Clin. Lab. 2025;71:XXX-XXX

ORIGINAL ARTICLE

Analytical Verification and Method Comparison of the Maglumi X8 for Thyroid Function Tests

Levent Deniz [1], Meltem Yardim [2], Hale Aral [3]

[1] *Department of Medical Biochemistry, University of Health Sciences Turkey, Istanbul Training and Research Hospital, Istanbul, Turkey*
[2] *Department of Medical Biochemistry, Yerkoy State Hospital, Yozgat, Turkey*
[3] *Department of Medical Biochemistry, Medicana Atakoy Hospital, Istanbul, Turkey*

SUMMARY

Background: **When transitioning to new analytical platforms, laboratories must assess the analytical performances of their methods. This study aimed to verify the precision and trueness of the Maglumi X8 device for thyroid-stimulating hormone (TSH) and free thyroxine (FT4) tests and to compare its performance with that of the Advia Centaur XP system previously used in our laboratory.**
Methods: **Precision and trueness verifications for TSH and FT4 were performed using three levels of Bio-Rad Quality Control (QC) materials following the CLSI EP15-A3 guidelines. Each day consisted of one run with five replicates, resulting in 25 analyses performed using three levels of QC material over five days. Passing-Bablok regression and Bland-Altman analyses were performed for method comparison analysis, following the CLSI EP09c guidelines.**
Results: **The repeatability coefficients of variation (CVs) of TSH for levels 1, 2, and 3 were 2.170, 1.945, and 2.567%, respectively, whereas the within-laboratory (WL) CVs were 2.720, 2.786, and 2.609%, respectively. The repeatability CVs of FT4 for levels 1, 2, and 3 were 3.262, 1.326, and 0.696%, respectively, whereas the WL CVs were 4.848, 4.309, and 4.879%, respectively. The bias or overall mean values obtained in the study for TSH and FT4 levels were within the verification targets. TSH levels were found to be lower on Maglumi X8 [1.770 (1.190 to 2.790)] than Centaur XP [1.975 (1.185 to 3.315)]. FT4 levels were found to be higher on Maglumi X8 [1.305 (1.200 to 1.450)] than Centaur XP [1.210 (1.090 to 1.460)]. The bias between the two methods obtained from the Bland–Altman analysis for TSH and FT4 was -3.76% and 6.68%, respectively. Although the calculated bias for TSH fell within the desirable targets based on biological variation, FT4 did not meet these targets.**
Conclusions: **The precision and trueness verification results for TSH and FT4 demonstrated acceptable performance under the CLSI EP15-A3. TSH results between the two analyzers were consistent and can be transferred; FT4 results from Maglumi X8 require careful interpretation and may need harmonization and standardization due to the observed bias.**
(Clin. Lab. 2025;71:xx-xx. DOI: 10.7754/Clin.Lab.2024.241017)

Correspondence:
Levent Deniz, MD
Department of Medical Biochemistry
University of Health Sciences Turkey
Istanbul Training and Research Hospital
Cerrahpaşa; Org. Abdurrahman Nafiz Gürman Cd., No. 24
34098, Fatih/Istanbul
Turkey
Phone: + 90 5058301741
Email: levent.deniz33@gmail.com
ORCID: 0000 0002 5444 9116

Manuscript accepted November 20, 2024

KEYWORDS

analytical performance, verification, method comparison, standardization, thyroid function tests

INTRODUCTION

The two most requested tests for the diagnosis, treatment, and post-treatment follow-up of endocrine pathologies of the thyroid gland are thyroid-stimulating hormone (TSH) and free thyroxine (FT4). These two parameters are sufficient to diagnose many thyroid dis-

eases without requiring the free triiodothyronine (FT3) test [1]. Accurate measurement of TSH and FT4 levels is crucial for obtaining reliable laboratory results in diagnosing thyroid disorders [2,3]. Lack of standardization among immunological measurement methods may lead to misinterpretation of TSH and FT4 results, and a patient with hypo/hyperthyroidism may be incorrectly diagnosed with euthyroid [4]. Efforts towards standardization have been led by organizations such as the International Federation of Clinical Chemistry and Laboratory Medicine, particularly the Committee for Standardization of Thyroid Function Tests, which aims to achieve equivalence in laboratory test results for FT4 and TSH [5]. Since there is no reference method among the immunoassay methods for measuring TSH, the measurement methods should be harmonized. Standardization and harmonization of FT4 and TSH testing are crucial for achieving comparability of results and improving the diagnosis and treatment of thyroid disorders [4]. The transition between diverse immunoassay methodologies presents significant challenges because of methodological disparities in each assay. According to the ISO 15189 standard, when implementing new analytical platforms or changing methods, it is necessary to verify the specifications provided by the manufacturer [6]. Method verification, a crucial aspect of laboratory testing, is essential to ensure that the performance of a laboratory method aligns with claims made by the manufacturer [7]. It is the user's responsibility to verify whether the methods validated by the manufacturer, according to international guidelines and standards, are as claimed. The Clinical and Laboratory Standards Institute (CLSI) provides essential traceability for method verification and comparison to evaluate analytical performance in laboratories [8].

This study aimed to evaluate the precision and trueness verification of the measurement procedure for TSH and FT4 according to CLSI EP15-A3 [9] on the Maglumi X8 system. Additionally, the results obtained from the Maglumi X8 system were evaluated and compared with those generated by the Advia Centaur XP system according to CLSI EP09c [10].

MATERIALS AND METHODS

This study was approved by the Istanbul Training and Research Hospital Clinical Research Ethics Committee (decision date and number: 06-16-2023/162). This study adhered to the principles outlined in the Declaration of Helsinki.

Assays

Maglumi X8 is a chemiluminescence immunoanalyzer designed for medium and large laboratories, with up to 600 tests/hour (single module), 300 sample positions, and 42 reagent positions. TSH assay is a sandwich chemiluminescence immunoassay. FT4 assay is a competitive chemiluminescence immunoassay. The same reagent lot was used for all the assays.

Verification of precision and trueness

A precision and trueness verification study was performed after familiarization with Maglumi X8 [Shenzhen New Industries Biomedical Engineering Co., Ltd. (SNIBE), Shenzhen, China] analyzer. Before the verification study, the instrument was calibrated, and its quality control (QC) values were within the expected range. Precision and trueness verifications were performed using third-party Bio-Rad QC materials (Bio-Rad Laboratories, Inc. Hercules, California, USA) in accordance with the CLSI EP15-A3 guidelines. For TSH and FT4, three levels of Lyphocheck Immunoassay Plus Control (lot number 40420) were utilized. Each day consisted of one run with five replicates per run. A total of 25 analyses were performed using three levels of QC materials over five days. The QC manufacturer provided the target values for QC materials specific to the SNIBE Maglumi series.

Method comparison

Procedures and analyses were performed according to CLSI EP09c guidelines. The instruments were calibrated before the method comparison study, and their internal QC values were within the expected range. Hemolysis, icteric, and lipemic serum samples were excluded in the study. Samples with overlapping analytical measurement ranges for both immunoassay analyzers were included in this study. A method comparison study was conducted for five days, analyzing TSH (n = 100) and FT4 (n = 70) serum samples collected from routine analyses on the Siemens Centaur XP and subsequently analyzed on the SNIBE Maglumi X8 in two hours. The Siemens Centaur XP method, which was previously routinely used, was accepted as a comparative method, and the SNIBE Maglumi X8 method was accepted as a candidate method.

Statistical analysis

Repeatability (within-run) and between-run variations were calculated using one-way analysis of variance. The total imprecision, known as within-laboratory (WL) variability, was calculated using these two factors. The calculated coefficients of variations (CVs) were evaluated against the values provided by the manufacturer. The precision is verified if the calculated values are less than or equal to the manufacturer's CV% values. If the calculated values exceed the manufacturer CV% values, the upper verification limit (UVL) is calculated and compared against this value. UVL represents the upper limit of the expected 95% confidence interval for the uncertainty obtained in the experiment. The allowable CV target calculated according to Clinical Laboratory Improvement Amendments (CLIA) was accepted as 1/3 of the total allowable error [11].

The overall means of the 25 results and the verification limits were calculated for each control level. The standard error of the overall mean was derived from the re-

Table 1. Precision and trueness verification results for the TSH according to the EP 15A3.

Manufacturer claims (SNIBE Maglumi)	Parameter		
	TSH (µIU/mL)		
	Level-1	Level-2	Level-3
Mean (insert)	0.620	6.543	17.21
User estimates (n = 25)	Level-1	Level-2	Level-3
Repeatability (CV%)	2.170	1.945	2.567
Claimed repeatability (CV%)	3.070	2.090	2.000
UVL for repeatability imprecision (CV%)	4.102	2.792	2.672
WL imprecision (CV%)	2.720	2.786	2.609
Claimed WL imprecision (CV%)	3.710	2.600	2.630
UVL for WL imprecision (CV%)	5.148	3.643	3.768
Target value of QC material	0.489	5.630	39.70
Overall mean	0.470	5.819	41.65
Verification interval	0.473 to 0.505	5.406 to 5.854	38.79 to 40.62
Allowed bias *	0.074	0.850	5.995
Bias	0.019	0.189	1.947

*** - EFLM biological variation database.**
TSH - thyroid-stimulating hormone, UVL - upper verification limit, WL - within-lab, QC - quality control.

Table 2. Precision and trueness verification results for the FT4 according to the EP 15A3.

Manufacturer claims (SNIBE Maglumi)	Parameter		
	FT4 (pg/mL)		
	Level-1	Level-2	Level-3
Mean (insert)	9.740	20.94	45.96
User estimates (n = 25)	Level-1	Level-2	Level-3
Repeatability (CV%)	3.262	1.326	0.696
Claimed repeatability (CV%)	4.990	4.210	2.940
UVL for repeatability imprecision (CV%)	6.667	5.625	3.928
WL imprecision (CV%)	4.848	4.309	4.879
Claimed WL imprecision (CV%)	7.940	5.400	4.700
UVL for WL imprecision (CV%)	12.13	7.646	7.182
Target value of QC material	10.90	28.90	58.20
Overall mean	10.62	28.71	59.73
Verification interval	10.17 to 11.63	26.79 to 31.01	53.08 to 63.32
Allowed bias *	0.382	1.012	2.037
Bias	0.278	0.192	1.528

*** - EFLM biological variation database.**
FT4 - free thyroxine, UVL - upper verification limit, WL - within-lab, QC - quality control.

peatability and WL CVs. For commercial QC material, as the standard error of the target value cannot be estimated, uncertainty is disregarded. Consequently, the standard error of the mean is considered equal to the combined standard error. The verification intervals for the target values of three QC materials have been deter-

Table 3. Analysis of the statistical and clinical significance of the difference between the two instruments.

Parameter	Centaur XP	Maglumi X8	p *	Bias% (CI)	Minimum bias%	Desirable bias%	Optimum bias%
TSH	1.98 (1.19 to 3.32)	1.77 (1.19 to 2.79)	< 0.05	-3.76 (-6.31 to -1.21)	15.1	10.1	5.0
FT4	1.21 (1.09 to 1.46)	1.31 (1.20 to 1.45)	< 0.05	6.68 (3.94 to 9.42)	3.5	2.3	1.2

*** - p-value for Wilcoxon signed-rank test. Variables are expressed as median and interquartile range (25th - 75th percentile). Bias targets were taken from EFLM biological variation targets.**
TSH - thyroid-stimulating hormone, FT4 - free thyroxine.

Table 4. The results of the Passing-Bablok regression and Spearman correlation analysis for TSH and FT4.

Parameter	Intercept	95% CI	Slope	95% CI	Correlation coefficient
TSH	0.0665	0.0123 to 0.1280	0.9143	0.8759 to 0.9500	0.988
FT4	0.3938	0.2389 to 0.5067	0.7500	0.6500 to 0.8649	0.848

CI - confidence interval, TSH - thyroid-stimulating hormone, FT4 - free thyroxine.

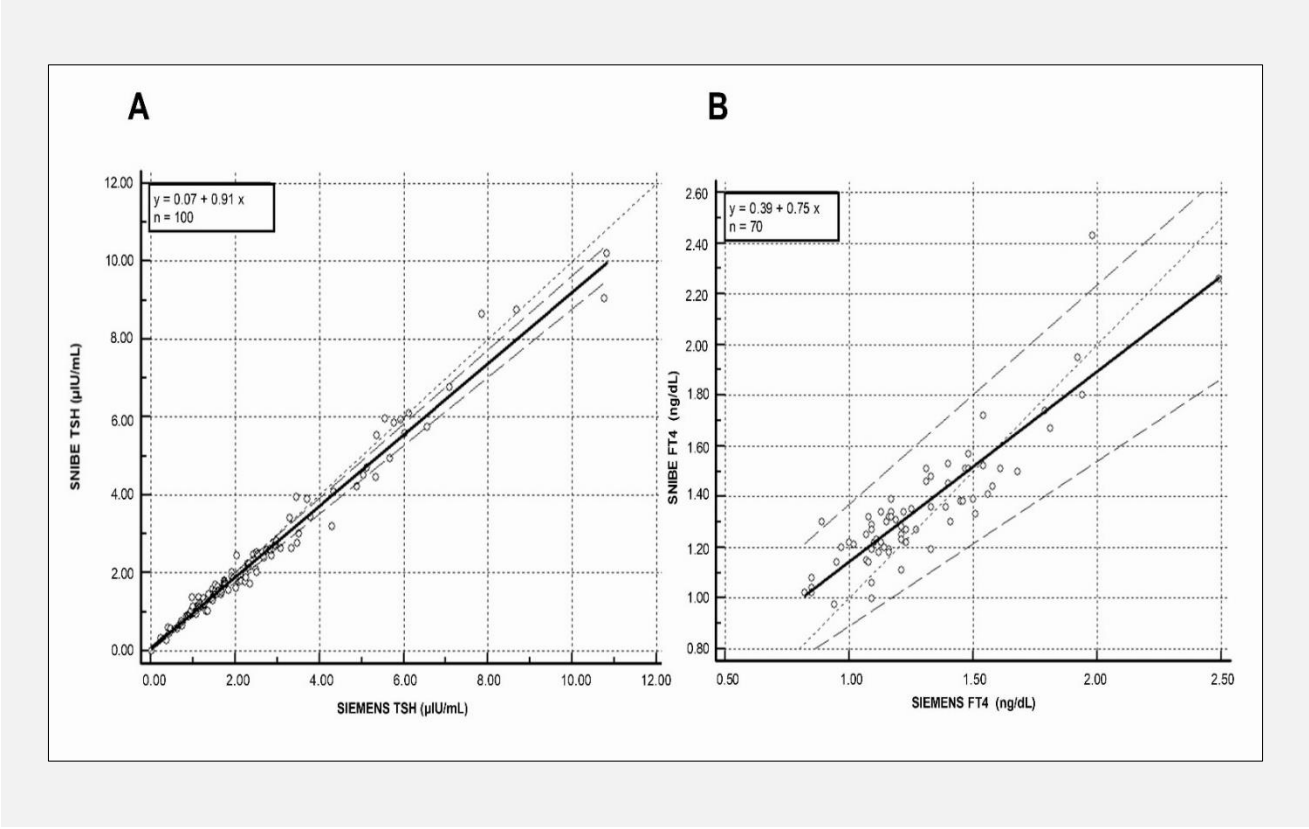

Figure 1. Passing-Bablok Regression Analysis for TSH (A) and FT4 (B).

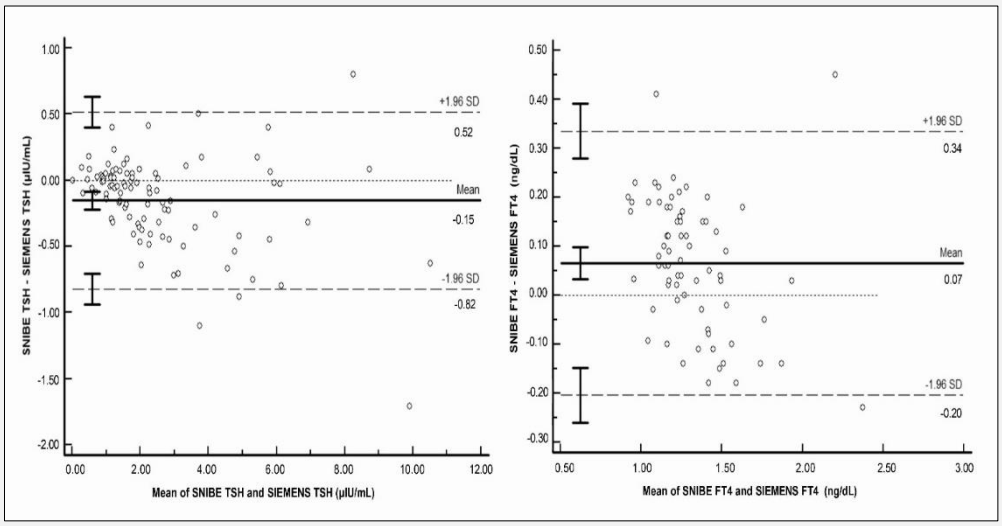

Figure 2. Bland-Altman Analysis of TSH (A) and FT4 (B) levels. Differences between the Maglumi X8 and Centaur XP measurements are expressed as absolute units.

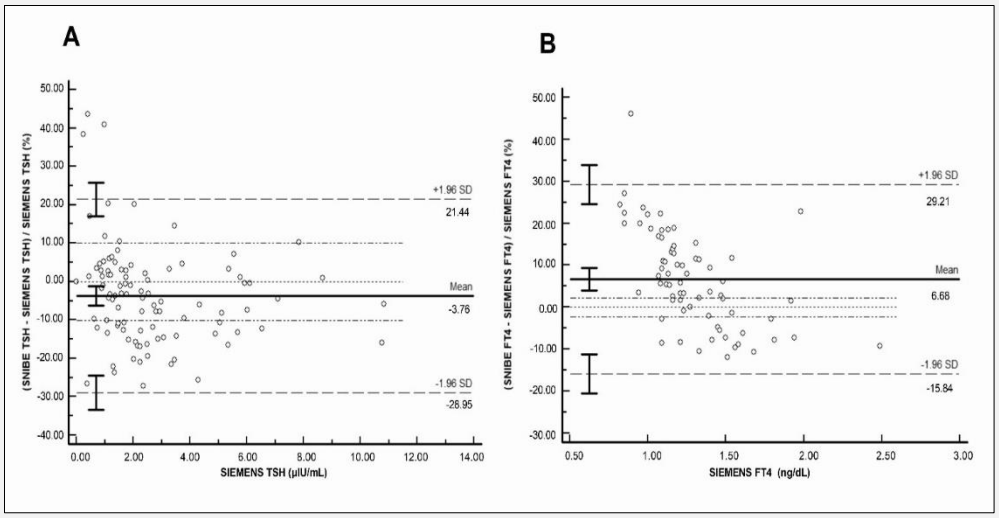

Figure 3. Bland-Altman Analysis of TSH (A) and FT4 (B) levels. Differences between the Maglumi X8 and Centaur XP measurements were expressed as a percentage of values.

mined, taking into account the degrees of freedom and combined standard error. The study utilized a formula to determine the observed bias (overall mean - target value). This calculation involved comparing the overall means obtained from the research with the target value of the QC materials. The trueness was verified if the overall mean value was within the verification limits. However, if the overall mean value exceeded the verification limit, the minimum bias goal for biological variation was considered. If the observed bias was less than the allowed bias, trueness was verified [12].

To assess whether the data conformed to a normal distribution, the Shapiro-Wilk test was employed. Since the data exhibited a non-normal distribution, they were

represented using the median and interquartile range (25th - 75th percentile). The Wilcoxon signed-rank test evaluated the statistically significant differences between the dependent groups. Spearman correlation analysis was conducted to assess the relationship between the two devices. Passing-Bablok regression analysis was performed to examine the proportional and constant systematic errors, including the intercept and slope values, along with their 95% CIs. Bland-Altman analysis was performed to evaluate both absolute and percentage mean of differences between the two methods, providing a visual assessment of agreement. The mean of the differences (percentage bias) and their 95% CIs obtained by the Bland-Altman analysis were compared with the optimum, desirable, and minimum bias targets based on biological variation. Statistical analyses were conducted using Microsoft Excel (Microsoft Corporation, Redmond, WA, USA) and MedCalc Statistical Software version 22.013 (MedCalc Software Ltd., Ostend, Belgium). p-values below 0.05 were regarded as statistically significant.

RESULTS

Precision verification

Tables 1 and 2 provide detailed data on the precision and trueness verification studies conducted according to the CLSI EP15-A3. The repeatability CVs of TSH for levels 1, 2, and 3 were 2.170, 1.945, and 2.567%, respectively, whereas the WL CVs were 2.720, 2.786, and 2.609%, respectively. The repeatability and WL CV results for level 1 were lower than those claimed by the manufacturer. The repeatability CV for level 2 was observed to be within the claimed imprecision target. In contrast, the WL CV for level 2 was within the UVL limits but exceeded the manufacturer's claims. For level 3, the repeatability CV exceeded the manufacturer's claims but met the UVL targets, and WL imprecision was observed to meet the manufacturer's claims (Table 1). The repeatability CVs of FT4 for levels 1, 2, and 3 were 3.262, 1.326, and 0.696%, respectively, whereas the WL CVs were 4.848, 4.309, and 4.879%, respectively. Repeatability CVs for FT4 levels 1, 2, and 3 were within the manufacturer's specified claims. The WL CVs for FT4 levels 1 and 2 also met the manufacturer's claims. However, the WL CV for FT4 level 3 was within the UVL, but exceeded the manufacturer's claims (Table 2).

Trueness verification

For the TSH test, although the overall mean values of levels 1 and 3 were outside the verification interval, bias was verified as the bias value fell within the allowed bias target. The overall mean value of level 2 was within the verification interval (Table 1). For the FT4 test, bias was verified, as the overall mean values obtained from the three levels of QC material were within the verification interval (Table 2).

Comparison of Maglumi X8 and Centaur XP

A statistically significant strong correlation ($r = 0.988$, $p < 0.05$) was observed between the TSH test results obtained from the Maglumi X8 and Centaur XP devices. Passing-Bablok regression analysis for TSH resulted in a slope of 0.914 (CI 0.876 - 0.950) and an intercept of 0.067 (CI 0.012 - 0.128) (Figure 1A, Table 4). TSH levels were found to be significantly lower on Maglumi X8 [1.770 (1.190 - 2.790)] than on Centaur XP [1.975 (1.185 - 3.315)], $p < 0.05$ (Table 3). A weaker correlation ($r = 0.848$, $p < 0.05$) was observed between the FT4 test results obtained using the Maglumi X8 and Centaur XP devices. Passing-Bablok regression analysis for FT4 resulted in a slope of 0.750 (CI 0.650 - 0.865) and an intercept of 0.394 (CI 0.239 - 0.507) (Figure 1B, Table 4). Analysis revealed that FT4 concentrations were significantly elevated when measured using Maglumi X8 [1.305 (1.200 - 1.450)] compared to Centaur XP [1.210 (1.090 - 1.460)], $p < 0.05$ (Table 3).

DISCUSSION

Over the past 30 years, the analytical capabilities of immunoassays have significantly improved; however, notable differences have emerged in the outcomes of various immunoassay techniques, including thyroid function tests. Manufacturers have attempted to address this issue by employing various reference ranges for their methods; however, this problem cannot be fully resolved [1]. Standardization and harmonization of these tests are crucial to guarantee the consistency and precision of laboratory results for serum TSH and free thyroid hormones, benefiting both primary care physicians and endocrinologists [4]. Continuous research and standardization efforts have significantly advanced thyroid function testing, contributing to improved precision diagnostics and clinical implications [13].

Precision and trueness verification are essential processes in laboratory testing to ensure the quality, accuracy, and reliability of test results. This process helps to determine the consistency and reproducibility of test results, which are crucial for making clinical decisions based on laboratory findings [13]. Bias verification evaluates systematic errors or deviations in test results by comparing them to a reference standard or known value [14]. The detection and correction of bias in laboratory testing are vital to ensure that the results are accurate and provide a true reflection of the analyte being measured. Laboratories are required to establish and document their performance specifications for tests, as mandated by regulatory standards such as CLIA [15]. This verification process ensures that the tests meet the predefined criteria for accuracy and precision before being used for patient care.

A new immunoassay device, Maglumi X8, was tested in our laboratory using the CLSI EP15-A3 and EP09c protocols. For TSH level 1, repeatability and WL CVs were lower than the manufacturer's claims. For TSH level 2,

the repeatability CV met the claims, but the WL imprecision exceeded the manufacturer's claims within the UVL limits. For TSH level 3, the repeatability CV exceeded the claims but met the UVL targets, and the WL imprecision met the manufacturer's claims. The repeatability CVs for FT4 levels 1, 2, and 3 were within the manufacturer's specified claims. For FT4 levels 1 and 2, the within-laboratory CV met the manufacturer's claims. However, for FT4 level 3, the within-laboratory CVs were within the UVL but exceeded the manufacturer's claims. Finally, we conducted a verification process to assess the repeatability of our findings using UVL as the acceptance criterion for CVs. The results demonstrated that the repeatability and WL CVs obtained from our study were within an acceptable range compared to UVL. The use of UVL is recommended by CLSI, particularly when the imprecision of the study exceeds the manufacturer's claims. This helps protect testing laboratories from failing inappropriately [8]. Moreover, the CVs obtained for TSH were lower than those obtained for European Federation of Clinical Chemistry and Laboratory Medicine (EFLM) biological variation [12], Westgard biological variation [16], and CLIA 2024 [17] performance goals (8.9%, 9.7%, and 6.7%, respectively). However, the CVs obtained for FT4 were outside the biological variation goals of EFLM and Westgard (2.4% and 2.9%, respectively), but were lower than the CLIA 2024 goals (5.0%). These targets, based on biological variation, were not achieved by the manufacturer Maglumi for FT4. This situation can be explained by the relatively limited range of biological variation targets. Consequently, we concluded that the laboratory imprecision of TSH and FT4 assays for Maglumi was consistent with the manufacturer's claims, according to the EP15A3 protocol.

The EP15-A3 protocol allows for convenient evaluation of precision and bias within a single experimental design. According to the EP15A3 protocol, bias was calculated by comparing the overall mean value obtained from the precision study (n = 25) using three different levels of Bio-Rad QC material against the target values specified by the manufacturer. For the Maglumi series, the target value was utilized since there were insufficient participants to obtain a peer group mean in the unity^{TM} Interlab-Bio-Rad. According to EP15-A3, the mean values must fall within the verification intervals for the bias verification to be acceptable. If this requirement is not fulfilled, the computed bias values should fall within specified bias targets. In the bias verification study for the TSH test, although the overall mean values calculated for the Level-1 and Level-3 QC materials were not within the verification intervals, the verification study was deemed acceptable, because the bias values were within the allowable limits. In the bias verification study of the FT4 test, the calculated overall mean values for all three QC material levels were within the verification intervals.

Correlation analysis between the two devices revealed that TSH demonstrated a stronger correlation than FT4. Statistical analysis indicated that the Maglumi analyzer tended to measure TSH levels lower and FT4 levels higher than the Centaur analyzer. The fact that the slope value did not include 1 and the intercept value did not include 0 in the Passing-Bablok regression analysis for TSH indicated the presence of both constant and proportional systematic errors. In the Bland-Altman analysis of the TSH test, the percentage mean of differences (%bias) and CI [-3.76 (-6.31 to -1.21)] were within the desirable bias (10.1%) and the minimum bias (15.1%) targets (Table 3, Figure 3A). Although the bias value remained within the optimum target (5.0%), the lower confidence interval exceeded this target. Although a statistically significant difference was found between the two devices in TSH measurements, the data suggest that this difference is within a clinically acceptable range. The fact that the slope value does not include 1 and the intercept value does not include 0 in the Passing-Bablok regression analysis for FT4 indicates the presence of both constant and proportional systematic errors. The fact that the obtained intercept and slope values are not as close to 1 and 0 as the TSH measurement values indicates that the discrepancy in FT4 measurements between the two devices results from their inconsistency. The observation that the obtained intercept and slope values deviated from 1 and 0 when compared to the TSH measurement values suggests a more pronounced statistical significance in the differences observed between the FT4 measurements of the two devices. In the Bland-Altman analysis of the FT4 test, % bias and CI [6.68 (3.94 to 9.42)] were not within the optimum bias (1.20%), desirable bias (2.30%), or even the minimum bias (3.50%) targets (Table 3, Figure 3B). Analysis of Passing-Bablok regression and Bland-Altman plots indicated that FT4 measurements exhibited a bias that varied with concentration levels. At lower concentrations, a positive bias was noted, whereas higher concentrations exhibited a negative bias (Figure 1, 2, and 3). The existing literature contains a limited number of studies examining the analytical capabilities of SNIBE Maglumi instruments, particularly for common immunoassay tests. In one study, an analysis was conducted in which the results obtained from the Maglumi 800 were compared with those generated by the Immulite 2000 device to evaluate the levels of TSH and FT4. For the Maglumi 800 TSH, the within-run CV was 1.7% at a low concentration of 0.48 mIU/L and 2.8% at a high concentration of 3.39 mIU/L. In addition, the between-run CV was 5.9% at a low concentration of 3.33 mIU/L and 2.1% at a high concentration of 37.5 mIU/L [18]. In our study, the within-run and WL-CV values were found to be below 3% for concentrations close to those employed in the study. In the study, between-run CV for TSH was relatively high at low concentrations, whereas better CV values were obtained in our study. The regression analysis and Bland-Altman plots showed no significant systematic or proportional differences between the TSH results from the two platforms, suggesting good agreement and transferability of the results be-

tween Maglumi 800 and Immulite 2000 [18]. Similar findings were observed in our study. No clinically significant difference was found, because the bias value obtained between the two devices in the TSH measurements was within the allowable bias limits.
For the Maglumi 800 FT4, the within-run CV was 4.9% at a low concentration of 13.7 pmol/L and 6.6% at a high concentration of 31.4 pmol/L. In addition, the between-run CV was 5.86 % at a low concentration of 15.27 pmol/L and 6.4% at a high concentration of 32.0 pmol/L [18]. Unlike TSH, significant differences were observed between FT4 results from the two platforms. Maglumi 800 consistently showed a bias with higher FT4 values than Immulite 2000. The regression analysis did not reveal significant differences, but the Bland-Altman plot indicated a substantial bias, suggesting poor agreement between the two methods for FT4 [18]. In our study, which was conducted using devices of the same brand, FT4 did not yield consistent results between the two devices.

CONCLUSION

The precision and bias verification results demonstrated acceptable performance under the CLSI EP15-A3. TSH results between the two analyzers are comparable and can be transferred confidently; FT4 results from Maglumi X8 require careful interpretation and may need harmonization and standardization due to the observed bias.

Source of Funds:
There was no grant support for this study.

Declaration of Interest:
There are no conflicts of interest for any of the authors. All the authors have read and approved the manuscript.

References:

1. Bozkaya G, Şişman AR. The method comparison and the verification of precision of Mindray CL-6000i thyroid function tests (TFTs). Turk J Biochem 2021;46(3):255-62. https://www.degruyter.com/document/doi/10.1515/tjb-2019-0506/html

2. Thienpont LM, Beastall G, Christofides ND, et al. Measurement of free thyroxine in laboratory medicine-proposal of measurand definition. Clin Chem Lab Med 2007;45(4):563-4. (PMID: 17439341)

3. Li Z-Z, Yu B-Z, Wang J-L, Yang Q, Ming J, Tang Y-R. Reference intervals for thyroid-stimulating hormone and thyroid hormones using the access TSH 3rd IS method in China. J Clin Lab Anal 2020;34(5):e23197. (PMID: 31912542)

4. 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(2):305-6. (PMID: 26201824)

5. Faix JD, Miller WG. Progress in standardizing and harmonizing thyroid function tests. Am J Clin Nutr 2016;104 Suppl 3(Suppl 3):913S-7S. (PMID: 27534642)

6. Lapic I, Kralik Oguic S, Rogic D. Preliminary evaluation of eight less frequent endocrine assays designed for MAGLUMI 800 chemiluminescence immunoanalyzer. Scand J Clin Lab Invest 2021;81(4):332-8. (PMID: 33798031)

7. Loh TP, Cooke BR, Markus C, et al. Method evaluation in the clinical laboratory. Clin Chem Lab Med 2023;61(5):751-8. (PMID: 36327459)

8. Chakravarthy SN, Ramanathan S, S S, Nallathambi T, S M. EP15A3 Based Precision and Trueness Verification of VITROS HbA1C Immunoassay. Indian J Clin Biochem 2019;34(1):89-94. (PMID: 30728678)

9. CLSI. User Verification of Precision and Estimation of Bias; Approved Guideline-Third Edition. CLSI document EP15-A3. Clinical and Laboratory Standards Institute 2014. https://clsi.org/standards/products/method-evaluation/documents/ep15/

10. CLSI. Measurement Procedure Comparison and Bias Estimation Using Patient Samples. 3rd ed. CLSI guideline EP09c. Clinical and Laboratory Standards Institute 2018. https://webstore.ansi.org/preview-pages/CLSI/preview_CLSI+EP09c.pdf?srsltid=AfmBOope4809PRu51sPHTIJcpcYaTTLr6Qzq9TxxibBNq1W829hgdlur

11. Ehrmeyer SS, Laessig RH, Leinweber JE, Oryall JJ. 1990 Medicare/CLIA final rules for proficiency testing: minimum intralaboratory performance characteristics (CV and bias) needed to pass. Clin Chem 1990;36(10):1736-40. (PMID: 2170059)

12. Aarsand AK, Fernandez-Calle P, Webster C, et al. The EFLM Biological Variation Database. https://biologicalvariation.eu/

13. Yazdaan HE, Jaya F, Sanjna F, et al. Advances in Thyroid Function Tests: Precision Diagnostics and Clinical Implications. Cureus 2023;15(11):e48961. (PMID: 38111456)

14. Topic E, Nikolac N, Panteghini M, et al. How to assess the quality of your analytical method? Clin Chem Lab Med 2015;53(11):1707-18. (PMID: 26408611)

15. Burd EM. Validation of laboratory-developed molecular assays for infectious diseases. Clin Microbiol Rev 2010;23(3):550-76. (PMID: 20610823)

16. Westgard QC. Desirable biological variation database specifications. Desirable specifications for total error, imprecision, and bias, derived from intra- and inter individual biologic variation. https://westgard.com/clia-a-quality/quality-requirements/biodatabase1.html

17. Westgard QC. 2024 CLIA Acceptance Limits for Proficiency Testing. https://westgard.com/clia-a-quality/quality-requirements/2024-clia-requirements.html

18. Kuzmanovska S, Miladinova D. Comparison of Thyroid -Stimulating Hormone and Free Thyroxine Immunoassays Performed on Immulite 2000 and Maglumi 800 Automated Analyzers. Open Access Macedonian Journal of Medical Sciences 2020;8(B):168-74. https://oamjms.eu/index.php/mjms/article/view/4318