Research article

Giray Bozkaya and Ali Rıza Şişman*

The method comparison and the verification of precision of Mindray CL-6000i thyroid function tests (TFTs)

<https://doi.org/10.1515/tjb-2019-0506>

Received December 4, 2019; accepted March 11, 2020;

published online July 20, 2020

Abstract

Objectives: Thyroid diseases are the most frequent endocrine disorders and thyroid function tests (TFTs) are the most commonly requested endocrine tests. The reliable measurements of these tests are quite important. The aim of our study was to determine the bias and to verify the precision of the newly introduced Mindray CL-6000i immunoassay system in the guidance of CLSI guidelines.

Methods: A precision and bias study was performed in Mindray CL-6000i analyzer for FT3, FT4, TSH, Anti-TG, and Anti-TPO tests by using BioRad quality control (QC) materials and serum samples, respectively. Bland–Altman difference plot and Passing-Bablok regression analysis was made for method comparison with Beckman Coulter DXI 800 analyzer.

Results: The repeatability coefficient of variations (CVs) of FT3, FT4, TSH, Anti-TG, and Anti-TPO tests were ≤ 2.36 , ≤ 1.66 , ≤ 2.38 , ≤ 3.48 , and $\leq 3.31\%$ while within laboratory CVs were ≤ 2.85 , ≤ 4.61 , ≤ 2.59 , ≤ 3.78 , and $\leq 3.60\%$, respectively. The mean differences between the two methods obtained from Bland–Altman analysis for FT3, FT4, TSH, Anti-TG, and Anti-TPO were defined to be -19% , 1.95% , -5.9% , -3.5% , and 7.3% , respectively.

Conclusions: Mindray CL-6000i had good precision in all tests, but the difference between the two methods in some tests shows that the harmonization and standardization of TFTs initiated globally is required.

Keywords: analytical performance; chemiluminescent immunoassay; method comparison; thyroid antibody; thyroid hormones; thyroid function tests.

*Corresponding author: Ali Rıza Şişman, Medical Biochemistry, Dokuz Eylül University Faculty of Medicine, Inciraltı, İzmir, Turkey, Phone: +90 232 4124415, E-mail: e-mail: aliriza.sisman@deu.edu.tr, <https://orcid.org/0000-0002-9266-0844>

Giray Bozkaya: Medical Biochemistry, İzmir Bozyaka Training and Research Hospital, İzmir, Turkey

Introduction

Thyroid gland dysfunctions can be regarded as the most frequent endocrine disorders and laboratory tests are essential for screening, diagnosing and follow-up of these conditions. Besides, thyroid auto-antibodies such as anti-thyroglobulin (Anti-Tg) and anti-thyroperoxidase (Anti-thyroperoxidase) are frequently used for diagnosing Hashimoto's thyroiditis and Graves' disease, in addition to thyroid function tests (TFTs) [1]. TFTs are the most commonly requested endocrine tests in clinical practice [1]. Serum thyroid-stimulating hormone (TSH) is used as a first-line test for detecting thyroid dysfunction and approximately 59 million TSH tests were performed while approximately 18 million free T4 tests (FT4) were ordered secondly after TSH in 2008 in the United States alone [2]. Free T3 (FT3) measurement is mostly unnecessary, compared to FT4 and TSH because they are sufficient for hypo- and hyper-thyroidism diagnosis [3–5].

When the prevalence of thyroid disorders and symptoms are taken into account, especially TSH measurement methods are important for reliable laboratory results [6, 7]. The measurement of TSH, as well as the other TFTs, is routinely performed using immunoassay methods in automated platforms in clinical laboratories [8, 9]. In the last 3 decades, the analytical performances of immunoassays have been developed gradually but remarkable distinctions in the results are met between different immunoassay methods, including TFT [9–11]. In order to attain comparability of measurement results between different methods, TFT methods should be standardized. One of the TFT tests, FT3 can be standardized since it has a reference method. However, there is no reference method option for TSH; therefore rather than standardization, harmonization is a more realistic approach for TSH [12, 13]. The International Federation of Clinical Chemistry and Laboratory Medicine (IFCC) initiated a phase IV study aiming to cover the benefits, risks and practical implementation of standardization and harmonization [10] in which some manufacturers re-calibrated their methods to participate in harmonization [14]. In harmonization,

reference material needs traceability but its' non-commutability causes the same patient sample to yield different TSH results on different platforms [15]. Thus comparing the results of the samples measured with different methods becomes problematic. Manufacturers try to solve this problem using different reference ranges for their own methods, but the problem cannot be solved completely.

Despite these problems, the analytical performances of the immunoassay methods are being improved by the manufacturers and new products are introduced to the market every day. The analytical performance of these products should be tested by clinical laboratories and checked if they meet the quality requirements. This study aimed to determine the bias and to verify the precision of the newly introduced Mindray CL-6000i TFTs in the light of the most current guidelines of the Clinical & Laboratory Standards Institute (CLSI) (EP15-A3 and EP09-A3).

Materials and methods

Assays

FT3, FT4, TSH (third generation), Anti-TG, and Anti-TPO measurements were made in two chemiluminescent systems as Beckman Coulter-DXI 800 (Beckman Coulter Diagnostics Division Headquarters, Brea, CA, USA) and Mindray CL-6000i (Mindray Bio-Medical Electronics Co., Ltd., Shenzhen, China). Mindray CL-6000i has a throughput of up to 480 tests per hour using a measurement principle of enhanced alkaline phosphatase with adamantyl 1, 2-dioxetane aryl phosphate (ALP-AMPPD) method. It can handle up to 300 samples in one batch with refrigerated 36 reagent positions. Beckman Coulter-DXI 800 analyzer was being used for routine immunoassay tests in our laboratory for many years while Mindray CL-6000i was used for the first time. Mindray CL-6000i analyzer uses CL-series FT3, FT4, TSH (third generation), Anti-TG, and Anti-TPO assays which are chemiluminescent immunoassays. The reference limits were; 1.8–4.2 pg/mL, 0.5–1.4 ng/dL, 0.35–5.1 µIU/mL, 0–4 IU/mL, and 0–9 IU/mL for FT3, FT4, TSH, Anti-TG, and Anti-TPO assays, respectively. Beckman Coulter Access FT3, FT4, TSH, Anti-TG, and Anti-TPO kits are used in Beckman Coulter-DXI 800 analyzer which is also chemiluminescent immunoassays for the quantitative determination of respective tests.

Verification of precision

After a one-month familiarization period for Mindray CL-6000i analyzer, a precision and bias study was performed. Analyzers were calibrated at the start of the study and both of them were in conformance with their quality control (QC) parameters. The third-party QC materials (BioRad, USA) were used for evaluating precisions by using CLSI EP15-A3 protocol [16]; Lypochek Immunoassay Plus Control (Lot no 40350) for FT3, FT4, and TSH; Liquichek Specialty Immunoassay Control (Lot no 60250) for Anti-Tg and Anti-TPO. Three levels of QC materials for FT3, FT4, TSH, and two levels of QC materials for Anti-TG and Anti-TPO were measured for five consecutive days as five

replicates. There were no outliers as assessed by two-sided Grubbs test. One-way analysis of variance (ANOVA) was used to calculate the repeatability (intra-day) and between-run variation. Within-laboratory (WL) variability (total imprecision) was then calculated from these two parameters. The estimated CVs were compared with the manufacturer's declared values. The statistical analyses for precision were made by using MedCalc Statistical Software version 19.1 (Oostende, Belgium).

As there are no QC target values for Mindray CL-6000i kits in BioRad QC samples, we could not estimate bias by CLSI EP15-A3 protocol. Instead, we used CLSI EP9-A3 [17] protocol for the determination of bias by using patient samples.

Method comparison and bias estimation using patient samples

The method comparison study was performed in 10 days by analyzing serum samples that were collected randomly from leftover serum samples after routine analysis in Beckman Coulter DXI 800 for whom FT3 (n=104), FT4 (n=342), TSH (n=308), Anti-TG (n=142) or Anti-TPO (n=228) test was requested. The samples were analyzed according to the CLSI EP9-A3 protocol. None of the samples were interfered with hemolysis, lipemia or icterus. Since the serum samples were collected after routine measurements were finished, the conducted research did not need ethical approval. After the routine measurement in Beckman Coulter DXI 800 system, the serum samples were collected and analyzed in Mindray CL-6000i in 2 h. In this method comparison study, the Beckman Coulter DXI 800 Access methods were accepted as comparative, and the Mindray CL-6000i methods were accepted as candidate methods. Duplicate measurements were made in both analyzer systems and the means of duplicate measurements were compared. Data were collected every day and the possible transcription errors in typing the data were prevented by taking all results online. Bland-Altman difference plot and Passing-Bablok regression analysis were made by MedCalc Statistical Software version 19.1. Bland-Altman difference plot was used to assess the distribution of the difference between the two methods throughout the measurement range. The mathematical relationship between the two methods was determined by using the Passing-Bablok regression analysis. The mean systematic difference (bias) between the two methods was defined at different medical decision levels (upper and lower reference limits, LRL and URL, of FT3, FT4, TSH, Anti-TG and Anti-TPO). In addition to reference limits, biases for TSH were calculated at 0.1 µIU/mL for evaluation of hyperthyroidism and 4.12 µIU/mL for evaluation of hypothyroidism according to data obtained from The National Health and Nutrition Examination Survey (NHANES III) [18].

Results

Table 1 presents the data of precision study performed according to the CLSI EP15-A3 protocol for Mindray CL-6000i system and manufacturer's claims for FT3, FT4, TSH, Anti-TG, and Anti-TPO tests. The repeatability and WL imprecision results were obtained by using the BioRad QC materials. The repeatability CVs of FT3, FT4, TSH, Anti-TG, and Anti-TPO tests were ≤2.36, ≤1.66, ≤2.38, ≤3.48,

and $\leq 3.31\%$, while WL CVs were ≤ 2.85 , ≤ 4.61 , ≤ 2.59 , ≤ 3.78 , and $\leq 3.60\%$, respectively. Since the results were lower than the manufacturer's claims, there was no need to calculate the Upper Verification Limit (UVL), as indicated in the CLSI EP15-A3 protocol.

The mean differences between the two methods obtained from Bland-Altman analysis for FT3, FT4, TSH, Anti-TG, and Anti-TPO were defined to be -19% , 1.9% , -5.9% , -3.5% , and 7.3% , respectively (Figure 1A–E). The mean differences determined by Passing-Bablok regression between the two methods at LRL and URL were presented in Table 2.

Table 2 presents the data of the method comparison study between Mindray CL-6000i and Beckman Coulter DXI-800 according to CLSI EP9-A3. It consists of Passing-Bablok regression results, 95% confidence intervals (CI) and bias% at medical decision limits according to reference lower and upper limits of FT3, FT4, TSH tests, and upper reference limits of Anti-TG and Anti-TPO tests. The additional biases for TSH were calculated at the concentrations of $0.1 \mu\text{IU/mL}$ and $4.12 \mu\text{IU/mL}$. Higher correlations were found between the two methods in all TFTs except FT3, $r=0.797$ (Figure 2A–E).

Discussion

Since TFTs are very important on the diagnosis and management of thyroid diseases, their analytical performance should meet the quality requirements. Unfortunately, there are problems in interpreting the results due to differences in the analytical performance of TFTs and interferences.

The IFCC Committee for Standardization of TFTs has been working for the standardization of FT4 and TSH testing since there is a need for comparable measurement results [10]. At present, new automated platforms come out to the market which makes the immunoassay methods more complicated because of the variability of the methods and different analytical performances. The introduction of new methods and systems brings the laboratory specialist a task to test the analytical performance of the platforms.

Precision and bias are two basic parameters of analytical performance that should be tested for every new analyzer. Mindray CL-6000i, which is a new immunoassay platform that was tested for this purpose in our laboratory in the light of CLSI EP15-A3 and CLSI EP9-A3 protocols. Although the CLSI EP15-A3 guide was prepared to estimate precision and bias at the same experiment, the lack of target values for Mindray CL-6000i TFT in BioRad QC materials, forced us to use CLSI EP9-A3 guide for bias calculation. This fact may be thought of as a deficiency of CLSI EP15-A3 usage in new platforms.

In this study, the repeatability and WL-CVs of the Mindray CL-6000i system for TSH were found as $\leq 2.38\%$ and $\leq 2.59\%$, respectively (Table 1). These CVs were lower than the manufacturer's claims and the desirable limit of biological variation studies in EFLM and Westgard websites (7.95% , 9.7% , respectively) [19, 20]. Besides, the Mindray CL-6000i system seems to have a good performance when compared with the results of other studies using different methods [9, 21, 22]. When evaluating a TSH result of a patient, the clinician first checks whether the result is within the reference range; however, cut off or medical decision limits below and above the reference

Table 1: The results of the Precision study of Mindray CL-6000i measuring's using BioRad QC material according to CLSI EP 15-A3.

Measurands	Level	Our study			Manufacturer's claims		
		Mean	Repeatability (CV%)	WL imprecision (CV%)	Mean	Repeatability (CV%)	WL imprecision (CV%)
FT3 (pg/mL)	1	2.27	2.36	2.57	4.60	4.20	8.54
	2	6.04	1.85	2.85	14.78	3.54	7.47
	3	10.23	1.86	2.43	*	*	*
FT4 (ng/dL)	1	0.98	1.66	4.61	1.28	4.47	9.07
	2	2.22	1.45	3.79	2.61	3.56	7.29
	3	3.99	0.91	3.93	*	*	*
TSH ($\mu\text{IU/mL}$)	1	0.28	2.38	2.59	0.58	3.53	8.14
	2	3.64	1.89	2.10	62.54	3.32	7.61
	3	24.34	1.36	2.14	*	*	*
Anti-TG Ab (IU/mL)	1	40.98	3.48	3.78	8.29	4.65	8.99
	2	80.01	3.08	3.09	261.32	4.70	8.94
Anti-TPO Ab (IU/mL)	1	30.13	3.30	3.35	10.14	4.72	8.77
	2	83.89	3.31	3.60	100.48	3.93	6.82

*Not determined.

Figure 1: The Bland-Altman plot for A) FT3, B) FT4, C) TSH, D) Anti Tg, and E) Anti TPO assays.

range are also taken into account. For example, in the diagnosis of subclinical hyperthyroidism, a TSH level of 0.1 μ IU/mL is used as a cut off level. Therefore, in the

evaluation of hyperthyroidism, the measurement method is expected to have good precision in the gray zone of 0.1–0.4 μ IU/mL (LRL for many methods). In our study, WL

Table 2: The Passing Bablok regression results and bias % at medical decision limits.

Measurands	Medical decision limits (reference limits)	Regression results		Bias at medical decision limits (%)	
		Slope (95%CI)	Intercept (95%CI)		
FT3 (pg/mL)	(2.5–3.9)	1.14 (1.03 to 1.26)	–1.12 (–1.53 to –0.71)	–30.57	–14.52
FT4 (ng/dL)	(0.61–1.12)	1.03 (1.00 to 1.07)	–0.02 (–0.05 to 0.01)	–0.23	1.37
TSH (μIU/mL)	(0.38–5.33)	0.98 (0.96 to 1.01)	–0.05 (–0.10 to –0.01)	–15.37	–2.69
Anti-TG (IU/mL)	(0–4)	1.01 (0.93 to 1.06)	–0.33 (–0.77 to 0.05)	*	–7.62
Anti-TPO (IU/mL)	(0–9)	1.19 (1.13 to 1.23)	–0.60 (–1.74 to –0.15)	*	12.30

*Not determined.

CV was found to be 2.59% at a mean level of 0.28 μIU/mL, which is a gray zone concentration. Clerico et al. reported that the between-methods variability is high especially at TSH concentrations <0.4 μIU/mL, because of lower analytical sensitivity of some immunoassay methods [23]. Besides, at TSH concentrations used in the evaluation of subclinical hypothyroidism, the Mindray CL-6000i method's WL CV was found to be 2.14%. When these results are taken into consideration, it is seen that the Mindray CL-6000i TSH method has the desirable precisions for evaluating both subclinical hyperthyroidism and subclinical hypothyroidism.

However, in patient-sample-based comparison of Mindray CL-6000i and Beckman Dxl 800 TSH methods, the bias at LRL and URL were –15.37%, and –2.69%, respectively (Table 2). Mindray CL-6000i TSH results were lower compared to Beckman Dxl 800 TSH results. When the Bland-Altman graph is examined, it is seen that the inconsistency between the two methods is partially higher, especially at low TSH levels (Figure 1C). Although the bias of –15.37% estimated from the regression analysis at LRL is lower than the Total Error allowable (TEa) of 22.0 and 23.7% obtained from biological variation studies [19, 20], it is higher than the desirable bias of 7.8% [19, 20]. Considering the biases obtained from the method comparison study and discrepancy of the results between two methods at low TSH levels in the Bland-Altman plot (Figure 1C), it can be concluded that it could cause misinterpretation of a patient with hyperthyroidism. These findings suggest that manufacturers should improve the analytical performance of the TSH methods especially at low levels while studies should continue for global harmonization. In fact, the analytical performance of TSH immunoassay tests has been gradually developed in recent years and TSH kits beginning from the first generation to the fourth generation were produced [24]. Thus detection limits of TSH progressively went down to 0.001–0.002 μIU/mL levels [25]. Standardization and/or harmonization of the assays is necessary for the comparison of the results between

laboratories but neither pure substance reference material and/or a reference measurement procedure currently exist for TSH. As standardization is very hard to carry out, harmonization of TSH is more realistic and it is thought this problem may be solved by efforts such as World Health Organization's third TSH International Standard (IS, International Reference Preparation 81/565) [12, 15]. With the help of these studies, harmonization issues between methods may be solved.

The repeatability and WL CVs of Mindray CL-6000i FT3 were found to be ≤2.36% and ≤2.85%, respectively (Table 1). These precision values were lower than both manufacturer's claims in the kit inserts and desirable precisions obtained from biological variation studies. In the comparison of Mindray CL-6000i and the Beckman Dxl 800 FT3 methods by serum samples, the biases at LRL of 2.5 pg/mL and URL of 3.9 pg/mL were found to be –30.37% and –14.59%, respectively (Figure 2A). Mindray's FT3 results were lower compared to Beckman Coulter's FT3 results. When the mean difference of 19% in Bland-Altman analysis (Figure 1A) and the difference of 30.57 % at LRL in the method comparison study (Table 2, Figure 2A) were taken into account together, we can say that there are important differences over TEa of 9.2 and 11% [19, 20], especially at LRL levels. However, the repeatability and WL precisions were 2.36 and 2.57%, respectively. When precision and bias for FT3 are evaluated together, it can be concluded that the source of difference between the two methods comes from bias rather than imprecision. These findings support the idea there is a need for global standardization and harmonization also for FT3, likewise TSH and FT4.

We found the repeatability and WL CVs of the Mindray CL-6000i system for FT4 to be ≤1.66% and ≤4.61%, respectively. These values are lower than the manufacturer's repeatability and WL imprecision claims (Table 1). The repeatability CVs were also lower than the desired imprecision obtained from biological variation studies of 2.9% but WL CVs were slightly higher. In the method comparison

Figure 2: The patient sample-based method comparison of Beckman Coulter Dxl 800 and Mindray CL-6000i A) FT3, B) FT4, C) TSH, D) Anti Tg, and E) Anti TPO assays with Passing-Bablok regression.

study of Mindray CL-6000i and Beckman Dxl 800 FT4 methods, estimated bias at LRL of 0.6 ng/dL and URL of 1.1 ng/dL were found to be -0.23 and 1.37% , respectively (Table 2, Figure 1B, Figure 2B). The difference between the two methods was lower than TEa of The Royal College of Pathologists of Australasia (RCPA) and TEa of the Biological Variation Study, 15 and 8%, respectively [20, 25]; which shows that there is a good agreement between the two methods in terms of FT4 results. However many methods are not as good as the Mindray CL-6000i system because, in a survey of 13 FT4 methods, it was reported that more than 50% the results did not meet the allowable inaccuracy criteria [22]. Also in laboratories where direct analog immunoassays are used for FT4 measurements, the results of FT4 are controversial and not standardized [2, 26].

Anti-TPO and Anti-TG antibodies are not only the markers of Hashimoto's thyroiditis and Graves' disease, according to the NHANES III study [18], they may be also found in 14.6 and 8.0% of euthyroid females and males, respectively [18]. Also in the Whickham survey, it was shown that the presence of Anti-TPO antibodies increased the rate of progression to overt hypothyroidism [27]. However, there are no defined acceptable analytical performance criteria for Anti-TPO and Anti-TG. In our study, the repeatability and WL CVs of both Anti-TPO and Anti-TG were lower than the manufacturer's claims in the kit inserts (Table 1). When compared to routine laboratory practice, these precisions can be accepted to be satisfactory. In the method comparison study, we found that Mindray CL-6000i Anti-TPO had 12.3% higher results than those of Beckman Dxl 800, while Mindray CL-6000i Anti-TG had 7.62% lower results (Table 2). In Bland-Altman graphs, there are discrepancies between the two methods at low analyte concentrations in both antibody tests, especially in Anti-TG (Figure 1D). The concentrations of thyroid antibodies higher than URL are clinically important, but these findings show that the manufacturers should improve the sensitivities of their methods.

Differences between analytical systems in immunoassay results, including TFT, are currently an important problem. Manufacturers recommend the use of reference ranges specific to their systems when evaluating the results produced from their systems. However, many patients are displaced or tested in another analytical system. Therefore, difficulties in evaluating the results continue to exist. Also, the development of modern public health standards such as clinical guidelines specifying fixed decision limits and the integration of electronic medical records into the health system is prevented by method-specific measurement results [28].

In conclusion;

- (1) The Mindray CL-6000i FT3, FT4, TSH, Anti-Tg and Anti-TPO methods have a good analytical performance according to the suggestions in good laboratory practice,
- (2) The difference between Mindray CL-6000i and Beckman Dxl 800 in FT3 methods exceeds TEa values,
- (3) While Mindray CL-6000i have good precisions in all tests, significant differences between the two methods in some tests show that the harmonization and standardization of TFTs initiated globally is required.

Acknowledgment: We thank Med-Kim and Mindray for providing the necessary reagents.

Research funding: None declared.

Author contributions: All authors have accepted responsibility for the entire content of this manuscript and approved its submission.

Competing interests: Authors state no conflict of interest.

References

1. Soh SB, Aw TC. Laboratory testing in thyroid conditions-pitfalls and clinical utility. *Ann Lab Med* 2019;39:3–14.
2. Thienpont LM, Van Uytendaele K, Poppe K, Velkeniers B. Determination of free thyroid hormones. *Best Pract Res Clin Endocrinol Metab* 2013;27:689–700.
3. Ladenson PW, Singer PA, Ain KB, Bagchi N, Bigos ST, Levy EG, et al. American Thyroid Association guidelines for detection of thyroid dysfunction. *Arch Intern Med* 2000;160:1573–75.
4. Singh RJ, Kaur P. Thyroid hormone testing in the 21st century. *Clin Biochem* 2016;49:843–45.
5. Sheehan MT. Biochemical testing of the thyroid: TSH is the best and, often times, only test needed-a review for primary care. *Clin Med Res* 2016;14:83–92.
6. Iervasi G, Clerico A. Harmonization of free thyroid hormone test: a mission impossible?. *Clin Chem Lab Med* 2011;49:43–8.
7. Spencer CA. Assay of thyroid hormones and related substances. *Thyroid disease manager*. In: De Groot LJ, Beck-Peccoz P, Chrousos G, Dungan K, Grossman A, Hershman JM, et al., editors. *Endotext* [Internet]. South Dartmouth (MA): MDText.com, Inc.; 2000–2013 Jan 1.
8. Rawlins ML, Roberts WL. Performance characteristics of six third generation assays for thyroid-stimulating hormone. *Clin Chem* 2004;50:2338–44.
9. Dittadi R, Rizzardi S, Masotti S, Prontera C, Ripoli A, Fortunato A, et al. Multicenter evaluation of the new immunoassay method for TSH measurement using the automated Dxl platform. *Clin Chim Acta* 2017;468:105–10.
10. Thienpont LM, Faix JD, Beastall G. Standardization of FT4 and harmonization of TSH measurements – a request for input from endocrinologists and other physicians. *Clin Endocrinol (Oxf)* 2016;84:305–6.

11. Padoan A, Cosma C, Plebani M. Evaluation of the analytical performances of six measurands for thyroid functions of Mindray CL-2000i system. *J Lab Precis Med* 2018;3:1–7.
12. Miller GW, Myers GL, Lou Gantzer M, Kahn SE, Schonbrunner ER, Thienpont LM, et al. Roadmap for harmonization of clinical laboratory measurement procedures. *Clin Chem* 2011;57:1108–17.
13. Thienpont LM, Van Houcke SK. Traceability to a common standard for protein measurements by immunoassay for in-vitro diagnostic purposes. *Clin Chim Acta* 2010;411:2058–61.
14. Thienpont LM, Van Uytendaele K, De Grande LAC, Reyniers 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;63:1248–1260.
15. Haugen BR, Alexander EK, Bible KC, Doherty GM, Mandel SJ, Nikiforov YE, et al. American Thyroid Association Management guidelines for adult patients with thyroid nodules and differentiated thyroid cancer: The American Thyroid Association guidelines task force on thyroid nodules and differentiated thyroid cancer. *Thyroid* 2015;26:1–133.
16. CLSI. User Verification of Precision and Estimation of Bias; Approved Guideline-Third Edition. CLSI Document EP15-A3. Wayne PA: Clinical and Laboratory Standards Institute; 2014.
17. CLSI. Method Comparison and Bias Estimation Using Patient Samples; Approved Guideline-Third Edition. CLSI Document EP09-A3. Wayne PA: Clinical and Laboratory Standards Institute; 2013.
18. Hollowell JG, Staehling NW, Flanders WD, Hannon WH, Gunter EW, Spencer CA, et al. Serum TSH, T4, and thyroid antibodies in the United States population (1988 to 1994): National Health and Nutrition Examination Survey (NHANES III). *J Clin Endocrinol Metab* 2002;87:489–499.
19. EFLM. EFLM biological variations. Available from: https://biologicalvariation.eu/bv_specifications [Accessed February 2020].
20. Westgard QC. Desirable biological variation database specifications. Desirable specifications for total error, imprecision, and bias, derived from intra- and interindividual biologic variation. Available from: <https://www.westgard.com/biodatabase1.htm> [Accessed December 2019].
21. Winston-McPherson GN, Samraj AN, Poster K, Yamaguchi D, Dickerson JA, Drees JC, et al. Performance characteristics of the Beckman Coulter UniCel Dxl 800 TSH (3rd IS) assay. *Clinica Chimica Acta* 2018;478:90–100.
22. Steel BW, Wang E, Klee GC, Thienpont LM, Soldin SJ, Sokoll LJ, et al. Analytic bias of thyroid function tests. *Arch Pathol Lab Med* 2005;129:310–7.
23. Clerico A, Ripoli A, Zucchelli GC, Plebani M. Harmonization protocols for thyroid stimulating hormone (TSH) thyroid immunoassays: different approaches based on the consensus mean value. *Clin Chem Lab Med* 2015;53:377–82.
24. Çalıcı E, Doğan HO, Sağlam F, Turhan T, Berker D. Comparison of the performance of second (fast TSH) and third (HYPERsensitive TSH) generation automated TSH immunoassays in healthy euthyroid subjects. *Erciyes Med J* 2019;41:46–9.
25. Clerico A, Trenti T, Aloe R, Dittadi R, Rizzardi S, Migliardi M, et al. Italian Section of the European Ligand Assay Society (ELAS). A multicenter study for the evaluation of the reference interval for TSH in Italy (ELAS TSH Italian Study). *Clin Chem Lab Med* 2018;57:259–67.
26. Westgard QC. RCPA Allowable Limits of Performance for Biochemistry. Royal College of Pathologists of Australasia Analytical Quality Requirements. Available from: <https://www.westgard.com/rcpa-biochemistry.htm> [Accessed December 2019].
27. Thienpont LM, Van Uytendaele K, Beustall G, Faix JD, Ileri 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. Vanderpump MP, Tunbridge WM, French JM, Appleton D, Bates D, Clark F, et al. The incidence of thyroid disorders in the community: a twentyyear follow-up of the Whickham Survey. *Clin Endocrinol (Oxf)* 1995;43:55–68.