

Critical evaluation of the newborn screening for congenital hypothyroidism in the Netherlands

Kevin Stroek¹, Annemieke C Heijboer^{1,2}, Marelle J Bouva³, Catharina P B van der Ploeg⁴, Marie-Louise A Heijnen⁵, Gert Weijman⁶, Annet M Bosch⁷, Robert de Jonge⁸, Peter C J I Schielen³, A S Paul van Trotsenburg⁹ and Anita Boelen¹

¹Endocrinology Laboratory, Department of Clinical Chemistry, Amsterdam Gastroenterology & Metabolism, Amsterdam UMC, University of Amsterdam, Amsterdam, the Netherlands, ²Endocrinology Laboratory, Department of Clinical Chemistry, Amsterdam Gastroenterology & Metabolism, Amsterdam UMC, Vrije Universiteit, Amsterdam, the Netherlands, ³Reference Laboratory Neonatal Screening, Center for Health protection, National Institute for Public Health and the Environment, Bilthoven, the Netherlands, ⁴Netherlands Organization for Applied Scientific Research TNO, Department of Child Health, Leiden, the Netherlands, ⁵Center for Population Screening, National Institute for Public Health and the Environment, Bilthoven, the Netherlands, ⁶Department of Vaccine Supply and Prevention, National Institute for Public Health and the Environment, Bilthoven, the Netherlands, ⁷Division of Metabolic Disorders, Department of Pediatrics, Amsterdam UMC, University of Amsterdam, Amsterdam, the Netherlands, ⁸Department of Clinical Chemistry, Amsterdam UMC, Vrije Universiteit & University of Amsterdam, Amsterdam, the Netherlands, and ⁹Department of Paediatric Endocrinology, Emma Children's Hospital, Amsterdam UMC, University of Amsterdam, Amsterdam, the Netherlands

Correspondence should be addressed to A Boelen
Email
a.boelen@amsterdamumc.nl

Abstract

Objective: Congenital hypothyroidism (CH) is defined as thyroid hormone deficiency at birth due to disorders of the thyroid gland (thyroidal CH, CH-T), or the hypothalamus or pituitary (central CH, CH-C). The Dutch Newborn Screening (NBS) strategy is primarily based on determination of thyroxine (T4) concentrations in dried blood spots followed, if necessary, by thyroid-stimulating hormone (TSH) and thyroxine-binding globulin (TBG) measurement enabling detection of both CH-T and CH-C. A calculated T4/TBG ratio serves as an indirect measure for free T4. A T4/TBG ratio ≤ 17 in a second heel puncture is suggestive of CH-C.

Design and methods: In the present study, we evaluated 11 years of Dutch CH NBS using a database of referred cases by assessing the contribution of each criterion in the unique stepwise T4-TSH-TBG NBS algorithm.

Results: Between 2007 and the end of 2017, 1 963 465 newborns were screened in the Netherlands. Use of the stepwise algorithm led to 3044 referrals and the identification of 612 CH cases, consisting of 496 CH-T, 86 CH-C, and 30 CH of unknown origin diagnoses. We detected 62.8% of CH-C cases by the T4/TBG ratio in the second heel puncture. The positive predictive value (PPV) of the stepwise T4-TSH-TBG NBS algorithm was 21.0%.

Conclusion: This evaluation shows that the Dutch stepwise T4-TSH-TBG NBS algorithm with a calculated T4/TBG ratio is of great value for the detection of both CH-T and CH-C in the Netherlands, at the cost of a lower PPV compared to TSH-based NBS strategies.

European Journal of
Endocrinology
(2020) **183**, 265–273

Introduction

Congenital hypothyroidism (CH), that is, thyroid hormone deficiency already present at birth, can be caused by defective thyroid gland development or

function (primary or thyroidal CH, abbreviated CH-T), or by insufficient thyroid gland control by the pituitary or hypothalamus (central CH or CH-C) (1). The prevalence

of CH-T and CH-C in the Netherlands has been estimated at 1 in 3017 and 1 in 16 404 live newborns, respectively (2). Worldwide, the incidence of CH has been reported to vary between 1:1400 and 1:2800 (3). Untreated, CH often results in irreversible mental and motor retardation. However, since the introduction of newborn screening (NBS) programs for CH in the 1970s, brain damage can be effectively prevented (4, 5, 6, 7).

Historically, most NBS programs employed a primary thyroxine (T4)-based strategy, aiming at detection of CH-T. In addition, T4-based strategies also detected CH-C. Following a markedly improved sensitivity of the thyroid stimulating hormone (TSH) assay, many existing programs switched to primary TSH testing, with many new programs adopting this strategy. By elimination of low T4 false-positives, this approach improved NBS specificity (8). However, newborns with CH-C – who typically have a low T4 in combination with a normal TSH concentration – were missed. Consequently, the reported CH-C incidence decreased (4). Currently, in the United States of America 22 states employ a primary TSH strategy, while 20 use a primary T4 with reflex TSH (performed in original sample if T4 is abnormal) testing strategy. The remaining nine measure T4 and TSH simultaneously (9).

In the Netherlands a unique NBS strategy is in use since 1981. The first step is T4 measurement, followed by TSH measurement in the lowest 20% T4 values of the day. Since 1995, T4-binding globulin (TBG) is measured in the lowest 5% daily T4 values. From these values, a modified T4/TBG ratio is calculated as an indirect measure of the free T4 (fT4) concentration (2). TBG measurement lowers the number of T4 false-positives due to (partial) TBG deficiency, while the T4/TBG ratio enables detection of newborns with possible CH-C. In an evaluation of the Dutch NBS covering the period between 1995 and 2000, it was shown that adding the T4/TBG ratio clearly improved the detection rate of CH-C at an acceptable cost. However, the downside was an increased false-positive rate (2), which is known to cause anxiety or depression in parents and may lead to a dysfunctional parent-child relationship (10). Recently, the importance of detection of CH-C was confirmed. Zwaveling *et al.* showed that a majority of CH-C patients have moderate CH with initial fT4 below 10 pmol/L (11, 12) and suggested a low mortality rate due to pituitary insufficiency in early-detected and early-treated CH-C patients (13).

Recently, the Dutch NBS experienced a further increase in the number of false-positive due to aberrant T4/TBG ratios. The aim of the present study was to evaluate the current Dutch NBS for CH by assessing the contribution

of each criterion in the stepwise T4-TSH-TBG algorithm to the yield of the program.

Methods

NBS algorithm for CH

Heel puncture blood samples are collected on filter paper between 72 and 168 h after birth and sent to one of the five regional screening laboratories in the Netherlands. In every sample, the total T4 concentration is measured in nmol/L blood and expressed as a s.d. relative to the daily mean (14). In samples with the 20% lowest T4 concentrations of the day ($T4 \leq -0.8$ s.d.), TSH is measured. In the 5% lowest T4 values ($T4 \leq -1.6$ s.d.), TBG is also measured (at time of implementation of the T4/TBG ratio it was established (unpublished national report) that this s.d. was the most discriminative for detection of CH-C, with an increasing risk of missed cases with lower s.d.). The T4/TBG ratio is calculated as follows: $((T4 \text{ (in s.d.)} + 5.1) \times 1000) / (TBG \text{ (nmol/L blood)})$ (2). All neonates with an abnormal NBS result are referred and evaluated by a pediatric endocrinologist or general pediatrician. Figure 1 gives a schematic overview of the referral criteria used in the Dutch NBS for CH. Historically, all newborns with a $T4 \leq -3.0$ s.d. were immediately referred to enable diagnosing severe CH as quickly as possible, without awaiting TSH and TBG results (original referral criterion-1). Because this led to unnecessary referral of children with a low total T4 due to TBG deficiency, since July 2012 a $T4 \leq -3.0$ s.d. only led to referral when the TBG concentration was >40 nmol/L blood (referred to as 'TBG rule'). Since this measure has eliminated false-positive referrals due to TBG deficiency completely and does not influence the number of true-positives, we applied the TBG rule to the entire evaluated period (2007–2017), creating a clear overview of the performance of the first referral criterion (criterion-1: $T4 \leq -3.0$ s.d. and TBG >40 nmol/L blood). Newborns with an abnormal TSH concentration are referred based on referral criterion-2: $T4 \leq -0.8$ s.d. and $TSH \geq 22$ mIU/L blood. A borderline TSH concentration ($8 \leq TSH < 22$ mIU/L blood) and/or a borderline T4 ($-3.0 < T4 \leq -1.6$ s.d.) in combination with a T4/TBG ratio ≤ 17 are considered inconclusive results, leading to a request for a second blood sample. In all second heel punctures T4, TSH and TBG are measured, and subsequently all newborns with a $T4 \leq -3.0$ s.d. are immediately referred (referral criterion-3: $T4 \leq -3.0$ s.d.). Newborns with a borderline ($8 \leq TSH < 22$ mIU/L blood) or abnormal TSH (≥ 22 mIU/L blood) concentration (referral

First heel puncture	TSH**								
T4 in all samples TSH in T4 values ≤ -0.8 SD (20%) TBG in T4 values ≤ -1.6 SD (5%)	Criterion-2 TSH ≥ 22			$8 \leq \text{TSH} < 22$			TSH < 8		
Screening result	CH-T	CH-C	FP	CH-T	CH-C	FP	CH-T	CH-C	FP
Criterion-1: T4 ≤ -3.0 SD (& TBG $> 40^*$)	222	0	3	6	0	12	10	13	714
$-3.0 < \text{T4} \leq -1.6$ SD & T4/TBG ratio ≤ 17	0	0	0	Inconclusive***			Inconclusive***		
$(-3.0 < \text{T4} \leq -1.6$ SD & T4/TBG ratio > 17) OR T4 SD > -1.6	165	0	24	Inconclusive***			Normal		

Abnormal	Inconclusive 1st heel puncture: recall for 2nd heel puncture	Normal: no further action
-----------------	--	----------------------------------

Second heel puncture	TSH**								
T4, TSH & TBG in all samples	Criterion-4 TSH ≥ 22			Criterion-4 $8 \leq \text{TSH} < 22$			TSH < 8		
Screening result	CH-T	CH-C	FP	CH-T	CH-C	FP	CH-T	CH-C	FP
Criterion-3: T4 ≤ -3.0 SD*	9	0	0	4	0	3	2	16	62
Criterion-5: $-3.0 < \text{T4} \leq -1.6$ SD & T4/TBG ratio ≤ 17	0	0	0	0	0	0	10	54	1196
$(-3.0 < \text{T4} \leq -1.6$ SD & T4/TBG ratio > 17) OR T4 SD > -1.6	12	0	10	48	0	155	Normal		

Figure 1

A schematic representation of the Dutch stepwise T4-TSH-TBG NBS algorithm for CH, with for all referral criteria the referral rates in the period 2007–2017 included. CH-C, central congenital hypothyroidism; CH-T, thyroidal congenital hypothyroidism; FP, false-positives; NBS, newborn screening; T4, total thyroxine (nmol/L blood); TBG, thyroxine-binding globulin (nmol/L blood); TSH, thyroid-stimulating hormone (mIU/L blood); Total number of referrals after an abnormal (1169) or inconclusive (15, not included in this figure) first heel puncture or abnormal second (1581) heel puncture: 2765. *A TBG $\leq 40 + \text{T4} \leq -3.0$ s.d. is a reason for not referring in only the first heel puncture, because of TBG deficiency (TBG rule). In the second heel puncture, all neonates with a T4 ≤ -3.0 s.d. are referred. **Newborns with a birth weight of ≤ 2500 g and a gestational age of ≤ 36 (+0 days) weeks and infants that received the heel puncture on or later than day 60 of life are subject to the decision rule for premature newborns and referred based on only TSH (only measured in newborns with a T4 ≤ -0.8 s.d.). These results are not included in this scheme. ***Inconclusive results are not reported, since a definitive result will be available from the second heel puncture.

criterion-4: TSH ≥ 8 mIU/L blood) and/or a borderline T4 ($-3.0 < T4 \leq -1.6$ s.d.) in combination with a T4/TBG ratio ≤ 17 (referral criterion-5: T4 $-3.0 < T4 \leq -1.6$ s.d. and T4/TBG ratio ≤ 17) are also referred. Premature newborns often have a low T4 concentration associated with prematurity or illness (12). Prematurity in the context of our CH NBS is defined as newborns with both a birth weight ≤ 2500 g and a gestational age of ≤ 36 (+0 days) weeks. These newborns are referred based on TSH (only measured in newborns with a T4 ≤ -0.8 s.d.), with the previously mentioned cut-off values. Low T4 values will not be taken into account. This decision rule also applies to infants that received the heel puncture \geq day 60 of life.

NBS assays

All assays were performed according to the manufacturer's protocol. Total T4 was measured by an immunoassay (Perkin Elmer®). TSH was measured using an immunoassay supplied by Perkin Elmer® or Brahms (Thermo Fisher), depending on the screening laboratory. The TBG measurement was performed with a radioimmunoassay (RIA, Eiken®, Tokyo, Japan) until the end of 2011. From January 1, 2012, this RIA was replaced by an enzyme-linked immuno sorbent assay (ELISA, Monobind Inc.®, Lake Forrest, CA, USA) (14). To gain insight into the performance of assays over the years, all results of all NBS parameters are regularly reported by the reference laboratory for neonatal screening at the National Institute for Public Health and the Environment (RIVM, Bilthoven). Our ISO15189-accredited laboratories actively participate in the Newborn Screening Quality Assurance Program (NSQAP) by the Centers of Disease Control and Prevention (Atlanta, Georgia) and the newborn screening quality assurance program of the Reference Institute for Bioanalytics, RfB (Bonn, Germany).

Data management

The Netherlands Organization of Applied Scientific Research (TNO, department of Child Health) registered results of the NBS for CH and subsequent diagnostic findings in a database for monitoring and quality improvement. Unfortunately, missing data are a recurring problem, since completeness of these data depend on the response of the treating pediatricians. Use of the data in this study was approved by the Privacy Committee of TNO. Newborns are classified as no CH (false-positives), CH-T, CH-C, or CH of unknown origin (true-positives). The latter classification is made when no clear CH-T

or CH-C diagnosis has been made. False-negatives are currently not centrally reported but estimated to be at least 2–3 per year. For analyses in the current study, data from the period January 1, 2007 to December 31, 2017 were used.

Statistical analysis

SPSS Statistics version 25 was used to calculate NBS characteristics and positive predictive value (PPV). Since there is no structural neither complete registration of false-negatives, the sensitivity and negative predictive value (NPV) could not be calculated. Specificity could not be calculated with reasonable accuracy due to the effect of missing data. Each individual referral criterion in the NBS algorithm was evaluated separately.

To assess the performance of the T4, TSH and TBG assays, measured concentrations per 3 months were collected from each laboratory from 2008 up to and including the second quarter of 2018. Medians of these quarterly concentrations were calculated, and from the means and standard deviations of these medians the coefficient of variation (CV%) of the median quarterly concentrations was determined. Also, the quarterly inter-laboratory variation (reported as a CV%) was collected, and for relevant periods the mean CV% was calculated.

Results

Population and NBS characteristics

Between January 1, 2007 and December 31, 2017, samples of 1 963 465 Dutch newborns were analysed for NBS for CH. In the analysis, 3264 newborns had an abnormal CH result in their first or second heel puncture and were referred to a pediatrician. In addition, 15 newborns had an inconclusive first heel puncture result and, because they were hospitalised, directly underwent further diagnostic evaluation (eight had CH-T, three CH-C and four did not have CH). Thus, in these children, a second heel puncture was not performed, but referral based on the algorithm is assumed. In total, 3279 children were referred. In 167 newborns pediatric diagnostic data were incomplete or missing. Hence, these results were not taken into account. Of the remaining 3112 referrals, 68 were referred following the decision rule for premature newborns (38 CH-T diagnoses and 30 false-positives). After exclusion of referrals following the decision rule for premature newborns, 3044 remaining newborns were

referred based on the stepwise T4-TSH-TBG algorithm, of whom 612 were diagnosed with CH. The total number of CH cases sums at 650, with a calculated 11-year incidence of CH of 1:3020. The incidences of CH-T and CH-C were 1:3677 and 1:22 829, respectively.

Referral rates from 2007 to 2017

Use of the stepwise algorithm led to 3044 referrals and the identification of 612 CH cases (20.1% of referrals), consisting of 496 CH-T, 86 CH-C, and 30 CH of unknown origin diagnoses (Fig. 2). The remaining 2432 referrals were false-positives (79.9%, including 249 TBG deficiencies referred before July 2012). After excluding TBG deficiency and CH of unknown origin, a total number of 2765 referrals (2183 false-positives, 496 CH-T and 86 CH-C) remained, consisting of 1169 abnormal and 15 inconclusive first heel puncture results and 1581 abnormal second heel puncture results. A total of 1490 neonates (53.9% of all referrals) were referred because of referral criteria 1–4 (T4 and/or TSH-based), while 1260 neonates (45.6% of all referrals) were referred because of referral criterion-5 (T4/TBG ratio-based). The 15 newborns (0.5% of all referrals) that received clinical evaluation

after an inconclusive first heel puncture result could not be related to a referral criterion of the stepwise algorithm. The individual contribution of each referral criterion is discussed subsequently and summarised in the scheme in Fig. 1.

True-positive results following referral criteria 1–4

Through the first and second heel puncture, 478 CH-T (96.4% of all CH-T) and 29 CH-C patients (33.7% of all CH-C) were referred based on T4 and/or TSH criteria. Of these, 460 CH-T patients (92.7% of all CH-T) were positive for a TSH-based criterion: criterion-2, $n=387$ (with 222 also positive for criterion-1), and criterion-4, $n=73$ (with 13 also positive for criterion-3), while 18 were referred based on only T4-criteria 1 or 3. The 29 CH-C patients had normal TSH results and were referred based on referral criteria 1 or 3.

False-positive results following referral criteria 1–4

In total, 983 neonates turned out to have a false-positive result due to T4 and/or TSH criteria (45.0% of all false-positives). Of them, 788 (36.1% of all false-positives) were referred because of T4-based referral criteria 1 ($n=726$) or 3 ($n=62$). The remaining 195 false-positives (8.9% of all false-positives) were referred because of TSH-based criterion-2, $n=27$ (with 3 also positive for criterion-1), and criterion-4, $n=168$ (with 3 also positive for criterion-3, and mostly ($n=158$) consisting of neonates with a TSH between 8 and 22 mIU/L blood).

True and false-positive results following referral criterion-5

Ten patients with CH-T (2.0% of all CH-T) and 54 with CH-C (62.8% of all CH-C) were detected by the T4/TBG ratio in the second heel puncture. The number of false-positive referrals was 1196 (54.8% of all false-positives). In the 6-year period (between January 1, 2012 and December 31, 2017) when TBG was measured by ELISA, 875 false-positive referrals occurred, compared to 321 in 2007–2011 when the RIA was used. In Table 1 the referral numbers of each year are displayed. True and false-positive numbers differed substantially per year with a rising number of false-positive referrals between 2012 and 2015 and a decrease in 2016 and 2017. In the year 2017, 72% of false-positive referrals following this criterion had a T4/TBG ratio between 15 and 17 and an average T4/TBG ratio of 15.2. In 2012–2016, these numbers were 41–53%

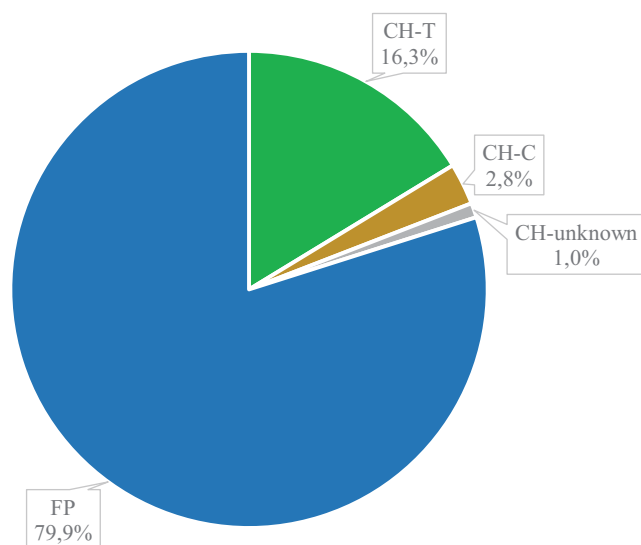


Figure 2

Conclusive referral rates of the stepwise T4-TSH-TBG NBS algorithm. CH-C, central congenital hypothyroidism; CH-T, thyroidal congenital hypothyroidism; CH-unknown, CH of unknown origin; FP, false-positives; NBS, newborn screening; T4, total thyroxine; TBG, thyroxine-binding globulin; TSH, thyroid-stimulating hormone. A full color version of this figure is available at <https://doi.org/10.1530/EJE-19-1048>.

Table 1 Referrals in the years 2007–2017 following the T4/TBG ratio.

TBG assay	RIA					ELISA						Total
	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	
CH-C	1	5	4	8	8	3	7	4	4	7	3	54
CH-T	0	1	0	4	2	0	0	0	0	1	2	10
False-positives	20	45	79	98	79	120	137	191	253	145	29	1196
Total referrals	21	51	83	110	89	123	144	195	257	153	34	1260
PPV (%)	4.7	11.8	4.8	10.9	11.2	2.4	4.9	2.1	1.6	5.2	14.7	3.4

CH-C, central congenital hypothyroidism; CH-T, thyroïdal congenital hypothyroidism; ELISA, enzyme-linked immuno sorbent assay; PPV, positive predictive value; RIA, radio-immunoassay; T4, total thyroxine; TBG, thyroxine-binding globulin.

false-positive referrals and average T4/TBG ratio of 14.4–14.9, respectively.

Positive predictive value of our NBS referral criteria

Based on the aforementioned numbers, in the 11-year period of NBS the PPV of the combined T4- and TSH-based referral criteria (criteria 1–4) is 34.0%. Because a number of neonates were positive for more than one referral criterion, PPVs per criterion are not reported. The PPV of T4/TBG ratio-based referral criterion-5 was 9.3% in the RIA-period and 3.4% in the ELISA-period, with an overall PPV of 5.1%. We emphasise that these PPV calculations do not contain the 15 neonates with an inconclusive first heel puncture result. All referral criteria combined (now including the 15 neonates) have a PPV of 21.0%.

Performance of the assays

The quarterly medians of the T4, TSH and TBG concentrations from 2008 up to and including the second quarter of 2018 are shown in Fig. 3. The variation over time and the inter-laboratory variation are summarised in Table 2.

The median quarterly T4 concentrations were consistently between 79 and 94 nmol/L blood (variation over time 4.3%), with also a low inter-laboratory variation (3.9%).

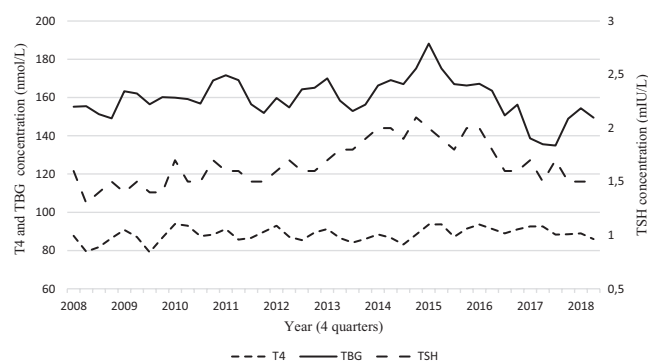
The median quarterly TSH concentrations varied between 1.3 and 2.1 mIU/L blood (variation over time 12.3%), with a higher inter-laboratory variation (13.8%). Since the fourth quarter of 2016, all five laboratories used the same immunoassay resulting in a lower inter-laboratory variation in this period (9.4%).

The median quarterly TBG concentrations varied from 135 to 188 nmol/L (variation over time 6.6%). From 2008 until the end of 2011, the median quarterly TBG concentrations varied between 149 and 172 nmol/L blood (Eiken® assay; variation over time 4.1%). From

2012 onwards, median quarterly TBG concentrations varied between 135 and 188 nmol/L (Monobind® assay; variation over time 7.8%). Inter-laboratory variation (4.8%) was lower in the Eiken® than in the Monobind® assay (6.7%).

Discussion

In this study we evaluated 11 years of the Dutch T4-based NBS program for CH. The incidence of both CH-T (1:3677) and CH-C (1:22 829) were lower than previously reported for the Netherlands by Lanting *et al.* (CH-T 1:3017; CH-C 1:16 404) (2). This may be explained by missing information in the database about some of the referrals, potentially leading to an underestimation of the incidence. From a national voluntary registration for pediatric disorders, it is known that at least 0.8 cases of CH-C and 0 to 2 cases of CH-T per year between 1995 and 2015 were missed by NBS (unpublished data). The actual false-negative count is likely to be higher, as also

**Figure 3**

Median of mean quarterly concentrations of T4, TBG (nmol/L blood) and TSH (mIU/L blood) of each laboratory in the NBS period 2008 until the end of the second quarter of 2018. NBS, newborn screening; T4, total thyroxine; TBG, thyroxine-binding globulin; TSH, thyroid-stimulating hormone.

Table 2 Performance of the T4, TSH and TBG assays in the period between 2008 and 2018.

	Variation over time ^a	Inter-laboratory variation ^b
T4		
2008–2018	4.3%	3.9%
TSH		
2008–2018	12.3%	13.8%
TBG		
2008–2018	6.6%	
2008–2011	4.1%	4.8%
2012–2018	7.8%	6.7%

^aCV% of median quarterly concentrations. Quarterly concentrations are expressed as the median of the mean concentrations of the five screening laboratories in the respective quarter. This quarterly variation of the measured concentrations is plotted in Fig. 3; ^bMean of quarterly CV% between laboratories.

CV%, Coefficient of variation; T4, total thyroxine; TBG, thyroxine-binding globulin; TSH, thyroid-stimulating hormone.

suggested in a study showing the occurrence of (probable) CH-C after initiation of growth hormone treatment in children with congenital growth hormone deficiency (15). Our unique stepwise T4-TSH-TBG NBS algorithm resulted in a PPV of 21.0%.

T4-based referrals (referral criteria 1 and 3)

In the Netherlands all neonates with a T4 ≤ -3.0 s.d. are immediately referred to a general pediatrician or pediatric endocrinologist, enabling early diagnosis and treatment and – with that – optimal care for severe CH. Between 2007 and 2017, this led to the fastest possible detection (first heel puncture) of 13 cases of CH-C and in the combined first and second heel puncture to 29 cases which is 33.7% of all CH-C cases detected. Moreover, 12 cases of CH-T with normal levels of TSH were also detected due to their T4 concentration. While this may seem an atypical result, CH-T cases with a delayed TSH rise have been described (16).

From 2007 until July 2012, 249 neonates with TBG deficiency were unnecessarily referred because of a low T4. The implementation of the TBG rule (no referral for low T4 when TBG ≤ 40 nmol/L blood) in July 2012 has proven to be of major additional value by preventing referral of approximately 50 false-positives per year. If the TBG rule had not been applied, the PPV for T4 and TSH criteria would have been 29.2% instead of 34.0% and the total PPV 19.3% instead of 21.0%.

Unfortunately, our T4 referral criterion still was the cause of 36.1% of all false-positive referrals. Many of these neonates have (very) low plasma T4 concentrations due to illness (17). Since other NBS programs that use a strategy

with T4 followed by a TSH measurement do not refer based on only the T4 concentration (8), it is optional to delay the referral and wait for the TSH concentration. A second heel puncture in case of normal TSH might result in a normal NBS result, but would result in a delayed diagnosis in true-positives with possible negative impact on their outcome.

TSH-based referrals (referral criteria 2 and 4)

As expected, the performance of the TSH criteria in the lowest 20% T4 concentrations was excellent, with detection of 92.7% of all CH-T cases and causing only 8.9% of all false-positives, as in primary TSH-based screening programs. A recent study from Canada showed a ~100% sensitivity with a TSH cut-off value of ≤ 17 mIU/L. However, the number of false-positives was higher than with our method (PPV 55%) (18). This can be explained by the fact that in the Dutch strategy neonates at risk for CH-T are pre-selected based on their T4 concentration and that 80% of all neonates do not receive a TSH measurement. This pre-selection might result in missing very mild CH-T cases with normal T4 concentrations and could therefore contribute to the lower incidence for CH-T in the Netherlands compared to other countries (3). Another factor that may explain the lower CH-T incidence might be the absence of a repeat sample at the age of 2 weeks in neonates with a delayed TSH rise (common in newborns with a very low birth weight (19)).

The added value of the T4/TBG ratio (referral criterion-5)

NBS programs with a T4-reflex TSH or a combined T4 and TSH approach generally report a higher recall rate and detect CH-C, which distinguishes them from TSH-based screening programs. While in many NBS programs detection of CH-C is not the primary objective, CH-C is a condition worthwhile to detect. The incidence of CH-C in the Netherlands (1:16 404) (2) was and currently is much higher than the incidence reported in the US (1:55 656) (8). This is mainly due to the use of the T4/TBG ratio as an indirect measure of fT4 (2). In our evaluation, we identified 62.8% of detected CH-C cases by the T4/TBG ratio only, proving that this approach is of significant value. Additionally, the T4/TBG ratio identified 2.0% of all CH-T cases that otherwise would have been missed by screening. Although these cases were classified as (probably mild) CH-T by the treating pediatricians, it not unlikely that these cases are misclassified CH-C cases (12).

The PPV of the T4/TBG ratio, however, was very low over the course of 11 years of NBS (5.1%). However, this high number of false-positives should be weighed against the benefits of detecting neonates with CH-C as part of multiple pituitary hormone deficiencies and severe isolated CH-C (20) and the proven cost-effectiveness of the Dutch NBS for CH (2).

T4 and TSH measurement variability

The T4 measurements performed well with a low variation over time and between laboratories. In contrast, the TSH measurements showed much more variation which is probably caused by the use of different immunoassays and temporary harmonisation factors by the various laboratories. Only from 2018 onwards all five laboratories in the Netherlands use the same method without any harmonisation factor for the measurement of TSH.

TBG assay and measurement variability

Our current evaluation shows that, since the introduction of the TBG Monobind® assay in 2012, the referral rate due to the T4/TBG ratio has increased tremendously, with a subsequent decrease in the most recent years. After the forced introduction of the TBG Monobind® assay, we observed an increased variation in the TBG concentrations compared to previous years, while T4 concentrations remained stable. The most likely explanation of the variation are the fluctuations in the concentration of the TBG standard materials between different lot numbers of the assay, since no changes of the used antibody or equipment have occurred. This is further supported by the fact that, in 2017, when TBG concentrations were the lowest, a larger proportion of false-positive referrals had a T4/TBG ratio close to the cut-off value of 17 and that the average T4/TBG ratio was higher as compared to 2012–2016. This suggests that a large number of false-positive referrals was prevented in that year, while three cases of CH-C were still detected. This observation has led to the introduction of intensive verification of batches since July 2019, resulting in a more constant TBG measurement and therefore T4/TBG ratio.

Alternatively, using a fT4 assay could be considered as a replacement of the current total T4 and TBG measurements. As an alternative for the Eiken® TBG RIA, the Neonatal fT4 immunoassay of Bio-Rad was validated in 2011 for the Dutch CH NBS. However, the performance of the fT4 assay was not superior to the used T4/TBG ratio, since the fT4 concentrations in blood spots

were also strongly dependent on TBG concentrations. Therefore, it was decided not to implement this assay (14). In Sapporo, Japan, the CH NBS program employs a simultaneous measurement of fT4 and TSH in order to detect both CH-T and CH-C (21). In 10 years of fT4 screening for CH-C in Japan, a PPV of 11.5% was reported, with a sensitivity of 59.1% (22). Unfortunately, this assay is currently not available outside of Japan, and to the best of our knowledge, no other fT4 assays are available for NBS.

With this study, we provide an evaluation of the unique stepwise T4-TSH-TBG NBS algorithm for CH in the Netherlands during 11 years. Several adaptations have been applied in the meantime and this study gives an overview of the effect of the most important ones. One of the limitations of our retrospective study is the use of a database with diagnostic information directly imputed by mostly general pediatricians but including unclear diagnoses and misclassified cases. Another limitation is that no solid documentation of diagnoses missed by NBS (false-negatives) is available, and the sensitivity of our NBS could therefore not be calculated. However, despite their incompleteness, the available data has given us insight into the performance of 11 years of NBS for CH. Importantly, any NBS method has the pitfall of false-negative results and this is not different for our algorithm. CH-T cases with a very low birth weight, causing a delayed TSH rise, or mild CH-T cases with a neonatal screening T4 in the normal range, aside from premature newborns with CH-C, are likely to be missed. Awareness of these manageable issues should lead to further improvement of our screening strategy in the near future.

In conclusion, this 11-year evaluation shows that the Dutch stepwise T4-TSH-TBG NBS algorithm with a calculated T4/TBG ratio is of great value for the detection of both CH-T and CH-C, although our approach leads to a lower PPV compared to TSH-based NBS strategies. While the TBG rule has resulted in a significant reduction of false-positives, there is still room for improvement: decreasing the variability of the TBG measurements could further improve the PPV. Alternatively, when available, for NBS in the Netherlands, it would be interesting to compare a fT4 measurement to our current T4 and TBG measurements.

Declaration of interest

Annet M Bosch has received a speakers fee from Nutricia and has been a member of advisory boards for Biomarin. The other authors have nothing to disclose.

Funding

This research did not receive any specific grant from any funding agency in the public, commercial or not-for-profit sector.

Ethical approval

This study has been approved with a signed statement by the Privacy Committee of TNO. All parents participating in the Neonatal screening program of the Netherlands gave consent for both participation and use of data for scientific research, unless they actively declined.

References

- Rastogi MV & LaFranchi SH. Congenital hypothyroidism. *Orphanet Journal of Rare Diseases* 2010 **5** 17. (<https://doi.org/10.1186/1750-1172-5-17>)
- Lanting CI, van Tijn DA, Loeber JG, Vulmsa 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** 168–173. (<https://doi.org/10.1542/peds.2004-2162>)
- Wassner AJ & Brown RS. Congenital hypothyroidism: recent advances. *Current Opinion in Endocrinology, Diabetes, and Obesity* 2015 **22** 407–412. (<https://doi.org/10.1097/MED.0000000000000181>)
- Ford G & LaFranchi SH. Screening for congenital hypothyroidism: a worldwide view of strategies. *Best Practice and Research: Clinical Endocrinology and Metabolism* 2014 **28** 175–187. (<https://doi.org/10.1016/j.beem.2013.05.008>)
- Brosco JP, Mattingly M & Sanders LM. Impact of specific medical interventions on reducing the prevalence of mental retardation. *Archives of Pediatrics and Adolescent Medicine* 2006 **160** 302–309. (<https://doi.org/10.1001/archpedi.160.3.302>)
- Simoneau-Roy J, Marti S, Deal C, Huot C, Robaey P & Van Vliet G. Cognition and behavior at school entry in children with congenital hypothyroidism treated early with high-dose levothyroxine. *Journal of Pediatrics* 2004 **144** 747–752. (<https://doi.org/10.1016/j.jpeds.2004.02.021>)
- Simons WF, Fuggle PW, Grant DB & Smith I. Intellectual development at 10 years in early treated congenital hypothyroidism. *Archives of Disease in Childhood* 1994 **71** 232–234. (<https://doi.org/10.1136/adc.71.3.232>)
- LaFranchi SH. Newborn screening strategies for congenital hypothyroidism: an update. *Journal of Inherited Metabolic Disease* 2010 **33** (Supplement 2) S225–S233. (<https://doi.org/10.1007/s10545-010-9062-1>)
- Kilberg MJ, Rasooly IR, LaFranchi SH, Bauer AJ & Hawkes CP. Newborn screening in the US may miss mild persistent hypothyroidism. *Journal of Pediatrics* 2018 **192** 204–208. (<https://doi.org/10.1016/j.jpeds.2017.09.003>)
- Hewlett J & Waisbren SE. A review of the psychosocial effects of false-positive results on parents and current communication practices in newborn screening. *Journal of Inherited Metabolic Disease* 2006 **29** 677–682. (<https://doi.org/10.1007/s10545-006-0381-1>)
- Zwaveling-Soonawala N, van Trotsenburg AS & Verkerk PH. The severity of congenital hypothyroidism of central origin should not be underestimated. *Journal of Clinical Endocrinology and Metabolism* 2015 **100** E297–E300. (<https://doi.org/10.1210/jc.2014-2871>)
- Zwaveling-Soonawala N, van Trotsenburg ASP & Verkerk PH. TSH and FT4 concentrations in congenital central hypothyroidism and mild congenital thyroidal hypothyroidism. *Journal of Clinical Endocrinology and Metabolism* 2018 **103** 1342–1348. (<https://doi.org/10.1210/jc.2017-01577>)
- Zwaveling-Soonawala N, Naafs JC, Verkerk PH & van Trotsenburg ASP. Mortality in children with early-detected congenital central hypothyroidism. *Journal of Clinical Endocrinology and Metabolism* 2018 **103** 3078–3082. (<https://doi.org/10.1210/jc.2018-00629>)
- Boelen A, van Veen M, Verkerk PH, Diependaal G, Loeber G, Elvers B & Endert E. Measuring free thyroxine levels in neonatal heel-prick samples. *Clinica Chimica Acta: International Journal of Clinical Chemistry* 2013 **423** 51–55. (<https://doi.org/10.1016/j.cca.2013.04.004>)
- 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? *Hormone Research in Paediatrics* 2018 **89** 98–107. (<https://doi.org/10.1159/000486033>)
- LaFranchi SH. Congenital hypothyroidism: delayed detection after birth and monitoring treatment in the first year of life. *Journal of Pediatrics* 2011 **158** 525–527. (<https://doi.org/10.1016/j.jpeds.2010.11.030>)
- van Wassenae AG, Kok JH, de Vijlder JJM, Briët JM, Smit BJ, Tamminga P, van Baar A, Dekker FW & Vulmsa T. Effects of thyroxine supplementation on neurologic development in infants born at less than 30 weeks' gestation. *New England Journal of Medicine* 1997 **336** 21–26. (<https://doi.org/10.1056/NEJM199701023360104>)
- Saleh DS, Lawrence S, Geraghty MT, Gallego PH, McAssey K, Wherrett DK & Chakraborty P. Prediction of congenital hypothyroidism based on initial screening thyroid-stimulating-hormone. *BMC Pediatrics* 2016 **16** 24. (<https://doi.org/10.1186/s12887-016-0559-0>)
- Larson C, Hermos R, Delaney A, Daley D & Mitchell M. Risk factors associated with delayed thyrotropin elevations in congenital hypothyroidism. *Journal of Pediatrics* 2003 **143** 587–591. ([https://doi.org/10.1067/S0022-3476\(03\)00332-9](https://doi.org/10.1067/S0022-3476(03)00332-9))
- van Tijn DA, de Vijlder JJ, Verbeeten Jr B, Verkerk PH & Vulmsa T. Neonatal detection of congenital hypothyroidism of central origin. *Journal of Clinical Endocrinology and Metabolism* 2005 **90** 3350–3359. (<https://doi.org/10.1210/jc.2004-2444>)
- Soneda A, Adachi M, Muroya K, Asakura Y, Yamagami Y & Hirahara F. Overall usefulness of newborn screening for congenital hypothyroidism by using free thyroxine measurement. *Endocrine Journal* 2014 **61** 1025–1030. (<https://doi.org/10.1507/endocrj.14-0143>)
- Adachi M, Soneda A, Asakura Y, Muroya K, Yamagami Y & Hirahara F. Mass screening of newborns for congenital hypothyroidism of central origin by free thyroxine measurement of blood samples on filter paper. *European Journal of Endocrinology* 2012 **166** 829–838. (<https://doi.org/10.1530/EJE-11-0653>)

Received 20 January 2020

Revised version received 9 June 2020

Accepted 18 June 2020