Establishing Laboratory-Specific Reference Intervals for TSH and fT4 by use of the Indirect Hoffman Method

Sylwia Płaczkowska (sylwia.placzkowska@umed.wroc.pl)
Wroclaw Medical University
Małgorzata Terpińska
Wroclaw Medical University
Agnieszka Piwowar
Wroclaw Medical University

Research Article

Keywords: Reference Values, Thyroid Hormones, Clinical Laboratory Information Systems, Indirect Hoffman Method

Posted Date: August 3rd, 2021

DOI: https://doi.org/10.21203/rs.3.rs-763059/v1

Version of Record: A version of this preprint was published at PLOS ONE on January 7th, 2022. See the published version at https://doi.org/10.1371/journal.pone.0261715.

Abstract

Background: The results of examinations of laboratory parameters are the basis of appropriate medical decisions. The availability of reliable and accurate reference intervals (RIs) for each laboratory parameter is an integral part of its appropriate interpretation. Each medical laboratory should confirm their RIs. Up to date reference intervals for thyroid function hormones are still a matter of ongoing controversy.

The aim of the study was application of the indirect Hoffman method to determine RIs for TSH and fT4 based on the large data pools stored in laboratory information systems, and the comparison of these RIs to generally used RIs.

Material and methods: The TSH and fT4 routine examination results of a hospitalized and outpatients population over five years (2015-2019) were collected, and reference limits were established by the improved Hoffmann method after the exclusion of outliers. Comparative verification of established RIs was conducted with the RIs values provided by test manufacturers and literature data.

Results: The various RIs in different age groups in examined populations were observed. For TSH RIs varied between different age groups, with a narrower range of RIs in the studied adult population, while for children, a shift of both reference boundaries toward higher values in comparison to manufacturers' data was observed. RIs estimated for fT4 were very similar to the manufacturer and literature data.

Conclusion: Thyroid hormone levels change during a person's lifetime and vary between sexes, but this difference does not always influence the clinical interpretation of laboratory results in the context of RIs. The use of indirect methods is justified due to the ease and low cost of their application.

Background

The results of examinations of laboratory parameters provide useful information for assessing the current health condition of patients. They are necessary for early detection and recognition of disturbances as well as for making appropriate medical decisions. For the purpose of interpretation of laboratory results, the use of the concept of reference intervals (RIs) is currently generally accepted in laboratory medicine [1]. The Clinical Laboratory Standards Institute (CLSI) has released a relevant guideline (C28-A3c) for the evaluation of RIs. According to the CLSI recommendations, an RI is defined as the interval between which 95% of values of a reference population fall into, and includes two extreme reference limits – boundaries derived from the distribution of reference values, which could be associated with good health but also with other physiological or pathological conditions [2, 3]. It is recommended that medical laboratories determine their own local reference intervals to embrace the variations in local populations and the methods and equipment used in particular laboratories. The confusion of RIs with clinical decision limits (CDLs) still remains an issue, especially in paediatric and geriatric age groups, where it presents a significant diagnostic problem [4].

CLSI currently recommends a direct method based on the collection of a minimum of 120 samples from members of a specific preselected reference population, making measurements, and then determining the range which includes 95% of the all measured values using a parametric (mean ± 2SD) or non-parametric method (2.5th and 97.5th percentile) [2]. The requirement that each laboratory determines its own reference intervals is virtually impossible to perform in practice because of the tremendous amount of time and money required to carry out additional laboratory tests and gather the appropriate reference group. Thus most laboratories adopt external sources for RIs, often without taking into account the problems of transferring values between different populations or laboratory methods [5]. Data provided by the test manufacturer is the most often used source of reference intervals, because such information is required from reagents suppliers by the ISO 15189:2013 standards [6]. At the same time, the methods and processes for determination of reference intervals using indirect methods have been in development for over 50 years, but still are not widely applied. This alternative approach is based on the statistical analysis of results generated as part of routine laboratory testing in hospital and outpatient clinics in order to determine reference intervals [3]. Indirect methods eliminate results that do not fit the assumed hypothetical model of the distribution of results, generally a normal distribution, and designate the RI as the central or marginal 95% of the selected results. The application of the indirect method has major potential advantages compared with direct methods. In particular, this process is faster and cheaper; it involves no inconvenience, discomfort, or any additional risk to patients; and laboratory staff need not examine any additional samples. Therefore, additional costs are avoided, which is important in the modern and effective management of the medical laboratory and the hospital [7].

Establishing RIs is particularly problematic for constituents with a large diversity of existing biological variation and inter-population differences, as for example is observed for thyroid hormones, especially thyrotropin – TSH (Thyroid Stimulation Hormone) as well as the thyroid hormones free triiodothyronine (fT3) and free thyroxine (fT4). The prevalence of thyroid dysfunction in the general world population is estimated to be between 1 and 2% [8]. There are still discrepancies between TSH, fT3 and fT4 reference values applied to the diagnosis of thyroid dysfunction not only between laboratories, but also in the scientific literature to date. It seems to erroneous to apply the concept of universal limits reference intervals for thyroid function hormones, especially for TSH [9]. Regarding this fact, it appears important and useful to establish reference intervals using a costless, optimized indirect statistical method.

At present, it is assumed that the value of TSH in a healthy general population is 0.4-4.0 mIU/L, which is the result of the fairly large inter-individual variability of this parameter. However, the variability of the value in an individual is much smaller, and the value determined in a state of hormonal equilibrium can be regarded as an individual's set-point. Due to the fact that slight changes in fT4 value correspond to a significant change in TSH, it is used in the screening of disorders of thyroid hormones. Therefore, it is advisable to find out about this individual point by measuring the level of TSH for each person in times of health. This allows for earlier detection of disturbances, even without direct comparison to the reference interval [10].

According to the current recommendations, in order to confirm thyroid primary dysfunction, TSH and fT4 determinations should be performed first, while fT3 determinations should be ordered only in specific cases. At the same time, it is not advisable to order isolated ft4 testing, it should always be combined with TSH [11]. For this reason, in our work, we focused primarily on determining the reference intervals for TSH and FT4.

Material And Methods

Objective

The aim of this study was to provide the possibility to utilise the large data pools of TSH and fT4 stored in laboratory information systems (LIS) to determine reference intervals using the indirect Hoffman method, and conducting a comparison of RIs to generally used reference limits.

Laboratory methods

The a third generation test of TSH and fT4 examinations were performed utilising a chemiluminescence method on the Atellica IM analyser. The linear range of this method was 0.008-150.000 mIU/L and 0.1–12.0 ng/dL for TSH and fT4, respectively. The laboratory intra-series analytical coefficient of variation (CVa) for TSH and fT4 was assessed as 6.4% and 2.5%, respectively.

Data gathering

The study was performed in accordance with the Declaration of Helsinki and consent was approved by the Wroclaw Medical University Bioethical Commission (decision No. 537/2018). Based on the bioethics committee decision patient informed consent was waived due the retrospective nature of the study conducted on a completely anonymous and only numerical values.

All the laboratory results of TSH and fT4 examinations, together with the patient's age, sex and date of examination, archived in the Laboratory Information System of Department of Laboratory Diagnostics and derived from patients hospitalized at the University Clinical Hospital in Wroclaw during the five-year observation (1st January 2015– 31st December of 2019 year) were included in the study, exported to an Excel spreadsheet, and used for statistical analysis. For TSH: 105 927 (65 163 from women and 40 764 from men) results were included in the study and for fT4–41 400 (26406 from women and 14 994 from men), respectively. Out of these, 5756 (1 962 from men and 3 802 from women) and 1 314 (486 from men and 785 from women) outliers of TSH and fT4 were eliminated as outliers by a two-sided Tukey's test. Eventually 100 171 (38 802 men and 61 361 women) and 40 086 (14 508 men and 25 621 women) results for TSH and fT4, respectively, were included in further calculations. Participants' age range for the analysed hormones was from 0 to 109 years.

Statistical analysis

Firstly, before performing any kind of analysis, all data were logarithmically transformed because of the strong right-skewness of the data distribution. Next, all data were divided into 8 age groups (< 1 y., ≥ 1 y. <

6, ≥ 6 y., < 12, ≥12 y. <18, ≥ 18 y. <40, ≥ 40 y. <65, ≥ 65 y. <90, ≥ 90 y.), which reflected the physiological changes associated with human ontogenetic development. The Tukey test was used to reject outliers separately for each age group. The Hoffman method – an indirect statistical method based on the graphic distribution of lnTSH and lnfT4 values – was applied in each age group for all participants and in regard to sex. In accord with this method, a Q-Q chart was created in each study age group. In the next step an elimination of outliners on the basis of Q-Q plot was visually conducted and the distribution of remaining data was used to calculate the regression equation using the least squares method. The regression line included the middle range of data, which was initially fitted by visual inspection. The linear line over the linear part of the Q-Q chart was described by the following equation:

$y = a * x + c$

where y = the lnTSH or lnfT4 value, respectively, x = normal theoretical quantile of the standard normal distribution (μ = 0, σ = 1), a = the slope of the regression line, c = the intercept.

In the next step, extrapolation the linear regression equation to the boundaries of the 95% confidence interval were conducted, as follows: Lower Reference Limit (LRI) = a * (-1,96) + b and Upper Reference Limit (URL) = a * 1,96 + b. All statistical procedures mentioned above were performed on logarithmically transformed data. The antilogarithm was applied in the last step for the calculation of RIs values on the basis of the linear regression equation. A typical Q-Q plot is presented in Fig. 1.

Linear least squares regression was applied to the middle part of the data distribution

Then Reference Change Value (RCV) was used in order to determine the clinical significance of the relationship between RIs in all selected age groups and manufacturers and published RI. RCV was calculated according to the formula [12]:

$RCV = 2^{1/2} * Z * (CVa^2 + CVi^2)^{1/2}$

where Z is the probability selected for significance, the chosen Z value of 1.96 corresponds to a significance level of 0.05, CVa – the analytic variation (inter-assay variation estimated from our laboratory data) and CVi – the within-subject biologic variation (data from Ricos et al. [13]).

Statistical analyses were performed using Statistica 13.1 PL .

Results

The RIs values obtained for TSH and fT4 in each separate age group for all study participants and by gender are summarized in Fig. 2.

For TSH the estimated values of RIs decreased with the age of the patients and simultaneously were tended to decrease ranges. Whereas, for the fT4 value, an inverse relationship was observed with regard to the range of RI, which increased with age, while the RI limits did not indicate significant shifts in

values. TSH detailed numerical data for all results with the comparison on LRI and URI for women and man are provided in the Table 1.

The comparison of LRI and URI between women and men in particular age groups showed the greatest difference for LRI of 11.7% TSH in the age group of < 1 years. The same analysis of URI for TSH indicated the highest differences in group > 90 years, but also for infants it was one of the highest. It is worth to mentioning the obtained differences in percentage for LRI and URI are smaller than the extra-individual variability for TSH – 24.9% [12]. Therefore, the results obtained for all study participants in a given age group were considered as general RIs values and sex differences were not taken into account.

The applied Hoffman method shows various TSH RIs in different age groups. The LRIs gradually decreased in subsequent age groups from children up to the age of ≥ 90, when for URI the same trend was observed up to < 65 and then increased in participants older than 65. Differences greater than those determined by RCV for comparisons between each age groups were revealed for LRI both in all subjects and in samples of women and men separately. In all cases, they concerned the results of the oldest patients and groups of children and adolescents.

The LRIs for participants aged < 40 established in this study were within TSH RIs provided by the manufacturer. For the three oldest participant groups, the LRIs were lower than reported by manufacturers, while, TSH were higher in children and adolescents group and groups ≥ 65 years. For adults URI estimated by Q-Q plots were lower them provided by manufacturer. Generally, the Hoffman method revealed a shift of both reference boundaries toward higher in children and adolescents, the narrower RIs for TSH in the studied adult population, while for seniors a wider RIs boundaries were observed values in comparison to manufacturers' data. The difference between LRI and URI established by the Hoffman method with RIs reported by the manufacturer did not exceed the acceptable 56 % RCV for TSH in any age group (Table 2).

RIs established for fT4 presented Table 3 were characterized by the low variability in comparison between sex and age groups.

The sex differences for fT4 LRI were less than 10% and the highest were observed form infants, ≥ 40 y. <65 and ≥ 90y. age groups. Except, the youngest group, the sex difference of URI was lower than even 5%, which is only a quarter of the RCV value established for fT4 in this study. Therefore, further analyses were carried out without regarding the sex.

The differences calculated for comparison between subsequent age groups was in all cases under the value of RCV 21.6%. It is worth to mentioning the obtained differences in percentage for LRI and URI are smaller than the extra-individual variability of fT4 -12.1% [12].

Comparison of calculated fT4 reference intervals with RIs reported by manufacturer are presented in the Table 4.

LRI for fT4 were almost the same with the ranges provided the manufacturers' and the higher differences were noticed for infants URI'. Generally, for infants and children LRI were identical, when URI was slightly higher than provided by manufacturer. For group adolescents were noticed the narrower RIs which were placed within the manufacturer' data. A minimal shift toward the lower values of RIs were revealed for age group adults ≥ 40 and < 65 years. The similar trend was observed for seniors but only for LRI value.

Discussion

Currently, medical decisions mainly are based on results of laboratory diagnostic tests, which are used to confirm, exclude, classify or monitor disease in order to guide treatment. Therefore, establishing appropriate reference ranges is crucial for the correct interpretation of laboratory results. Unfortunately, our observations show that the main source of reference intervals in Poland is data from literature, often based on research carried out among the general population or with a different genetic and/or cultural profile [14]. The second source of information about the expected values is provided by the laboratory reagent manufacturer. The analysis of the RI values in these materials shows that although the manufacturers declare that their procedures comply with the CLSI recommendations, the reference groups are very often too small in number and very poorly characterized in many important aspects, such as race, disease, or body weight [15–17]. Very rarely, a laboratory procedure is also used in order to verify the reference intervals provided by the manufacturer. This approaches involve a comparison of the values of the data approved by CLSI with a laboratory reference group of at least 20 people, which is a substitute for the standard method based on the analysis of the results of at least 120 people accepted as reference [1]. Considering the above-mentioned common weaknesses or even shortcomings in determining reference intervals, indirect methods and their possibility of implementing computer applications [18], may will be the best alternative for regional hospitals and field laboratories serving the same population of a given region. This is especially useful in terms of accessing hard-to-reach paediatric or geriatric populations.

There are a limited number of publications available using the Hoffman for RIs estimation in the general population, and particularly in the paediatric population. Initiatives are being taken around the world to harmonize paediatric reference intervals to improve paediatric diagnostics with a range of effects from national to global. Reported thyroid hormone RIs in childhood vary widely and knowledge on the matter is sparse. Reported RIs are mainly based on direct methods, which, combined with limited availability to the child population, makes them difficult to implement for the average laboratory [19]. We showed that paediatric values of RI for TSH were higher than those provided by our reagents manufacturer, and the greatest differences were observed for LRI (Table 2). The literature overview indicated, the lower limit for TSH ranged between 0.32 and 1.30 mIU/L in children aged ≥ 1 years, and the highest value is similar to the value observed in our study for children aged ≥ 1 (LRI for children ≥ 1 y. <6: 1.03 mIU/L and 0.79 mIU/L for children ≥ 12 y. <18). The upper limit for TSH in an overview by Onsesveren et al. [9] ranged from 2.36 to 6.57 mIU/L, as reviewed by their studies.

In CALLIPER project the RI values determined by the direct CLSI method were compared with the values determined by the Hoffman method showed a much wider range of RIs, determined by the Hoffman method for TSH and fT4. This study indicate the limited usefulness of the values determined by the Hoffman method, especially if it is carried out on a limited number of results from centres providing highly specialized care for children, which may cause a large percentage of pathological results [20]. The values determined in our study by the indirect RI method for TSH were very similar to those obtained in the CALIPER study, although direct comparison is very difficult due to the use of different age ranges in the CALIPER study for the three different immunochemical analysers applied. However, in both studies a significant TSH URL decrease was observed in participants aged 12–14 years to values similar to those of adults. As highlighted by the authors themselves, establishing assay-specific reference intervals for immunoassays is a need of great importance [21]. The observed differences may be due to the recognition of different epitopes of TSH by monoclonal antibodies used in various kits and microheterogeneity TSH molecules [22]. In our opinion, this approach is far more feasible when using the indirect method on a hospitalised population.

An estimation of TSH RIs for a Mediterranean population over 15 years old was performed by Lo Sasso et al. [23]. The TSH RIs estimated in this study, which had enrolled over 22 602 individuals, were 0.18–3.54 mIU/L in the general population – 0.19–3.23 mIU/L in men and 0.18–3.94 mIU/L in women. Moreover, their results showed a significant interaction between sex and age, suggesting that the effect of age on TSH between the sexes. However, this comparison was based on a statistical test, not on RCV like our study. Our results in the adult population indicated significantly higher values of LRI and much more similar values for URI. This difference may be due to the use of a different reagent kit and analyser in Lo Sasso's et al. [23] study, for which the RIs values provided by the reagent manufacturer were significantly lower than for the Siemens kit used in our study (0.20–4.20 mIU / L Cobas e801 vs. 0.55–4.78 mIU / L Atellica IM). This fact again indicates the necessity to determine the assay-specific RIs for the TSH. Moreover, our study also confirmed the trend observed by Lo Sasso et al. [23] of higher URI for adult women and a decrease of both LRI among participants from 18 to < 90 years old. Similar results to ours were also obtained among the Turkish adult population by Inal et al. [7] from results of TSH and fT4 from hospitals and health centres; data were prepared only by logarithmic transformation, with repeated exclusion of outliers and computation of non-parametric 2.5th and 97.5th percentiles. Indirect RIs were then compared with RIs established by the IFCC direct method (RIs currently used in Inal laboratory) and manufacturers' recommendations. This comparison revealed the largest difference in LRI, where the manufacturer’s value was 1.6 mIU/L lower than established by the indirect method and 1.3 mIU/L lower than estimated through use of the direct method. The differences between direct and indirect methods both for LRI and URI were much less visible, with 0.40 vs. 0.43 for LRI and 3.93 vs. 4.2 for URI, respectively [7].

The study closely related to our research were conducted by Katayev et al. [24] in 2008 on both, female and male adult TSH results collected in 6 laboratories from the Laboratory Corporation of America. Without taking age groups into account, the estimated RIs were in the range of 0.44–3.05 mIU/L for all 129443 results. Comparison of our study to Katayev et al. results [24], the greatest differences were also

seen for the LRI, and ranged from 0.19 mIU/L (30.1%) for the ≥ 18 y < 40 group to 0.05 mIU/L (-12.8%) for the ≥ 90 years. The higest absolute difference (2.15 mIU/L) was observed for URI in the ≥ 90 age group, but the percentage difference was as high as 41.3%. Even though Katayev et al. [24] found a substantial (though still below the RCV) difference between their RIs and the literature data, they emphasised that the proper use of statistical techniques and the large number of observations allowed to estimate RIs with accuracy using the indirect method.

Apart from the TSH tests, the second most important parameter in the diagnosis of thyroid disorders is the concentration of fT4 [25]. As we indicated above, the literature data shows large differences in RIs regardless of the method used for its estimation. According to a literature overview by Onsesveren et al. [9], the LRIs for fT4 range from 0.54 to > 0.78 ng/dL (from7.0 to > 10 pmol/L), while the URIs vary from 1.20 to > 2.33 ng/dL (from 15.5 to > 30.0 pmol/L). In the same paper, the authors presented RIs derived from their prospective study – 1.07–1.62 ng/dL (13.8–20.8 pmo/L) assessed as 2.5th and 97.5th results obtained among 4273 thyroid disorder-free children at a median age of 6 years [9].

The CALIPER study defined the childhood fT4 RIs as 0.89–1.70 ng/dL for infants and 0.89–1.37 ng/dL for children under 19 years by Abbott ARCHITECT analyser [21] – which figures similar to results obtained by our indirect method and using another analyser.

The fT4 RIs discussed above take into account values obtained by the direct method and are generally consistent with our results, even though they are potentially influenced by many factors, such as weight, sex, ethnicity, and the laboratory method of determination. However, fT4 RI values among children derived from the indirect method are absent in literature. The most similar study was published by Kapelari et al. [26] for the Austrian child population. In this study, a posteriori direct methods were applied to child hospital populations in five subsequent age groups. Authors concluded that the variation of fT4 concentration was the highest in first months of life and then RIs values decreased in a continuous fashion. LRI were lower in infants, at 0.71 ng/dL (9.17 pmol/L), and then increased up to about 0.82 ng/dL (10.6 pmol/L) in older age groups. In our study we observed similar values, but there were no real differences in values. We noticed the same trend for URI (1.61 ng/dl for infants and 1.33 ng/dl for adolescents). This was generally consistent with the Kapelari et al. [26] study, although their values were higher: 1.97 ng/dL for infants and 1.78 ng/dL for adolescents. For adults, fT4 RIs were estimated using the direct method in euthyroid populations; 0.78–1.55 ng/dL (10–20 pmol/L) were found in a British study [5] and 0.70–1.56 ng/dl (12.29–20.03 pmol/L) among the population of the Republic of Srpska [27]. These results are generally in compliance with those established in our study (with minimal higher values) and data from the manufacturers of the reagents we used (minimal lower values). Moreover, our results corresponded with sparse studies using the indirect method, such as the study by Jakubowski et al. [28] in Poland, which was conducted over 16 years. In this study, the authors applied many exclusion criteria linked to the pre-analytical phase and clinical factors that suggested severe or endocrinological disease and to testing the same patient more than once. The 2.5th and 97.5th percentile of that prepared data were adopted as the low and upper RIs for fT4: 0.86–1.23 ng/dl (11.13–15.87 pmol/L) and were

narrower than provided by their reagent manufactures at 9.01–19.05 pmol/L. The results of ours study indicated a substantial shift in URI to higher values than in Jakubowski et al. [28].

The determination of reference intervals for thyroid hormones is a subject of ongoing debate due to poor standardisation of immunochemical methods and the lack of unambiguous reference materials [29] as well as large inter-population differences, along with variability related to the age and gender of patients within the same population. The legitimacy of determining reference intervals for thyroid hormones with an indirect method based on the results of hospital tests is additionally supported by the fact that these tests are screened, and therefore a large portion of the results comes from people without thyroid disorders. The use of such an approach additionally eliminates the differences in values that may result from different methods of sampling in different hospital wards and collection points [30].

Encouraged by the hints contained in Jones' publication [3], we searched for the values of reference intervals characteristic for the population of our hospital and the methods and apparatus used in our laboratory. At the same time, we do not forget that no intervals are perfect and final, and the results obtained by the indirect method, even if they are not absolutely accurate, are closer to the actual state of the population of a given region, because they take into account the analytical and biological variability of the analysed parameter.

Conclusions

TSH and fT4 levels change during a person's lifetime and vary between sexes, but the difference does not always have an influence on the clinical interpretation of laboratory results in the context of reference intervals. The discussed differences in RI estimated in this study regarding to literature data and manufacturers' information, reflect the actual distribution of results in the population of a given region and are consistent with the idea of screening the functioning of the thyroid gland firstly by TSH and fT4. Differences between our RIs in comparison to other direct and indirect studies were caused by analytical (different antibody characteristics used in reagents, lack of harmonization) and epidemiological factors (different populations, socio-economical status and undefined geographic covariates). Comparing the RIs obtained by the indirect method can also be used to confirm compliance with the RIs values provided by the manufacturer of the reagent kits. Considering the differentiation of reference values determined both by direct and indirect methods in different populations, in our opinion, the use of indirect methods is justified due to their low cost and ease of application, and above all due to the possibility of imaging the distribution of results in the population of a given region rather than reliance on data from other geographic and cultural areas.

The greatest limitation of this study was using unselected data from our LIS. We did not have the possibility to download the data with the ICD code and delete the results that were repeatedly performed on the same patient. However, the applied Hoffman method is based on the use of values located in the middle of the distribution to estimate the reference intervals, which was preceded by the rejection of outliers, i.e. pathological values. With such an approach, the Hoffman method is relatively robust in the

occurrence of outliers. The another important limitation was a lack of verification of the indirect RIs established using the direct method and which would have enabled us to confirm that the exclusion of patients with thyroid disorders is not required to obtain proper reference intervals from hospital populations with many thousands of records.

Abbreviations

CDLs: Clinical Decision Limits

CLSI: The Clinical Laboratory Standards Institute

CVa: the intra-series analytical coefficient of variation

CVi – the within-subject biologic coefficient of variation

fT3: free Triiodothyronine

fT4: free Thyroxine

LIS: Laboratory Information Systems

LRL: Lower Reference Limit

RCV: Reference Change Value

RIs: Reference Intervals

TSH: Thyroid Stimulation Hormone

URL: Upper Reference Limit

Declarations

Ethics approval and consent to participate

The study was performed in accordance with the Declaration of Helsinki and consent was approved by the Wroclaw Medical University Bioethical Commission (decision No. 537/2018). Based on the bioethics committee decision patient informed consent was waived due the retrospective nature of the study conducted on a completely anonymous and only numerical values.

Availability of data and materials

The datasets generated and analysed during the current study are not publicly available due to they are medical records but are available from the corresponding author on reasonable request.

Competing interests

The authors declare that they have no competing interests.

Contribution statement

SP and AP conceived the concept of the study. All authors contributed to the design of the research. MT and SP were involved in data collection. All authors analysed the data. All authors edited and approved the final version of the manuscript

References

1. Ozarda Y. Reference intervals: current status, recent developments and future considerations. Biochem Med. 2016; 26: 5–16.
2. CLSI and IFCC. C28-A3 document; Defining, establishing and verifying reference intervals in the clinical laboratory: approved guideline-third edition. 2008; 28: 1–76.
3. Jones GRD, Haeckel R, Loh TP, et al. Indirect methods for reference interval determination – review and recommendations. Clin Chem Lab Med. 2018; 57: 20–9.
4. Barth JH, Spencer JD, Goodall SR, Luvai A. Reference intervals for thyroid hormones on Advia Centaur derived from three reference populations and a review of the literature. Ann Clin Biochem. 2016; 53: 385–9.
5. Arzideh F, Brandhorst G, Gurr E, et al. An improved indirect approach for determining reference limits from intra-laboratory data bases exemplified by concentrations of electrolytes. J Lab Med. 2009; 33: 52–66.
6. International Organization for Standardization. ISO 15189:2013-05 Medical laboratories – Requirements for quality and competence, version in Polish. http://sklep.pkn.pl/pn-en-iso-15189-2013-05p.html
7. Inal TC, Serteser M, Coskun A, et al. Indirect reference intervals estimated from hospitalized population for thyrotropin and free thyroxine. Croat Med J. 2010; 51: 124–30.
8. Vanderpump MPJ, Tunbridge WMG. Epidemiology and prevention of clinical and subclinical hypothyroidism. Thyroid. 2002; 12: 839–47.
9. Onsesveren I, Barjaktarovic M, Chaker L, et al. Childhood Thyroid Function Reference Ranges and Determinants: A Literature Overview and a Prospective Cohort Study. Thyroid. 2017; 27: 1360–9.
10. Pearce SHS, Brabant G, Duntas LH, et al. 2013 ETA Guideline: Management of Subclinical Hypothyroidism. Eur Thyroid J. 2014; 2: 215–28.
11. LeFevre ML. Screening for Thyroid Dysfunction: U.S. Preventive Services Task Force Recommendation Statement. Ann Intern Med [Internet]. http://annals.org/article.aspx?doi=10.7326/M15-0483. Accessed 27 Apr 2021.

12. Fraser CG. Biological Variation: From Principles to Practice. Washington, USA: AACC Press; 2001: 67–90.
13. Ricos C, Alvarez V, Cava F, et al. Desirable Biological Variation Database specifications. 2014. https://www.westgard.com/biodatabase1.htm. Accessed 15 May 2021.
14. Placzkowska S, Terpinska M, Piwowar A. The importance of establishing reference intervals - is it still a current problem for laboratory and doctors? Clin Lab. 2020;66. doi: 10.7754/Clin.Lab.2020.191120
15. Reagent leaflet Atellica IM Thyroid Stimulating Hormone 3-Ultra (TSH3UL), Siemens Healthcare Diagnostics Inc., rev.01, 2017-05.
16. Reagent leaflet Atellica IM Free Triiodothyronine (FT3), Siemens Healthcare Diagnostics Inc., rev.01, 2017-06
17. Reagent leaflet Atellica IM Free Thyroxine (FT4), Siemens Healthcare Diagnostics Inc., rev.01, 2017-12.
18. Hoffmann G, Lichtinghagen R, Wosniok W. Simple estimation of reference intervals from routine laboratory data. J Lab Med. 2016; 39. doi: https://doi.org/10.115/labmed-2015-0104
19. Tahmasebi H, Higgins V, Fung AWS, et al. Pediatric Reference Intervals for Biochemical Markers: Gaps and Challenges, Recent National Initiatives and Future Perspectives. EJIFCC. 2017; 28: 43–63.
20. Shaw JLV, Konforte D, Binesh-Marvasti T, et al. Validity of calculating pediatric reference intervals using hospital patient data: A comparison of the modified Hoffman approach to CALIPER reference intervals obtained in healthy children. Clin Biochem. 2012; 45: 1117–8.
21. Adeli K, Higgins V, Trajcevski K, White-Al Habeeb N. The Canadian laboratory initiative on pediatric reference intervals: A CALIPER white paper. Crit Rev Clin Lab Sci. 2017; 54: 358–413.
22. Kratzsch J, Fiedler GM, Leichtle A, et al. New reference intervals for thyrotropin and thyroid hormones based on National Academy of Clinical Biochemistry criteria and regular ultrasonography of the thyroid. Clin Chem. 2005; 51: 1480–6.
23. Lo Sasso B, Vidali M, Scazzone C, et al. Reference interval by the indirect approach of serum thyrotropin (TSH) in a Mediterranean adult population and the association with age and gender. Clin Chem Lab Med. 2019; 57: 1587–94.
24. Katayev A, Balciza C, Seccombe DW. Establishing Reference Intervals for Clinical Laboratory Test Results: Is There a Better Way? Am J Clin Pathol. 2010; 133: 180–6.
25. de Carvalho GA, Perez CLS, Ward LS. Utilização dos testes de função tireoidiana na prática clínica. Ar Bras Endocrinol Metab. 2013; 57: 193–204.
26. Kapelari K, Kirchlechner C, Högler W, et al. Pediatric reference intervals for thyroid hormone levels from birth to adulthood: a retrospective study. BMC Endocr Disord. 2008; 8: 15. doi: 10.1186/1472-6823-8-15
27. Mirjanic-Azaric B, Avram S, Stojakovic-Jelisavac T, et al. Direct Estimation of Reference Intervals for Thyroid Parameters in the Republic of Srpska. J Med Biochem. 2017; 36: 137–44.

28. Jakubowski Z, Zagórska M, Robakowska M, et al. Method of establishing reference intervals of thyroid profile parameters. Post N Med. 2018; 4: 206–10.
29. Thienpont LM, Faix JD, Beastall G. Standardization of Free Thyroxine and Harmonization of Thyrotropin Measurements: A Request for Input from Endocrinologists and Other Physicians. Thyroid. 2015; 25: 1379–80.
30. Drees JC, Huang K, Petrie MS, et al. Reference Intervals Generated by Electronic Medical Record Data Mining with Clinical Exclusions: Age-Specific Intervals for Thyroid-Stimulating Hormone from 33038 Euthyroid Patients. J Appl Lab Med. 2018; 3: 231–9.

Tables

Due to technical limitations, tables 1 to 4 is only available as a download in the Supplemental Files section.

Figures

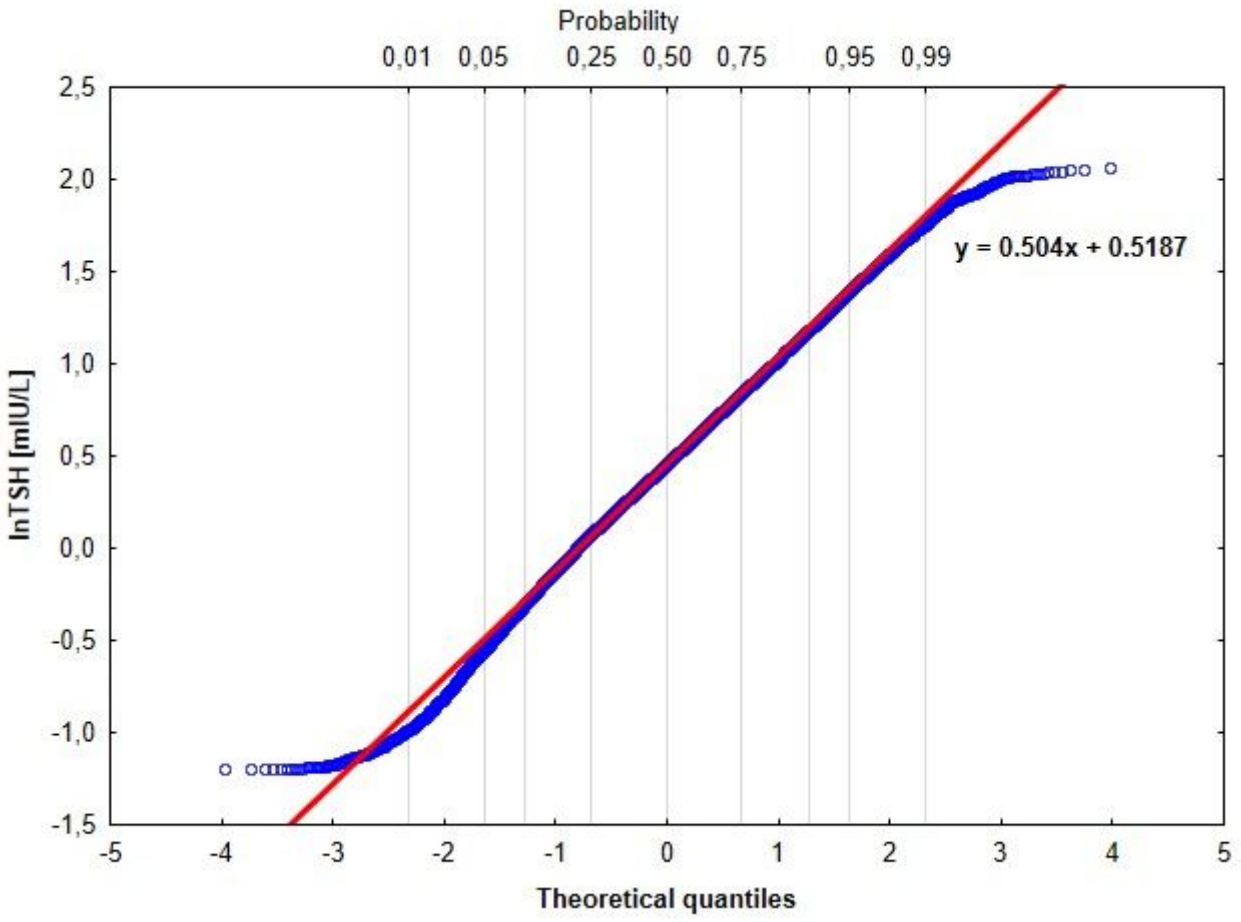

Figure 1

A representative Q-Q plot of distribution of lnTSH for participants aged ≥18 and <40.

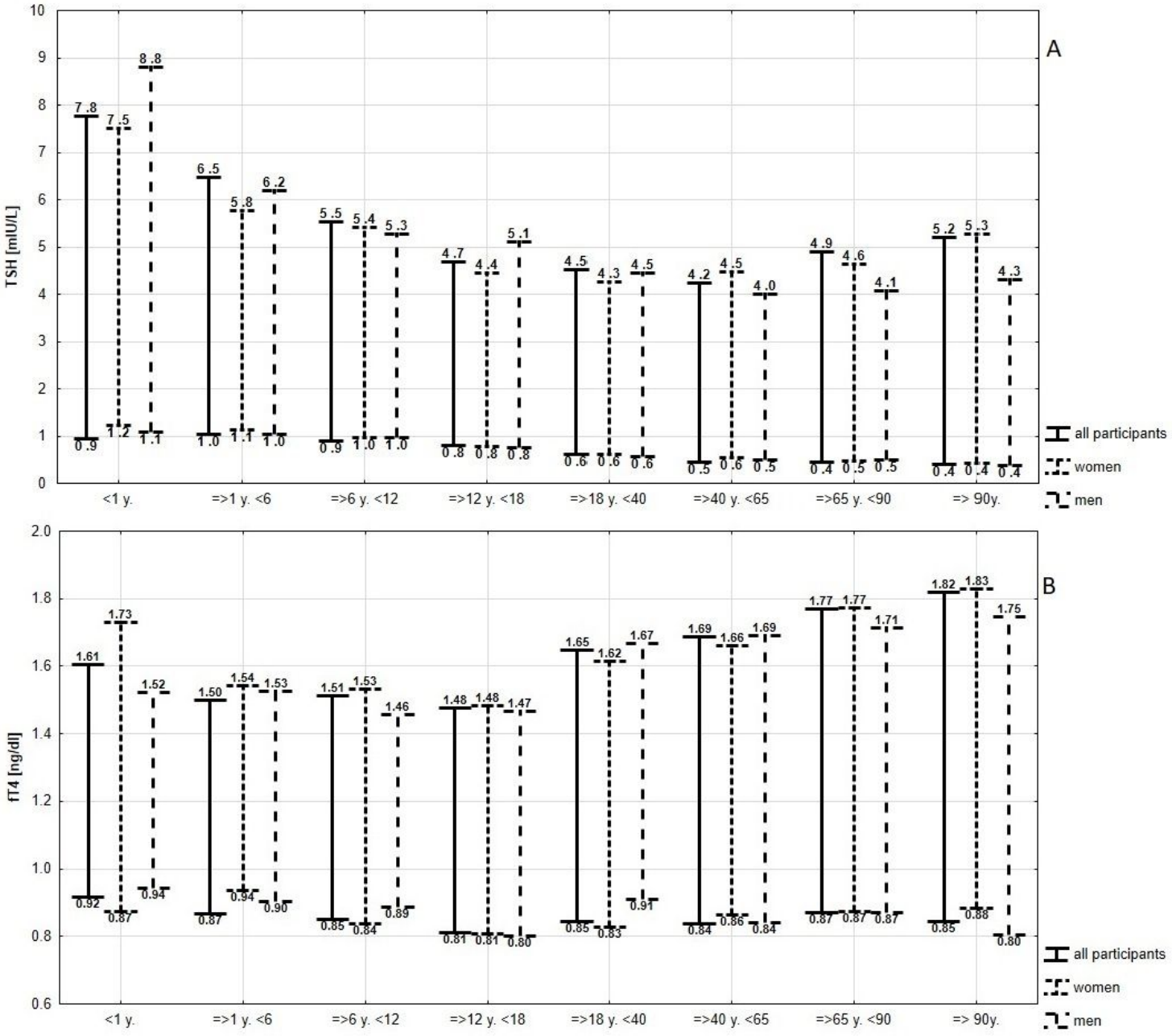

Figure 2

RIs for TSH (panel A) and fT4 (panel B) in all participants included in the analysis groups and among males and females in each age group.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Table1.xlsx
- Table2.xlsx
- Table3.xlsx

- Table4.xlsx