Heleen I. Jansen, Rob van der Steen, André Brandt, André J. Olthaar, Hubert W. Vesper, Eri Shimizu, Annemieke C. Heijboer, Katleen Van Uytfanghe and Antonius E. van Herwaarden*

Description and validation of an equilibrium dialysis ID-LC-MS/MS candidate reference measurement procedure for free thyroxine in human serum

https://doi.org/10.1515/cclm-2022-1134
Received November 8, 2022; accepted March 15, 2023;
published online March 31, 2023

Abstract

Objectives: Free thyroxine (FT4) in serum is routinely measured in clinical practice to diagnose and monitor thyroid disease. Due to its concentration in picomolar range and the delicate equilibrium of free and protein-bound T4, accurate measurement is challenging. As a consequence, large inter-method differences in FT4 results exists. Optimal method design and standardization of the FT4 measurement is therefore necessary. The IFCC Working Group for Standardization of Thyroid Function Tests proposed a reference system with a conventional reference measurement procedure (cRMP) for FT4 in serum. In this study, we describe our FT4 candidate cRMP and its validation in clinical samples.

Methods: This candidate cRMP is based on equilibrium dialysis (ED) combined with determination of T4 with an isotope-dilution liquid chromatography tandem mass-spectrometry (ID-LC-MS/MS) procedure and was developed according to the endorsed conventions. Its accuracy, reliability, and comparability was investigated using human sera.

Results: It was shown that the candidate cRMP adhered to the conventions and its accuracy, precision, and robustness were adequate in serum of healthy volunteers.

Conclusions: Our candidate cRMP measures FT4 accurately and performs well in serum matrix.

Keywords: conventional reference measurement procedure; free thyroxine; LC-MS/MS; standardization

***Corresponding author: Dr. Antonius E. van Herwaarden**, Department of Laboratory Medicine, Radboud University Medical Center, Geert Grooteplein Zuid 10, 6525 GA, Nijmegen, The Netherlands, E-mail: teun.vanherwaarden@radboudumc.nl

Heleen I. Jansen, Department of Clinical Chemistry, Endocrine Laboratory, Amsterdam UMC Location Vrije Universiteit Amsterdam, Amsterdam, The Netherlands; Department of Clinical Chemistry, Endocrine Laboratory, Amsterdam UMC Location University of Amsterdam, Amsterdam, The Netherlands; and Amsterdam Gastroenterology Endocrinology & Metabolism, Amsterdam, The Netherlands. https://orcid.org/0000-0003-2359-2865

Rob van der Steen, André Brandt and André J. Olthaar, Department of Laboratory Medicine, Radboud University Medical Center, Nijmegen, The Netherlands

Hubert W. Vesper, Division of Laboratory Sciences, Centers for Disease Control and Prevention, Atlanta, GA, USA

Eri Shimizu, Reference Material Institute for Clinical Chemistry Standards, Yokohama, Kanagawa, Japan

Annemieke C. Heijboer, Department of Clinical Chemistry, Endocrine Laboratory, Amsterdam UMC Location Vrije Universiteit Amsterdam, Amsterdam, The Netherlands; Department of Clinical Chemistry, Endocrine Laboratory, Amsterdam UMC Location University of Amsterdam, Amsterdam, The Netherlands; Amsterdam Gastroenterology Endocrinology & Metabolism, Amsterdam, The Netherlands; and Amsterdam Reproduction & Development Research Institute, Amsterdam, The Netherlands

Katleen Van Uytfanghe, Department of Bioanalysis, Faculty of Pharmaceutical Sciences, Laboratory of Toxicology, Ref4U, Ghent University, Ghent, Belgium

Introduction

Thyroid hormones are important for the regulation of many processes regarding development, growth and metabolism in the human body [1, 2]. The measurement of serum thyroid stimulating hormone (TSH) and free thyroxine (FT4) is critical for diagnosis and monitoring treatment of thyroid diseases. In human serum T4 is for 99.98% bound to thyroxine binding globulin (TBG), albumin and transthyretin, leaving approximately 0.02% FT4 in the circulation. This balance is dependent upon the concentration and binding affinity between T4 and its binding proteins. According to the free hormone hypothesis, it is the unbound, free fraction of thyroid hormone that is able to interact with its receptor and exerts biological activity [3]. As T4 binding protein concentrations are also affected by causes independent of thyroid disease free and total hormone concentrations may not be concordant and measurement of FT4 is preferred over total T4 to assess thyroid status. To

achieve accurate quantitation of serum FT4, its measurement must be performed while minimally disturbing protein binding of T4 *in vitro* to correctly determine FT4 concentrations *in vivo*.

Depending on its design, FT4 measurement methods may variably cause disturbances in protein binding to T4 and thereby cause shifts in FT4 concentration. The disturbance may be counterbalanced by additives in the reagents. This may result in acceptable FT4 concentrations measured in serum samples of healthy controls but may cause larger discrepancies in samples with aberrant protein constitution, such as in pregnancy, familial dysalbuminemic hyperthyroxinemia or low albumin concentration in non-tyroidal illness [4]. The FT4 measurement by immunoassays (IAs) is therefore challenging and not as accurate or reliable as desired [5–8].

FT4 is usually measured using automated IAs in clinical laboratories but these methods demonstrate large between-method differences in FT4 quantitation. De Grande et al. [7] showed that most tested IAs had considerable negative biases in a range of −30 to −73% compared to the conventional reference measurement procedure (cRMP). This is consistent with a study including 3,900 clinical laboratories that reported a range of −40 to −97% [6]. These between-method biases complicate the use of common clinical decision points and hamper follow-up on thyroid status in individual patients when different FT4 IAs are used over time. To allow accuracy, reliability, and comparability between FT4 measurement methods the International Federation of Clinical Chemistry and Laboratory Medicine (IFCC) committee for Standardization of Thyroid Function Tests (C-STFT) advocated a reference system with a cRMP for FT4 in serum based on equilibrium dialysis (ED) combined with determination of T4 in dialysate with a trueness-based RMP [9–13]. FT4 cRMP measurement capacity allows the implementation of standardization and comparative studies in patient cohorts with various clinical serum matrix constitutions.

In this study, we described a candidate cRMP for FT4 in serum based on ED according to the defined conventions combined with determination of T4 in the dialysate by an isotope-dilution liquid chromatography tandem mass-spectrometry (ID-LC-MS/MS) procedure using certified primary reference material as calibrator [13]. Furthermore, we investigated its accuracy, reliability and comparability using human sera.

Materials and methods

The candidate cRMP consisted of ED followed by a liquid-liquid extraction (LLE) and subsequent 2D-ID-LC-MS/MS measurement of T4 and is compliant with ISO15193 [13]. All volumetric steps during the cRMP with standards, reagents, controls, and samples were gravimetrically performed and concentrations expressed as *mass* fractions. The cRMP operated at Radboudumc strictly adhered to the conventions for equilibrium dialysis as endorsed by the IFCC [12, 13]. FT4 concentrations were displayed and calculated in pmol/L by dividing T4 in the sample (pg) by the sample dialysate weight (g) and then divided by its molecular weight 776.87 g/mol and multiplied by 1.0058 g/L density to convert to pmol/L.

Calibration

For single-point calibration [14], Thyroxine (3, 3′, 5, 5′-tetraiodo-L-thyronine; T4) certified reference material IRMM-468 (VWR International, NL) with a purity of 98.6 ± 0.70% (expanded uncertainty, k=2) was used [15] and concentrations of calibration solutions were adjusted for impurity. As internal standard, [$^{13}C_6$]-T4 (ISO-Sciences; King of Prussia; art. nr. 5031) was used.

Three independently prepared standard stock solutions were diluted to a working solution. All volumetric steps were gravimetrically performed (Sartorius SE2-OCE). To prepare T4 standard stock solution, 15 mg T4 standard was weighted, transferred to a glass vial (Perkin Elmer) and dissolved in methanol + 0.125% (v/v) NH_4OH (Biosolve; Acros) with sonification and stored overnight at −30 °C. Intermediate T4 stock solutions A (80 µg/g), B (2.5 µg/g), C (60 ng/g) and working solution (1.1 ng/g) were prepared in methanol. [$^{13}C_6$]-T4 internal standard stock solution (25 µg/g) was used to prepare solution A (2.63 µg/g), solution B (8.0 ng/g) and working solution (0.2 ng/g) in methanol.

A 3-iodo-L-tyrosine (MIT; Sigma, art. nr. I8250) solution was used to prevent adsorption of T4 to vials and added to T4 standard working solutions, internal standard solution B and working solution at a concentration of 5,000 times higher than the T4 concentration.

Equilibrium dialysis

ED was performed by the defined conventions as described [13]. They required biochemical composition of dialysis buffer resembling the ionic environment of serum; samples to be buffered to a pH of 7.4 ± 0.03 (at 37 °C) during dialysis; thermostatic control to be maintained during dialysis at 37.0 ± 0.5 °C; the use of an identical volume of dialysand/dialysate compartments and these compartment should be separated by a membrane of regenerated cellulose and adequate cut-off [13].

Frozen samples were thawed at room temperature and 1 mL serum was buffered to pH of 7.4 ± 0.03 at 37 °C by adding 1/10 (v/v) concentrated HEPES (0.776 M (N-(2-hydroxyethyl) piperapiperazine-N9-(2-ethylsulfonic acid)) (VWR International, NL) and 0.22 mol/L NaOH (sodium hydroxide; Merck, art. nr. 1.06498)). Samples were dialyzed against HEPES buffer (52.74 nM HEPES, 22.5 mmol/L NaOH, 91.6 mmol/L NaCl (Merck), 1.65 mmol/L KH_2PO_4, 2.68 KCl (Merck), 1.12 mmol/L $MgSO_4.7H_2O$ (Merck), 5.0 mmol/L urea (Merck), 1.90 mmol/L $CaCl_2.2H_2O$ (Merck), 8 mmol/L NaN_3 (Merck)) set at pH 7.4 ± 0.03 at 37 °C with 10 M NaOH and when necessary adjusted just before use.

Regenerated cellulose dialysis membranes (Dianorm, ref 10.14, Harvard Apparatus, art. nr. 74–2100) with a diameter of 63 mm and 5 kDa cut-off value were pretreated with deionized water and dialysis buffer at 37 °C. 1.0 mL dialysis cells (PTFE) were used and consisted of two identical halves between which the dialysis membrane was placed. 1.0 mL pH-adjusted serum was injected in one half of the dialysis cell and

an equal amount of dialysis buffer was injected in the other half. Assembled dialysis cells were then placed in the Dianorm® Equilibrium dialyser in a water bath at 37 ± 0.5 °C for 4 h and continually rotated. An appropriate amount of [$^{13}C_6$]-T4 was added to a borosilicate vial. The required amount of [$^{13}C_6$]-T4 was estimated using an IA (FT4 on a E801 random access analyzer, Roche diagnostics) or based on the previous single measurement result with the cRMP to achieve ratios of T4/[$^{13}C_6$]-T4 close to 1:1. After dialysis, dialysate was emptied from the cells, added to the borosilicate vial with [$^{13}C_6$]-T4 and equilibrated for at least 1 h before proceeding with LLE [16]. Whereas the JCTLM listed RMP uses a solid phase extraction step to purify the dialysate for LC-MS/MS measurement of T4 [13], the cRMP described here uses a liquid-liquid extraction (LLE) step [16]. In brief, 65 µL of 35% formic acid solution (CHO_2H; Biosolve; cat. nr. 000 69 141 A8) was added to the sample followed by three wash steps with 900 µL C_6H_{12} (cyclohexane; Merck; cat. nr. 1.02817.1000). After an additional 35 µL of 35% formic acid (v/v) solution T4 and [$^{13}C_6$]-T4 were extracted twice in 1 mL $CH_3CO_2C_2H_5$ (ethyl acetate; VWR International, NL; cat. nr. 83621.290). After evaporation on a nitrogen flow at 40 °C (Techne Dryblock DB-3D) extracts were reconstituted in 1 mL of ethyl acetate and dried twice to reduce residual formic acid, subsequently reconstituted into 80 µL reconstitution solution (25% acetonitrile (CH_3CN), 0.02% (v/v) formic acid) and transferred to an injection vial. Every sample was processed and measured at least in duplicate in independent experiments.

Isotope-dilution liquid chromatography tandem mass-spectrometry (ID-LC-MS/MS)

ID-LC-MS/MS analyses of T4 were performed using a Waters Acquity 2D UPLC system coupled to a Waters Xevo TQ-XS mass spectrometer (Milford, MA, USA). Mobile phases A and B consisted of 0.02% formic acid in water and 0.02% formic acid in acetonitrile (Biosolve; cat. nr. 000 1204 101), respectively. Gradient program and mass spectrometer settings were described in Table 1. Acquisition and data processing were performed using Waters MassLynx Software. Calculations were made in Excel (Supplementary Table 1).

Table 1: FT4 LC-MS/MS settings.

First-dimension column: BEH protein C4 (2.1 × 50 mm, 1.7 µm) 22 °C				Second-dimension column: BEH C18 (2.1 × 50 mm, 1.7 µm) 22 °C			
Time, min	Flow-rate, mL/min	A%	B%	Time, min	Flow-rate, mL/min	A%	B%
0.0	0.250	75.0	25.0	0.0	0.250	70.0	30.0
7.0	0.250	69.0	31.0	6.5	0.250	70.0	30.0
8.0	0.250	30.0	70.0	7.0	0.250	65.0	35.0
10.0	0.250	30.0	70.0	10.0	0.250	65.0	35.0
11.0	0.250	75.0	25.0	11.0	0.250	0.0	100
15.0	0.250	75.0	20.0	13.0	0.250	0.0	100
				14.0	0.250	70.0	30.0
Columns in series from 4 to 7 min				15.0	0.250	70.0	30.0
Retention time: 8.65 min				Total run time: 15 min			

MS-settings ionization mode	Positive electrospray
Capillary, kV	3
Cone, V	40
Source temperate, °C	150
Desolvation temperature, °C	600
Cone gas flow, L/h	150
Desolvation gas flow, L/h	1,100
Collision gas flow, mL/min	0.15 (Argon)

	Dwell time, ms	Mass transitions Q1>Q3	Collision energy, eV
T4 qual	500	777.8>731.82	25 eV
T4 quant	500	777.8>633.8	25 eV
[$^{13}C_6$]-T4 qual	500	783.8>737.8	25 eV
[$^{13}C_6$]-T4 quant	500	783.8>639.6	25 eV

Evaluation of dialysis conventions of accurate pH, temperature and buffer composition

HEPES buffers were prepared as specified within maximum allowed deviation of 1.5% (w/w) (HEPES, NaCl, KH_2PO_4, KCl, urea, $CaCl_2.2H_2O$, NaN_3), 5% (w/w) NaOH or 1% (w/w) $MgSO_4.7H_2O$ and pH was set at 7.4 ± 0.03 at 37 °C. pH stability during ED in dialysis buffer and serum was tested (n=4) at 15 min and after 4 h of dialysis (pH-meter; Metrohm 744, Applikon) at 37 °C. Continuous temperature measurements were performed using a temperature logger (Tinytag Plus 2 TGP-4020; Gemini data loggers) and probe (PB5001-1M5 10K NTC). Accuracy of the temperature reading was verified with a second calibrated thermometer (Fluke 53 II, FLUKE USA).

Specificity and sensitivity

Potential interferences on T4 quantitation were evaluated in each individual serum sample and compared to (matrix free) calibrators by assessment of peak shape, ratio of T4 qualifier to quantifier of mass transitions (Q1/Q2) of both T4 and [$^{13}C_6$]-T4, peak width at half its height (W1/2) and retention times (RT). An accurate T4 RT compared to the RT of the T4 calibrator and compared to the sample [$^{13}C_6$]-T4 RT, a normal Gaussian peak shape and less than 5% deviation of Q1/Q2 compared to the mean calibrator Q1/Q2 was used to confirm absence of interferences in serum samples. The limit of detection (LOD) was calculated in three independent experiments by measuring blank dialysis buffer and dialysis buffer to which 1.31 pmol/L T4 was added, each in triplicate (n=9). LOD was calculated as follows: average T4 quantification in the blank + 3* standard deviation (SD) of the 1.31 pmol/L samples. The lower limit of quantitation (LLOQ) was determined by serial measurement of a serum pool, composed of sera from patients with severe hypothyroidism, at a concentration with an expected 20% CV. Ion suppression was assessed by continuous infusion of labeled standard. The abundance was compared between a serum dialyzed matrix and mobile phase at the RT of T4.

Linearity

Linearity of the measurement range was evaluated following principles as described in CLSI document EP6-A [17]. In short, a dialyzed serum pool of patients with severe hypothyroidism and hyperthyroidism were

used to prepare a series of dialysates with varying concentrations of T4 with known dilution ratios relative to another. T4 in these samples was measured after LLE by ID-LC-MS/MS in quadruplets and a-linearity was assessed [18].

Evaluation of accuracy and precision

Accuracy of the LLE–ID-LC-MS/MS procedure was evaluated using serum-based T4 certified reference material (CRM) (Certified Reference control sample, Thyroxine 103.7 nmol/L RfB DGKL, Germany, D-K-15117-01-00; 40122) diluted in dialysis buffer with MIT. Precision was evaluated by analyzing serum pools with levels at the lower and upper reference limits and at clinical levels of strong hypo- and hyperthyroidism. Controls were aliquoted and kept at −80 °C until analysis. FT4 was measured in duplicate and for up to 20 separate experiments over 15 months. Imprecision was expressed as % coefficient of variance (CV), where both within-run CV (CV_{wr}) and total CV (CV_T) were calculated based on CLSI document EP5.

Accuracy of the cRMP was evaluated by an interlaboratory comparison with the IFCC endorsed FT4 cRMP at Ref4U, Laboratory of Toxicology, Faculty of Pharmaceutical Sciences, Ghent University, Ghent, Belgium; (JCTLM DB identification number: C8RMP1) [13], at the National Center for Environmental Health, Centers for Disease Control and Prevention, Atlanta, GA, USA (CDC) and at the Reference Material Institute for Clinical Chemistry Standards, Yokohama, Kanagawa, Japan (ReCCS). For this comparison, serum samples in euthyroid state (n=20) were collected (Solomon Park Research Laboratories, Kirkland WA) with informed consent, serum was centrifuged and aliquoted and distributed to participating centers. Samples were analyzed in 3–6 independent sample preparations using one replicate per run and a Deming regression analysis was performed to compare our FT4 cRMP with the mean of four laboratories and a bias was calculated (Bland-Altman plot). In addition, imprecision of these samples was calculated.

Measurement uncertainty was assessed at three levels. Type A uncertainty was obtained from imprecision of repeated measurements in serum pools; Type B uncertainty was estimated from uncertainties in purity of primary reference material, inaccuracy in weighing of standards, measurement of serum dialysate density, sample-related effect and unattributed variances.

Robustness of the FT4 candidate cRMP

To demonstrate robustness of the cRMP an experiment in which serum was diluted with HEPES buffer was performed. Four sera were serially diluted up to 8-fold (excluding the 1:1 dilution in the dialysis procedure) and FT4 was measured. Results were calculated as percentages compared to undiluted sera.

Results

Several experiments were conducted to verify performance of the developed candidate cRMP according to predefined conventions and were described below [13]. Concentrated HEPES buffer and dialysis buffer were prepared according to convention and within the allowed deviation (data not shown) [13]. pH in serum and dialysis buffer was measured after 15 min of dialysis (mean pH serum 7.39 ± 0.00 (SD) at 37 °C, n=4; mean pH dialysis buffer 7.38 ± 0.00 (SD) at 37 °C, n=4) and after 4 h of dialysis (mean pH serum 7.39 ± 0.01 (SD) at 37 °C, n=5; mean pH dialysis buffer 7.40 ± 0.00 (SD) at 37 °C, n=5). pH was stable during dialysis in both serum and dialysis buffer conform required convention of pH 7.40 ± 0.03. Measurement accuracy of the temperature logger during dialysis was within temperature convention of 37 ± 0.5 °C; as a representative sample of five dialysis series mean temperature was 36.92 ± 0.06 °C (SD). Samples from five different pools were submitted to 4 and 6 h of dialysis time (n=3) and all five pools showed no significant difference between these two settings (paired samples t-test; p=0.108, 0.438, 0.940, 0.503, 0.214) in correspondence with experiments of van Houcke et al. [13]; therefore, dialysis time of 4 h was maintained.

Specificity and sensitivity

Sample and (matrix free) calibrator peak characteristics for T4 and [$^{13}C_6$]-T4 from four independent experiments were summarized in Table 2. No interferences based on peak characteristics were observed in samples. Ion suppression up to 30% was found (data not shown). The use of an carbon-13 isotope labeled T4 internal standard ([$^{13}C_6$]-T4) resulted in an identical chromatographic behavior compared to the unlabeled standard. Both the use of [$^{13}C_6$]-T4 and an adjusted ratio of T4/IS at 1:1 minimizes potential inaccuracy due to ion suppression.

Linearity was evaluated in serum dialysates over a range of 1.0–113 pmol/L and a non-significant a-linearity (p=0.54) was observed. Furthermore, limit of detection (LOD) and limit of quantification (LOQ) were assessed and estimated at 0.39 and 1.39 pmol/L (CV 23.9%), respectively.

Table 2: Mean T4 and [$^{13}C_6$]-T4 peak characteristics (standard deviation), n=4 experiments.

	Compound	Rt, min	W 1/2	Q1/Q2
Calibrator	T4	8.96 (0.013)	15.15 (0.054)	0.141 (0.001)
	[$^{13}C_6$]-T4	8.96 (0.005)	15.15 (0.064)	0.118 (0.001)
Serum	T4	8.98 (0.018)	14.50 (0.113)	0.140 (0.002)
	[$^{13}C_6$]-T4	8.98 (0.012)	14.47 (0.116)	0.119 (0.002)

Rt, retention time; W, peak width at half its height; Q, quantifier of mass transitions.

Table 3: Imprecision of serum pools and T4 CRM.

Serum pool			
FT4 concentration	**n (duplicates)**	**CV_{wr}**	**CV_T**
13.1 pmol/L	20	3.7%	3.9%
32.0 pmol/L	20	3.1%	3.1%
113 pmol/L	12	1.0%	3.3%
T4 CRM			
FT4 concentration	**n**	**CV_{wr}**	**CV_T**
9.7 pmol/L	20	1.8%	2.9%
38.8 pmol/L	20	0.8%	2.6%

CV_{wr}, within-run CV; CV_T, total CV.

Accuracy and imprecision

Table 3 shows the imprecision of serum pools and T4 CRM. The mean CV of the individual 20 serum samples in the euthyroid range used for the method comparison was 4.1%. The relative Measurement Uncertainty was assessed at three levels. Type B uncertainty was minimized by using primary IRMM-468 standards and using gravimetric measurements. For the recommended measurement protocol of n=3 (singlicate measurement of a sample on three independent occasions), the relative expanded uncertainty was estimated at ≤7.6% (Supplementary Table 2), which is comparable to the JCTLM listed RMP [13].

Accuracy of the FT4 candidate cRMP

The accuracy of the candidate cRMP was evaluated with a method comparison of serum samples from 20 volunteers and measured by the cRMP described here (laboratory at Radboudumc) as well as the other participating centers. The Radboudumc cRMP has a slope of 1.07 (1.02–1.13 95% CI), an intercept of −0.835 (−1.9 to 0.3 95% CI), and a correlation coefficient (r^2) of 0.98 (Figure 1A) with an average bias of 2.9% (1.35–4.39 95% CI, p=0.0009) compared to the overall mean (Figure 1B). Results from all participating laboratories (UGhent, CDC, ReCCS, Radboudumc) compared to the overall mean are displayed in Figure 2.

Robustness of the FT4 candidate cRMP

Robustness of the FT4 candidate cRMP in sera with different thyroid hormone binding capacities was investigated by measuring serum FT4 in a dilution series. Based upon concentrations of T4 and binding proteins and their dissociations constants, FT4 in euthyroid sera remains constant upon dilution with an inert buffer [19]. A dilution of sera up to 16-fold in sera with FT4 concentrations in the range of 16.5–35.7 pmol/L displayed an average deviation of 1.9% demonstrating its robustness in samples with low T4 binding protein concentrations.

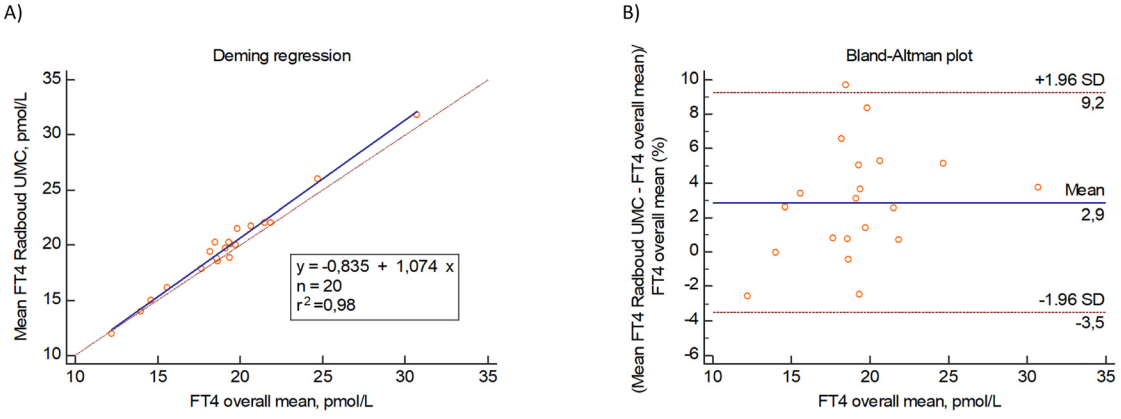

Figure 1: Accuracy of the FT4 candidate cRMP. (A) Deming regression analysis; on the y-axis the FT4 concentration measured at the Radboudumc and on the x-axis the mean FT4 concentration measured at the University of Ghent, CDC ReCCS and Radboudumc (overall mean). (B) Bland-Altman plot; on the y-axis the relative difference (mean FT4 Radboudumc- FT4 overall mean/FT4 overall mean; %) and on the x-axis the mean FT4 concentration measured at Ghent University, CDC, ReCCS and Radboudumc (overall mean).

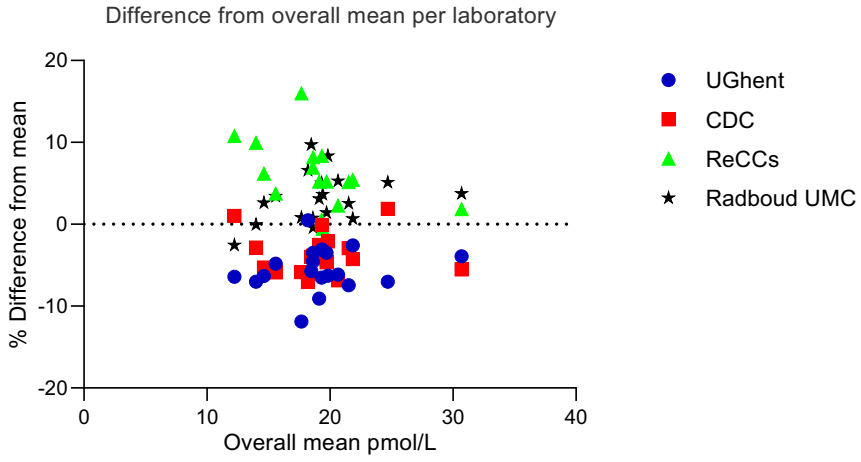

Figure 2: FT4 scatterplot of all reference network laboratories. Scatter plot representing the bias of the reference network laboratories (UGhent, CDC, ReCCS, Radboudumc) compared to the overall mean of four participating laboratories. On the x-axis the overall mean was shown and on the y-axis the % difference per laboratory from the overall mean was depicted.

Discussion and conclusions

In our study, we presented the development and validation of the cRMP for accurate and reliable measurement of FT4 in serum. Importantly, we validated the measurement method with human serum samples assuring clinical applicability.

Validation experiments were conducted in the clinical relevant range far exceeding the reference values of 13.5–24.3 pmol/L [7]. Linearity analysis showed that the measurement range of our candidate cRMP is compatible with the desirable range of 25% of the lower euthyroid limit to 400% of the upper limit of the euthyroid range [19]. In addition, our FT4 candidate cRMP showed results strongly comparable with the other laboratories in the reference network that consists of one established and three candidate RMPs. Its high accuracy is possible due to the use of certified primary reference material IRMM-468 as calibrators in line with the ISO 17511:2020 standard [20] and the use of gravimetric measurements. Our imprecision is well within the limits defined previously for desired imprecision (<5%) of the FT4 cRMP, but our bias (2.9%) is slightly higher than the desired bias (<2.5%). We note that the results from the reference measurement procedures at UGhent and CDC are systematically lower than those from ReCCS and Radboudumc. This systematic deviation may arise most probably from differences in preparation procedures of the T4 standard solutions. Pre-analytical or systematic differences due to variation in the performance of the conventions within their allowed band-width are suspected to play less of a role. The laboratories that operate the measurement procedure collaborate to elucidate causes of systematic deviation between the FT4 candidate cRMPs and further improve the current state of art [21]. The latter should also result in a reduction of the relative expanded measurement uncertainty which will further close the gap between desirable and achievable quality specifications for bias, precision, and measurement uncertainty.

Accurate measurement of FT4 is technically complicated due to its low concentration in picomolar range and the delicate equilibrium of free and protein-bound T4. FT4 concentrations in serum remained constant upon high dilution (without correction for the dilution) in our cRMP, as expected because of the high binding affinity of T4 to binding proteins. This demonstrates robustness of our cRMP design.

In conclusion, the FT4 candidate cRMP can be reliably used as a reference method for the measurement of FT4 and performs well in complex serum matrix. The cRMP is therefore suitable for standardization efforts in serum matrices and allows investigational studies in specific patient cohorts.

Research funding: This research project was funded by the Dutch Foundation of Quality Assessment in Medical Laboratories (SKML) and Landelijke Werkgroep Bindingsanalyse (LWBA). The SKML and LWBA played no role in the study design; in the collection, analysis, and interpretation of data; in the writing of the report; or in the decision to submit the report for publication.

Author contributions: All authors have accepted responsibility for the entire content of this manuscript and approved its submission. The findings and conclusions in this report are those of the author(s) and do not necessarily represent the official position of the Centers for Disease Control and Prevention/the Agency for Toxic Substances and Disease Registry. Use of trade names is for identification only and does not imply endorsement by the Centers for Disease

Control and Prevention, the Public Health Service, and the US Department of Health and Human Services.

Competing interests: Authors state no conflict of interest. Certain commercial equipment, instruments, materials, and companies are identified in this report to adequately specify the experimental procedure. Such identification does not imply recommendation or endorsement, nor does it imply that the equipment, instruments, materials, and companies identified are the best available for the purpose.

Informed consent: Not applicable.

Ethical approval: The local Institutional Review Board deemed the study exempt from review.

References

1. Mullur R, Liu YY, Brent GA. Thyroid hormone regulation of metabolism. Physiol Rev 2014;94:355–82.
2. Brent GA. Mechanisms of thyroid hormone action. J Clin Invest 2012; 122:3035–43.
3. Ekins R. The free hormone hypothesis and measurement of free hormones. Clin Chem 1992;38:1289–93.
4. Docter R, van Toor H, Krenning EP, de Jong M, Hennemann G. Free thyroxine assessed with three assays in sera of patients with nonthyroidal illness and of subjects with abnormal concentrations of thyroxine-binding proteins. Clin Chem 1993;39:1668–74.
5. Sapin R, d'Herbomez M. Free thyroxine measured by equilibrium dialysis and nine immunoassays in sera with various serum thyroxine-binding capacities. Clin Chem 2003;49:1531–5.
6. 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 3,900 clinical laboratories. Arch Pathol Lab Med 2005;129:310–7.
7. De Grande LAC, Van Uytfanghe K, Reynders D, Das B, Faix JD, MacKenzie F, et al. Standardization of free thyroxine measurements allows the adoption of a more uniform reference interval. Clin Chem 2017;63:1642–52.
8. Lee RH, Spencer CA, Mestman JH, Miller EA, Petrovic I, Braverman LE, et al. Free T4 immunoassays are flawed during pregnancy. Am J Obstet Gynecol 2009;200:260.e1–e6.
9. 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;56:912–20.
10. 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;3:109–16.
11. Tractenberg RE, Jonklaas J, Soldin SJ. Agreement of immunoassay and tandem mass spectrometry in the analysis of cortisol and free T4: interpretation and implications for clinicians. Int J Anal Chem 2010; 2010:234808.
12. Thienpont LM, Beastall G, Christofides ND, Faix JD, Ieiri T, Jarrige V, et al. Proposal of a candidate international conventional reference measurement procedure for free thyroxine in serum. Clin Chem Lab Med 2007;45:934–6.
13. Van Houcke SK, Van Uytfanghe K, Shimizu E, Tani W, Umemoto M, Thienpont LM. IFCC international conventional reference procedure for the measurement of free thyroxine in serum: International Federation of Clinical Chemistry and Laboratory Medicine (IFCC) Working Group for Standardization of Thyroid Function Tests (WG-STFT). Clin Chem Lab Med 2011;49:1275–81.
14. Thienpont LM, Van Nieuwenhove B, Stöckl D, De Leenheer AP. Calibration for isotope dilution mass spectrometry—description of an alternative to the bracketing procedure. J Mass Spectrom 1996;31:1119–25.
15. Toussaint B, Klein C, Wiergowski M. The certification of the mass fraction of thyroxine in a CRM intended for calibration: certified reference material IRMM-468. Luxembourg: Office for Official Publications of the European Communities; 2006.
16. De Brabandere VI, Hou P, Stöckl D, Thienpont LM, De Leenheer AP. Isotope dilution-liquid chromatography/electrospray ionization-tandem mass spectrometry for the determination of serum thyroxine as a potential reference method. Rapid Commun Mass Spectrom 1998; 12:1099–103.
17. McEnroe RJ. CLSI-EP06: evaluation of linearity of quantitative measurement procedures. CLSI guideline EP06, 2nd ed. Clinical and Laboratory Standards Institute; 2020.
18. Ross HA, Sweep FCGJ. An improved procedure for testing for assay linearity. Ann Clin Biochem 2003;40:75–8.
19. CLSI. Clinical Laboratory and Standards Institute. Measurement of free thyroid hormones; approved guideline. Wayne, PA: CLSI Document C45-A; 2004.
20. ISO17511:2020. In vitro diagnostic medical devices—requirements for establishing metrological traceability of values assigned to calibrators, trueness control materials, and human samples. Geneva: International Organization for Standardization; 2020.
21. Van Uytfanghe K. Standardization and harmonization laboratory updates. In: IFCC C-STFT 2020 meeting; 2020.

Supplementary Material: This article contains supplementary material (https://doi.org/10.1515/cclm-2022-1134).