Borczyk A, Schoonheyt WE, Clarke JT, Feigenbaum A. Newborn phenylketonuria (PKU) Guthrie (BIA) screening and early hospital discharge. Early Hum Dev. 1997; 47(1):87–96. [PubMed: 9118832]

5. Panny, SR. In: Pass, KA.; Levy, HL., editors. Problems associated with early discharge experienced by the Maryland Newborn Screening Program: effectiveness and limitations of voluntary repeat screening; Early Hospital Discharge: Impact on Newborn Screening: Proceedings of a Conference Held; Washington, D.C. March 31–April 1, 1995; Atlanta, GA: Council of Regional Networks for Genetic Services, Emory Univeristy School of Medicine; 1996. p. 213-232.

6. Brosnan CA, Brosnan P, Therrell BL, Slater CH, Swint JM, Annegers JF, Riley WJ. A comparative cost analysis of newborn screening for classic congenital adrenal hyperplasia in Texas. Public Health Rep. 1998; 113(2):170–178. [PubMed: 9719819]

7. Doyle DL, Sanderson M, Bentvelzen J, Fineman RM. Factors which influence the rate of receiving a routine second newborn screening test in Washington State. Am J Med Genet. 1995; 59(4):417–420. [PubMed: 8585558]

8. Hunter MK, Mandel SH, Sesser DE, Miyabira RS, Rien L, Skeels MR, LaFranchi SH. Follow-up of newborns with low thyroxine and nonelevated thyroid-stimulating hormone-screening concentrations: results of the 20-year experience in the Northwest Regional Newborn Screening Program. J Pediatr. 1998; 132(1):70–74. [PubMed: 9470003]

9. LaFranchi SH, Hanna CE, Krainz PL, Skeels MR, Miyahira RS, Sesser DE. Screening for congenital hypothyroidism with specimen collection at two time periods: results of the Northwest Regional Screening Program. Pediatrics. 1985; 76(5):734–740. [PubMed: 3932955]

10. Maniatis AK, Taylor L, Letson GW, Bloch CA, Kappy MS, Zeitler P. Congenital hypothyroidism and the second newborn metabolic screening in Colorado, USA. J Pediatr Endocrinol Metab. 2006; 19(1):31–38. [PubMed: 16509526]

11. Therrell BL Jr, Berenbaum SA, Manter-Kapanke V, Simmank J, Korman K, Prentice L, Gonzalez J, Gunn S. Results of screening 1.9 million Texas newborns for 21-hydroxylase-deficient congenital adrenal hyperplasia. Pediatrics. 1998; 101(4 Pt 1):583–590. [PubMed: 9521938]

12. Secretary's Advisory Committee. Heritable Disorders and Genetic Diseases in Newborns and Children, Summary of Eighth Meeting; June 5–6, 2006; Washington, D.C.

13. Levine GD, Therrell BL Jr. Second testing for hypothyroidism. Pediatrics. 1986; 78(2):375–376. [PubMed: 3737315]

14. Grant DB, Smith I. Survey of neonatal screening for primary hypothyroidism in England, Wales, and Northern Ireland 1982–4. Br Med J. 1988; 296(6633):1355–1358. [PubMed: 3134984]

15. Harris KB, Pass KA. Increase in congenital hypothyroidism in New York State and in the United States. Mol Genet Metab. 2007; 91(3):268–277. [PubMed: 17512233]

16. Hinton CF, Harris KB, Borgfeld L, Drummond-Borg M, Eaton R, Lorey F, Therrell BL, Wallace J, Pass KA. Trends in incidence rates of congenital hypothyroidism related to select demographic factors: data from the United States, California, Massachusetts, New York, and Texas. Pediatrics. 2010; 125(2 suppl):S37–S47. [PubMed: 20435716]

17. Lorey FW, Cunningham GC. Birth prevalence of primary congenital hypothyroidism by sex and ethnicity. Hum Biol. 1992; 64(4):531–538. [PubMed: 1644422]

18. Schoen EJ, Clapp W, To TT, Fireman BH. The key role of newborn thyroid scintigraphy with isotopic iodine (^{123}I) in defining and managing congenital hypothyroidism. Pediatrics. 2004; 114(6):e683–e688. [PubMed: 15574601]

19. Waller DK, Anderson JL, Lorey FW, Cunningham GC. Risk factors for congenital hypothyroidism: an investigation of infant's birth weight, ethnicity, and gender in California, 1990–1998. Teratology. 2000; 62(1):36–41. [PubMed: 10861631]

20. Medda E, Olivieri A, Stazi MA, Grandolfo ME, Fazzini C, Baserga M, Burroni M, Cacciari E, Calaciura F, Cassio A, Chiovato L, Costa P, Leonardi D, Martucci M, Moschini L, Pagliardini S, Parlato G, Pignero A, Pinchera A, Sala D, Sava L, Stoppioni V, Tancredi F, Valentini F, Vigneri R, Sorcini M. Risk factors for congenital hypothyroidism: results of a population case-control study (1997–2003). Eur J Endocrinol. 2005; 153(6):765–773. [PubMed: 16322381]

21. Oakley GA, Muir T, Ray M, Girdwood RW, Kennedy R, Donaldson MD. Increased incidence of congenital malformations in children with transient thyroid-stimulating hormone elevation on neonatal screening. J Pediatr. 1998; 132(4):726–730. [PubMed: 9580779]

22. Sorcini M, Fazzini C, Olivieri A, Grandolfo ME, Medda E, Stazi MA, Balestrazzi P, Giovannelli G, Carta S. Neonatal screening in congenital hypothyroidism in Italy. The National Registry. Ann Ist Super Sanita. 1994; 30(3):275–287. [PubMed: 7879993]

23. Murphy N, Hume R, van Toor H, Matthews TG, Ogston SA, Wu SY, Visser TJ, Williams FL. The hypothalamic-pituitary-thyroid axis in preterm infants; changes in the first 24 hours of postnatal life. J Clin Endocrinol Metab. 2004; 89(6):2824–2831. [PubMed: 15181064]

24. Rapaport R. Thyroid function in the very low birth weight newborn: rescreen or reevaluate? J Pediatr. 2002; 140(3):287–289. [PubMed: 11953723]

25. Rapaport R, Rose SR, Freemark M. Hypothyroxinemia in the preterm infant: the benefits and risks of thyroxine treatment. J Pediatr. 2001; 139(2):182–188. [PubMed: 11487741]

26. Knobel M, Medeiros-Neto G. An outline of inherited disorders of the thyroid hormone generating system. Thyroid. 2003; 13(8):771–801. [PubMed: 14558921]

27. Moreno JC, de Vijlder JJ, Vulsma T, Ris-Stalpers C. Genetic basis of hypothyroidism: recent advances, gaps and strategies for future research. Trends Endocrinol Metab. 2003; 14(7):318–326. [PubMed: 12946874]

28. Polak, M.; Szinnai, G. Chapter 84 - thyroid disorders. In: Rimoin, DL.; Pyeritz, RE.; Korf, B., editors. Emery and Rimoin's Principles and Practice of Medical Genetics. Academic Press; London: 2013. p. 1-24.

29. Rastogi MV, LaFranchi SH. Congenital hypothyroidism. Orphanet J Rare Dis. 2010; 5:17. [PubMed: 20537182]

30. Van Vliet, G.; Deladoey, J. Hypothyroidism in infants and children: congenital hypothyroidism. In: Braverman, LE.; Cooper, D., editors. Werner & Ingbar's The Thyroid: a Fundamental and Clinical Text. Lippincott Williams and Wilkins; Philadelphia: 2012. p. 790-802.

	2003	2004	2005	2006	2007	2008	2009	2010	2011
AL						■	■	■	■
CA			■	■	■				
DE	■	■	■	■	■				
MD	■	■	■	■	■				
OR		■	■	■	■	■			
TX		■	■	■	■	■			
WI	■	■	■	■	■				

Fig. 1.
Source of cases of congenital hypothyroidism by year.

	Both T4 and TSH	T4 and, if below specified cutoff, then TSH	TSH only
One-Screen States			
CA			
WI			
Two-Screen States			
AL			
DE			
MD			
OR			
TX			

Fig. 2.

Congenital hypothyroidism screening algorithm for each state.

T4=thyroxine; TSH=thyroid stimulating hormone
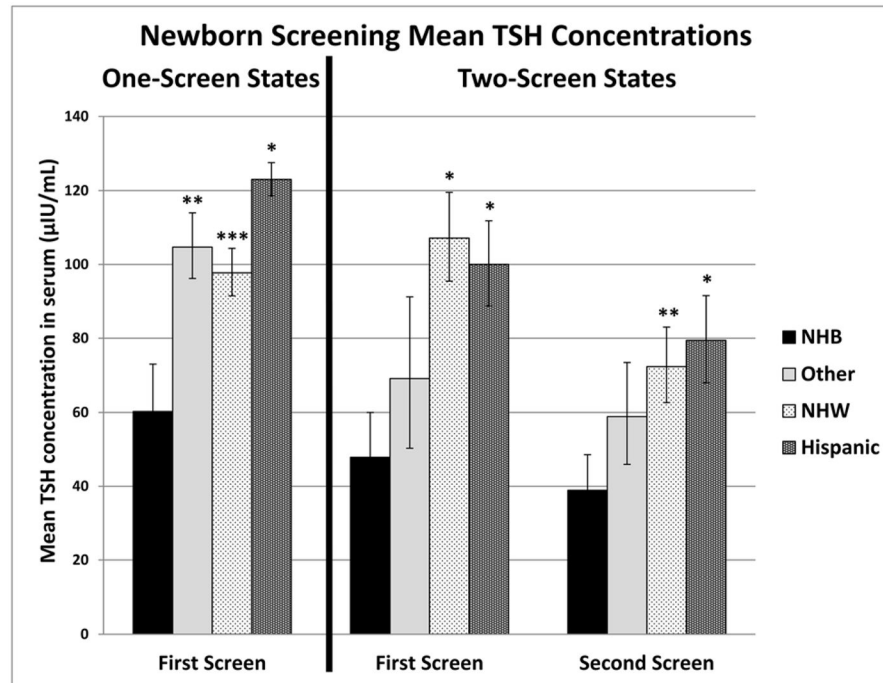
Fig. 3.

Adjusted[a] mean newborn screening TSH screening concentrations by race/ethnicity for primary CH cases identified on the first screen in one-screen states, and on the first or the second screen in two-screen states.

[a]Adjusted for state, sex, birth weight, blood transfusion prior to first screen, and age of newborn at first screen specimen collection

TSH=thyroid stimulating hormone; CH=congenital hypothyroidism; NHB=non-Hispanic black; NHW=non-Hispanic white

Other includes Asian/Pacific Islander, Middle Eastern, Native American, and Other, non-specified

Error bars specify the standard deviation

P-values are shown for non-Hispanic black compared with the other racial/ethnic groups

*P < 0.001; **P = 0.002; ***P = 0.006

Table 1

Primary congenital hypothyroidism cases detected in the one-screen and two-screen states

Initial Abnormal Screen	One-Screen States N (%)	Two-Screen States N (%)
First Screen	1027 (97.2)	1041 (87.1)
Second Screen	NA	137 (11.5)
Targeted Second Screen	19 (1.8)	NA
Unknown[a]	2 (0.2)	12 (1.0)
Not Detected by NBS	8 (0.8)	5 (0.4)
TOTAL	1056 (100)	1195 (100)

[a]Unknown cases had an initial specimen submitted for NBS that was deemed unsatisfactory by the laboratory; the case was detected on a subsequent screening specimen, but because of the initial unsatisfactory specimen, it is unknown whether the case would have been detected on the first screen.

N=number; NA=not-applicable; NBS=newborn screening

Table 2

Detection rate of primary congenital hypothyroidism in the one-screen and two-screen states, based on case characteristics

Characteristic	Category	One-Screen States Detection Rate of Primary CH (1 in N)	Two-Screen States Detection Rate of Primary CH (1 in N)	P-value[a]
Race/Ethnicity	Non-Hispanic White	2240	2621	**0.038**
	Non-Hispanic Black	3441	3932	0.499
	Hispanic	1573	1767	0.056
	Asian/Pacific Islander	1586	968	**0.001**
	Other	1893	2637	0.168
Sex	Male	2988	2987	0.998
	Female	1375	1734	**<0.001**
Birth Weight	<1000 g	2101	324	0.209
	1000–2499 g	3431	1247	0.373
	2500–3999 g	1714	2314	**0.009**
	4000 g	1623	1965	0.893
NICU Admission	Yes	915	908	0.942
	No	2062	2374	**0.003**
Blood Transfusion	Yes	560	225	**0.023**
	No	1766	2029	**0.002**
Age at Specimen Collection (first screen)	0 days (<24 hours)	1767	2493	**0.010**
	1 day (24–47 hours)	2063	3478	**<0.001**
	2 days (48–71 hours)	1802	1508	0.087
	3 days (72–95 hours)	1884	1445	0.248
	4 days (96 hours)	968	554	**<0.001**
Overall		1904	2201	**0.001**

[a]Based on Z-test for 2 proportions; significant p-values are shown in bold font

CH=congenital hypothyroidism; N=number; NICU=neonatal intensive care unit

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Table 3

Mean age of first specimen collection for first screen positive cases of primary congenital hypothyroidism in the one-screen and two-screen states, based on race/ethnicity

Race/Ethnicity	One-Screen States		Two-Screen States		
	N	Mean First Screen Age (SD)[a]	N	Mean First Screen Age (SD)[a]	P-Value[b]
Non-Hispanic White	282	1.46 (1.66)	364	2.03 (1.88)	**<0.001**
Non-Hispanic Black	30[c]	1.07 (0.74)	71	3.25 (4.97)	**0.019**
Hispanic	558	1.13 (1.41)	469	2.32 (5.04)	**<0.001**
Other[d]	130	1.34 (3.94)	76	3.14 (7.19)	**0.021**
Overall	1000	1.25 (1.98)	980	2.34 (4.40)	**<0.001**

[a] Mean age in days

[b] Based on 2-tailed t-test; significant p-values are shown in bold font

[c] Excludes one outlier case (615 gm birth weight infant with initial screen at day of life 68)

[d] Other includes Asian/Pacific Islander, Middle Eastern, Native American, and Other, non-specified

N=number; SD=standard deviation

Table 4

Odds of primary congenital hypothyroidism cases being identified on the first versus the second screen in the two-screen states

		Identified on the First vs. Second Screen in the Two-Screen States		
		Univariate	Full Model	Predictive Model
Characteristic	**Primary CH Cases N (%)**	**OR (95% CI)**	**OR (95% CI)**	**OR (95% CI)**
State				
Texas	868 (73.7%)	Reference	Reference	
Alabama	94 (8.0%)	1.35 (0.61–3.01)	2.02 (0.81–5.04)	
Delaware	21 (1.8%)	0.65 (0.19–2.26)	0.93 (0.23–3.82)	
Maryland	89 (7.5%)	**0.33 (0.19–0.56)**	0.60 (0.28–1.31)	
Oregon	106 (9.0%)	**0.44 (0.26–0.75)**	0.72 (0.32–1.64)	
Race/Ethnicity				
Non-Hispanic White	405 (36.2%)	Reference	Reference	Reference
Non-Hispanic Black	90 (8.0%)	**0.40 (0.22–0.73)**	**0.45 (0.23–0.88)**	**0.40 (0.22–0.73)**
Hispanic	512 (45.8%)	1.33 (0.84–2.13)	1.09 (0.66–1.80)	1.33 (0.84–2.13)
Asian/Pacific Islander	73 (6.5%)	**0.21 (0.11–0.37)**	**0.22 (0.12–0.40)**	**0.21 (0.11–0.37)**
Other	39 (3.5%)	0.49 (0.20–1.18)	0.45 (0.17–1.19)	0.49 (0.20–1.18)
Sex				
Male	436 (37.7%)	Reference	Reference	
Female	722 (62.3%)	**1.51 (1.05–2.16)**	1.26 (0.84–1.88)	
Feeding Status				
Breastfeeding Only	404 (38.1%)	Reference		
Breastfeeding and Formula	277 (26.1%)	1.22 (0.77–1.96)		
Formula Only	278 (26.2%)	1.38 (0.85–2.24)		
TPN	83 (7.8%)	1.13 (0.55–2.32)		
Other	19 (1.8%)	0.43 (0.15–1.25)		
Birth Weight				
<1000 g	43 (3.8%)	**0.42 (0.20–0.88)**	1.05 (0.36–3.07)	
1000–2499 g	134 (11.7%)	0.88 (0.51–1.52)	0.97 (0.52–1.83)	
2500–3999 g	886 (77.3%)	Reference	Reference	
4000 g	83 (7.2%)	1.99 (0.79–5.02)	1.63 (0.62–4.26)	
NICU Admission				
No	939 (80.7%)	Reference		
Yes	225 (19.3%)	0.69 (0.45–1.05)		
Blood Transfusion				
No	1063 (97.0%)	Reference	Reference	
Yes	33 (3.0%)	**0.43 (0.19–0.97)**	0.60 (0.21–1.67)	
Age of Newborn at Collection				
0 days	66 (5.6%)	Reference	Reference	

| Characteristic | Primary CH Cases N (%) | Identified on the First vs. Second Screen in the Two-Screen States | | |
| | | Univariate | Full Model | Predictive Model |
		OR (95% CI)	OR (95% CI)	OR (95% CI)
1 day	444 (38.1%)	**2.02 (1.03–3.98)**	1.50 (0.77–2.91)	
2 days	465 (39.9%)	1.87 (0.96–3.65)	1.38 (0.69–2.77)	
3 days	191 (16.4%)	1.88 (0.89–4.00)	1.08 (0.49–2.38)	
Time from Collection to Assay				
0–3 days	251 (21.3%)	Reference	Reference	
4–5 days	356 (30.2%)	**2.27 (1.40–3.69)**	1.54 (0.80–2.97)	
6–7 days	368 (31.3%)	**2.01 (1.26–3.20)**	1.26 (0.64–2.50)	
8 days	203 (17.2%)	**1.76 (1.02–3.01)**	0.93 (0.43–2.02)	

CH=congenital hypothyroidism; OR= odds ratio; CI=confidence interval; N=number; TPN=total parenteral nutrition; NICU=neonatal intensive care unit

Significant odds ratios are shown in bold font