Research Article

Hale Aral*, Ömer Faruk Özer, Hatice Onur, Ahmet Mete Çilingirtürk, İlker Tolga Özgen and Mevlude Ayyıldız

Pediatric reference intervals of free thyroxine and thyroid stimulating hormone in three different hospitals

Üç farklı hastanede serbest tiroksin ve tiroid stimulan hormon pediyatrik referans aralıkları

https://doi.org/10.1515/tjb-2017-0342
Received December 20, 2016; accepted December 14, 2017; previously published online January 30, 2018

Abstract

Objective: Using data retrieved from three different hospitals, we established indirect reference intervals of free thyroxine (FT4) and thyroid stimulating hormone (TSH) for the Centaur XP or the Immulite 2000 instruments, in separate reference limits at each subset.

Methods: We categorized children into seven age groups: 4–7 days, 8–15 days, 16–23 days, 24–61 days, 3–6 months, 7–36 months and 4–6 years. After a Box-Cox transformation was employed, we followed the Horn algorithm to eliminate the extreme values.

Results: The remaining FT4 (11,230) and TSH (11,274) tests were statistically analyzed. We determined separate reference limits at each subset with their own 2.5th and 97.5th percentiles. The interaction effect of both hospital and age grouping on FT4 was meaningful, but there was no interaction effect on TSH.

Conclusions: Pediatric FT4 and TSH test results should be interpreted via narrowed age groups, especially in the first 3 weeks of neonatal period. Our reference limits may be recommended in pediatric follow-ups, considering the conditions of prematurity, birth-weight or multiple births. Preanalytical and analytical variations related with complex molecular structure of FT4 should be taken into consideration to ensure the validity of the result.

Keywords: Immunoassays; pediatric reference intervals; hospital information system; thyroid stimulating hormone; thyroxine.

Özet

Amaç: Üç farklı hastanede elde edilen verileri kullanarak, her bir alt-grup için ayrı ayrı referans sınırları olmak üzere Centaur XP veya Immulite 2000 cihazlarında serbest tiroksin (FT4) ve tiroid stimulan hormon (TSH) indirek referans aralıklarını tespit ettik.

Yöntem: Çocukları 7 farklı yaş grubuna ayırdık: 4–7 gün, 8–15 gün, 16–23 gün, 24–61 gün, 3–6 ay, 7–36 ay and 4–6 yıl. Box-Cox transformasyon uygulamasını takiben, Horn algoritması sayesinde uç değerleri uzaklaştırdık.

Bulgular: Elımizde kalan FT4 (11,230) ve TSH (11,274) test sayıları ile istatistiksel inceleme gerçekleştirilmiştir. Her bir alt-gruba ait kendi 2,5 ve 97,5 persentil değerleri ile referans sınırlarını belirlendik. Hastane ve yaş gruplarının her ikisinin de FT4 üzerinde etkileşim etkisi anlamlı bulundu, ancak TSH üzerinde anlamlı etki görülmemiş.
Introduction

Reference intervals of thyroid function tests are known to be method- and population-dependent (i.e. due to differences in ethnicity, population or geographically derived covariates such as lifestyle, salt iodination, selenium status, nutrition or other unknown geographic covariates). Unfortunately, manufacturer’s recommendations/inserts are unsatisfactory in population based and age-adjusted cut-off levels for the methods used on venous samples, especially in diagnosis of congenital hypothyroidism (CH) [1].

The neonatal and the pubertal periods are examples of periods for which there is no known reference ranges related to thyroid hormones [2]. Ideally, direct reference intervals should be determined using blood samples obtained from a large cohort of healthy subjects [3]. However, due to ethical and practical considerations in children, indirect reference interval determination often relies on large hospital databases [4]; e.g. difficulties in acquiring parental consent to take blood samples and making technical trouble in blood sampling. In addition, symptoms of hypothyroidism are relatively nonspecific, diverse and often mild; many children with symptoms of attention-deficit disorder or hyperactivity, fatigue, obesity, constipation and hair loss have thyroid function tests ordered to rule out hypothyroidism or thyroid hormone resistance as the cause of these symptoms. So, defining the real ‘reference individual’ is another challenge.

In this study, selecting three close hospitals, wage gap, life-style, environmental and ethnic effects of preanalytical factors were all supposed to be minimalized. The aim of the study was basically to establish indirect reference intervals of serum FT4 and TSH levels in children younger than 7 years for the Centaur XP or the Immulite 2000 instruments. It was decided to determine separate reference limits at each subset in heterogeneous hospital and age categories. Moreover, we examined the impact of the age and the hospital, where the values retrieved, on FT4 and TSH (both clinically and methodologically) with the other variable controlled.

Materials and methods

Patients aged 4 days to 72 months made up our study population. Age was estimated in accordance with the difference between the ‘birthday’ and the ‘test requisition’ times in units of (postnatal) day or month, and patients were categorized according to the following age groups: 4–7 days, 8–15 days, 16–23 days, 24–61 days, 3–6 months, 7–36 months and 4–6 years.

Accumulated data of serum FT4 and TSH levels was retrieved from three different hospitals in the same region, Fatih – an administrative district in Istanbul, Turkey. Children for whom simultaneous FT4 and TSH testing requested from out-patient clinics were included our study. But, tests requested from departments of pediatric endocrinology and medical genetics were excluded. Patients with repeated test requisitions were also excluded, due to the likelihood of thyroid dysfunction. In all three hospitals, the total (between-day) coefficients of variation (CV) of each instrument were satisfactory for FT4 and TSH (third generation). Percents of exclusion of the patients and imprecision values of the instruments can be reached through the Supplementary data (exclusion and imprecision). This study has been cleared by the Bezmialem Vakif University Ethics Review Board for human studies.

We used ‘pmol/L’ as the unit in FT4 values, so we converted the data collected from Hospitals A and C from ‘ng/dL’ by multiplying by a factor of 12.9. The unit ‘mU/L’ was used in all TSH measurements.

The same kind of instrument (Advia Centaur XP, Siemens Healthcare Diagnostics Inc., Tarrytown, NY, USA) was used at both the Istanbul Research and Training Hospital (Hospital A, three instruments) and the Bezmialem Vakif University (Hospital B, two instruments). At the Suleymaniye Women’s Disease and Children’s Research and Training Hospital (Hospital C). The instrumentation differed (only one instrument, Immulite 2000, Siemens Healthcare Diagnostics Inc., Tarrytown, NY, USA).

Exploratory graphical presentation of the raw data clearly expressed the presence of extreme values and we observed with normal Q-Q plots that FT4 and TSH measurements did not constitute a Gaussian distribution. Furthermore, the Kolmogorov-Smirnov goodness-of-fit test confirmed this finding. Therefore, a Box-Cox transformation was employed before further analysis [5].
power ‘λ’ was estimated for factored data according to age groups because the distributions were expected to differ among these groups. Accordingly, we followed the Horn algorithm to eliminate the extreme values [6]. The Horn algorithm is based on the computation of the lower and upper quartiles of the transformed data. The details of analysis can be reached through Supplementary data (exclusion and imprecision).

A univariate general linear model (GLM) was applied to detect the impact of age and measurement mean on FT4 and TSH with the other variable controlled. Thus, the effect of each factor represents its own discrimination ability, taking other effects into consideration. Although the GLM has three assumptions – normality, homogeneity and independence of errors – and requires the absence of outliers, analysis of variance (ANOVA) is said to be a fairly robust analysis. Homogeneity was tested according to Levene’s test. Conducted according to the type III sum of squares with constant intercept effect, these analysis tests hypothesize the differences in sub-population (or marginal) means for ANOVA designs with unequal numbers of sub-populations. These least squares means are the best linear-unbiased estimates of the marginal means for the design [7]. Tests of differences in least squares means have the important property of being invariant to the choice of the coding of effects for categorical independent variables. Statistical analysis was performed by using Spss11 and MS Excel 2013.

Results

The percent of extreme values in three different hospitals (A, B and C) is shown in Table 1. Without extreme values, the remaining FT4 (11,230) and TSH (11,274) test results were statistically analyzed.

| Age Group  | Hospital | 4–7 days | 8–15 days | 16–23 days | 24–61 days | 3–6 months | 7–36 months | 4–6 years |
|------------|----------|----------|-----------|------------|------------|-------------|-------------|-----------|
|            | FT4      | TSH      | FT4       | TSH        | FT4       | TSH         | FT4         | TSH       |
| A          | N        | 95       | 95        | 83         | 83         | 51          | 51          | 244       | 244       | 171       | 171       | 1624      | 1624      | 1111      | 1111      |
|            | n        | 95       | 91        | 83         | 83         | 48          | 49          | 239       | 241       | 167       | 170       | 1609      | 1608      | 1093      | 1105      |
|            | %        | 0.0%     | 4.2%      | 0.0%       | 0.0%       | 5.9%        | 3.9%        | 2.0%      | 1.2%      | 2.3%      | 0.6%      | 0.9%      | 1.0%       | 1.6%       | 0.5%      |
| B          | N        | 93       | 93        | 171        | 171        | 111         | 111         | 419       | 420       | 516       | 518       | 2429      | 2441      | 1476      | 1481      |
|            | n        | 91       | 91        | 166        | 161        | 110         | 105         | 408       | 414       | 503       | 508       | 2389      | 2411      | 1452      | 1466      |
|            | %        | 2.2%     | 2.2%      | 2.9%       | 5.8%       | 0.9%        | 5.4%        | 2.6%      | 1.4%      | 2.5%      | 1.9%      | 1.6%      | 1.2%       | 1.6%       | 1.0%      |
| C          | N        | 296      | 296       | 682        | 682        | 336         | 336         | 472       | 472       | 223       | 223       | 519       | 519       | 308       | 308       |
|            | n        | 293      | 288       | 668        | 657        | 328         | 330         | 462       | 468       | 218       | 215       | 508       | 510       | 300       | 303       |
|            | %        | 1.0%     | 2.7%      | 2.1%       | 3.7%       | 2.4%        | 1.8%        | 2.1%      | 0.8%      | 2.2%      | 3.6%      | 2.1%      | 1.7%       | 2.6%       | 1.6%      |

Age related changes in the 2.5th, 5th, 10th, 25th, 50th, 75th, 90th, 95th and 97.5th percentiles of FT4 and TSH are shown graphically in Figures S2 and S3 via the Supplementary data (figures). The measurements of both variables (age and hospital) were found homogenous and had statistically equal central tendency according to sex. Although, some age groups and hospitals had statistically different means, they were found to have heterogeneous distributions; we determined separate reference limits at each subset (except sex) with their own 2.5th and 97.5th percentiles.

Between heterogeneous hospital and age categories, we found significant differences of main effects in FT4 (p < 0.001) and TSH mean values (p < 0.001) using the ANOVA method, as shown in Figure 1A,B, respectively.

Discussion

Indirect reference intervals of FT4 and TSH

In this study, we achieved pediatric FT4 and TSH reference intervals in serum samples for the sake of accurate diagnosis of CH of the newborn. Mild and non-specific clinical findings in newborns, later development of endocrine and metabolic disorders in unrecognized CH, false negative results in screening, variation in reference ranges according to the method used in the laboratory, and the lack of standardization in this field led us to analyse reference limits of FT4 and TSH. Specimens from umbilical cord were used to measure FT4 and TSH for CH screening [8, 9]; the 2.5th and 97.5th cord blood FT4 were reported as 11.48 pmol/L and 19.74 pmol/L [8], 13.80 pmol/L and 26.06 pmol/L [9], and the 97.5th cord blood TSH was reported as 27.56 mU/L [8] as shown in
Table 2. Nevertheless, an abrupt increase in pituitary TSH secretion occurs soon after birth and may reach a peak of about 60–80 mU/L at the first 25–30 min of life [23]. Infants receiving heparin or triglyceride emulsions – or even drugs with potential protein displacement activity – and possible structural changes in thyroid-binding globulin, which decrease the affinity of thyroxine (T4) binding, could have caused an increase in serum free thyroxin levels in the first 24 h [24]. We designed our study group starting from the fourth day of life to avoid these high (within-subject) biological variations in the first days of life (4–7 days). As the pediatric endocrinologist (co-author) took initiative, we worked on our exclusion criteria and had age groups classified according to approximate developmental periods of hypothalamus–hypophysis axis in the childhood.

In our preliminary study (presented as a poster in IFCC WorldLab 2014, Istanbul, Turkey), a Dunnett T3 test revealed no significant difference in FT4 between the age groups of 24–30 days and 2–6 months (p = 1.000). With respect to our previous experience, in this study we decided to determine the values for the age groups of 24–61 days in our population. On the other hand, some studies have demonstrated that FT4 and TSH levels continue to decline from birth to the end of adolescence; this decline was more rapid in neonatal period [4, 19]. In this study, serum FT4 and TSH levels were statistically different regarding age groups. Therefore, pediatric FT4 and TSH test results should be interpreted via narrowed age groups as shown in our study; especially in the first 3 weeks of neonatal period, accurate cut-off values should be determined for each week.

We found the 97.5th reference limits for TSH in the age groups of 4–7 days, 8–15 days, 16–23 days, respectively, as 18.02 mU/L, 12.87 mU/L, 7.67 mU/L in Hospital A, 14.72 mU/L, 11.85 mU/L, 8.76 mU/L in Hospital B, and 16.27 mU/L, 9.91 mU/L, 7.05 mU/L in Hospital C. In a study conducted in Turkey [9], 97.5th reference limits for TSH were reported similar in the 3rd, 10th, 14th days of life, respectively, as 18.58 mU/L, 10.72 mU/L, 7.87 mU/L via electrochemiluminescence immunoassay in E-170 analyzer (Table 2).

Especially in the neonatal period, we had the greatest number of patient so far, hospitals using the same model of instruments (Hospitals A and B) showed very close values of TSH with the exception of the age group of 4–7 days (Table 2, Figure 1B). Indirect reference intervals of pediatric thyroid function tests were performed in six other studies [4, 11, 14, 18, 19, 21], as shown in Table 2. Zurakowski et al. [21] found TSH upper limits similar to ours. For the Immulite 2000, it was reported that the median concentrations of TT4, FT4 and TSH to be up to 3.2-fold higher during the first 2 weeks [12]. By examining the scatter plot of the Canadian Laboratory Initiative in Pediatric Reference Intervals (CALIPER) data and hospital based data with FT4 on the x-axis and TSH on the y-axis, one can interpret (statistically) the manufacturer’s overly broad FT4 reference intervals [14, 16]. Hübner et al. [10] calculated the reference range by performing the piecewise linear regression analysis, using specified square root functions.
Table 2: Comparison of our data to previously published reference intervals.

| Author (year) place, estimation, instrumentation | Age | No. of samples | FT4 (pmol/L) | TSH (mU/L) |
|-----------------------------------------------|-----|---------------|--------------|------------|
|                                               |     |               | 2.5<sup>th</sup> | 97.5<sup>th</sup> | 2.5<sup>th</sup> | 97.5<sup>th</sup> |
| Hübner et al. (2002) Homburg-Bremerhaven, direct, Advia Centaur XP [10] | 1 day to 3 days | 32 | 10.8 | 26.8 | 0.13 | 9.23 |
|                                                | 4 days to 30 days | 40 | 11.0 | 24.3 | 0.19 | 7.78 |
|                                                | 31 days to 60 days | 61 days to 12 months | 11.4 | 20.9 | 0.30 | 5.88 |
|                                                | 1 year to 5 years | 76 | 11.4 | 19.0 | 0.42 | 4.79 |
| Kapelari et al. (2008) | 1 day to 1 month | 23 | 8.50 | 30.50 | 22 | 0.70 | 18.10 |
| Innsbruck, indirect, Advia Centaur XP [4] | 2 months to 12 months | 45 | 9.17 | 25.28 | 42 | 1.12 | 8.21 |
| Aral, İstanbul-Hospital A, indirect, Advia Centaur XP, this study | 4 days to 7 days | 95 | 13.93 | 26.08 | 91 | 1.10 | 18.02 |
|                                                | 8 days to 15 days | 83 | 12.53 | 23.46 | 83 | 1.21 | 12.87 |
|                                                | 16 days to 23 days | 48 | 11.68 | 19.15 | 49 | 1.38 | 7.67 |
|                                                | 24 days to 61 days | 239 | 11.74 | 18.97 | 241 | 1.28 | 8.11 |
|                                                | 3 months to 6 months | 167 | 11.61 | 20.09 | 170 | 0.80 | 6.50 |
|                                                | 7 months to 36 months | 1609 | 10.97 | 18.32 | 1608 | 0.77 | 5.54 |
|                                                | 4 to 6 years | 1093 | 11.48 | 18.19 | 1105 | 0.96 | 6.04 |
| Aral, İstanbul-Hospital B, indirect, Advia Centaur XP, this study | 4 days to 7 days | 91 | 14.94 | 24.82 | 91 | 1.02 | 14.72 |
|                                                | 8 days to 15 days | 166 | 12.69 | 21.76 | 161 | 1.24 | 11.85 |
|                                                | 16 days to 23 days | 110 | 12.85 | 20.23 | 105 | 1.51 | 8.76 |
|                                                | 24 days to 61 days | 408 | 12.38 | 19.61 | 414 | 1.16 | 7.15 |
|                                                | 3 months to 6 months | 503 | 12.38 | 19.88 | 508 | 0.81 | 6.83 |
|                                                | 7 months to 36 months | 2389 | 11.55 | 18.64 | 2411 | 0.77 | 5.46 |
|                                                | 4 to 6 years | 1452 | 11.93 | 19.08 | 1466 | 0.82 | 5.53 |
| Aral, İstanbul-Hospital C, indirect, Immulite 2000, this study | 4 days to 7 days | 293 | 13.33 | 24.99 | 288 | 1.21 | 16.27 |
|                                                | 8 days to 15 days | 668 | 12.38 | 21.93 | 657 | 1.12 | 9.91 |
|                                                | 16 days to 23 days | 328 | 11.75 | 19.59 | 330 | 1.01 | 7.05 |
|                                                | 24 days to 61 days | 462 | 11.55 | 18.64 | 468 | 0.87 | 6.05 |
|                                                | 3 months to 6 months | 218 | 11.93 | 19.08 | 215 | 0.60 | 5.49 |
|                                                | 7 months to 36 months | 508 | 10.72 | 17.67 | 510 | 0.64 | 5.18 |
|                                                | 4 to 6 years | 300 | 12.06 | 18.52 | 303 | 0.60 | 4.12 |
| Omuse et al. (2016) Nairobi, indirect, Roche e601 [11] | 0 to 7 days | 552 | 13.62 | 34.75 | 4.79 F | 0.56 | 11.00 |
|                                                | 8 to 14 days | 145 | 13.45 | 30.17 | 415 M | 0.59 | 12.84 |
|                                                | 15 to 22 days | 465 | 14.16 | 24.80 | 745 | 0.90 | 7.46 |
|                                                | 23 to 30 days | 167 | 13.26 | 23.37 | 0.13 | 9.23 |
|                                                | 0 to 14 days | 479 F | 29.6 | 97.2 | 1.79 | 9.69 |
|                                                | 0 to 14 days | 479 F | 29.6 | 97.2 | 1.79 | 9.69 |
| Emlinger et al. (2001)                                   | 1 day to 7 days | 45 | 29.6 | 79.2 | 1.79 | 9.69 |
| Tübingen, direct, Immulite 2000 [12]               | 8 days to 15 days | 40 | 18.0 | 63.6 | 1.80 | 7.97 |
|                                                | 16 days to 3 years | 86 | 11.1 | 27.3 | 0.63 | 4.12 |
|                                                | 4 years to 6 years | 51 | 12.9 | 23.9 | 0.53 | 2.94 |
| Najam et al. (2003) Islamabad, direct, Immulite [13]      | 0 day to 4 days | 992 | 2.42 | 27.5 | 0.12 | 5.50 |
|                                                | 5 days to 7 days | 36 | 0.4 | 13.95 |
|                                                | 8 days to 1 months | 38 | 0.7 | 17.56 |
|                                                | 2 months to 1 year | 38 | 0.31 | 14.51 |
|                                                | 2 years to 5 years | 38 | 0.39 | 19.86 |
| Henderson and Grey (2011) Hamilton, indirect, Roche Modular E-170 [14] | 1 year to 6 years | 46 | 14.5 | 19.8 | 1.3 | 5.5 |
| Soldin et al. (2010) Virginia, direct, Abbott Architect ci8200 [15] | <2 months | 160 F | 8.90 | 23.87 | 171 F | 1.12 | 5.50 |
|                                                | 2 months to 11 months | 138 F | 8.90 | 19.22 | 152 F | 1.12 | 4.47 |
|                                                | 2 months to 11 months | 125 M | 9.16 | 20.12 | 138 M | 0.96 | 4.90 |
|                                                | 12 months to 24 months | 202 F | 9.42 | 18.19 | 203 F | 1.00 | 4.37 |
|                                                | 12 months to 24 months | 190 M | 10.58 | 17.03 | 197 M | 0.93 | 4.79 |
| Author (year) place, estimation, instrumentation | Age | FT4 (pmol/L) | TSH (mU/L) |
|---------------------------------------------|-----|--------------|------------|
|                                            |     | 2.5th Percentile | 97.5th Percentile | No. of samples | 2.5th Percentile | 97.5th Percentile |
|                                            |     |                |              |                |                |                |
| 3 years to 5 years                          | 253 F | 10.19          | 17.80        | 264 F          | 0.85           | 4.07           |
| 3 years to 5 years                          | 286 M | 10.32          | 17.03        | 297 M          | 0.83           | 4.37           |
| 5 days to 14 days                           | 66   | 13.55          | 41.41        |                |                |                |
| 15 days to 29 days                          | 55   | 8.77           | 32.64        |                |                |                |
| 30 days to 1 year                           | 270  | 11.48          | 21.93        |                |                |                |
| 1 year to 19 years                          | 952  | 11.48          | 17.67        |                |                |                |
| 4 days to 5 months                          | 139  | 0.73           | 4.77         |                |                |                |
| 6 months to 14 years                        | 640  | 0.7            | 4.17         |                |                |                |
| 3 years to 5 years                          | 286 M | 10.32          | 17.03        | 297 M          | 0.83           | 4.37           |
| 3 years to 5 years                          | 253 F | 10.19          | 17.80        | 264 F          | 0.85           | 4.07           |
| 3 years to 5 years                          | 286 M | 10.32          | 17.03        | 297 M          | 0.83           | 4.37           |
| 5 days to 14 days                           | 66   | 13.55          | 41.41        |                |                |                |
| 15 days to 29 days                          | 55   | 8.77           | 32.64        |                |                |                |
| 30 days to 1 year                           | 270  | 11.48          | 21.93        |                |                |                |
| 1 year to 19 years                          | 952  | 11.48          | 17.67        |                |                |                |
| 4 days to 5 months                          | 139  | 0.73           | 4.77         |                |                |                |
| 6 months to 14 years                        | 640  | 0.7            | 4.17         |                |                |                |

*Pediatric reference ranges were determined with use of the arithmetic mean ± 2 SD for FT4 and the geometric mean ± 2 SD for TSH, after outliers were identified by the Grubbs test and deleted. 2.5th and 97.5th reference limits of FT4 and TSH (in bold) in children aged <7 years.
F, female; M, male; y, year(s); m, month(s); w, week(s); d, day(s); IRMA, immunoradiometric assay.
The interaction effect of both the hospital and age groupings on FT4

The interaction effect of both the hospital and age groupings on FT4 was meaningful. Using different model of instruments (or number of instruments of the same model), there were possible analytical variations in immunoassay methodology among the hospitals. Some technical specificities of the instruments can be reached through the Supplementary data (exclusion and imprecision).

In the FT4 measurements, a competitive immunoassay (one step) using direct chemiluminescent technology, during which a labeled structural analog (labeled antigen considered as a tracer) competes with serum T4 for antibody binding sites, is performed in both types of instrumentation [26]. It has been reported that measurements of FT4 based on immunoassay did not give data clinically consistent with the TSH measurements made on the same platform; such problems were avoided when the FT4 was measured by ultra-filtration tandem mass spectrometry [15]. However, the immunoassay procedures meet requirements for clinical use by giving the results in under than an hour [4].

Having the same type of instrumentation Hospitals A and B had similar results of TSH, but there was discrepancy in FT4 values. Defining the analytical variations between the systems (including instrument, water system, reagents, calibration process, after sales service, quality control procedure and quality control program scores) and a better method of choice (e.g. tandem mass spectrometric measurements) should give us advantages in studying traceability. Values of immunoassay calibrators are produced special for the instrumentation. The professional experience of the operator may be another factor influencing the test results. Most of the techniques show acceptable performance when the assay is performed on euthyroid patient samples [27]. Attempts toward standardization and harmonization of thyroid function tests are going on [28, 29]. Moreover, evaluation of the analytical bias of the methods or showing the traceability was not within the scope of this study.

In our study, duration of experience (time intervals of data retrieved) also differed in three hospitals; 4 years in Hospital A, 2 years in Hospital B and 3 years in Hospital C. Immunoassays are particularly prone to lot-to-lot differences since antibodies and detection reagents used on these systems tend to be more sensitive to matrix differences [30, 31]. Moreover, the data retrieved from Hospital B (using only two instruments of the same model) was greater than that retrieved from Hospital A (using three instruments of the same model), as presented in Table 2.

When we scrutinize the analyte FT4, the formation of thyroxine (T4) illustrates a great diversity in the thyroid gland. It requires rare element iodine for bioactivity (iodine oxidation); it is synthesized as part of a very large precursor molecule (thyroglobulin); it is stored in an intracellular reservoir (colloid) for several weeks; degradation in lysosome occurs to supply T4 and triiodothyronine (T3). The T4 circulates in the plasma mainly (99.98%) in the form linked to proteins: thyroxine binding globulin (TBG, 75–80% of the bound fraction), transthyretin (thyroxine-binding prealbumin, TBPA, 15–20%) and albumin (5–10%) [32, 33]. It has been shown that lipoproteins contribute to the transport of thyroid hormones but in lower proportions: high density lipoprotein (3%), very low density lipoprotein (0.03%) and low density lipoprotein (0.2%) [26]. After transportation, there is peripheral conversion of T4 to T3 in target tissues. In this sense, T4 can be thought of as a prohormone [32]. Serum T4 and T3 levels are markedly decreased without a compensatory rise in the serum TSH level during severe illness and nutritional deficiency, leading to a significant decline in basal metabolic rate along with a decrease in protein and fat catabolism [34]. So, FT4 has a more complex structure than TSH (glycoprotein).

Within the scope of age effect on FT4 (Figure 1A), measurements were kept unchanged (horizontal line) in the age group of 24 days to 6 months measurements in all the hospitals. Then, measurements came closer and presented a plateau following a decrease. When we examined the hospital effect on FT4, higher measurements were noticed in Hospital B, in general. Similar measurements were achieved in Hospital A and Hospital C, but measurements in the age group of 24 days to 6 months seemed higher in Hospital A.

Within the scope of age effect on TSH (Figure 1B), similar measurements were seen in Hospitals A and B (the same kind of instrumentation) after 16 days of age. And after 3 years of age, a slight increase similar to the other was also seen in both Hospitals A and B. Furthermore, at least one age group differed significantly than others after neonatal period, while some sequential groups did not differ statistically.
On behalf of the pre-analytical variations, the serum concentration of FT4 seems stable during diurnal rhythm, menstrual cycles and seasons, but varies during pregnancy. It has a plasma half-life of about 7 days. The direct measurement of serum FT4 (without prior separation) requires no affect in the pre-existing fragile balance between the free and bound forms of thyroxine. This equilibrium is generally broken when there is a significant variation in the concentration of the carrier proteins (genetic/illness/drug effects) or an alteration in their binding affinity [26, 35]. On the other hand, TSH is secreted in a pattern that is both pulsatile and circadian, and it has a plasma half-life of about 65 min [36, 37].

In a study conducted in three ethnic groups (whites, East Asians and South Asians), both sexes and three sampling times (morning, afternoon and evening), significant differences were observed in the concentrations of FT4, total triiodothyronine and total thyroxine (TT4) across ethnic groups [23]. In another study, TSH increases were commonly observed in obese children [38]; anthropometric measurements (body mass index and/or waist circumference) might be taken into consideration in decision making. Sampling strategy, sample size, analytical factors such as instrumentation, gender, age and other demographic and lifestyle factors and the statistical methods (outlier detection, partitioning, reporting standard error and confidence intervals) used to construct reference intervals are all factors affecting the resulting intervals. However, these are often overlooked as factors that might have a considerable effect on the validity of the reference intervals [39].

In studies included adults, within-subject and between-subject biological variations in serum TSH levels (19.3% vs. 24.6%) were shown higher than FT4 levels (5.7% vs. 12.1%, respectively) [40]. Nearly three times lower within-subject biological variations of FT4 (5.7%) in comparison with TSH (19.3%) may cause estimation of lower reference change values (RCV) in FT4 measurements of each instrument. Having lower RCVs, FT4 may have greater clinical significance effect than that of TSH, so analytical variations in FT4 of each instrument should be monitored separately.

In our study, individuals in the early days of life in Hospital B showed lower TSH and higher FT4 levels (Figure 1A), possibly as a result of frequency of prematurity (or multiple births). It is known that prematurity complicates screening for primary hypothyroidism due to the resulting a developmental delay [8, 24, 41]. TSH levels are shifted toward lower values in premature babies, lowering of the TSH cut-off point from 20 to 10 mU/L and resulting in a significant increase in the number of recalled newborns with a birth weight less than 2500 g (37.5%). With regard to transient CH (usually iodine-, drug- or maternal antibody-induced), the overall percentage was 11.2%, and it was significantly higher in premature infants than in full-term infants (25.8 and 7.5%, respectively) [42]. In another study, 26 patients (9.1%) were diagnosed with thyroid dysfunction whereas routine newborn screening identified only three patients, on day of life 30 in premature infants born at <30 weeks gestation. CH with delayed TSH elevation was diagnosed in 20 patients (6.9%) and was significantly associated with multiple gestation, lower birth weight, higher gestational age and lower 5 min APGAR score [43]. Due to the increased risk of primary and secondary hypothyroidism in preterm and low birth weight babies, the determination of TSH and FT4 (between days 3 and 5) is recommended, irrespective of the screening test [44].

**Limitations**

The primary study limitation is the fact that ethnicity, anthropometric measurements, prematurity according to gestation age of less than 37 weeks (or low birth weight and very low birth weight neonates), multiple births (particularly in cases of same-sex twins), evidence of infection, thyroid auto antibodies and a family history of autoimmune thyroid disease were not taken into consideration. Children diagnosed with known metabolic abnormalities – such as diabetes mellitus, CH, protein-energy malnutrition, and constipation – were not excluded in accordance with the International Classification of Diseases codes.

This study was designed in three different hospitals, in the same region, in order to exclude external effects of ethnic variation, environment, income/occupation/life status of the family, etc. This was strong side of our study. Accepting other hospitals’ databases far from this region, different results may be obtained in connection with the possible environmental and ethnic variations, as potential preanalytical factors. Moreover, the hospitals mentioned in this study have been using different instrumentation due to the tender process, since the study was performed. So, evaluation together with some other hospitals (using the same type of instrumentation) was not possible. This side of our study may be considered as another limitation.

**Conclusion**

Laboratory staff should gain proper and sufficient information about the specificities of the analytes, the principles of the instrumentation and methodology in detail.
Preanalytical and analytical variations related with complex molecular structure of FT4, the equilibrium between T4 and carrier proteins, and other interferences in competitive immunoassay methodology should be taken into consideration to ensure the validity of the result. The interaction effects of both the hospital and age groupings on TSH were no meaningful. We believe that this finding supports the consensus of measuring TSH in guidelines for CH [35] or in annual health assessments when the accurate cut-off values were determined in children with narrowed age groups, for the purpose of careful neurodevelopmental and neurosensory evaluations started early in life and taking into account disease severity at diagnosis and providing appropriate interventions as required.

In this study, FT4 and TSH reference limits differed statistically regarding narrowed age groups; accurate cut-off values were determined for each week, especially in the first 3 weeks of neonatal period. And these reference limits presented similar profiles among all three hospitals. Our reference limits may be recommended in pediatric follow-ups using the Centaur XP or the Immulite 2000, considering the conditions of prematurity, birth-weight, multiple births or infection.

**Conflict of interest statement:** There is no conflict of interest in this study.

**References**

1. Demers LM, Spencer CA. Laboratory medicine practice guidelines: laboratory support for the diagnosis and monitoring of thyroid disease. Clin Endocrinol (Oxf) 2003;58:138–40.
2. Colantonio DA, Kyriakopoulou L, Chan MK, Daly CH, Brinc D, Venner AA, et al. Closing the gaps in pediatric laboratory reference intervals: a CALIPER database of 40 biochemical markers in a healthy and multiethnic population of children. Clin Chem 2012;58:854–68.
3. Clinical and Laboratory Standards Institute. Defining, Establishing, and Verifying Reference Intervals in the Clinical Laboratory. Third approved guideline, EP28A3c. Wayne, PA: Clinical and Laboratory Standards Institute, 2010.
4. Kapelari K, Kirchlechner C, Högl W, Schweitzer K, Virgolini I, Moncayo R. Pediatric reference intervals for thyroid hormone levels from birth to adulthood: a retrospective study. BMC Endocr Disord 2008;8:15. Available at: http://www.biomedcentral.com/1472-6823/8/15.
5. Box GE, Cox DR. An analysis of transformations. J R Stat Soc 1964;B26:211–52.
6. Horn PS, Feng L, Li Y, Pesce AJ. Effect of outliers and nonhealthy individuals on reference interval estimation. Clin Chem 2001;47:2137–45.
7. Milliken GA, Johnson DE. Analysis of Messy data. Nonreplicated experiments, Vol. 2. New York: Van Nostrand Reinhold, 1989.
8. Mehari A, Challa F, Gebreyesus G, Alemayehu D, Seifu D. Establishment of reference intervals of thyroid function tests from cord blood of neonates in two selected hospitals, Addis Ababa, Ethiopia. BMC Pediatr 2016;16:118.
9. Mutlu M, Karagüzel G, Akyaziçiçlioğlu Y, Eyüpoğlu I, Okten A, Aslan Y. Reference intervals for thyrotropin and thyroid hormones and ultrasonographic thyroid volume during the neonatal period. J Matern Fetal Neonatal Med 2012;25:120–4.
10. Hübner U, Englisch C, Werkmann H, Butz H, Georgs T, Zabransky S, et al. Continuous age-dependent reference ranges for thyroid hormones in neonates, infants, children and adolescents established using the ADVIA Centaur Analyzer. Clin Chim Lab Med 2002;40:1040–7.
11. Omuse G, Kassim A, Kiigu F, Hussain SR, Limbe M. Reference intervals for thyroid stimulating hormone and free thyroxine derived from neonates undergoing routine screening for congenital hypothyroidism at a university teaching hospital in Nairobi, Kenya: a cross sectional study. BMC Endocr Disord 2016;16:23.
12. Elmlinger MW, Kühnel W, Lambrecht HG, Ranke MB. Reference intervals from birth to adulthood for serum thyroxine (T4), triiodothyronine (T3), free T3, free T4, thyroxine binding globulin (TBF) and thyrotropin (TSH). Clin Chem Lab Med 2001;39:973–9.
13. Najam Y, Khan M, Ihali F, Alam A. Distribution of T4 TSH values in children – the Shifa experience. J Pak Med Assoc 2003;53:26–8.
14. Henderson MP, Grey V. Establishing and evaluating pediatric thyroid reference intervals on the Roche Modular Analytics E170 using computational statistics and data-mining techniques. Clin Biochem 2011;44:767–70.
15. Soldin SJ, Cheng LL, Lam LY, Werner A, Le AD, Soldin OP. Comparison of FT4 with log TSH on the Abbott Architect ci8200: pediatric reference intervals for free thyroxine and thyroid-stimulating hormone. Clin Chim Acta 2010;411:250–2.
16. Bailey D, Colantonio D, Kyriakopoulou L, Cohen AH, Chan MK, Armbruster D, et al. Marked biological variance in endocrine and biochemical markers in childhood: establishment of pediatric reference intervals using healthy community children from the CALIPER cohort. Clin Chem 2013;59:1393–405.
17. Aldrimer M, Ridefelt P, Rödöö P, Niklasson F, Gustafsson J, Hellberg D. Reference intervals on the Abbot Architect ci8200 for serum thyroid hormones, lipids and prolactin in healthy children in a population-based study. Scand J Clin Lab Invest 2012;72:326–32.
18. Chafer EA, Fiorenzano R, Chilelli C, Linares V, Arey G, Herzovich V, et al. Age-specific thyroid hormone and thyrotropin reference intervals for a pediatric and adolescent population. Clin Chim Lab Med 2012:50:885–90.
19. Djemli A, Van Vliet G, Belgoudj I, Lamberti M, Delvin EE. Reference intervals for free thyroxine, total triiodothyronine, thyrotropin and thyroglobulin for Quebec newborns, children and teenagers. Clin Biochem 2004;37:328–30.
20. Marks AG, La Franchi SH. Assessing thyroid function in infants and children. In: Brent GA, editor. Thyroid function testing. New York: Springer, 2010:173–85.
21. Zurakowski D, Di Canzio J, Majzoub JA. Pediatric reference intervals for serum thyroxine, triiodothyronine, thyrotropin, and free thyroxine. Clin Chem 1999;45:1087–91.
22. Nelson JC, Clark SJ, Borut DL, Tomei RT, Carlton EI. Age-related changes in serum free thyroxine during childhood and adolescence. J Pediatr 1993;123:899–905.
23. Lahoti A, Frank GR. Laboratory thyroid function testing: do abnormalities always mean pathology? Clin Pediatr 2013;52:287–97.
24. Murphy N, Hume R, Toor H, Matthews TG, Ogston SA, Wu S, et al. The hypothalamic-pituitary-thyroid axis in preterm infants; changes in the first 24 hours of postnatal life. J Clin Endocrinol Metab 2004;89:2824–31.
25. Soldin SJ, Morales A, Albolar F, Lenherr S, Riffai N. Pediatric reference ranges on the Abbott IMx for FSH, LH; prolactin, TSH, T4, T3, free T4, free T3, T-uptake, IgE, and ferritin. Clin Biochem 1995;28:603–6.
26. Doggai R, Ingrand J. Free thyroxine immunoassay: analytical review. Ann Biol Clin 2015;7:199–211.
27. Thienpont LM, Van Uytfanghe K, Beastall G, Faix JD, Leiris T, Miller WG, et al. Report of the IFCC Working Group for Standardization of Thyroid Function Tests; part 2: free thyroxine and free triiodothyronine. Clin Chem 2010;56:912–20.
28. De Grande LA, Van Uytfanghe K, Reyners D, Das B, Faix JD, MacKenzie F, et al. Standardization of free thyroxine measurements allows the adoption of a more uniform reference interval. Clin Chem 2017;63:164–2.
29. Thienpont LM, Van Uytfanghe K, De Grande LA, Reyners D, Das B, Faix JD, et al. Harmonization of serum thyroid-stimulating hormone measurements paves the Way for the adoption of a more uniform reference interval. Clin Chem 2017;63:1248–60.
30. Katzman BM, Ness KM, Algeciras-Schimnich A. Evaluation of the CLSI EP26-A protocol for detection of reagent lot-to-lot differences. Clin Biochem 2017;50:768–71.
31. Liu J, Tan CH, Loh TP, Badrick T. Detecting long-term drift in reagent lots. Clin Cem 2015;61:1292–8.
32. Weil PA. The diversity of the endocrine system. In: Rodwell VW, Bender DA, Botham KM, Kennelley PJ, Weil PA, editors. Harper’s illustrated biochemistry. US, McGraw-Hill, 2015:698–517.
33. Ömer B. Protein yapılı ve amino asit türevi hormonlar. In: Gürdöl F, editor. Tibbi Biyokimya. İstanbul, Nobel Tip Kitabevi 2015;405–34.
34. Tarim O. Thyroid hormones and growth in health and disease. J Clin Res Ped Endo 2011;3:51–5.
35. Leger J, Olivieri A, Donaldson M, Torresani T, Krude H, van Vliet G, et al. European society for paediatric endocrinology consensus guidelines on screening, diagnosis, and management of congenital hypothyroidism. J Clin Endocrinol Metab 2014;99:363–84.
36. Jialal I, Winter WE, Bertholf RL. Disorders of the pituitary. In: Burtis CA, Bruns D, editors. The fundamentals of clinical chemistry and molecular diagnostics. Missouri: Elsevier Saunders, 2015:769–84.
37. Semple RK. Biochemical endocrinology. In: Baynes JW, Dominiczak MH, editors. Medical biochemistry. China: Elsevier Saunders, 2014:513–35.
38. Reinehr T. Thyroid function in the nutritionally obese child and adolescent. Curr Opin Pediatr 2011;23:415–20.
39. Daly CH, Liu X, Grey VL, Hamid JS. A systematic review of statistical methods used in constructing pediatric reference intervals. Clin Biochem 2013;46:1220–7.
40. Quality requirements. Desirable specifications for total error, imprecision, and bias, derived from intra- and inter-individual biologic variation. https:/ /www.westgard.com/biodatabase1.htm. (Last accessed: May 2017).
41. Zhu L, Zhang X, He X, Yang X, Wang Y, Wang C, et al. Reference intervals for serum thyroid hormones in preterm hospitalized infants. J Pediatr Endocrinol Metab 2013;26:463–7.
42. Mengreli C, Kanaka-Gantenbein C, Girginoudis P, Magiakou M, Christakopoulou I, Giannouli-Karantana A, et al. Screening for congenital hypothyroidism: the significance of threshold limit in false-negative results. J Clin Endocrinol Metab 2010;95:4283–90.
43. Kaluarachchi DC, Colaizy TT, Pesce LM, Tansey M, Klein JM. Congenital hypothyroidism with delayed thyroid-stimulating hormone elevation in premature infants born at less than 30 weeks gestation. J Perinatol 2017;37:277–82.
44. Kucharska AM, Beń-Skowronek I, Walczak M, Ołtarzewski M, Szalecki M, Jackowska T, et al. Congenital hypothyroidism – Polish recommendations for therapy, treatment monitoring, and screening tests in special categories of neonates with increased risk of hypothyroidism. Endokrynol Pol 2016;67:536–47.

Supplementary Material: The online version of this article offers supplementary material (https://doi.org/10.1515/tjb-2017-0342).