Improving the Dutch Newborn Screening for Central Congenital Hypothyroidism by Using 95% Reference Intervals for Thyroxine-Binding Globulin

Kevin Stroek[a]   Annemieke C. Heijboer[a, b]   Marja van Veen-Sijne[a]
Annet M. Bosch[c]   Catharina P.B. van der Ploeg[d]   Nitash Zwaveling-Soonawala[e]
Robert de Jonge[f]   A.S. Paul van Trotsenburg[e]   Anita Boelen[a]

[a]Endocrine Laboratory, Department of Clinical Chemistry, Amsterdam Gastroenterology, Endocrinology & Metabolism, Amsterdam UMC, University of Amsterdam, Amsterdam, The Netherlands; [b]Endocrine Laboratory, Department of Clinical Chemistry, Amsterdam Gastroenterology, Endocrinology & Metabolism, Amsterdam UMC, Vrije Universiteit, Amsterdam, The Netherlands; [c]Division of Metabolic Disorders, Department of Pediatrics, Emma Children's Hospital, Amsterdam Gastroenterology, Endocrinology & Metabolism, Amsterdam UMC, University of Amsterdam, Amsterdam, The Netherlands; [d]Department of Child Health, Netherlands Organization for Applied Scientific Research TNO, Leiden, The Netherlands; [e]Department of Paediatric Endocrinology, Emma Children's Hospital, Amsterdam Gastroenterology, Endocrinology & Metabolism, Amsterdam UMC, University of Amsterdam, Amsterdam, The Netherlands; [f]Department of Clinical Chemistry, Amsterdam UMC, Vrije Universiteit & University of Amsterdam, Amsterdam, The Netherlands

Abstract
Introduction: Newborn screening (NBS) for congenital hypothyroidism (CH) in the Netherlands consists of thyroxine (T4), thyroid-stimulating hormone (TSH), and T4-binding globulin (TBG) measurements to detect thyroidal CH and central CH (CH-C). CH-C is detected by T4 or a calculated T4/TBG ratio, which serves as an indirect measure of free T4. TSH and TBG are only measured in the lowest 20 and 5% of daily T4 values, respectively. A recent evaluation of the Dutch NBS for CH showed that the T4 and T4/TBG ratio contribute to the detection of CH-C but also lead to a low positive predictive value (PPV). Dried blood spot (DBS) reference intervals (RIs) are currently unknown and may contribute to improvement of our NBS algorithm. ***Materials and Methods:*** RIs of T4, TSH, TBG, and the T4/TBG ratio were determined according to Clinical & Laboratory Standards Institute guidelines in heel puncture cards from routine NBS in both sexes and at the common NBS sampling ages. Scatter plots were used to compare the healthy reference population to previously published data of CH-C patients and false positives. ***Results:*** Analyses of 1,670 heel puncture cards showed small differences between subgroups and led to the formulation of total sample DBS RIs for T4 (56–118 nmol/L), TSH (<2.6 mIU/L), TBG (116–271 nmol/L), and the T4/TBG ratio (>20). 46% of false-positive referrals based on T4 alone had a TBG below

Anita Boelen
Endocrine Laboratory, Department of Clinical Chemistry
Amsterdam UMC/University of Amsterdam
Meibergdreef 9, NL–1105 AZ Amsterdam (The Netherlands)
a.boelen @ amsterdamumc.nl

the RI, indicating preventable referral due to partial TBG deficiency. One case of CH-C also had partial TBG deficiency (TBG 59 and T4 12 nmol/L blood). ***Discussion/Conclusion:*** Established DBS RIs provided possibilities to improve the PPV of the Dutch CH NBS algorithm. We conclude that by taking partial TBG deficiency into account, approximately half of T4 false-positive referrals may be prevented while maintaining NBS sensitivity at the current level.

Introduction

Congenital hypothyroidism (CH) is thyroid hormone deficiency at birth mostly caused by impaired development of the thyroid gland or hormone biosynthesis (primary or thyroidal CH, CH-T). A less frequent, but equally important cause is hypothalamic or pituitary dysfunction, named secondary or central CH (CH-C) [1]. Early detection and treatment of CH prevent neurodevelopmental disabilities. Therefore, CH has been included in many newborn screening (NBS) programs worldwide since the 1970s [2]. Most programs employ a thyroid-stimulating hormone (TSH)-based strategy aiming at detection of CH-T [3]. However, newborns with CH-C – who typically have a low T4 in combination with a normal TSH concentration – will be missed [2]. In the Netherlands, the NBS strategy for CH is primarily based on the measurement of dried blood spot (DBS) T4 in order to screen for both CH-T and CH-C. Unique for the Dutch approach is that to detect CH-T, TSH is measured in the lowest 20% of T4 concentrations, and in order to detect CH-C, thyroxine-binding globulin (TBG) is measured in the lowest 5% of T4 concentrations enabling calculation of a modified T4/TBG ratio that serves as an indirect measure of free T4 (FT4). With this approach, low T4 values due to TBG deficiency can be excluded, and only neonates at risk for CH-C, that is, those with a low T4 value combined with a low T4/TBG ratio are referred to the pediatrician. In 2005, we showed that the Dutch NBS for CH is clinically and cost-effective with respect to detecting CH-T and CH-C [4]. In a recent evaluation of 11 years of the Dutch NBS for CH, we affirmed that T4 and the calculated T4/TBG ratio are very useful in detection of CH-C at the cost of a lower positive predictive value (PPV) compared to TSH-based screening programs [5].

Based on the premise that NBS for CH-C is useful and desirable, that false positives need to be prevented and that reliable assays for measurement of FT4 in DBS are currently not available in the Netherlands [6], optimization of the Dutch NBS for detection of CH-C is warranted to improve its specificity. Reference intervals (RIs) for diagnostic tests are crucial for the interpretation of the measurements. A RI is based on measurement of a laboratory test in a presumed healthy population and assessment of the 95% central range of all measurements. Neonatal RIs for DBS T4, TSH, and TBG concentrations and for the T4/TBG ratio as used in the Dutch NBS are currently not available. The aim of the present study was to establish these neonatal RIs and to investigate whether they could be used to reduce the number of false positives by improving the NBS algorithm, specifically for CH-C.

Methods

Sample Size and Subgroup Categories

According to the Clinical & Laboratory Standards Institute (CLSI) guidelines, a sample size of at least 120 reference subjects is required per category [7]. In order to establish RIs divided by sex (categories: males and females) and NBS sampling age after birth (categories: days 4, 5, 6, and 7), a population of at least 960 reference subjects is required for this study. RIs based on gestational age or weight would include premature neonates, but since this group is small compared to the group of full-term neonates, this category (defined as ≤36 weeks and ≤2,500 g) was not included in the current study.

Study Procedures

In the Netherlands, blood samples obtained by heel puncture are collected on filter paper between 72 and 168 h (days 4–7) after birth. For the current study, DBS T4, TSH, and TBG were measured prospectively during 18 subsequent days in all dried blood samples that were sent to the neonatal screening laboratory (part of the Endocrine Laboratory) in Amsterdam for daily NBS routine in a fixed period (February/March 2019).

Routine NBS

In the Dutch NBS algorithm for CH, all newborns with a T4 ≤−3 SD of the daily mean are immediately referred to the pediatrician, unless TBG is ≤40 nmol/L blood, indicating complete TBG deficiency. Newborns with a T4 ≤−0.8 SD and TSH ≥22 mIU/L blood or a TSH ≥8 mIU/L blood in a second screen are at risk for CH-T and are referred. TBG is measured in DBS when T4 concentrations are ≤−1.6 SD of the daily mean, and subsequently, a calculated T4/TBG ratio ≤17 requires a second heel puncture. If a second screen yields the same result, the newborn is considered at risk for CH-C and is referred to the pediatrician. For a detailed explanation of the NBS algorithm and the used referral criteria, we refer to our recent publication [5].

Laboratory Measurements

All measurements were performed with the same methods as used in the Dutch routine NBS. T4 and TSH were measured using

Table 1. Calculated DBS 95% RIs for T4, TSH, TBG, and the T4/TBG ratio in identified subgroups

Subgroup	T4, nmol/L							
	female days 4–5		male days 4–5		female days 6–7		male days 6–7	
95% RIs	lower	upper	lower	upper	lower	upper	lower	upper
	60	126	53	118	57	111	55	113
90% CIs	56–62	119–132	49–57	116–125	55–60	108–120	51–58	109–119
Characteristics								
Sample size	409		433		382		446	
Lowest value	52		44		50		41	
Highest value	141		136		148		137	
Arithmetic mean	88		85		84		82	
Median	87		85		84		82	
SD	16		16		14		15	

Subgroup	TSH, mIU/L	
	day 4	days 5–7
95% RIs	<3.4	<2.4
Characteristics		
Sample size	347	1,323
Lowest value	0.3	0.1
Highest value	7.6	5
Arithmetic mean	1.5	1.2
Median	1.3	1.1
SD	1.0	0.7

Subgroup	TBG, nmol/L				T4/TBG ratio	
	female		male		days 4–5	days 6–7
95% RIs	lower	upper	lower	upper	>21	>20
	125	274	113	270		
90% CIs	118–128	264–285	109–116	263–277	20–21	19–20
Characteristics						
Sample size	791		879		842	828
Lowest value	91		70		17	14
Highest value	359		351		53	48
Arithmetic mean	189		183		29	27
Median	185		180		28	27
SD	39		40		5	5

Measurements were performed in (full) blood. CI, confidence interval of the lower and the upper limit of the RI; RI, reference interval; SD, standard deviation; T4, thyroxine; TBG, thyroxine-binding globulin; TSH, thyroid-stimulating hormone; DBS, dried blood spot.

an immunoassay (GSP®, Perkin Elmer®, Waltham, MA, USA), and TBG was measured using an ELISA (Monobind Inc®, Lake Forrest, CA, USA). The total-assay variation (CV%) of the T4 immunoassay was 7.1% at the lower level (mean 21.5 nmol/L), 4.6% at the level of the medium control (mean 54.5 nmol/L), and 4.4% at the level of the high control (mean 92.2 nmol/L). The TSH immunoassay had a total-assay variation of 6.1% at the lower level (mean 16.3 mIU/L) and 5.1% at the higher level (mean 67.9 mIU/L), and the TBG ELISA 7.7% at the lower level (mean 76 nmol/L), 9.1% at the level of the medium control (mean 138 nmol/L), and 8.6% at the level of the high control (mean 234 nmol/L). All measurements were performed as duplicates.

Data Collection

Permission for the use of leftover heel prick blood from the NBS for this research project was given by the Dutch Research Committee on Neonatal Screening at the National Institute for Public Health and the Environment (RIVM). Data on sex and

Table 2. Calculated DBS 95% RIs for T4, TSH, TBG and the T4/TBG ratio in the complete reference population

Overall 95% RIs	T4, nmol/L		TSH, mIU/L	TBG, nmol/L		T4/TBG ratio
	lower	upper	<2.6	lower	upper	>20
	56	118		116	271	
90% CIs	55–58	117–121	2.5–2.8	113–119	267–277	20–21
Sample size	1,670		1,670	1,670		1,670
Lowest value	41		0.1	70		14
Highest value	148		7.6	359		53
Arithmetic mean	85		1.3	186		28
Median	84		1.1	183		28
SD	15		0.7	40		5

Measurements were performed in (full) blood. CI, confidence interval of the lower and the upper limit of the RI; RI, reference interval; SD, standard deviation; T4, thyroxine; TBG, thyroxine-binding globulin; TSH, thyroid-stimulating hormone; DBS, dried blood spot.

NBS sampling age were collected directly and anonymously from the heel puncture cards. NBS sampling age was defined in hours after birth and reported in days (NBS day 1 = 0–24 h after birth). The NBS Privacy Committee (at the Netherlands Organization for Applied Scientific Research [TNO], department of Child Health) gave permission to use available CH NBS data of referrals (2007–2017).

Statistical Analyses

IBM SPSS Statistics version 25 was used for descriptive statistics and comparisons. Outliers were detected by the D/R ratio, with a cutoff value (COV) of 1/3. To calculate the T4/TBG ratios in the reference population, we used the established formula {(T4 [in SD] + 5.1) × 1,000}/(TBG [nmol/L blood]) [4]. The average and the spread were presented as mean and SD for normally distributed variables and as median and interquartile range (IQR) for non-normal distributions. To approach normality in heavily skewed distributions, a log-transformation was applied before proceeding. The effects of sex and NBS sampling age on NBS parameters T4, TSH, TBG, and the T4/TBG ratio were assessed by conducting a two-way ANOVA, allowing for unequal sample sizes of subgroups, followed by the Tukey post hoc test to compare the means of relevant subgroups. In MedCalc® version 18.5, 95% RIs for T4, TSH, TBG, and the T4/TBG ratio of the total reference population and of the identified subgroups were calculated with 90% confidence intervals using the percentile method. Subgroup RIs were considered relevant when the difference between the subclass means was at least 25% as large as the 95% RI estimated in the total reference population, according to the CLSI-guidelines [7, 8]. Finally, we compared the DBS T4 concentrations, and when available, the TBG concentrations and the T4/TBG ratios of the first heel puncture samples of 2,067 full-term babies who were referred because of a very low T4 or the T4/TBG ratio, collected during 11 years (2007–2017) of CH NBS in the Netherlands [5], with the healthy reference population displayed as scatter plots.

Results

Reference Population

Neonatal heel puncture cards of 1,720 newborns were included in the study. None of these cards had a result leading to referral. After duplicate measurement of T4, TSH, and TBG, results with a high duplo variation (*n* = 50) were excluded, thus leaving 1,670 cards for further analyses. Descriptive statistics of the measured T4, TSH, and TBG concentrations, and the calculated T4/TBG ratio in the reference population, as well as divided by sex and NBS sampling age are presented in online suppl. Table 1; see www.karger.com/doi/10.1159/000513516 for all online suppl. material.

Reference Intervals and Subgroup Comparisons

Calculated RIs of T4, TSH, TBG, and the T4/TBG ratio in DBS for identified subgroups are reported in Table 1. Boxplots representing the distribution of the values of each subgroup are displayed in online suppl. Figures 1–4. Multiple comparisons with post hoc analyses identified homogeneous subsets within NBS sampling age for all parameters (shown in online suppl. Table 2). Total sample RIs of all parameters are reported in Table 2.

T4

T4 was significantly higher in females compared to males. Also, a statistically significant difference between NBS sampling ages was found, leading to identification of two homogeneous subsets: (1) days 4 and 5, and (2) days 6 and 7. Calculated RIs for T4 were 60–126 (female

days 4 and 5), 53–118 (male days 4 and 5), 57–111 (female days 6 and 7), and 55–113 (males days 6 and 7) nmol/L blood. The total sample RI was 56–118 nmol/L blood.

Thyroid-Stimulating Hormone

TSH was not different between females and males, but concentrations were significantly different between NBS sampling ages, forming two homogeneous subsets: (1) day 4 and (2) days 5, 6, and 7. The RIs for TSH were <3.4 and <2.4 mIU/L blood, respectively. The total sample 95% RI was <2.6 mIU/L blood.

TBG

TBG was significantly different between females and males, but concentrations were not significantly different between NBS sampling ages. The RIs were 125–274 (females) and 113–270 (males) nmol/L blood. The total sample 95% RI was 116–271 nmol/L blood.

T4/TBG Ratio

The T4/TBG ratio was not different between females and males, but a statistically significant difference was found between NBS sampling ages, leading to identification of two homogeneous subsets: (1) days 4 and 5, and (2) days 6 and 7. The RIs for the T4/TBG ratio were >21 and >20, respectively. The total sample RI was >20.

Reference Population Compared to Historical NBS Referrals

In 11 years (2007–2017) of NBS for CH, the T4 and T4/TBG ratio-based referral criteria led to referral of 83 CH-C patients and 1,984 false positives [5].

T4 and T4/TBG Ratio Referrals versus Reference Population

All detected CH-C patients were referred based on a T4 concentration ≤−3 SD or a T4/TBG ratio ≤17 (TBG is only measured when T4 is ≤−1.6 SD). The highest T4 concentration in CH-C patients was 60 nmol/L blood, overlapping with the lower end of the T4 RI. However, the T4/TBG ratio in these patients separated them from the healthy reference population (blue CH-C patients, shown in Fig. 1a). In all but one of the T4-based referrals, the T4/TBG ratio was below the RI and current COV. False-positive referrals due to T4 or T4/TBG ratio are displayed together with the healthy reference population in Figure 1b. A substantial number of T4/TBG ratio false-positive referrals showed a large overlap with T4 and T4/TBG ratio values of CH-C patients, but 10% (n = 118) of this

group with a T4 SD value between −1.6 and −3.0 had a T4 >60 nmol/L blood.

All CH-C patients had a TBG concentration within or above the TBG RI, except for one patient with a TBG of 59 nmol/L blood, a T4 of 12 nmol/L blood (shown in Fig. 1c), and a T4/TBG ratio of 20.3. False-positive referrals due to a T4 ≤−3 SD (n = 788), however, showed a partial overlap with CH-C patients and had a variety of T4/TBG ratios. Figure 1d shows that 364 (46%) of this group false-positives referred because of a T4 ≤−3 SD, had TBG concentrations below the TBG RI, indicating possible partial TBG deficiency.

Discussion

With this study, we established neonatal RIs for DBS T4, TSH, and TBG concentrations, and the calculated T4/TBG ratio as used in the Dutch NBS for CH. We established RIs based on sex and NBS sampling age for each parameter.

Separate RIs for TSH could be considered for newborns screened on day 4 versus days 5, 6, and 7, as an age-based difference in TSH concentrations was observed. This difference can be explained by the TSH surge that occurs in newborns directly after birth due to thyrotropin-releasing hormone secretion induced by cold. Subsequently, the neonatal TSH surge causes an increase of the serum FT4 in the first days of life, peaking at day 3 [9], and probably is the reason for the slightly higher DBS T4 concentrations on days 4 and 5. The COV for DBS TSH in our NBS is ≥8 mIU/L, which makes a day specific RI irrelevant for daily practice. None of the subjects in the reference population had a TSH ≥8 mIU/L blood.

TBG concentrations were higher in females compared to males, resulting in higher total T4 concentrations in females, an effect that is likely to be mediated by estrogen [10]. Although statistically significant, the differences between subgroups were small for T4, TBG, and the T4/TBG ratio, and therefore, one RI for the whole reference population seems sufficient for these parameters [8]. Due to the nature of our CH NBS algorithm, TSH and TBG are currently only determined in the population with relatively low T4 values. This study provides previously lacking information on these parameters in a complete population of healthy newborns.

In our recent evaluation of the Dutch CH NBS, we observed that the majority of full-term false-positive referrals were due to the (separate) T4 and T4/TBG ratio referral criteria [5]. With the current study, we visualized the
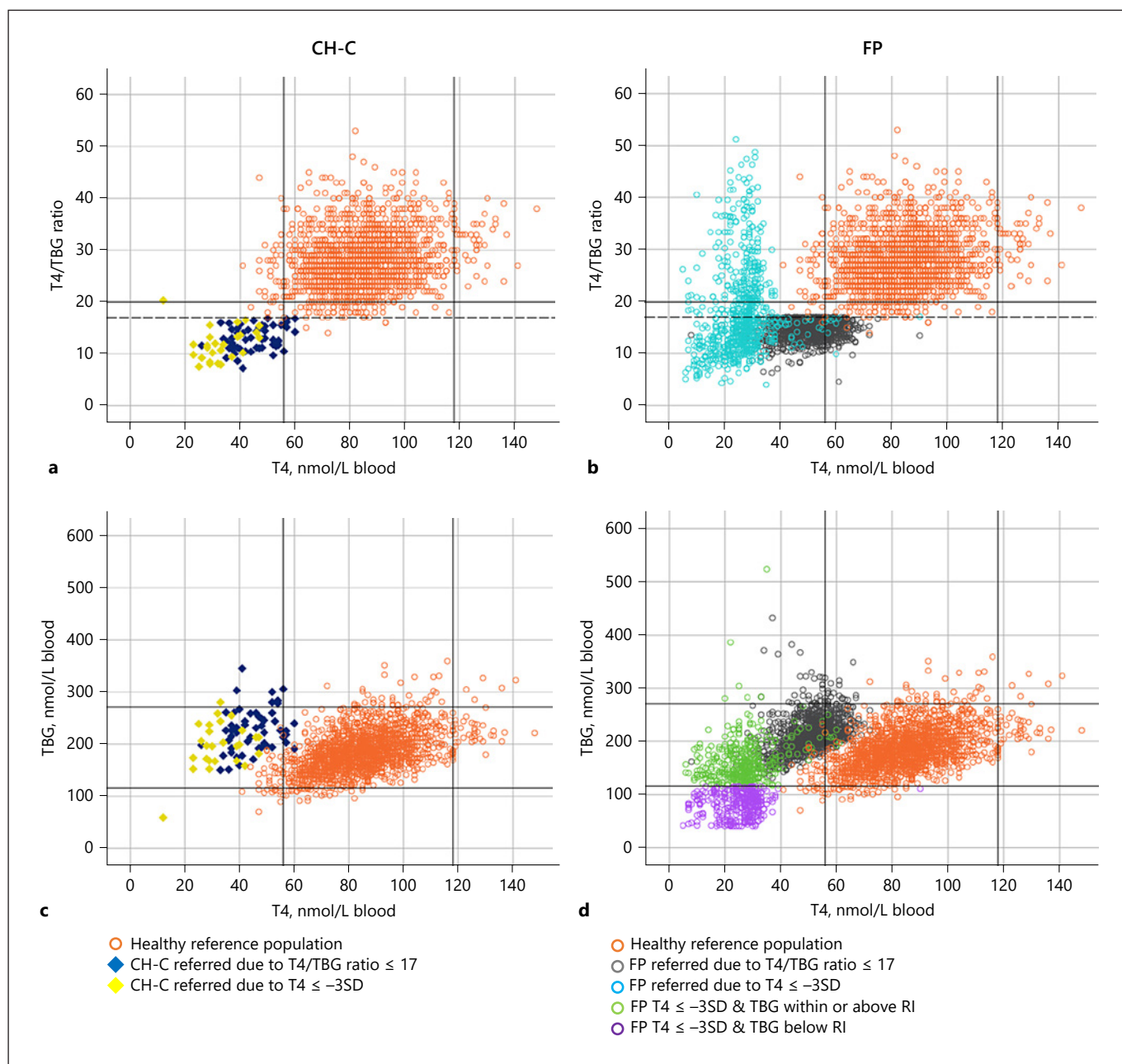
Fig. 1. T4 concentrations versus T4/TBG ratios of CH-C patients (**a**) and FP (**b**), and T4 versus TBG concentrations in CH-C patients (**c**) and FP (**d**) in 11 years CH NBS, compared to the healthy reference population. 95% RIs for T4 (56–118 nmol/L blood) and TBG (116–271 nmol/L blood) are displayed as black lines. The current COV (≤17) for the T4/TBG ratio is displayed as a dashed line. Orange circles (**a**–**d**): the healthy reference population ($n = 1,670$). Blue diamonds (**a**, **c**): CH-C patients that were referred due to a T4/TBG ratio ≤17 ($n = 54$). Yellow diamonds (**a**, **c**): CH-C patients that were referred due to a T4 ≤−3 SD ($n = 29$). Gray circles (**b**, **d**): FP that was referred due to a T4/TBG ratio ≤17 ($n = 1,196$). Turquoise circles (**b**): FP that was referred due to a T4 ≤−3 SD ($n = 788$). Green and purple circles (**d**): same group as the turquoise referrals, but now divided by a TBG concentration within or above the RI (green, $n = 424$) or below the RI (purple, $n = 364$). CH-C, central congenital hypothyroidism; FP, false-positive; NBS, newborn screening; RI, reference interval; T4, thyroxine; TBG, thyroxine-binding globulin.

DBS T4 and TBG concentrations and the calculated T4/TBG ratio of CH-C patients and false-positive referrals, and compared these values to those of healthy newborns. CH-C patients can clearly be distinguished from the reference population based on their T4 concentration and T4/TBG ratio, confirming the usefulness of this approach. Although the large number of false-positive referrals (low specificity) forms the major drawback of the current NBS algorithm to detect also CH-C patients, it should be noted that between 1995 and 2015 per year on average 0.8 patients with CH-C were missed by NBS (unpublished data from a national voluntary registration for pediatric disorders). In addition, in a recent study, we showed that this number may be even higher (2–3 missed CH-C patients per year) [11]. Hence, the sensitivity of the Dutch screening algorithm to detect CH-C patients is not 100%. The large overlap between CH-C patients and false-positive referrals was confirmed (shown in Fig. 1), but overlap between these groups and the reference population barely existed as the reference population did not contain any false-positive referrals and obviously no patients. We postulate that false-positive referrals with a DBS T4 concentration in the lower end of the RI could be prevented by altering the limit at which the T4/TBG ratio is calculated. Currently, a TBG is measured when T4 is ≤ -1.6 SD (T4 median 59, IQR 55–62 nmol/L blood). Expressing the DBS T4 concentration as a SD value has served for years as a means for harmonization of different radioactive T4 immunoassays between the regional Dutch laboratories. This, however, is not necessary anymore as the same T4 immunoassay is in use in all Dutch screening laboratories since 2003. Although two of the detected CH-C patients had a T4 of 60 nmol/L blood, our data show that no newborns with an absolute T4 concentration of >60 nmol/L blood had CH, suggesting that a T4 COV of ≤ 60 nmol/L blood might replace the currently used ≤ -1.6 SD as the limit below which a TBG measurement is required, ultimately preventing 10% of T4/TBG ratio false-positive referrals.

Also, a large group false-positives were referred based on their T4 alone (T4 ≤ -3 SD, TBG >40 nmol/L blood). Currently, TBG does not play a direct role in the referral process of these newborns (except if TBG <40 nmol/L blood). We, however, observed that in 46% ($n = 364$) of these false-positive referrals the TBG concentration was below the TBG RI (<116 nmol/L blood; shown in Fig. 1d), but >40 nmol/L, which is suggestive of the relatively common benign partial TBG deficiency (prevalence 1:4,000). A partial or complete TBG deficiency results in (relatively) low total T4 concentrations [12]. In contrast, only one CH-C patient with a TBG concentration below the RI was referred (shown in Fig. 1c) based on T4 ≤ -3 SD (and exhibited a normal T4/TBG ratio of 20.3). This seems to be an exceptional case of combined CH-C and partial TBG deficiency. We, therefore, conclude that by taking partial TBG deficiency into account, approximately half of T4 false-positive referrals may be prevented at the cost of missing one CH-C patient (in 11 years of NBS). The remaining group of false-positive referrals unfortunately is very similar to CH-C (normal TBG and a low to subnormal T4), and its number cannot be reduced with currently available information from our evaluation study of the CH NBS algorithm and this neonatal RIs study.

In conclusion, this study provides newly established RIs for the NBS parameters T4, TSH, TBG, and T4/TBG ratio as used in the Dutch NBS. Separate RIs for sex and NBS sampling age are not of added value for the evaluated screening parameters DBS T4, TSH, and TBG concentrations, and the T4/TBG ratio. The main finding of our study is that by using the TBG RIs to recognize partial TBG deficiency, approximately half of the low T4 false-positive referrals (due to partial TBG deficiency) may be prevented. This suggested adaptation of the Dutch NBS for CH may lead to an improvement of the PPV while maintaining the current sensitivity for CH-C detection. Regular evaluation and further research for improvement of the algorithm are required to further improve the PPV.

Acknowledgements

Statement of Ethics

This study has been approved with a signed statement by the NBS Privacy Committee of TNO. Parents participating in the neonatal screening program of the Netherlands give consent for both participation and anonymous use of leftover heel puncture blood for scientific research unless they actively declined.

Conflict of Interest Statement

Annet M. Bosch has received a speaker's fee from Nutricia and has been a member of advisory boards for Biomarin. All other authors have no conflicts of interest to declare.

References

1 LaFranchi SH. Newborn screening strategies for congenital hypothyroidism: an update. J Inherit Metab Dis. 2010;33(Suppl 2):S225–33.

2 Ford G, LaFranchi SH. Screening for congenital hypothyroidism: a worldwide view of strategies. Best Pract Res Clin Endocrinol Metab. 2014;28(2):175–87.

3 Kilberg MJ, Rasooly IR, LaFranchi SH, Bauer AJ, Hawkes CP. Newborn screening in the US may miss mild persistent hypothyroidism. J Pediatr. 2018;192:204–8.

4 Lanting CI, van Tijn DA, Loeber JG, Vulsma T, de Vijlder JJ, Verkerk PH. Clinical effectiveness and cost-effectiveness of the use of the thyroxine/thyroxine-binding globulin ratio to detect congenital hypothyroidism of thyroidal and central origin in a neonatal screening program. Pediatrics. 2005;116(1):168–73.

5 Stroek K, Heijboer AC, Bouva MJ, van der Ploeg CPB, Heijnen MA, Weijman G, et al. Critical evaluation of the newborn screening for congenital hypothyroidism in the Netherlands. Eur J Endocr. 2020;183(3):265–73.

6 Boelen A, van Veen M, Verkerk PH, Diependaal G, Loeber G, Elvers B, et al. Measuring free thyroxine levels in neonatal heel-prick samples. Clin Chim Acta. 2013;423:51–5.

7 Horowitz GL, Altaie S, Boyd JC, Ceriotti F, Garg U, Horn P, et al. Clinical and laboratory standards institute (CLSI). Defining, establishing, and verifying reference intervals in the clinical laboratory; approved guideline-third edition. CLSI document EP28–A3c (ISBN 1-56238-682-4). Wayne, PA: Clinical and Laboratory Standards Institute; 2008.

8 Sinton TJ, Cowley DM, Bryant SJ. Reference intervals for calcium, phosphate, and alkaline phosphatase as derived on the basis of multi-channel-analyzer profiles. Clin Chem. 1986;32(1 Pt 1):76–9.

9 Fisher DA, Nelson J, Carlton EI, Wilcox R. Maturation of human hypothalamic-pituitary-thyroid function and control. Thyroid. 2000;10:229–34.

10 Ain K, Mori Y, Refetoff S. Reduced clearance rate of thyroxine-binding globulin (TBG) with increased sialylation: a mechanism for estrogen-induced elevation of serum TBG concentration. J Clin Endocrinol Metabol. 1987;65:689–96.

11 van Iersel L, van Santen HM, Zandwijken GRJ, Zwaveling-Soonawala N, Hokken-Koelega ACS, van Trotsenburg ASP. Low FT4 concentrations around the start of recombinant human growth hormone treatment: Predictor of congenital structural hypothalamic-pituitary abnormalities? Horm Res Paediatr. 2018;89(2):98–107.

12 Pappa T, Ferrara AM, Refetoff S. Inherited defects of thyroxine-binding proteins. Best Pract Res Clin Endocrinol Metab. 2015;29(5):735–47.
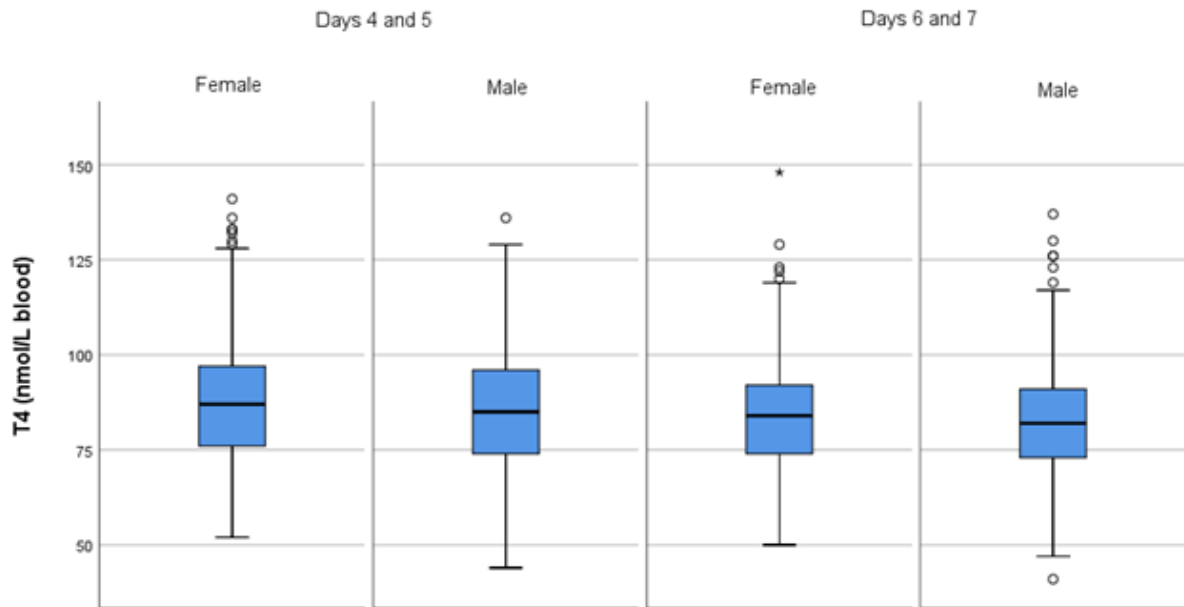
Fig. S1. T4 concentrations of reference subjects divided by subgroups: females and males with

newborn screening sampling days 4 and 5, and days 6 and 7.

Measurements were performed in (full) blood.

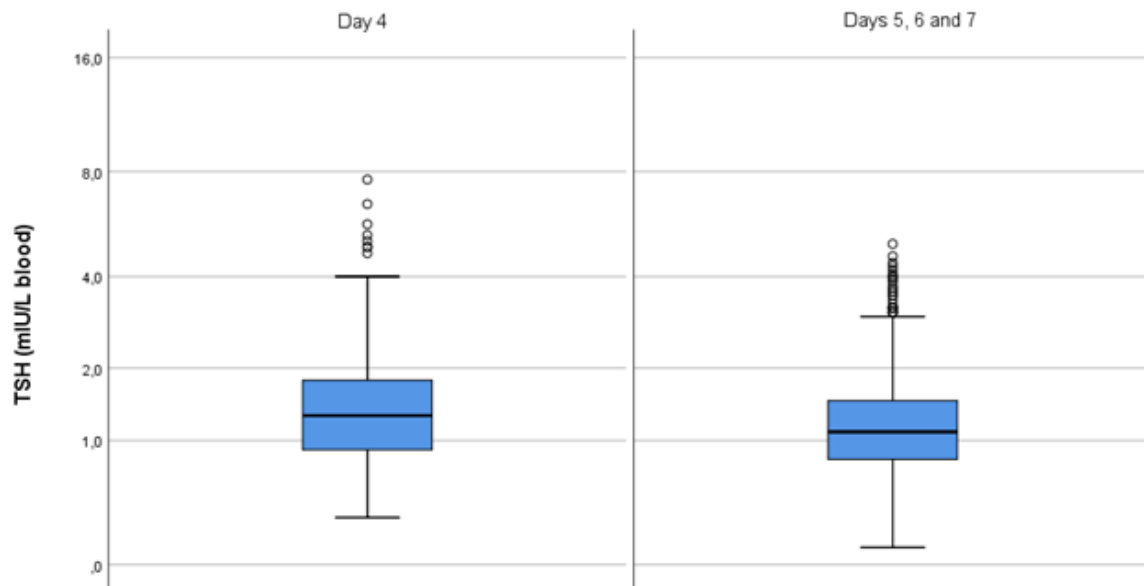Data are presented as boxplots.

Fig. S2. TSH concentrations of reference subjects divided by subgroup: newborn screening sampling

day 4 and days 5, 6 and 7.

Measurements were performed in (full) blood.

Data are presented as boxplots.

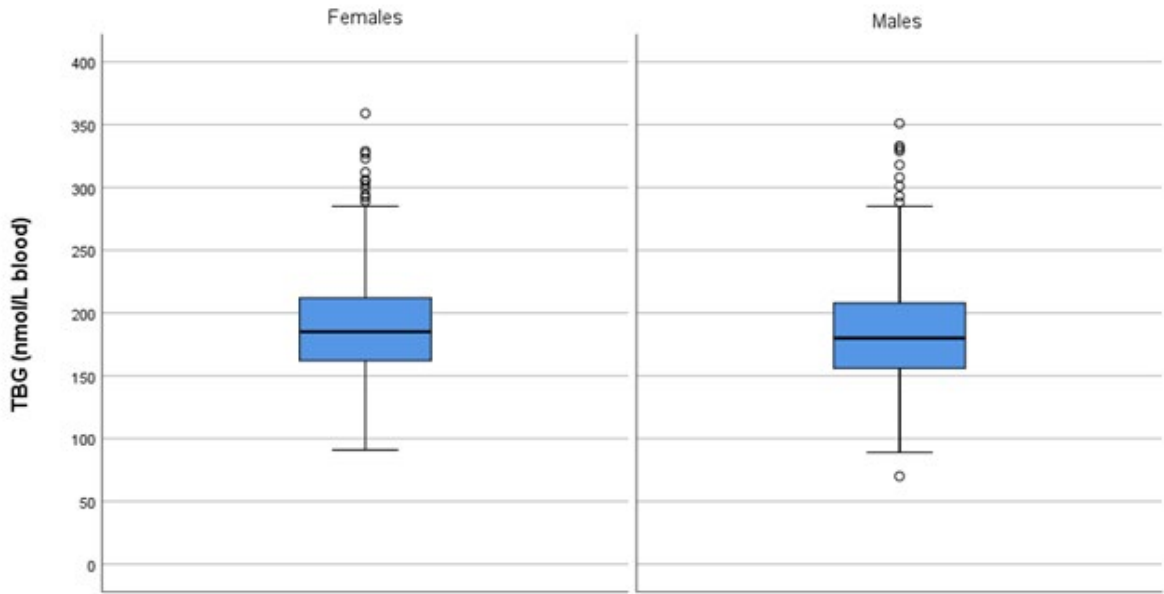TSH concentrations are displayed on a logarithmic scale.

Fig. S3. TBG concentrations of reference subjects divided by subgroup: females and males.

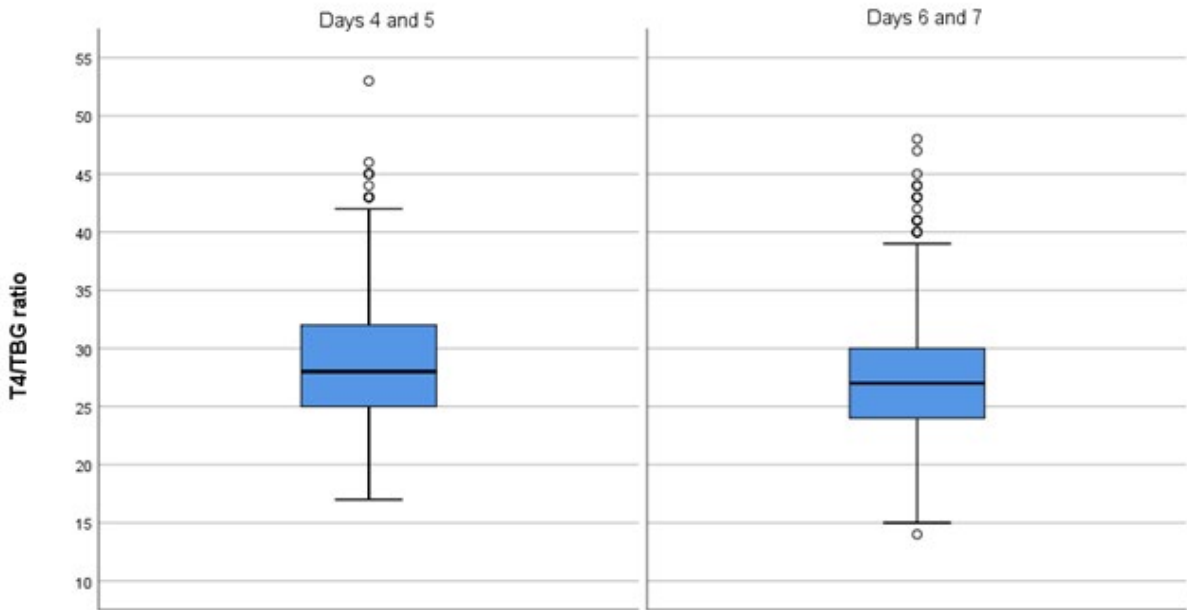Measurements were performed in (full) blood.

Data are presented in boxplots.

Fig. S4. T4/TBG ratios of reference subjects divided by subgroup: NBS sampling days 4 and 5, and

days 6 and 7.

	Sampling age (days)	Female			Male			Total		
		n	Average	Spread	n	Average	Spread	n	Average	Spread
T4*	4	167	88	16	180	86	15	347	87	16
	5	242	88	16	253	85	17	495	87	17
	6	219	85	13	253	82	15	472	83	14
	7	163	83	14	193	82	14	356	82	14
	Total	791	86	15	879	84	15	1670	85	15
TSH*	4	167	1.3	0.9-1.8	180	1.3	0.9-2.0	347	1.3	0.9-1.8
	5	242	1.2	0.8-1.7	253	1.1	0.7-1.6	495	1.1	0.8-1.6
	6	219	1.1	0.8-1.5	253	1.1	0.8-1.6	472	1.1	0.8-1.5
	7	163	1.1	0.8-1.4	193	1.1	0.8-1.5	356	1.1	0.8-1.5
	Total	791	1.2	0.8-1.6	879	1.1	0.8-1.6	1670	1.1	0.8-1.6
TBG*	4	167	186	38	180	178	35	347	182	37
	5	242	189	40	253	183	41	495	186	41
	6	219	189	39	253	185	44	472	187	42
	7	163	190	36	193	187	39	356	188	38
	Total	791	189	39	879	183	40	1670	186	40
T4/ TBG ratio *	4	167	29	6	180	30	6	347	29	6
	5	242	29	5	253	29	5	495	29	5
	6	219	28	5	253	28	6	472	28	5
	7	163	27	5	193	27	5	356	27	5
	Total	791	28	5	879	28	5	1670	28	5

Data are presented in boxplots.

Table S1. Descriptive statistics of measured T4, TSH and TBG concentrations, and the calculated T4/TBG ratio for all reference subjects combined and divided by sex and NBS sampling age.

Abbreviations: NBS = newborn screening, T4 = thyroxine (nmol/L blood), TBG = thyroxine binding globuline (nmol/L), TSH = thyroid stimulating hormone (mIU/L blood).

*The average and spread are presented as mean and standard deviation for T4, TBG and T4/TBG ratio, and as median and interquartile range (Q1 and Q3) for TSH.

Table S2. Effect of sex and NBS sampling age on measured concentrations of T4, TSH and TBG, and the calculated T4/TBG ratio.

	df	F-statistic	p-value[a]	Partial η^2
T4				
Sex	1	8.107	0.004	0.005
NBS sampling age	3	9.101	0.000[b]	0.016
Sex*NBS sampling age	3	0.227	0.878[c]	0.000
TSH (Log$_{10}$)				
Sex	1	0.39	0.844	0.000
NBS sampling age	3	11.854	0.000[d]	0.021
Sex*NBS sampling age	3	1.467	0.222[c]	0.003
TBG				
Sex	1	7.350	0.007	0.004
NBS sampling age	3	1.529	0.205	0.003
Sex*NBS sampling age	3	0.336	0.799[c]	0.001
T4/TBG ratio				
Sex	1	0.762	0.383	0.000
NBS sampling age	3	13.592	0.000[e]	0.024
Sex*NBS sampling age	3	0.299	0.826[c]	0.001

NBS = newborn screening, T4 = thyroxine (nmol/L blood), TBG = thyroxine binding globuline (nmol/L),

TSH = thyroid stimulating hormone (mIU/L blood).

[a]Two-way ANOVA

[b]Post-hoc analysis (Tukey HSD) shows a significant difference of T4 concentrations between day 4

and days 6 (p = 0.006) and 7 (p = 0.000), and between day 5 and days 6 (p = 0.005) and 7 (p = 0.000);

Identified homogeneous subsets: 1. days 4 and 5 (p = 0.722); 2. days 6 and 7 (p = 0.995).

[c]A P-value >0.05 states that no interaction effects between both variables was found.

[d]Post-hoc analysis (Tukey HSD) shows a significant difference of TSH (Log$_{10}$) concentrations between

day 4 and days 5, 6 and 7 (p = 0.000); Identified homogeneous subsets: 1. days 4; 2. days 5, 6 and 7 (p

= 0.500).

[e]Post-hoc analysis (Tukey HSD) shows a significant difference of the T4/TBG ratio between day 4 and

days 6 (p = 0.000) and 7 (p = 0.000), and between day 5 and days 6 (p = 0.030) and 7 (p = 0.000).

Identified homogeneous subsets: 1. days 4 and 5 (p = 0.334); 2. days 6 and 7 (p = 0.179).