 doi:10.20944/preprints202210.0053.v1

Correspondence

Uncertainty of serum TSH and thyroxine on Abbott Architect i1000sr from internal quality control data: Use in results interpretation and QC frequency planning.

Khelil Mohamed Mokhtar

Pasteur Institute of Algeria, Department of Immunology, Annex of M'sila, Algeria

Keywords: Measurement uncertainty, thyroid function tests, quality control, sigma-metric.

ABSTRACT

The minimum requirement for uncertainty estimation is to use only intermediate precision (R_w), especially for measurands lacking a reference measurement system such as thyroid functions tests (TFT). In this study, measurement uncertainty (MU) for TSH and FT4 from long-term internal quality control (IQC) data was estimated while reference change values (RCV) were calculated from estimated MU. Furthermore, intermediate precision (R_w) was used to establish appropriate risk-based QC frequency. Twenty fore months of third party IQC data were collected retrospectively, on the Abbott ARCHITECT i1000sr analyzer from INSTITUT PASTEUR OF MSILA laboratory, ALGERIA. The MU, RCV and sigma-metric were estimated simply from the intermediate precision (R_w), while a nomogram relating sigma performance to run size was used to establish QC frequency. The MU for the TSH and FT4 was 12% and 8% respectively. The $U_{\text{one-sided}}$ for the TSH and FT4 was 10%, 6.6% respectively. MU and $U_{\text{one-sided}}$ of TSH and FT4 met quality requirements for permissible uncertainty (pU %) and allowable total error (ATE %). When monitoring thyroid replacement therapy, an upward minimum change (RCV) of 54% and 22% or a downward of 35% and 18% in serum TSH and FT4 respectively, would be considered significant. Optimal QC strategy for serum TSH was selected to run 4 QC materials every 190 patients sample and to use a multi-rule ($1_{3s}/2_{2s}/R_{4s}/4_{1s}$). Our results suggest that MU estimation from long-term IQC alone may be acceptable for TFT to assist physician in results interpretation and to establish appropriate QC frequency.

INTRODUCTION:

Medical laboratories play a pivotal role in the diagnosis and management of thyroid conditions [1]. Thyroid function tests (TFT) represented by TSH and thyroxine (Free T4) are the most frequently ordered endocrine tests at our outpatient institution (INSTITUT PASTEUR OF MSILA, ALGERIA), they constitute more than 80 % of ordered endocrine tests.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 6 October 2022 doi:10.20944/preprints202210.0053.v1

The accuracy of TFT depends on both the trueness (bias) and imprecision (standard deviation) of the immunoassay procedure because only a single measurement is made to produce a test result [2]. The true value of the single measurement is then unknown and remains within the so called "uncertainty boundaries" [3]. The ISO 15189 has recommended that measurement uncertainty (MU) should be determined for all measurement procedures, with no restriction about which calculation method to use [4]. In the absence of a standard guide, MU could be reliably estimated from only long-term (> 6 months) intermediate precision (R_w) [5, 6, 7]. Internal QC data could be used to (R_w) estimation provided that QC materials are commutable and target values are near clinical decision limit values. Moreover, due to the unavailability of certified reference material (CRM) or reference method to calculate the systematic component (bias) for serum TSH and thyroxine , some sources of uncertainty due to bias (all short time bias and some long term bias) could be included in the random component (R_w) of MU [6]. Despite conceptual differences between MU and Total error (TE) models, they can be mathematically identical in special cases [8]. Thus, long-term intermediate precision (R_w) can be used for sigma-metric calculation and subsequently selection of risk-based quality control strategy [9]. In this study, based on long term IQC data, we estimated MU and RCV for TSH, FT4 on Abbott Architect i1000sr analyzer. Also, we used long term imprecision (R_w) to design risk-based quality control strategy for our TSH procedure.

MATERIALS AND METHODS

ARCHITECT TSH® (Abbott; Cat # 7K62-25) and ARCHITECT Free T4® (Abbott; Cat # 7K65), were run immunometrically on the "Load on the fly" random access analyzer (Abbott Architect i1000SR; USA) at the laboratory of INSTITUT PASTEUR OF M'SILA, ALGERIA.

Commutable third party quality control (QC) materials were obtained from Technopath Multichem IA plus® (Ref 05P76-10). We retrospectively collected data of clinically relevant levels of QC over 24 months (from July 2019 to July 2022). QC results were examined separately according to the lots of each level. Coefficient of variation (CV %) of each lot is calculated then pooled by the formula:

$$\text{Pooled CV\%} = \sqrt{\frac{(N_{lot1}-1)xCV_{lot1}^2+(N_{lot2}-1)xCV_{lot2}^2+\cdots+(N_{loti}-1)xCV^2 i}{(N_{lot1}+N_{lot2}+\cdots+N_{loti})-N_{periods}}} \qquad (1)$$

Intermediate precision (R_W) was just the long term within-laboratory reproducibility (S_{RW}) and it was determined by calculating the pooled coefficient of variation from Level 1 and Level 2 as indicated below:

$$u_c = R_w = CV_{pooled} = \sqrt{\frac{(N_{Level1}-1)xCV_{level1}^2+(N_{level2}-1)xCV_{level2}^2}{(N_{level1}+N_{level2})-2}} \qquad (2)$$

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 6 October 2022 doi:10.20944/preprints202210.0053.v1

The expanded uncertainty ($U_{\%}$) was then obtained by multiplying the intermediate precision (R_w) by the factor k=2 covering 95% confidence interval as follows:

$U_{\%} = 2 \times R_w$ (3)

Coverage factor =1.65 should be used to compare with the ATE% [10].Therefore, standard uncertainties was multiplied by 1.65 as one-sided estimation ($U_{\text{one-sided}}$).

Permissible expanded uncertainty (pU%) was estimated by the algorithm proposed by Haeckel et al. Reference limits and medians estimated locally by an indirect method were entered in the excel table to get pU% [11].

Reference change value (RCV) between two measurements with a confidence of 95% was calculated by use of Fokkema et al approach [11]. This method is convenient for measurands with logarithmic distribution and procedures with CV% > 5%. The EFLM website via an interactive tool enable the calculation of RCVs by entering the procedure CV% [12].

Sigma metric was calculated by the formula:

Sigma = (ATE% / R_W %)

To obtain the appropriate run size (number of patients samples between two QC events),Sigma values were plotted on the x-axis of the nomogram from **Fig.1** and a vertical line was drawn to intersect the lines on the graph that represent different control rules and different numbers of QC measurements (N). Then the appropriate run sizes were read on the y-axis.

RESULTS

MU for TSH and FT4 was 12%, 8% respectively. The $U_{\text{one-sided}}$ for the TSH and FT4 was 10% and 6.6% respectively. Because of the clinical irrelevance of TSH and FT4 QC Level (1), their CV% was not included to the final MU calculation. If TSH and FT4 level 1 CV% was considered, an unnecessary increase in MU would be observed. All MU were below the limits of pU%. All $U_{\text{one-sided}}$ were below the corresponding limits of allowable total error. The estimated RCVs were asymmetrical for all parameters. The sigma-metric showed a fair performance ($\sigma = 4$) for TSH and marginal performance ($\sigma = 3$) for FT4. Results are summarized in **Table 1.** Appropriate TSH run size was 190 patient samples for a multi-rule ($1_{3s}/2_{2s}/R_{4s}/4_{1s}$, N = 4) with a probability of error detection (P_{ed}) = 0.91, and 40 patient samples with only N=2 but with an unacceptable Ped=0.48.

DISCUSSION:

In the present study, we have estimated MU and RCV of TSH and thyroxine on Abbott architect i1000sr from long term IQC data. Then, we have planned QC strategy based on sigma quality. To see if our estimated MU met quality requirements we compared TSH and thyroxine MU with calculated permissible uncertainty limits pU% and with Allowable total errors (ATE%) using an appropriate coverage factor (k= 1.65). MU of TSH and FT4 were in compliance with pU% and

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 6 October 2022 doi:10.20944/preprints202210.0053.v1

ATE % targets. The Haeckel et al pU% estimation was based on the so called CV_E –empirical variation- which was derived from intra-laboratory reference intervals as a surrogate of the biological variation [13] .The performance goals of this approach were set for intra-laboratory use which explain the stringent limits compared to ATE% goals which were set for inter-laboratory use.

When serum TSH and FT4 are used to monitor thyroid hormone replacement therapy, endocrinologists needs an objective tools to know what magnitude of changes between two determinations should be considered significant i.e. not due to analytical or biological variations. With our MU estimations, the RCVs for TSH and FT4 were determined to be 54% and 22% respectively for increase concentration monitoring , and -35% , -18% respectively for decrease concentrations . This means that RCV are asymmetric depending on the direction of concentration change which allow more accurate patient monitoring.

The CLSI's new guideline for statistical quality control C24-Ed4 recommended the application of risk-based SQC for optimizing the frequency of QC, on the basis of the expected number of erroneous results when an undetected out-of-control error conditions occurs [14]. Graphical tools have been developed to support laboratory applications relating the sigma performance of the testing process to run size for continuous mode process [15]. Selecting an optimum QC strategy is to seek the strategy with the highest probability of error detection (P_{ed}) and lowest QC utilization. For our TSH procedure, we have selected the run size of 190 patient samples for a multi-rule procedure ($1_{3s}/2_{2s}/R_{4s}/4_{1s,\ N}$=4). However, additional factors have been described that influence decisions about the appropriate QC frequency [16]: (1) The frequency of out-of control events (method reliability), (2) the likelihood of inappropriate medical decision and the severity of harm of this decision. For example if our TSH procedure will give an out of control condition once a month it will be unacceptable, causing unnecessary treatment (e.g., L-thyroxine or antithyroid drugs) , stopping or modifying ongoing treatment and using unnecessary complementary tests - I^{123}thyroid scans or TRH stimulation tests-. Thus, the QC strategy must be fine-tuned to more frequent QC.

In summary, we have provided evidences that long-term IQC data (> 1 year) may be sufficient for MU estimation for thyroid function tests and may be applied for other immunoassay tests. Moreover, long-term intermediate precision (Rw) may be used to calculate sigma-metric and subsequently designing an appropriate risk-based QC frequency plan.

REFERENCES

1. Soh, S. B., & Aw, T. C. (2019). Laboratory Testing in Thyroid Conditions - Pitfalls and Clinical Utility. Annals of Laboratory Medicine, 39(1), 3–14.
2. Westgard, J. O., & Westgard, S. A. (2017). Measuring Analytical Quality. Clinics in Laboratory Medicine, 37(1), 1–13.
3. Dallas Jones, G. R. (2016). Measurement uncertainty for clinical laboratories – a revision of the concept. Clinical Chemistry and Laboratory Medicine (CCLM), 54(8).

4. 15189:2012, I., 2017. ISO 15189:2012 - Medical laboratories -- Requirements for quality and competence. [online] Iso.org. Available at: <https://www.iso.org/standard/56115.html> [Accessed 24 August 2022].
5. Ćelap, I., Vukasović, I., Juričić, G., & Šimundić, A. M. (2017). Minimum requirements for the estimation of measurement uncertainty: Recommendations of the joint Working group for uncertainty of measurement of the CSMBLM and CCMB. Biochemia Medica, 27(3).
6. Ceriotti, F. (2018). Deriving proper measurement uncertainty from Internal Quality Control data: An impossible mission? Clinical Biochemistry, 57, 37–40.
7. Coskun, A., Theodorsson, E., Oosterhuis, W. P., & Sandberg, S. (2022). Measurement uncertainty for practical use. Clinica Chimica Acta, 531, 352–360.
8. Oosterhuis, W. P., & Theodorsson, E. (2016). Total error vs. measurement uncertainty: revolution or evolution? Clinical Chemistry and Laboratory Medicine (CCLM), 54(2).
9. Westgard, M. S. S. (n.d.-b). Graphic Report of TSH Quality - Westgard. WESTGARD QC. Retrieved August 24, 2022, from https://www.westgard.com/2010-tsh-quality.htm
10. Çubukçu, H. C. (2018). Improper coverage factor to compare expanded uncertainty with allowable total error. Turkish Journal of Biochemistry, 43(5), 563.
11. Haeckel, R., Wosniok, W., Gurr, E., & Peil, B. (2015). Permissible limits for uncertainty of measurement in laboratory medicine. Clinical Chemistry and Laboratory Medicine (CCLM), 53(8).
12. Fokkema, M. R., Herrmann, Z., Muskiet, F. A., & Moecks, J. (2006). Reference Change Values for Brain Natriuretic Peptides Revisited. Clinical Chemistry, 52(8), 1602–1603.
13. EFLM Biological Variation. EFLM. Retrieved August 24, 2022, from https://biologicalvariation.eu/
14. CLSI C24-Ed4. Statistical quality control for quantitative measurement procedures: principles and definitions. 4th Ed. Wayne (PA): Clinical and Laboratory Standards Institute; 2016.
15. Bayat, H., Westgard, S. A., & Westgard, J. O. (2017). Planning Risk-Based Statistical Quality Control Strategies: Graphical Tools to Support the New Clinical and Laboratory Standards Institute C24-Ed4 Guidance. The Journal of Applied Laboratory Medicine, 2(2), 211–221.
16. Parvin, C. A., PhD, & Mscs, Y. J. C. (n.d.). Don't Rely Solely on Sigma Values to Determine QC Frequency. QC NET. Retrieved August 24, 2022, from https://www.qcnet.com/resources/qc-articles/dont-rely-solely-on-sigma-values-to-determine-qc-frequency