Gender disparities in screening for congenital hypothyroidism using thyroxine as a primary screen

Lenore DeMartino¹, Rebecca McMahon¹, Michele Caggana¹ and Norma P Tavakoli¹,²
¹Division of Genetics, Wadsworth Center, New York State Department of Health, Albany, New York, USA and
²Department of Biomedical Sciences, School of Public Health, State University of New York, Albany, New York, USA

Abstract

Objective: Newborn screening for congenital hypothyroidism (CH) is based on testing for the markers thyroxine (T4) and/or thyroid-stimulating hormone (TSH). Diagnosis of CH is complicated because many factors affect the levels of these hormones including infant birth weight, prematurity and age at specimen collection. We investigated whether the sex of the newborn affected the levels of T4 and TSH and consequently the outcome of newborn screening.

Design: In New York State, the Newborn Screening program initially tests all infants for T4 and any baby with a result in the lowest 10% is triaged for TSH screening. We analyzed data from 2008 to 2016 to determine mean and median T4 and TSH values and how these results correlate with the sex of infants who are reported as borderline, referred and confirmed with CH.

Methods: T4 and TSH concentrations in dried blood spots were measured using commercially available fluoroimmunoassays.

Results: From 2008 to 2016, of the 2.4 million specimens tested for thyroxine, 51.5% were from male and 48.5% were from female infants. Male infants constituted 60% of specimens triaged for TSH testing, 64.9% of repeat requests and 59.6% of referrals, but only 49% of confirmed CH cases. The mean and median T4 values were lower (a difference of approximately 0.8–1.1 μg/dL each year) and the median TSH values were higher in male compared to female infants.

Conclusions: Natural differences in thyroid hormone levels in male and female infants leads to male infants being disproportionately represented in the false-positive category.

Introduction

Newborn screening for congenital hypothyroidism (CH) is crucial in detecting this common endocrine disorder that untreated may result in intellectual disability and growth retardation. Screening strategies are based on detecting low thyroxine (T4), or, elevated thyroid-stimulating hormone (TSH) or, a combination of both (1, 2, 3). Factors other than CH disease that affect T4 and TSH values will contribute to false-positive and -negative results if they are not considered in the screening algorithm. For example, within the first 1–2 days after birth, there is a physiological surge in TSH which then normalizes (2, 4, 5). Consequently, if a specimen is collected within the first 24 h after birth and no consideration is made for time of collection, an elevated TSH will indicate CH and the baby will be referred unnecessarily. As a result, many programs have different cut-offs for TSH if the specimen was collected within the first 24 h after birth. Other factors that affect T4 and TSH values are birth weight and prematurity (6). Low birth weight (LBW) and preterm infants have low total T4 values (7, 8, 9). A significant number of
infants who are admitted to NICU, especially extremely LBW, ill or premature infants, exhibit a delayed TSH rise (10, 11, 12, 13, 14) due to the immaturity of the hypothalamic–pituitary axis. Because TSH is lower in premature infants than in full-term infants (15), the initial newborn screen cannot be relied upon to detect CH in premature infants. This leads to the recommendation that thyroid function should be monitored in LBW, preterm and babies admitted to the NICU subsequent to the initial newborn screen for at least 4–6 weeks, irrespective of weight (10, 12, 16).

The New York State (NYS) Newborn Screening Program (NBSP) initially screens all infants for T4 and any baby with a result in the lowest 10% is then screened for TSH. Infants with low T4 and elevated TSH are referred for follow-up diagnostic testing. An additional specimen is requested for infants with low T4 and normal TSH or normal T4 and slightly elevated TSH. We investigated whether in addition to the time of specimen collection, birth weight and prematurity, whether the sex of the infant would also affect T4 and TSH values and how this would influence referral rates of male and female infants. Any factor that influences T4 and TSH values in infants should be a consideration in selecting an algorithm for CH screening and setting cut-off values.

Subjects and methods

From 2008 to 2016, the NYS NBSP screened approximately 2.2 million babies for CH. In NYS, it is recommended that a blood specimen be collected via a heel stick from newborns on a Guthrie card 24–48h after birth and sent to the NBSP with mother and infant demographic information. Specimens are required to be shipped overnight at ambient temperature. Three millimeter dried blood spots were punched into 96-well plates. T4 concentration in dried blood spots was measured using the AutoDELFIA neonatal thyroxine kit (Perkin Elmer). Each T4 assay was comprised of two 96-well plates, each containing 87 patient specimens. The specimens with the lowest 10% values for T4 for each assay were triaged for TSH testing. Approximately 298 000 infants were screened for TSH using the AutoDELFIA neonatal hTSH kit (Perkin Elmer). Infants with abnormal results were referred for follow up diagnostic testing or if a borderline result was obtained, a repeat specimen was requested. Figure 1 is a schematic of the current CH testing algorithm.

Statistical analysis

Differences in T4 mean values for male and female infants were assessed using one-way ANOVA, and differences in the numbers of male and female infants referred reported as borderline or triaged for TSH testing were assessed using the Chi-square test. A P value of <0.05 was considered statistically significant.

Human subjects

The manuscript is a retrospective case report that does not require institutional review board approval at our institution (Wadsworth Center, NYSDOH). All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

Results

Amongst the 2.4 million specimens from 2.2 million babies tested for T4, 51.5% were from male and 48.5% were from female infants (Table 1). First tier screening was performed for T4 and the specimens with the lowest 10% T4 values were triaged for TSH testing. Of the specimens triaged to be screened for TSH, 60% were from male and 40% were from female infants. Repeat specimens were requested for 21 407 male and 11 569 female infants due to borderline T4 or TSH results (Table 1). Following receipt of repeat specimens for these infants, 625 male and 340 female infants were diagnosed with CH, or ‘other CH’. The ‘other CH’ category included: possible CH disease, maternal medication and maternal TSH-blocking antibodies, thyroid disease of other etiology (e.g. Down syndrome which is associated with thyroid dysfunction) and a disease not on the newborn screening panel.

From 2008 to 2016 based on the NYS screening algorithm, 4114 male and 2791 female infants were referred (Table 1). Follow-up of referred infants confirmed CH in 718 (49.0%) male and 747 (51.0%) female infants over the 9-year period (Table 2). Confirmed CH cases included, primary CH (generally with low T4, elevated TSH), secondary CH (low T4, normal TSH) and compensated CH (slightly low to normal T4 and elevated TSH). Of the 889 cases of primary CH, 370 (41.6%) were male and 518 (58.3%) were female infants.

Table 1: Display of CH cases

<table>
<thead>
<tr>
<th>Group</th>
<th>Number of Cases</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Primary CH</td>
<td>889</td>
<td></td>
</tr>
<tr>
<td>Secondary CH</td>
<td>112</td>
<td></td>
</tr>
<tr>
<td>Compensated CH</td>
<td>21</td>
<td></td>
</tr>
<tr>
<td>Other CH</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>889</td>
<td></td>
</tr>
</tbody>
</table>
and of the 49 cases of secondary CH, 27 (55.1%) were male and 22 (44.9%) were female infants. Five hundred fifty four (64.3%) male and 307 (35.7%) female cases were closed as ‘other CH’ (Table 2).

Four hundred sixty five (64.9%) male cases were closed as possible CH, whereas only 252 (35.1%) female cases were closed as possible CH (Table 2). Cases are closed as possible CH due to hypothyroxinemia or hyperthyrotopinemia.

Table 1  Numbers of different categories of specimens tested, and babies reported as borderline or referred for CH from 2008 to 2016.

<table>
<thead>
<tr>
<th>Category</th>
<th>Total</th>
<th>Male</th>
<th>Female</th>
</tr>
</thead>
<tbody>
<tr>
<td>Specimens tested</td>
<td>2,366 384</td>
<td>1,219 602 (51.5%)</td>
<td>1,146 530 (48.5%)</td>
</tr>
<tr>
<td>Babies, n</td>
<td>2,178 705</td>
<td>1,116 419 (51.2%)</td>
<td>1,062 097 (48.7%)</td>
</tr>
<tr>
<td>Specimens from NBW babies</td>
<td>1,994 133</td>
<td>1,031 141 (51.7%)</td>
<td>962 853 (48.3%)</td>
</tr>
<tr>
<td>Specimens from LBW babies</td>
<td>278 174</td>
<td>135 902 (48.9%)</td>
<td>142 200 (51.1%)</td>
</tr>
<tr>
<td>DOB specimens tested</td>
<td>46 385</td>
<td>24 865 (53.6%)</td>
<td>21 502 (46.4%)</td>
</tr>
<tr>
<td>Borderline</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low T4/normal TSH</td>
<td>23 963</td>
<td>15 109 (63.1%)</td>
<td>8 854 (36.9%)</td>
</tr>
<tr>
<td>Normal T4/slightly elevated TSH</td>
<td>9023</td>
<td>6298 (69.8%)</td>
<td>2724 (30.2%)</td>
</tr>
<tr>
<td>Total</td>
<td>32 986</td>
<td>21 407 (64.9%)</td>
<td>11 569 (35.1%)</td>
</tr>
<tr>
<td>Total referred</td>
<td>6907</td>
<td>4114 (59.6%)</td>
<td>2791 (40.4%)</td>
</tr>
</tbody>
</table>

DOB, day of birth (<24h); LBW, low birth weight (<2500g); NBW, normal birth weight (≥2500g; ≥24h, ≤14 day old babies); T4, thyroxine; TSH, thyroid-stimulating hormone.

Figure 1
Current NYS screening algorithm for CH. *A repeat specimen is requested when a specimen collected on day of birth (DOB) is received but the specimen is nevertheless tested. If the TSH value of a DOB specimen is ≥150 μU/mL, the infant is immediately referred for CH. For a non-DOB specimen, an infant with a TSH value ≥100 μU/mL is immediately referred. T4, thyroxine; TSH, thyroid-stimulating hormone.
Gender disparities in screening for CH

L DeMartino and others

from L DeMartino and others

of prematurity or persistent hypothyroxinemia or hyperthyrotropinemia.

From 2008 to 2016, thyroxine-binding globulin (TBG) deficiency was reported in 275 male infants and 34 female infants (Table 2). The expectation is that there is significant underreporting of TBG deficiency to the NBSP because this condition is generally non-harmful.

We hypothesized that the reason more male infants were triaged for TSH testing was reported with borderline results and were referred was because there were differences between T4 and TSH concentrations in male and female infants. Therefore, we calculated yearly mean and median values for T4 and TSH for male and female infants. The mean and median T4 values were lower in male than female infants (a difference of approx. 0.8–1.1 μg/dL each year) as evidenced by the mean yearly values (Fig. 2A) (P<0.0000001) and the normal distribution curves (Fig. 2B is an example from 2014). The lower mean and median T4 values in male infants were not due to low birth weight since male infants had higher mean and median birth weights than female infants, and there were fewer low birth weight (LBW) male infants than female infants (Tables 1 and 3). LBW was defined as babies who weighed less than 2500 g. The lower mean and median T4 values were also not due to more specimens from male babies being collected within the first 24 h after birth because when T4 values from these babies were not included in T4 calculations, the mean and median T4 values remained lower in male than female infants (data not shown). Mean TSH values were similar in male and female infants but median values for male infants were higher (Fig. 2C).
Table 3  Mean and median birth weight of male and female infants from 2008 to 2016.

<table>
<thead>
<tr>
<th>Year</th>
<th>Mean weight (g)</th>
<th>Median weight (g)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Male</td>
<td>Female</td>
</tr>
<tr>
<td>2008</td>
<td>3323.0</td>
<td>3208.5</td>
</tr>
<tr>
<td>2009</td>
<td>3321.8</td>
<td>3204.5</td>
</tr>
<tr>
<td>2010</td>
<td>3321.3</td>
<td>3203.4</td>
</tr>
<tr>
<td>2011</td>
<td>3321.9</td>
<td>3201.3</td>
</tr>
<tr>
<td>2012</td>
<td>3326.0</td>
<td>3206.4</td>
</tr>
<tr>
<td>2013</td>
<td>3328.0</td>
<td>3212.1</td>
</tr>
<tr>
<td>2014</td>
<td>3331.6</td>
<td>3216.1</td>
</tr>
<tr>
<td>2015</td>
<td>3330.8</td>
<td>3213.2</td>
</tr>
<tr>
<td>2016</td>
<td>3326.5</td>
<td>3206.6</td>
</tr>
</tbody>
</table>

The specificity and positive predictive value (PPV) of the CH screening method were calculated. We did not attempt to calculate the sensitivity or negative predictive value of the screening method as the program rarely receives feedback regarding false-negative cases. Although the specificity of the CH screening method is similar in male and female infants (99.7% vs 99.8%), the PPV is notably lower in male infants than female infants (17.54% vs 26.83%).

Discussion

Having tested 2.2 million infants from 2008 to 2016 for CH, our data show that male infants have a lower mean (and median) T4 and higher median TSH than female infants despite having a higher mean and median birth weight. If there was no difference between the sexes, the expectation would be that female infants would have lower T4 values and higher TSH values, since their mean and median birth weight are lower and birth weight is known to affect T4 and TSH values. Higher median TSH has been reported previously in male babies (17), as well as, in male adults (18).

As a result of the difference between male and female infants in their T4 and TSH hormone levels, in our program a higher percentage of male than female infants were triaged for TSH testing (60% vs 40%) \((P<0.0001)\), a higher number of repeat specimens from male infants were requested (64.9% vs 35.1%) \((P<0.0001)\) and a higher number of male infants were referred (59.6% vs 40.4%) \((P<0.0001)\) (Table 1). Yet, primary CH has a higher incidence in female infants and overall there are a higher number of confirmed CH cases in females (Table 2). However, a higher number of cases of male infants were closed as possible CH (Table 2). These infants have hypothyroxinemia or hyperthyrotropinemia of prematurity or persistent hypothyroxinemia or hyperthyrotropinemia and are either on treatment or continue to be monitored. Both hypothyroxinemia and hyperthyrotropinemia of prematurity resolve with advancing postnatal age, although they may require T4 replacement for up to 1 year (19). In NYS a larger number of male infant cases are closed as possible CH seemingly due to the borderline results in these babies indicating subclinical hypothyroidism/mild disease or transient hypothyroidism. Babies with possible CH are monitored regularly and some are on treatment. Long-term follow-up of these babies including information on their physical and cognitive development would be useful to determine whether these cases are true cases of CH and require treatment or whether they are transient. The current recommendation for babies whose CH status is not certain but who are on treatment is to discontinue treatment at three years of age and determine whether they remain hypothyroid (1, 20, 21). Unfortunately, our program does not have the ability to follow these babies until three years of age and therefore the outcome of the trial-off therapy is not available to us. In addition, it is not evident whether the recommendation is followed.

Our program detected a total of 309 cases of TBG deficiency (Table 2). Male babies constituted 89% of these cases. TBG deficiency is an X-linked condition and is fairly common in male infants (1 in 2400) and approximately ten-fold less frequent in females (22, 23). The incidence in our male populations was lower (1 in 4059) because there is underreporting of this condition to the program: either TBG testing was not done or if it was done, the result was not reported to the program. TBG deficiency does not cause thyroid disease but it does hinder the interpretation of thyroid function tests because these babies have low total T4 but normal free T4 and TSH (15). It is necessary to rule out TBG deficiency so that the baby is not unnecessarily evaluated and treated for CH. Unlike programs that use T4 as the primary screen for CH, programs that use TSH miss TBG deficient infants (1).
Using our algorithm, male infants are disproportionately represented in the false positive category even though incidence of CH in male infants is marginally lower. In a national data set from 1991 to 2000, the female to male ratio of CH was reported to be 1.56 (24). Our data shows that the female-to-male ratio of primary CH is 1.47 while the overall CH ratio is 1.09 and the ratio of possible CH is 0.57. Our results indicate that in addition to considering different cut-offs based on time of specimen collection and infant’s birth weight, the sex of the infant may also be a consideration. It would be interesting to determine whether programs that use a first tier TSH test to screen for CH also have a higher number of false positive results for males. Our expectation is that they will have a higher false positive rate because median TSH values for males are higher in male infants than female infants. Therefore, if the same cut-off is used for males and females, a higher number of male infants will cross the threshold value. From 2008 to 2016 our program requested a repeat specimen for 9023 specimens that were borderline because of normal T4 values and slightly elevated TSH values. Of the 9023 infants, 6298 (69.8%) were male and 2724 (30.2%) were female infants (Table 2). Five thousand nine hundred and eighty three of the male borderline cases and 2589 of the female borderline cases were closed as no disease indicating that even in infants with normal T4, a higher number of male infants fell into an abnormal category requiring an additional specimen.

The consequence of lower T4 and higher TSH values in male infants is that significantly increased number of recalls and follow-up tests are required for male infants. Most of the male babies were in the borderline range for either or both T4 and TSH and many cases were closed as possible disease (Table 1). Unfortunately, the status of these cases is unclear. Because, they may be on treatment or are just being followed and those who are on treatment may have potentially transient or mild CH which may resolve by 3 years of age. Additional data are needed to determine whether early treatment in these babies is beneficial. Our data are consistent with programs who have decreased their TSH cut-off value and then found an increased incidence of CH with the female to male ratio of presumptive positive cases decreasing (25). Cases of infants with transient hypothyroidism have a lower female-to-male ratio than cases of permanent CH (26). Additional studies are needed to determine whether these borderline T4 and TSH values in male infants indicate treatment requiring CH disease or whether male infants have naturally different ranges, in which case NBS programs should investigate setting different cut-off ranges for T4 and TSH for male and female infants. Although setting different cut-offs will complicate the CH algorithm, it will decrease the number of false positive screens for male infants.

Declaration of interest
The authors declare that there is no conflict of interest that could be perceived as prejudicing the impartiality of this study.

Funding
Funding for newborn screening in NYS is provided by the State of NY.

Author contribution statement
L D and R M were involved in acquisition of data and revising the manuscript. N P T was involved in design of the work, analysis and interpretation of data and drafting and finalizing the manuscript.

Acknowledgements
The authors would like to thank past and present members of the NYS’ NBSP’s Immuноassay and Accessioning Laboratory for performing T4 and TSH testing and the follow-up unit for performing follow-up of infants with CH.

References
7 Kok JH, Hart G, Endert E, Koppe JG & De Vijlder JIM. Normal ranges of T4 screening values in low birthweight infants. Archives of Disease in Childhood 1983 58 190–194. (https://doi.org/10.1136/adc.58.3.190)
8 Van Wassenaer AG, Kok JH, Briet JM, Pijning AM & de Vijlder JI. Thyroid function in very preterm newborns: possible implications. Thyroid 1999 9 85–91. (https://doi.org/10.1089/thy.1999.9.85)


14 Heather NL, Hofman PL, de Hora M, Carll J, Derraik JG & Webster D. Extreme premature and/or very low birth weight newborns: the importance of a specific protocol. Neonatal Inheritance Metabolism 2016 2010 83–92. (https://doi.org/10.1159/000323165)


25 Pollitt RJ. Evidence or enthusiasm? Why yields from UK newborn screening programmes for congenital hypothyroidism are increasing. Archives of Disease in Childhood 2016 101 120–123. (https://doi.org/10.1136/archdischild-2015-309546)


Received 7 May 2018
Accepted 26 June 2018