Comparison of Thyroid -Stimulating Hormone and Free Thyroxine Immunoassays Performed on Immulite 2000 and Maglumi 800 Automated Analyzers

Sonja Kuzmanovska*, Daniela Miladinova

Faculty of Medicine, Institute of Pathophysiology and Nuclear Medicine, Ss. Cyril and Methodius University of Skopje, Skopje, Republic of Macedonia

Abstract

OBJECTIVE: Our goal was to evaluate the analytical performance of the novel immunoassay platform and to compare the agreement between thyroid-stimulating hormone (TSH) and FT4 results, obtained by novel and currently used platform.

MATERIALS AND METHODS: Both immunoassay platforms, current Immulite 2000 and novel Maglumi 800, are based on chemiluminecsence immunoassay method. Analytical performance was evaluated by the use of serum pools and commercial quality control samples. The comparison study was carried out with 80 serum samples. Obtained results were analyzed by descriptive statistics, Mann–Whitney U-test, and Paired t-test. Method comparison was performed with Passing-Bablok regression analysis and Bland-Altman plots.

RESULTS: TSH Maglumi 800 showed better within-run precision for both concentration ranges (1.7–2.8 CV%) in comparison to Immulite 2000 (4.4–5.7 CV%). FT4 Maglumi 800 imprecision was higher compared with Immulite 2000 FT4 in both within-run (3.5–3.9 CV% vs. 4.9–6.6 CV%) and between-run (3.6–4.2 CV% vs. 4.6–5.9 CV%) tests. Mann–Whitney U-test for TSH revealed non-significant difference between data (p = 0.9011). Regression analysis showed no systematic (intercept = 0.01), nor proportional (slope = 0.9781) differences. Non-significant bias was observed in Bland-Altman Plots. For FT4, we found significant differences between methods using paired t-test (t_{39} = 10.5, p < 0.0001) and significant difference (p = 0.00745) with Mann–Whitney U-test. Bland-Altman plot revealed 22.8% average bias.

CONCLUSION: TSH evaluation showed good precision and close agreement between Maglumi 800 and Immulite 2000 methods, which assures transferability of results. However, FT4 performance evaluation revealed higher imprecision of Maglumi 800 platform and significant differences of test results.

Introduction

Thyroid function tests (TFTs) are important tool in diagnosis and monitoring of thyroid disorders. The assay menu includes thyroid hormones – thyroxine (T4) and triiodothyronine (T3), thyroglobulin, thyroid antibodies, and thyroid-stimulating hormone (TSH) from the pituitary. Serum TSH is the most important single measurement that can detect both overt and subclinical thyroid disorders [1]. The secretion is regulated by negative feedback mechanism from the free fraction of serum T4 (FT4) and free T3 (FT3). Even minor, non-pathological changes in FT4 and FT3 serum concentrations could influence TSH secretion, which confirms it as a sensitive biomarker for subclinical thyroid disorders [2]. Furthermore, TSH is useful in managing hormone replacement therapy in hypothyroidism and suppressive therapy in follicular cell thyroid cancer [3]. In regard with thyroid hormones assessment, FT4 has nowadays replaced total T4 measurement, which is highly influenced by the amount of serum carrier proteins. Therefore, the majority of clinical chemistry laboratories perform TSH and FT4 as first-line TFTs [4], [5], [6] as demanded by clinicians.

Immunoassays have been dominant analytical techniques for these tests for several decades, starting with radioimmunoassay and improving the sensitivity and specificity by introducing new labels and techniques over time. Despite of automation of assay platforms, the test results exhibit substantial bias performed with reagents from different manufacturers and on different analyzers [7], [8], [9]. They implement various types of labels and antibodies, different standard reference material, as well as diversity of fraction separation techniques. Thus, a change of immunoassay platform and reagent manufacturer requires performance of a certain evaluation process to assess comparability between methods. The basic evaluation procedures are imprecision (within-run and between-run) assessment and method agreement testing.

The purpose of our study was to compare the agreement of the results for TSH and FT4 obtained with

different immunoassays, performed on novel Maglumi 800 and the current Immulite 2000 automated analyzers. Close agreement would ensure the transferability of test results between platforms and their reporting within the previously established reference intervals for both analytes. Evaluation of analytical performance of TSH and FT4 immunoassays measured on Maglumi 800 automated platform would reveal the extent of imprecision and reliability of test results.

Materials and Methods

Serum samples

The samples used for comparison study were 80 sera of anonymous patients admitted for thyroid function examination at the Institute of Pathophysiology and Nuclear Medicine, Faculty of Medicine in Skopje. The samples were collected in evacuated tubes with cloth activator (CDRICH, China) and after serum separation, hemolytic, icteric, and lipemic samples were excluded from the study. After the initial assessment of TSH and FT4 from primary tubes with the routine immunoassay, tubes with results reflecting euthyroid, and hypothyroid and hyperthyroid status were selected for further analysis on the novel immunoassay platform. Collection of the samples was performed within 2 weeks, and finally two groups of 40 samples [10] were obtained for the comparison study: One group for TSH and the other for FT4.

Quality control material

For within-run imprecision assessment of the novel method, we prepared two concentration levels of in-house pooled, quality control serum samples for each measurand. Ten replicate measurements were performed in one run for each concentration level. Between-run imprecision for median concentration level was assessed using commercial quality control sample, provided by the manufacturer (Snibe, China). For pathological, high concentration level, we used commercial Lyphochek Immunoassay Plus Control (Bio-Rad, USA). The data were obtained for 10 consecutive days. Both of the imprecision data are expressed as coefficient of variation (CV %). The with-in and between-run imprecision assessment for the routine method was performed using two concentration levels of Lyphochek Immunoassay Plus Control.

Methods

Both of the automated immunoassay analyzers – Immulite 2000 (Siemens Healthineers, USA) and Maglumi 800 (Snibe Co. Ltd., China) are based on CLIA methods. However, they differ in labeling and separation technology. Brief description of analyzers methodology and immunoassay tests, principles are presented below.

Immulite 2000 platform utilizes enzyme (alkaline phosphatase) labeled reagent, which reacts with the analyte from the serum sample in the test unit tube. The immune complex is captured on antibody coated bead placed in the test unit tube. After washing and centrifugation step, luminogenic substrate (adamantyl dioxetane phosphate) is added to the labeled immune complex, producing unstable intermediate dioxetane, which stabilizes by light emission, measured by the instrument in relative light units (RLU).

Immulite 2000 TSH is the third generation solid phase, two-site chemiluminescent immunometric assay, performed in a single incubation cycle. TSH from the sera is sandwiched between the capture monoclonal murine anti-TSH antibody, coated on the bead surface and polyclonal goat anti-TSH antibody, labeled with alkaline phosphatase. The signal, generated after substrate addition to the immune complex, is proportional to the serum TSH concentration. The assay calibrators are traceable to the WHO second International Reference Preparation 80/558.

Immulite 2000, FT4 is solid phase, competitive immunoassay. Patient serum is incubated in the test tube with a bead coated with monoclonal murine anti-T4 antibody and alkaline phosphatase conjugated to T4. Unbound fractions are removed from the tube with centrifuge washes and luminogenic substrate is added to immune complex to generate a light signal, inversely proportional to the concentration of free T4.

Maglumi 800 platform applies non-enzyme ABEI labeled reagents. ABEI (N-(4-aminobutyl)-N-ethylisoluminol) is a small molecule which produces a flash of light in the presence of signaling reagents (NaOH and H_2O_2). The separation of labeled immune complex is performed by coated magnetic microbeads in magnetic field and unbound fraction is aspirated from the reaction micro wells.

Maglumi TSH is solid phase, two-site immunometric assay, performed using two different monoclonal anti-TSH antibodies. Serum TSH, ABEI labeled mouse anti-TSH monoclonal antibody, and magnetic microbeads coated with monoclonal anti-TSH antibody, are incubated in a single cycle. Subsequently, the sandwich immune complex is precipitated in magnetic field, afterward the signaling reagents is added to produce light, proportional to the TSH concentration. The assay calibrators are traceable to the WHO third International Reference Preparation 81/565.

Maglumi 800 FT4 is solid phase competitive immunoassay. The reaction occurs in micro wells after addition of serum, ABEI labeled anti-T4 monoclonal antibody and T4 antigen coated magnetic microbeads. After formation of immune-complex with the labeled

antibody and precipitation in magnetic field, the unbound liquid fractions are aspirated. The signaling reagents are introduced into the well, thus initiating a light that is detected as RLU, inversely proportional to the FT4 serum concentration.

Statistical analysis

TSH and FT4 within-run and between-run imprecision were analyzed by descriptive statistics (mean, standard deviation and CV %) using Microsoft Excel 2007. The significance of differences between paired groups was assessed with Mann–Whitney U-test and Paired t-test. Method comparison was performed using Passing-Bablok regression analysis [11] and Bland-Altman plots [12]. Passing-Bablok defined regression equitation, including 95% confidence intervals (CIs) and proportional and constant errors were calculated. For the TSH analysis, the results were log–transformed. For method comparison analysis, statistical package MedCalc v.14 was used.

Results

The data from the imprecision tests for each measurand are displayed on the tables. Different control materials were used for both tests and measurands, as mentioned previously. Nevertheless, the concentration ranges were selected to be comparable between analyzers. The results from the within-run test are expressed as mean values out of 10 replicates, and the imprecision as coefficient of variation.

Table 1: Within-run imprecision evaluation for TSH

Analyzer	QC$_1$		QC$_2$	
	Mean±SD	within-run	Mean±SD	within-run
	mIE/L	CV%	mIE/L	CV%
Maglumi 800	0.48^1±0.008	1.7	3.39^2±0.09	2.8
Immulite 2000	0.57^3±0.025	4.4	5.08^4±0.29	5.7

QC: Quality control sample; [1]Low range pooled control serum; [2]Upper range pooled control serum; [3]Low range Lyphochek control (Bio-Rad); [4]Upper range Lyphochek control (Bio-Rad). SD: Standard deviation.

From the data presented on Table 1, it is evident that TSH performed on Maglumi 800 showed better within-run precision in both, normal and high concentration ranges (1.7–2.8 CV%) in comparison to Immulite 2000 (4.4–5.7 CV%). Between-run imprecision, tested during 10 consecutive days, showed better results for the high pathological control performed on Maglumi 800, but for the borderline upper range level, Immulite 2000 TSH was more precise (5.0 CV% vs. 5.9 CV%) (Table 2).

Table 2: Between-run imprecision evaluation for TSH

Analyzer	QC$_1$		QC$_2$	
	Mean±SD	between-run	Mean±SD	between-run
	mIE/L	CV%	mIE/L	CV%
Maglumi 800	3.33^1±0.193	5.9	37.5^3±0.819	2.1
Immulite 2000	5.0^2±0.25	5.0	31.3^3±2.27	7.2

[1]Upper range snibe control; [2]Upper range Lyphochek control (Bio-Rad); [3]Pathological high Lyphochek control (Bio-Rad). SD: Standard deviation.

The results for FT4 imprecision evaluation, depicted on Tables 3 and 4, suggest higher precision obtained with Immulite 2000 FT4 in both within-run and between-run testing.

Table 3: Within-run imprecision evaluation for FT4

Analyzer	QC$_1$		QC$_2$	
	Mean±SD	within-run	Mean±SD	within-run
	pmol/L	CV%	pmol/L	CV%
Maglumi 800	13.7^1±0.67	4.9	31.4^2±2.08	6.6
Immulite 2000	11.6^3±0.42	3.5	34.3^4±1.5	3.9

QC: Quality control sample; [1]Low range pooled control serum; [2]High range pooled control serum; [3]Low range Lyphochek control (Bio-Rad); [4]High range Lyphochek control (Bio-Rad). SD: Standard deviation.

In regard to method comparison, the first data were obtained using non-parametric Mann–Whitney U-test. For TSH group, the results revealed non-significant difference between means of data (p = 0.9011), while for FT4 group the difference was significant (p = 0.00745).

Table 4: Between-run imprecision evaluation for FT4

Analyzer	QC$_1$		QC$_2$	
	Mean±SD	between-run	Mean±SD	between-run
	pmol/L	CV%	pmol/L	CV%
Maglumi 800	15.27^2±0.89	5.86	32.0^1±2.15	6.4
Immulite 2000	16.3^2±0.59	3.6	34.1^3±1.45	4.2

QC: Quality control sample; [1]High range snibe control; [2]Normal range Lyphochek control (Bio-Rad); [3]High range Lyphochek control (Bio-Rad). SD: Standard deviation.

According to the paired t-test, we again did not found significant average differences between log-transformed TSH values obtained by two analyzers (t_{39}= –0.85, p = 0.39), as depicted on Figure 1.
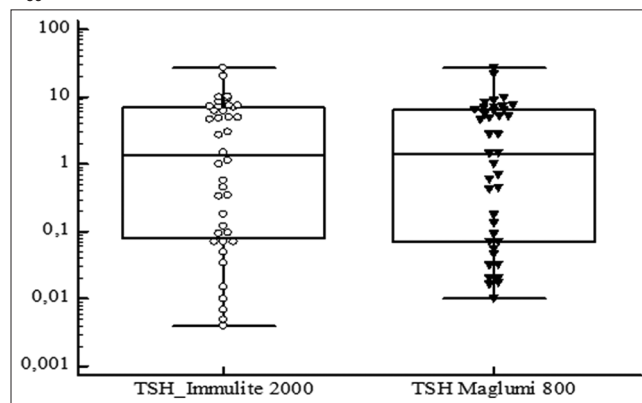

Figure 1: Paired t-test for thyroid-stimulating hormone (TSH) Immulite 2000 and TSH Maglumi 800 measurements

Passing-Bablok regression analysis for TSH group showed no systematic (intercept = 0.01, 95% CI = –0.0059–0.0196) nor proportional (slope = 0.9781, 95% CI = 0.9367–1.0100) differences. The regression equitation and regression line are presented on Figure 2.

The agreement between two TSH methods was evaluated by Bland-Altman plots, with differences expressed as absolute units and percentage. The bias was estimated using limits of agreement. As depicted in Figure 3, the limits of agreement were narrow and the mean was close to the line of equality, which suggests good agreement between TSH Immulite 2000 and TSH Maglumi 800 results.

Analyzing FT4 results, we found significant average differences between methods, obtained by
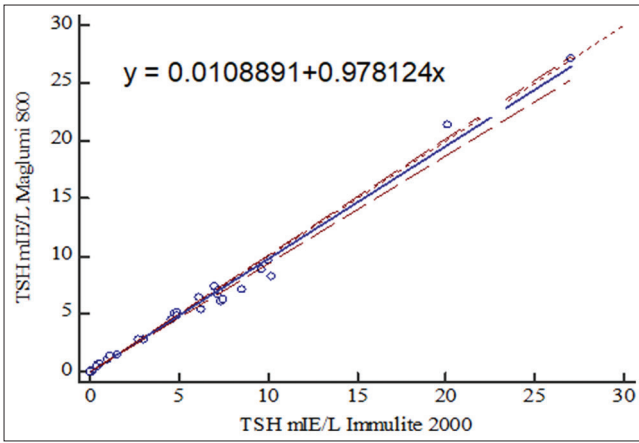
Figure 2: Passing-Bablok regression line between thyroid-stimulating hormone (TSH) Immulite 2000 and TSH Maglumi 800 measurements

paired t-test (t_{39}= 10.5, p < 0.0001), as displayed on Figure 4. Passing-Bablok regression analysis did not reveal significant systematic nor proportional differences.

Figure 3: Bland-Altman plots for thyroid stimulating hormone measurements: (a) Differences between Immulite 2000 and Maglumi 800 measurements expressed as absolute units; (b) differences between Immulite 2000 and Maglumi 800 measurements expressed as percentage of values)

Regression equitation was y=0.527812+1.234375x, intercept value included zero (0.528, 95% CI from -2.4464–3.1568), and slope included one (1.2344, 95% CI from 1.0943 to 1.4013). Regression line is presented in Figure 5.
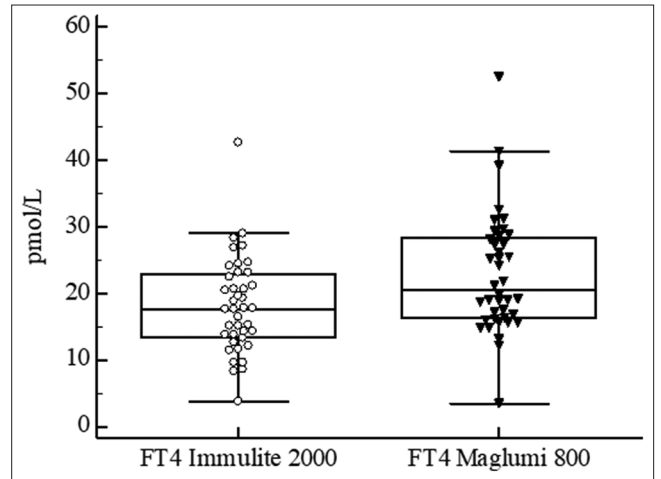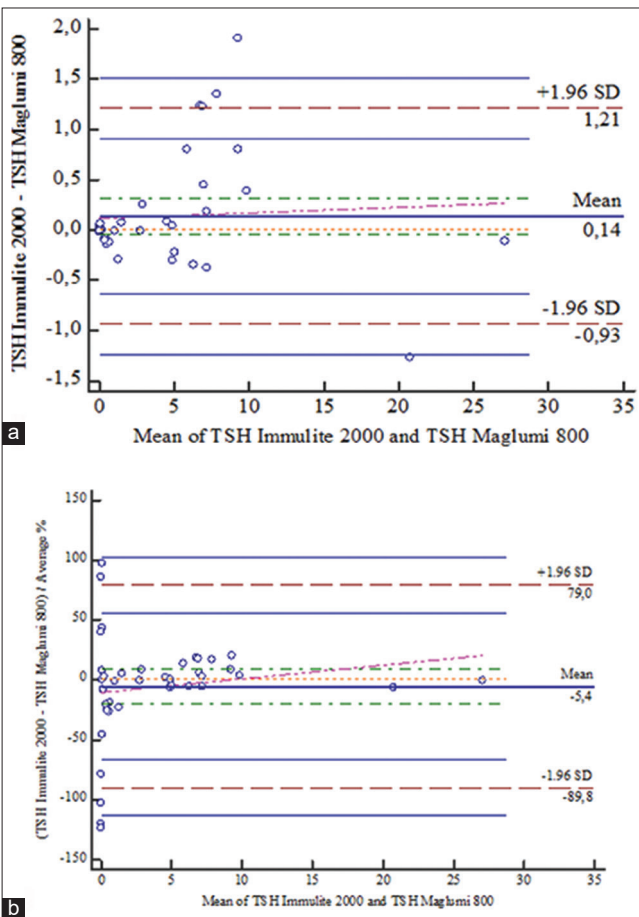

Figure 4: Paired t-test for FT4 Immulite 2000 and FT4 Maglumi 800 measurements

However, when we tested method agreement using Bland-Altman plot, we found overt bias of –4.7 pmol/L and wide limits of agreement, expressed as absolute units (Figure 6a).
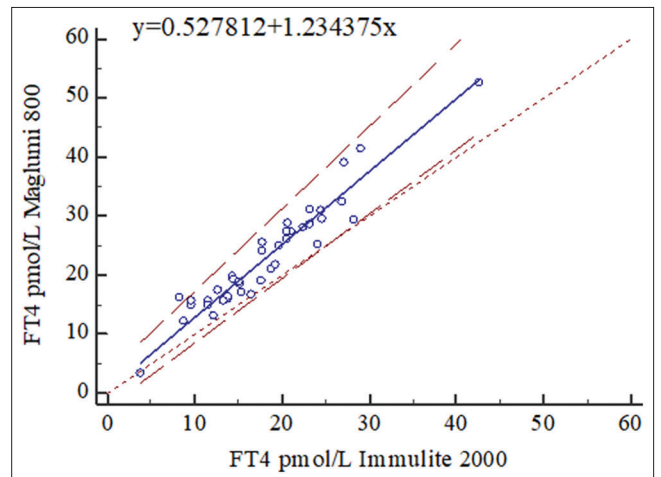

Figure 5: Passing-Bablok regression line between FT4 Immulite 2000 and FT4 Maglumi 800 measurements

Calculated as average percentage, the value obtained was below the line of equality, as well (Figure 6b). These foundings suggest that FT4 results performed on Maglumi 800 are higher in comparison to those performed on Immulite 2000, respectively.

Discussion

Initial evaluation of thyroid function can be performed by single TSH assessment and if TSH is out of reference range, with additional FT4 testing. This two-step approach has been recommended by different thyroid guidelines [13], [14]. However, in usual clinical practice, one-step assessment which includes both measurands simultaneously is often preferred [15]. This approach is generally practiced at IPMN as well.
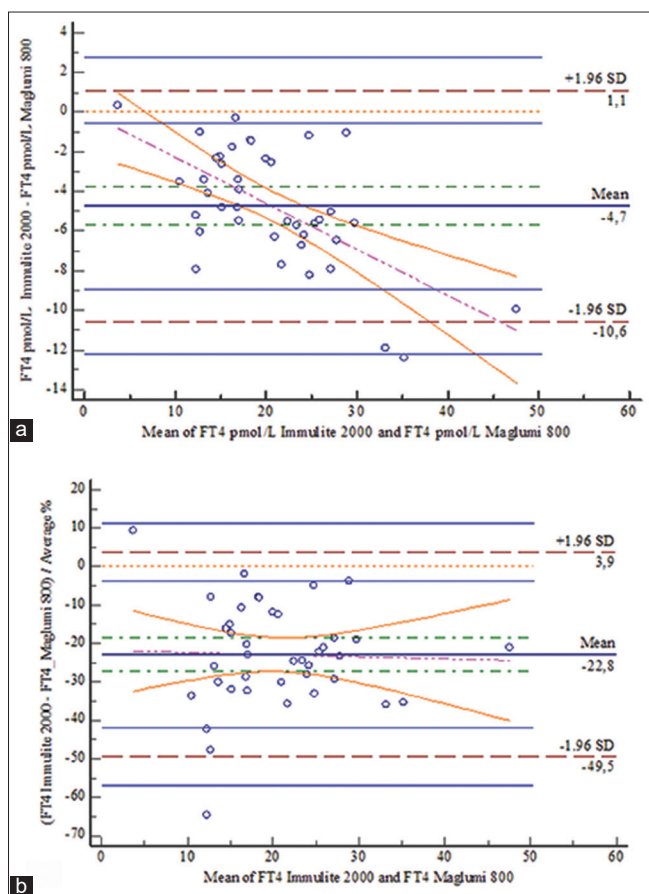
Figure 6: Bland-Altman plots for FT4 measurements: (a) Differences between Immulite 2000 and Maglumi 800 measurements expressed as absolute units; (b) differences between Immulite 2000 and Maglumi 800 measurements expressed as percentage of values)

Therefore, while introducing the new immunoassay test platform, we performed basic analytical performance evaluation and method comparison analysis for TSH and FT4 assays.

Initially, we verified the imprecision of both immunoassays and compared the results. TSH Maglumi 800 assay performed by our laboratory showed excellent within-run precision (1.7–2.8 CV%) compared to manufacturer's data for the similar concentration range (1.76–3.07 CV%). The between-run precision was lower compared to snibe results (2.1–5.9 CV% vs. 1.4–2.04 CV%), but within the desirable imprecision limit of 9.7 CV%, as declared by Westgard for this measurand [16]. It is noteworthy that both reagent manufacturers use the extensive CLSI EP5-A2 protocol [17] for the imprecision testing, including 80 test samples. For TSH Immulite 2000 assay, we obtained similar within-run imprecision with the manufacturer (4.4–5.7 CV% vs. 3.8–5.3 CV%), and 5–7.2 CV% for the between-run imprecision, a data which were not presented by manufacturer. Imprecision verification of both TSH assays performed in our laboratory showed better within-run precision for TSH Maglumi 800 assay.

For FT4 Maglumi 800 assay, our imprecision performance results were higher compared to manufacturer's for both within and between-run (4.9–6.6

CV% and 5.8–6.4 CV% respectively vs. 2.7–4.99 CV% and 1.51–6.17 CV%). The imprecision was higher when compared to FT4 Immulite 2000 assay as well. Our FT4 Immulite 2000 imprecision results were as follows: 3.5–3.9 CV% for within and 3.6–4.2 CV% for between-run imprecision, which were in agreement with data from manufacturer (3.0–7.8 CV%). However, if we apply the strict criteria for allowable imprecision updated by Westgard and based on biological variation [18], [19], the desirable imprecision for FT4 should be 2.9 CV%, a value that has not been achieved by FT4 Immulite 2000 assay manufacturer either, and for certain concentration levels, nor by FT4 Maglumi 800 assay manufacturer. Similar findings for FT4 Immulite 2000 assay were declared by Hendriks *et al.* [7]. A recent Chinese study revealed lower precision rates for several FT4 assays compared to precision performance of TSH assays on different analytical platforms [20].

The other important issue, while introducing a novel immunoassay analyzer, is comparability of results obtained by candidate and routine method. From the available literature reviewed, immunoassays are prone to analytical bias due to majority of factors. For this reason, in 2005, International Federation of Clinical Chemistry (IFCC) established a Working Group on Standardization of Thyroid Function Tests (WG-STFT) with an ultimate goal to standardize and harmonize TFTs [21]. It was restructured in 2012 to Committee for Standardization of Thyroid Function Tests (C-STFT) [22]. Regarding TSH immunoassays, to obtain comparable test results between methods, manufacturers are obliged to recalibrate their methods by adjusting the master calibrators to the all-method mean values, to harmonize assays [23]. These activities were expected to gain less test variability, as was confirmed by Faix and Miller [24]. Our results from TSH comparison tests revealed good agreement between TSH Maglumi 800 and TSH Immulite 2000 test results, despite of using different standard calibrators (WHO third IRP 81/565 and WHO second IRP 80/558, respectively). Mann–Whitney U-test and Paired t-test did not show significant differences and 95% CIs included 0 for intercept and 1 for slope, suggesting acceptable method agreement. The average percentage of difference was 5.4%, which was below 7.8% as a criterion for acceptable bias for serum TSH [16].

For the FT4 immunoassays, IFCC WG-STFT and later the C-STFT, proposed standardization against reference method based on equilibrium dialysis-isotope dilution-mass spectrometry [21]. In addition, small biological variation of FT4 serum concentration demands high performance from the analytical method. Recent criteria for acceptable performance of FT4 immunoassays are aforementioned 2.9% for imprecision and 3.3% for bias. Nonetheless, despite of standardization procedures, in routine laboratory settings, these criteria are not always accomplished [20], [25]. Furthermore, significant bias

and marked variations between methods have been reported [26], with substantial impact of reference range intervals. In concordance with, the data from literature were our findings from two FT4 method comparison studies. Mann–Whitney U-test and paired t-test revealed significant differences between FT4 Maglumi 800 and FT4 Immulite 2000 test results, obtained from 40 serum samples covering reference range, low and high FT4 concentration. The bias of –22.8% obtained from Bland-Altman Plot suggested poor agreement between methods and higher FT4 Maglumi 800 results.

Conclusion

The results of our study signify the importance of verification process that should be performed while introducing a novel immunoassay method. For TSH, we obtained good precision and in comparison study, we found close agreement between Maglumi 800 and Immulite 2000 methods, which assured transferability of results and reporting within the same reference interval. However, FT4 assay performance evaluation revealed higher imprecision of Maglumi 800 platform and significant differences of test results. Thus, clinical interpretation of FT4 results is method-dependent and requires establishment of new reference range for Maglumi 800 analytical platform.

References

1. Ross DS, Burch HB, Cooper DS, Greenlee MC, Laurberg P, Luiza Maia A, *et al*. 2016 American thyroid association guidelines for diagnosis and management of hyperthyroidism and other causes of thyrotoxicosis. Thyroid. 2016;26(10):1343-421.
   PMid:27521067

2. Spencer CA, Lo Presti JS, Patel A, Guttler RB, Eigen A, Shen D, *et al*. Applications of a new chemiluminometric thyrotropin assay to subnormal measurement. J Clin Endocrinol Metab. 1990;70(2):453-60.
   PMid:2105333

3. Garber JR, Cobin RH, Gharib H, Hennessey JV, Klein I, Mechanick JI, *et al*. Clinical practice guidelines for hypothyroidism in adults: Cosponsored by the American association of clinical endocrinologists and the American thyroid association. Thyroid. 2012;22(12):1200-35.
   PMid:23246686

4. Vaidya B, Ukoumunne OC, Shuttleworth J, Bromley A, Lewis A, Hyde C, *et al*. Variability in thyroid function test requests across general practices in south-west England. Qual Prim Care. 2013;21(3):143-8.
   PMid:23968263

5. Daucourt V, Saillour-Glenisson F, Michel P, Jutand MA, Abouelfath A. A multicenter cluster randomized controlled trial of strategies to improve thyroid function testing. Med Care. 2003;41(3):432-41.
   PMid:12618646

6. Willis EA, Datta BN. Effect of an educational intervention on requesting behaviour by a medical admission unit. Ann Clin Biochem. 2013;50(Pt2):166-8.
   PMid:23345588

7. Hendriks HA, Kotlandt W, Verweij WM. Analytical performance comparison of five new generation immunoassay analyzers. Ned Tijdschr Klin Chem 2000;25:170-7.

8. Steele BW, Wang E, Klee GG, Thienpont LM, Soldin SJ, Sokoll LJ, *et al*. Analytic bias of thyroid function tests: Analysis of a college of American pathologists fresh frozen serum pool by 3900 clinical laboratories. Arch Pathol Lab Med. 2005;129(3):310-7.
   PMid:15737023

9. Rawlins ML, Roberts WL. Performance characteristics of six third-generation assays for thyroid-stimulating hormone. Clin Chem. 2004;50(12):2338-44.
   PMid:15472032

10. Clinical and Laboratory Standards Institute. Measurement Procedure Comparison and Bias Estimation Using Patient Samples; Approved Guideline. 3rd ed. CLSI Document EP09-A3. Wayne, PA: Clinical and Laboratory Standards Institute; 2013.

11. Bilić-Zulle L. Comparison of methods: Passing and Bablok regression. Biochem Med (Zagreb) 2011;21(1):49-52.
    PMid:22141206

12. Giavarina D. Understanding bland altman analysis. Biochem Med (Zagreb). 2015;25(2):141-51.
    PMid:26110027

13. LeFevre ML, U.S. Preventive Services Task Force. Screening for thyroid dysfunction: U.S. preventive services task force recommendation statement. Ann Intern Med. 2015;162(9):641-50.
    PMid:25798805

14. Biondi B, Bartalena L, Cooper DS, Hegedus L, Laurberg P, Kahaly GJ. The 2015 european thyroid association guidelines on diagnosis and treatment of endogenous subclinical hyperthyroidism. Eur Thyroid J. 2015;4(3):149-63.
    PMid:26558232

15. Schneider C, Feller M, Bauer DC, Collet TH, Costa BR, Auer R, *et al*. Initial evaluation of thyroid dysfunction are simultaneous TSH and fT4 tests necessary? PLoS One. 2018;13(4):1-12.
    PMid:29709030

16. Minchinela J, Ricós C, Perich C, Fernández-Calle P, Alvarez V, Doménech MV, *et al*. Desirable Specifications for Total Error, Imprecision, and Bias, Derived from Intra and Inter-individual Biologic Variation; 2014. Available from: http://www.westgard.com/biodatabase-2014-update.htm. [Last accessed on 2019 Sep].

17. NCCLS. Evaluation of Precision Performance of Quantitative Measurement Methods; Approved Guideline. 2nd ed. NCCLS Document EP5-A2. Wayne, Pennsylvania; NCCLS; 2004.

18. Andersen S, Bruun NH, Pedersen KM, Laurberg P. Biologic variation is important for interpretation of thyroid function tests. Thyroid. 2003;13(11):1069-78.
    PMid:14651790

19. Tetteh TA, Wijeratne S, Swaminathan R. Intraindividual variation in serum thyroid hormones, parathyroid hormone and insulin-like growth factor 1. Ann Clin Biochem. 2008;45(Pt 2):167-9.
    PMid:18325180

20. Zhang S, Wang W, Zhao H et al. Status of internal quality control for thyroid hormones immunoassays from 2011 to 2016 in China. J Clin Lab Anal. 2018;32(1):1-10.
    PMid:28205257

21. Thienpont LM, Van Uytfanghe K, Beastall G, Faix JD, Ieiri 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(6):912-20.
PMid:20395624

22. Thienpont LM, Van Uytfanghe K, Van Houcke S, Das B, Faix JD, MacKenzie F, *et al*. A progress report of the IFCC committee for standardization of thyroid function tests. Eur Thyroid J. 2014;3(2):109-16.
PMid:25114874

23. Stockl D, Van Uytfanghe K, Van Aelst S, Thienpont LM. A statistical basis for harmonization of thyroid stimulating hormone immunoassays using a robust factor analysis model. Clin Chem Lab Med. 2014;52:965-72.
PMid:24566365

24. Faix JD, Miller WG. Progress in standardizing and harmonizing thyroid function tests. Am J Clin Nutr. 2016;104(Suppl):913S-7S.
PMid:27534642

25. Thienpont LM, Faix JD, Beastall G. Standardiztion of free T4 and harmonization of TSH measurements: A request for input from endocrinologists and other physicians. Eur Thyroid J. 2015;4(4):271-2.
PMid:26835432

26. Barth JH, Luvai A, Jassam N, Mbagaya W, Kilpatrick ES, Narayanan D, *et al*. Comparison of method-related reference intervals for thyroid hormones: Studies from a prospective reference population and a literature review. Ann Clin Biochem. 2018;55(1)107-12.
PMid:28081637