Received: 20 July 2020 | Revised: 30 September 2020 | Accepted: 5 October 2020

DOI: 10.1111/cen.14356

ORIGINAL ARTICLE

WILEY

Biological variation and analytical goals of four thyroid function biomarkers in healthy European volunteers

Antoine Mairesse[1] | Loris Wauthier[1] | Louisiane Courcelles[1] | Urszula Luyten[1] | Maria-Cristina Burlacu[2] | Diane Maisin[1] | Julien Favresse[3,4] | Marie-Astrid van Dievoet[1] | Damien Gruson[1,5]

[1]Clinical Biology Department, Cliniques Universitaires St Luc, Université catholique de Louvain, Brussels, Belgium

[2]Department of Endocrinology and Nutrition, Cliniques Universitaires St-Luc, Universite Catholique de Louvain, Brussels, Belgium

[3]Department of Laboratory Medicine, Clinique St-Luc Bouge, Namur, Belgium

[4]Department of Pharmacy, Namur Research Institute for Life Sciences, University of Namur, Belgium

[5]Pôle de recherche en Endocrinologie, Diabète et Nutrition, Institut de Recherche Expérimentale et Clinique, Cliniques Universitaires St-Luc and Université Catholique de Louvain, Brussels, Belgium

Correspondence
Damien Gruson, Pôle de recherche en Endocrinologie, Diabète et Nutrition, Université catholique de Louvain, Tour Claude Bernard, 54 Avenue Hippocrate, B-1200 Brussels, Belgium.
Email: damien.gruson@uclouvain.be

Abstract

Background: Interpretation of thyroid function tests by means of biological variation (BV) data is essential to identify significant changes between serial measurements at an individual level. Data on thyroid parameters in adults are limited.

Objectives: We aimed at determining the BV of four thyroid function test (thyroid-stimulating hormone (TSH), free thyroxin (FT4), free triiodothyronine (FT3) and thyroglobulin (Tg)) by applying recent recommendations to acquire BV data on a latest generation of immunoassay.

Methods: Nineteen healthy volunteers (8 males and 11 females) were drawn every week during 5 consecutive weeks. Samples were analysed in duplicate on the Cobas 602 analyzer (Roche Diagnostics). After normality assessment, outlier exclusion and homogeneity of variance analysis, analytical variation (CV_A), within-subject biological variation (CV_I) and between-subject biological variation (CV_G) were determined using nested ANOVA.

Results: CV_A, CV_I and CV_G were 0.9%, 19.7% and 37.6% for TSH; 3.6%, 4.6% and 10.8% for FT4; 2.2%, 6.0% and 8.6% for FT3; and 0.9%, 15.4% and 84.9% for Tg. Index of individuality (II) for all parameters was between 0.2 and 0.7. The percentage above which the change between two measures is truly significant (reference change value) was 54.7% for TSH, 16.2% for FT4, 17.7% for FT3 and 42.8% for Tg.

Conclusion: Based on recent international recommendations, our study provides updated BV data for four thyroid function tests in European healthy volunteers. Reliable BV characteristics, and especially RCV, can facilitate the interpretation of consecutive thyroid function tests in an individual and therefore have the potential to efficiently support clinical decisions regarding thyroid diseases.

KEYWORDS
biological variation, Free T3, Free T4, reference change value, thyroglobulin, thyroid

Antoine Mairesse and Loris Wauthier contributed equally to this work.

1 | INTRODUCTION

Thyroid disorders have become the world most common endocrine conditions and both their prevalence and the spectrum of thyroid-associated diseases have been increasing the last decades.[1] Symptoms and clinical signs are often non-specific, and the diagnosis and follow-up of treatment are mainly dependent on laboratory assessment of thyroid-stimulating hormone (TSH) and thyroid hormones (TH).[2] Leading laboratory tests for TH determination are mainly based on immunoassay technology. Major advances in assays standardization and analytical performances have been performed in the field the last decades, with latest generations of TSH immunoassays achieving functional sensitivity of 0.01 mIU/L which is 100 times lower than first generation assays.[3]

The high inter-method and inter-population variability of these tests have led to the recommendation that each laboratory should establish its own reference intervals.[4] Thyroid function in healthy subjects is known to present a narrow intraindividual variability: each individual tends to have its own homeostatic set point of thyroid regulation.[5,6] On the other hand, the interindividual variability is high, leading to broader laboratory reference ranges.[5] In clinical laboratory settings, the determination of biological variation (BV) data using routine analytical methods **allows** the documentation of this variability. In fact, the laboratory test results vary over time due to three main factors: pre-analytical changes (sampling, type of tube, fasting, meals,...), analytical variation (degree of dispersion of results (precision) and systematic error in relation to a value considered to be true (bias)) and inherent BV that represents the natural fluctuations and physiological components of biological constituents around an individualized homeostatic point.[7] The latter comprises the within-individual variation (CV_I; the random fluctuation around a homeostatic setting point) and the between-individual variation (CV_G; the variation in homeostatic setting points of different subjects).[7]

Currently, a reference database exists for BV data: the 'European Federation for Laboratory Medicine (EFLM) Biological Variation Database'. The aim is to refer the different studies of BV and to classify them according to the quality of the data generated.[8]

Interpretation of thyroid function tests by means of the BV model could establish the significance of a change in a serial of measurements within an individual and then facilitate clinical decision-making. Two parameters are of particular interest in demonstrating the use of BV: the index of individuality (II) which can help to establish the usefulness of the conventional population-based benchmark intervals[9] and the reference change value (RCV) that can establish the significance of a change in a series of measures within an individual.[10]

Clinically, II and RCV are critical in interpreting the thyroid test results. A low II indicates that a parameter varies widely in the same patient **and that** physicians need to keep this in mind when interpreting tests based on previous results. When a physician is facing a biological parameter with a low II, he should use the RCV or more traditionally the *critical difference* which allows to evaluate if the difference between two consecutive laboratory results is significant. A high RCV indicates that the difference between two results must be important to consider the result as significant and to incorporate it into the differential diagnosis. RCV may be reported as a percentage or as an absolute value with same unit as the analyte measured. RCV reported in **percentage** is constant in the measurement range but by looking at the data in absolute values, large difference depending on the baseline patient value can be visualized. Interpreting patient results based on reference intervals alone may mask clinically significant changes at the individual level given that reference intervals for thyroid hormones are larges.

The objective of our study was to assess the BV of TSH, free thyroxin (FT4), free triiodothyronin (FT3) and thyroglobulin (Tg) with latest generation immunoassay in a population of 19 healthy volunteers without known thyroid conditions.

2 | MATERIAL AND METHODS

2.1 | Subjects

Nineteen European healthy volunteers (HV), among which 8 males (25-59 years, median age: 33.5 years ± 10.8 years) and 11 females (22-60 years, median age: 39 years ± 11.4 years) were included in this study. All participants signed an informed consent form and the study protocol was approved by our institution's Ethics Committee (Cliniques Universitaires Saint-Luc, Brussels, Belgium; 2019/04SEP/388). For the recruitment of healthy subjects, inclusion and exclusion criteria were formulated according to Carobene et al.[11] Participants were required to be healthy, over 18 years of age, not suffering from diabetes, not on anticoagulant nor antiaggregant medication and not suffering from serious internal pathology. None of them were smokers or had substantial (>10g/day) alcohol consumption. None had any known thyroid disease and none had a visible goitre. Four females were on hormonal contraceptive treatment. During the study, all participants maintained their usual lifestyle and every drug/substance taken was inventoried. A short questionnaire was therefore completed prior to each blood draw to highlight drugs taken, exercise and health status.

2.2 | Sample sampling

Blood samples were collected by the same trained phlebotomist under standardized conditions to minimize pre-analytical variation. Samples were collected either on Wednesday or Thursday during the morning between 9 and 11 AM. The collection of samples followed recent guidelines for the collection of blood by antecubital venipuncture.[12] A discard tube was systematically drawn before collection.

Samples were collected in lithium heparinate tubes (S-Monovette 2.6 mL, lithium heparinate 20 IU/mLn Sarstedt, Numbrecht, Germany) and homogenized immediately after

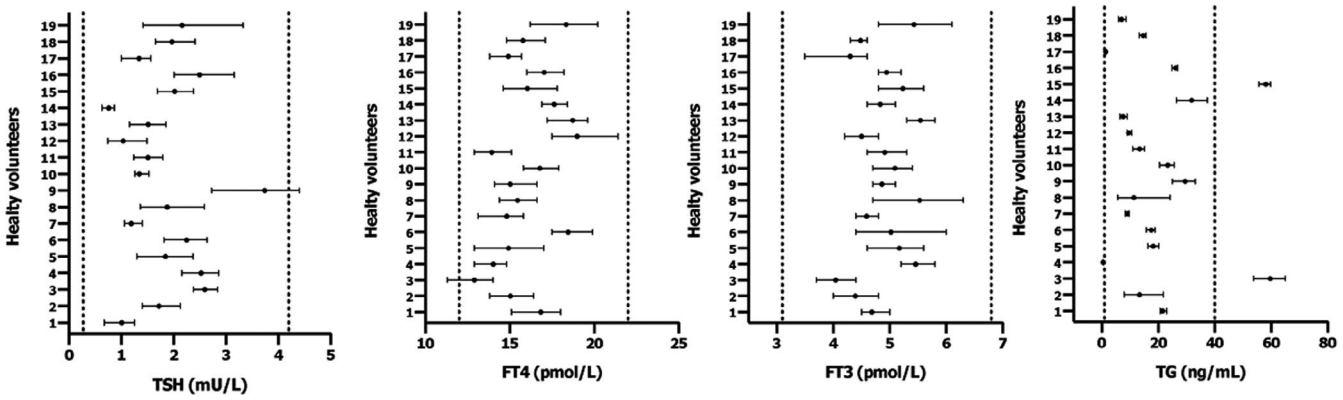

FIGURE 1 Variation of four thyroid biomarkers in healthy volunteers over the five weeks

collection. Tubes were centrifuged for 10 minutes at 2,465 g 30 minutes after venipuncture and immediately aliquoted into cryovials. The aliquots were then frozen and stored at −80°C until use. Samples were thawed at 37°C for 15 minutes before analysis and were analysed within 30 minutes.

2.3 | Laboratory measurement procedure

TSH, FT3, FT4 and Tg were measured on Cobas e602 electrochemiluminescence immunoassay system (Roche Diagnostics, Mannheim, Germany) in the same analytical run, in duplicate and in random order. TSH and Tg were measured using sandwich electrochemiluminescence immunoassay methods (Elecsys TSH and Elecsys Tg II; Roche Diagnostics, Mannheim, Germany), and FT3 and FT4 were measured using competition electrochemiluminescence immunoassay methods (Elecsys FT3 III and Elecsys FT4 III; Roche Diagnostics, Mannheim, Germany), according to the manufacturer's instructions. Internal quality control (PreciControl Universal (Roche Diagnostics, Mannheim, Germany)) was run twice a day for all parameters.

2.4 | Statistical analysis

2.4.1 | Outlier, normality and subgroup analysis

Outliers detection was performed using the statistical XLSTAT software Version 2019.2.2 (Addinsoft, Paris, France). Cochran's C was performed to identify an outlier from replicate measurements of individuals tested and on the distribution of the variance of each individual (S^2_{I+A}) to see if any individual's dispersion was smaller or larger compared to the whole group. Reed criteria were used to benchmark the mean values of each individual. At each step, the outliers were excluded from the analysis.[13] Normality was evaluated using the Shapiro-Wilk **test**. In case of rejection of the hypothesis of normality, we repeated the normality evaluation with another test (Kolmogorov-Smirnov test). Student's *t* test and F test were used for subgroup analysis (males vs females).

2.4.2 | Analysis of variance

Nested analysis of variance (ANOVA) was performed to assess the components of BV using JMP Pro Version 14.3.0 (SAS Institute Inc, Cary, NC, USA).[14] Sources of variability were calculated using a nested model with 3 factors (individuals, sampling occasions and replicates). Within-individual, between-individual and analytical variances were obtained to calculate within-individual coefficient of variation (CV_I), between-individual coefficient of variation (CV_G) and analytical variation (CV_A).

2.4.3 | Analytical performance specifications

Index of individuality (II) was calculated according to the following formula: $(CV_A^2 + CV_I^2)^{1/2}/CV_G$. Reference change value (RCV) was determined using the formula: $2^{1/2} \times Z \times (CV_A^2 + CV_I^2)^{1/2}$, where Z value of 1.96 represents a probability of 95%. Desirable, analytical goals for imprecision and bias for each parameter were obtained from BV components.[15]

3 | RESULTS

All results were situated within our laboratory normal reference ranges except for the Tg value of two HV (TSH: 0.27-4.2 mU/L; FT4: 12.0-22.0 pmol/L; FT3: 3.1-6.8 pmol/L; Tg: 1.0-40.0 ng/mL). Figure 1 shows the variation of four thyroid biomarkers in HV over the five weeks. For three analytes (TSH, FT4 and FT3), no patients have been excluded by outlier analysis with both Cochran's test among observations (derived from duplicate measurements) and S^2_{I+A} values. For Tg, two duplicates were excluded after analysis (HV15: week 2 and HV16: week 3). Finally, Reed's criterion applied to the mean concentration values of all subjects was successfully passed. Shapiro-Wilk test accepted the hypothesis of normality for the distribution of mean concentration values while the intra-subject data distribution normality was accepted for more than 50% of the 19 patients.

There was no significant difference in mean values of the test for any of the analytes between males and females allowing us the derivation of sex-independent CV_I and CV_G values.

3.1 | Within-individual and between-individual variation

The summary of the results is presented in Table 1. CV_I, CV_G and CV_A were obtained for TSH, FT4, FT3 and Tg. The CV_G ranged from 10.8% (CI95%: 8.0%-16.3%) for FT4 to 84.9% (CI95%: 75%-97.7%) for Tg and CV_I from 4.6% (CI95%:3.8%-5.8%) for FT4 to 19.7% (CI95%: 17.0%-23.5%) for TSH. Minimal precision desirable was 9.9% for TSH; 2.3% for FT4; 3.0% for FT3; and 7.7% for Tg and minimal bias desirable was 10.6% for TSH; 2.9% for FT4; 2.6% for FT3; and 21.7% for Tg. Index of individuality (II) was low (0.2-0.7) for all analytes. Table 2 illustrates the usefulness of RCV reported in absolute change at the extremes of the reference interval of TSH, FT4, FT3 and Tg.

4 | DISCUSSION

Robust and high-quality BV data can facilitate clinical decision-making and patient care by identifying significant changes between measurements and by establishing analytical performance specifications goals to ensure the quality of the results. Therefore, the aim of our study was to provide updated and rigorous BV data for thyroid function tests through the establishment of the CV_I, CV_G and the main APS for FT4, FT3, TSH and Tg in 19 European HV.

To our knowledge, no studies have been published on BV regarding thyroid parameters since the 2014 conference issued strict protocols for BV studies.[16] The previous publications on BV in thyroid parameters referenced in EFLM Biological Variation Database included fewer participants than our present study or were focused on a single gender.[17-19] Our study included both genders but no significant difference was observed between the two subgroups, allowing a 'combined' analysis. Furthermore, this is the first BV study performed on the latest generations of immunoassays for TSH, FT4, FT3 and Tg which are widely used today in European laboratories.[20]

Our results were compared to the EFLM Biological Variation Database to allow the comparison with previously established data.[8] On every step of our study, we followed the latest recommendations for acquisition, calculation and report of biological variation data.[11,14,16] A descriptive comparison of our data with the BV Database suggests that CV_I and CV_G are very similar to what was previously reported [18] for all studied parameters except for a fairly higher CV_G for Tg, a narrower CV_G for FT3 (8.6% *vs* 17.6%), and a narrower CV_I for FT4 (4.6% vs 7.4%).[8] Our results also showed an interestingly lower CV_I (6.0%) for FT3 than previously reported by the BV database (7.9%).[8]

The literature on BV for Tg remains quite poor and divergent. In one study CV_G was reported lower in man (25%) and women (35%) [17] whereas much higher CV_G (128%) was observed in a population of 24 women in a another study.[21] Of note, while our study was performed on healthy volunteers, iodine status was not determined. It was indeed shown that Tg is a good biomarker of iodine deficiency[22] and a hypothesis for the large between-subject variability observed in Tg could be explained by disparities in iodine status, which are commonly described in the European population.[23] More precisely, none of the HV had a visible goitre and the more recent study in the Belgian adult population confirm that mild iodine deficiency is endemic in Belgium.[24]

The low II (<0.7) observed in our study for all parameters means that clinicians should not solely base their interpretation of thyroid results on conventional population-based reference intervals but should also take into account the evolution in time of an individual's results.[4] Our findings are consistent with previous observations[5,17,18] showing that changes in thyroid hormones levels for an individual are more adequately identified using the II index rather than laboratory reference population values.

Comparison of serial results can be realized using the RCV reported as percentage or as absolute value.[10-26] The RCV represents an objective tool capable of assessing significant differences in thyroid function tests between serial measurements and allows a better test interpretation of thyroid function monitoring. However, according to the models proposed in the 1st EFLM Strategic Conference Consensus Statement to establish APS, the measurands we tested should be allocated to model 1 (outcome model) which suggests

TABLE 1 Biological variation, individuality index (II) and reference change values (RCV) of TSH, FT4, FT3 and Tg

Analyte[a]	Group (n)	Total No. of results	Mean value (95% CI)	CV_A[b] (%) (95% CI)	CV_I (%) (95% CI)	CV_G (%) (95% CI)	II[c]	RCV[d](%)
TSH	19	190	1.8 (1.7-1.9)	0.9 (0.8-1.1)	19.7 (17.0-23.5)	37.6 (28.0-57.1)	0.5	54.7
FT4	19	190	16.1 (15.8-16.4)	3.6 (3.2-4.2)	4.6 (3.8-5.8)	10.8 (8.0-16.3)	0.4	16.2
FT3	19	190	4.9 (4.8-5)	2.2 (1.9-2.5)	6 (5.1-7.2)	8.6 (6.3-13.3)	0.7	17.7
Tg	19	186	19.5 (17.2-21.8)	0.9 (0.8-1.1)	15.4 (13.3-18.3)	84.9 (75.0-97.7)	0.2	42.8

[a]TSH is in mU/L; FT4 is in pmol/L; FT3 is in pmol/L; Tg is in ng/mL.

[b]CV_A estimates are based on nested ANOVA of duplicate analysis of all study samples. CV_A measuring the dispersion of results around the mean.

[c]Index of individuality (II) = $(CV_A^2 + CV_I^2)^{1/2}/CV_G$

[d]Reference change value (RCV) = $2^{1/2} \times Z \times (CV_A^2 + CV_I^2)^{1/2}$. where Z value of 1.96 represents a probability of 95%. Represent the percentage above which the change between two measures is truly significant.

TABLE 2 Illustration of RCV in absolute change at the extremities of the reference interval of TSH, FT4, FT3 and Tg

	TSH[a]	FT4[a]	FT3[a]	Tg[a]
Lower Reference limit	0,27	12,00	3,10	1,00
Critical difference[b]	0,15	1,94	0,55	0,43
Value + RCV[c]	0,42	13,94	3,65	1,43
Value - RCV[c]	0,12	10,06	2,55	0,57
Upper Reference limit	4,20	22,00	6,80	40,00
Critical difference[b]	2,30	3,56	1,20	17,12
Value + RCV[c]	6,50	25,56	8,00	57,12
Value - RCV[c]	1,90	18,44	5,60	22,88

[a]TSH is in mU/L; FT4 is in pmol/L; FT3 is in pmol/L; Tg is in ng/mL.

[b]Absolute value of the RCV calculated on the Lower Reference limit and the Upper Reference limit.

[c]Absolute change above/below which a consecutive result is significant (95% CI).

using outcome-based APS to consider results, as opposed to model 2 (BV model). Although the model based on reference intervals is the standard for the interpretation of thyroid function tests to characterize a patient as having thyroid disease, we believe that the integration of BV and RCV values offers real added value for clinicians.

It would be key to know if a single measurement of TSH, FT4, FT3 and Tg may reliably be used to assess thyroid disorders or if repeated measurement should be performed over time to take into account the individuals' natural fluctuations. The outcome of our study has important clinical implications. In Table 2, we illustrated the importance of RCV in absolute change applied to the extreme values of the reference interval for each analyte. This is helpful in case of doubt: a value outside the reference interval but below the critical difference compared to the previous result can be interpreted by the clinician as non-pathological with a confidence interval of 95%.

The following illustrates situations where the clinician's knowledge of BV and RCV allows for better patient management.

- Subclinical primary hypothyroidism

Commonly asymptomatic or with unspecific symptoms presentation, subclinical primary hypothyroidism is classically identified in case of increased TSH along with normal FT4 concentration. An early decrease in FT4 concentration (but still within reference interval) may help the clinician to identify a thyroid disorder.

A knowledge of the RCV (TSH: 54.7%; FT4: 16.2%) can help the clinician in the follow-up of an individual patient, to conclude to a significant difference between two values and thus to suspect a subclinical hypothyroidism.

Moreover, TSH could be above the laboratory reference range but close to the intraindividual homeostatic point (CV_I: 19.7%) and could lead to results misinterpretation and concluding to a false subclinical hypothyroidism.

- Monitoring TH therapy in hypothyroidism

Even though the demand for thyroid testing is increasing, there is an under-screening in patients who require regular monitoring.[27]TH replacement is used to treat primary hypothyroidism by restoring normal TSH concentration. For that purpose, repeated thyroid function tests are needed. An interpretation of a TSH elevation that may indicate non-compliance can be balanced by the RCV (54.7%), which could tell us whether the difference between two measurements is significant or not. This finding supports the necessity to personalize TH replacement therapy in hypothyroid patients.

- Predict long-term remission in patients with differentiated thyroid cancer

Differentiated thyroid cancer (DTC) is the most common endocrine cancer and is responsible for 90% of malignant thyroid tumours.[28] During follow-up after the initial treatment, the risk of recurrence of the DTC is regularly reassessed using, along with imaging studies, the Tg level.[29,30] Most patients with thyroid cancer, thyroidectomized or receiving radioactive iodine, will have Tg levels below 1 ng/mL or even indeterminate values of 2 or 3 ng/mL. In this study, we generated data on a healthy population (mean Tg concentration of 19.5 ng/mL). We found that thyroglobulin is highly individualized, but a similar study on patients with thyroid cancer would be required to elucidate the RCV issue in this context.

4.1 | Limitations of our study

Despite the strict adherence to study protocol, our work has some limitations. As the samples were collected in the morning, we were unable to analyse 24-hour life cycle fluctuations. Furthermore, we included a relatively young study population.

5 | CONCLUSION

Our study provides updated data on the BV of four thyroid function tests (TSH, FT4, FT3 and Tg) in European HVs. The integration of updated and reliable BV characteristics for TSH, FT4, FT3 and Tg can facilitate the test interpretation and monitoring of thyroid function at the individual level.

ACKNOWLEDGEMENTS

We thank all the HVs who provided their blood several times to perform this study. We especially thank Mr Helder Lopes Dos Santos for conducting all the blood draws on his working time for five weeks.

DISCLOSURE

The authors have nothing to declare.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the corresponding author upon reasonable request.

ORCID

Antoine Mairesse https://orcid.org/0000-0002-0181-3076
Urszula Luyten https://orcid.org/0000-0003-0878-8798
Maria-Cristina Burlacu https://orcid.org/0000-0002-8437-3933

REFERENCES

1. Vanderpump MPJ. Epidemiology of Thyroid Disorders. In: Luster M, Duntas LH, Wartofsky L, eds. *The Thyroid and Its Diseases: A Comprehensive Guide for the Clinician*. Cham: Springer International Publishing; 2019:75-85.
2. Ladenson PW, Singer PA, Ain KB, et al. American Thyroid Association guidelines for detection of thyroid dysfunction. *Arch Intern Med*. 2000;160(11):1573-1575.
3. Favresse J, Burlacu M-C, Maiter D, Gruson D. Interferences with thyroid function immunoassays: clinical implications and detection algorithm. *Endocr Rev*. 2018;39(5):830-850.
4. Horowitz GL. Reference intervals: practical aspects. *EJIFCC*. 2008;19(2):95-105.
5. Andersen S, Pedersen KM, Bruun NH, Laurberg P. Narrow individual variations in serum T4 and T3 in normal subjects: a clue to the understanding of subclinical thyroid disease. *J Clin Endocrinol Metab*. 2002;87(3):1068-1072.
6. Hollowell JG, Staehling NW, Flanders WD, et al. Serum TSH, T(4), 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(2):489-499.
7. Fraser CG. *Biological variation: from principles to practice*. Amer. Assoc. for Clinical Chemistry. 2001.
8. Aarsand AKF-CP, Webster C, Coskun A, et al. The EFLM Biological Variation Database. https://biologicalvariation.eu/[timeofaccess]
9. Ricós C, Álvarez V, Perich C, et al. Rationale for using data on biological variation. *Clin Chem Lab Med*. 2015;53(6):863-870.
10. Fraser CG. Reference change values. *Clin Chem Lab Med*. 2011;50(5):807-812.
11. Carobene A, Strollo M, Jonker N, et al. Sample collections from healthy volunteers for biological variation estimates' update: a new project undertaken by the Working Group on Biological Variation established by the European Federation of Clinical Chemistry and Laboratory Medicine. *Clin Chem Lab Med*. 2016;54(10):1599-1608.
12. Simundic A-M, Bölenius K, Cadamuro J, et al. Joint EFLM-COLABIOCLI Recommendation for venous blood sampling. *Clin Chem Lab Med*. 2018;56(12):2015-2038.
13. Westgard JO, Barry PL, Hunt MR, Groth T. A multi-rule Shewhart chart for quality control in clinical chemistry. *Clin Chem*. 1981;27(3):493-501.
14. Braga F, Panteghini M. Generation of data on within-subject biological variation in laboratory medicine: an update. *Crit Rev Clin Lab Sci*. 2016;53(5):313-325.
15. Fraser CG, Hyltoft Petersen P, Libeer JC, Ricos C. Proposals for setting generally applicable quality goals solely based on biology. *Ann Clin Biochem*. 1997;34(Pt 1):8-12.
16. Bartlett WA, Braga F, Carobene A, et al. A checklist for critical appraisal of studies of biological variation. *Clin Chem Lab Med*. 2015;53(6):879-885.
17. Feldt-Rasmussen U, Hyltoft Petersen P, Blaabjerg O, Hørder M. Long-term variability in serum thyroglobulin and thyroid related hormones in healthy subjects. *Acta Endocrinol*. 1980;95(3):328-334.
18. Browning MC, Ford RP, Callaghan SJ, Fraser CG. Intra- and interindividual biological variation of five analytes used in assessing thyroid function: implications for necessary standards of performance and the interpretation of results. *Clin Chem*. 1986;32(6):962-966.
19. Ankrah-Tetteh T, 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):167-169.
20. von Eckardstein A, Roth H, Jones G, et al. cobas 8000 modular analyzer series evaluated under routine-like conditions at 14 sites in Australia, Europe, and the United States. *J Lab Automat*. 2013;18.
21. Jensen E, Petersen PH, Blaabjerg O, Hegedüs L. Biological variation of thyroid autoantibodies and thyroglobulin. *Clin Chem Lab Med*. 2007;45(8):1058-1064.
22. Ma ZF, Skeaff SA. Thyroglobulin as a biomarker of iodine deficiency: a review. *Thyroid*. 2014;24(8):1195-1209.
23. Delange F. Iodine deficiency in Europe and its consequences: an update. *Eur J Nucl Med Mol Imaging*. 2002;29(Suppl 2):S404-S416.
24. Moreno-Reyes R, Carpentier YA, Macours P, et al. Seasons but not ethnicity influence urinary iodine concentrations in Belgian adults. *Eur J Nutr*. 2011;50(4):285-290.
25. Ricós C, Alvarez V, Cava F, et al. Integration of data derived from biological variation into the quality management system. *Clin Chim Acta*. 2004;346(1):13-18.
26. Fraser CG. Making better use of differences in serial laboratory results. *Ann Clin Biochem*. 2012;49(Pt 1):1-3. https://doi.org/10.1258/acb.2011.011203
27. Fryer A, Hanna F. Managing demand for pathology tests: financial imperative or duty of care? *Ann Clin Biochem*. 2009;46:435-437.
28. Davies L, Welch HG. Increasing incidence of thyroid cancer in the United States, 1973–2002. *JAMA*. 2006;295(18):2164-2167.
29. Schlumberger MJ. Papillary and follicular thyroid carcinoma. *N Engl J Med*. 1998;338(5):297-306.
30. Haugen BR, Alexander EK, Bible KC, et al. 2015 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*. 2016;26(1):1-133.

How to cite this article: Mairesse A, Wauthier L, Courcelles L, et al. Biological variation and analytical goals of four thyroid function biomarkers in healthy European volunteers. *Clin Endocrinol*. 2021;94:845–850. https://doi.org/10.1111/cen.14356