Clinica Chimica Acta 430 (2014) 121–124

Contents lists available at ScienceDirect

Clinica Chimica Acta

journal homepage: www.elsevier.com/locate/clinchim

A pilot study: Subclinical hypothyroidism and free thyroid hormone measurement by immunoassay and mass spectrometry

Verena Gounden [a], Jacqueline Jonklaas [b], Steven J. Soldin [a,b,*]

[a] Department of Laboratory Medicine, NIH Clinical Center, Bethesda, MD, USA
[b] Division of Endocrinology, Georgetown University Medical Center, Washington, DC, USA

ARTICLE INFO

Article history:
Received 22 July 2013
Received in revised form 23 December 2013
Accepted 23 December 2013
Available online 31 December 2013

Keywords:
Mass spectrometry
Free thyroid hormone
Subclinical hypothyroidism

ABSTRACT

Background: The diagnosis of subclinical hypothyroidism is defined as the presence of an elevated thyroid stimulating hormone (TSH) with a normal free thyroxine (FT4) level. The commonly used direct analogue immunoassays for the measurement of FT4 have been shown to have poor performance at the upper and lower limits of the FT4 reference interval.
Purpose: The purpose of this pilot study was to investigate the percentage of individuals classified as having subclinical hypothyroidism with a standard immunoassay, that actually have low free thyroid hormone levels by mass spectrometry measurements.
Design: Outpatient samples with elevated TSH values and normal FT4 concentrations as per standard immunoassay methods were collected. FT4 and free triiodothyronine (FT3) analyses were performed on these samples using liquid chromatography–tandem mass spectrometry (LC–MS/MS).
Results: Sixty five percent ($n = 26$) of patients ($n = 40$) had (LC–MS/MS) FT4 or FT3 or both FT4 and FT3 values below mass spectrometry reference limits.
Conclusions: Our findings indicate that the direct analogue immunoassay method for FT4 measurement results in a significant proportion of patients being misclassified as having subclinical hypothyroidism.
Published by Elsevier B.V.

1. Introduction

Subclinical hypothyroidism is defined as increased serum thyroid stimulating hormone (TSH) levels with normal serum free thyroxine (FT4) concentrations [1]. The prevalence of subclinical hypothyroidism in the population without known thyroid disease has been reported to be 4% to 10% [2,3]. Debate still exists with regards to the clinical importance of and therapy for elevation of serum TSH (in particular elevated levels <10 mIU/l) and the exact upper limit of normal for the serum TSH level that also varies with age [1]. Most clinical laboratories perform TSH and FT4 measurements on immunoassay (IA) platforms [4]. Whilst TSH analyses on immunoassay platforms are considered quite reliable, the validity of FT4 analysis by direct analogue immunoassay has been questioned for many years [5,6]. Significant limitations of the currently used FT4 immunoassays that have been described are the influence of changes in binding protein concentrations and a weak inverse linear log relationship to TSH in hypo- and hyperthyroid individuals [7–11]. These assays appear to perform best in euthyroid individuals but poorly in individuals with thyroid disease [12]. FT4 immunoassays' poor performance at low concentrations may lead to misclassification of patients as having subclinical hypothyroidism, when in fact they have FT4 levels lower than the reference interval when measured by tandem mass spectrometry. Liquid chromatography–tandem mass spectrometry (LC–MS/MS) following ultrafiltration of the sample has been shown to perform better in such circumstances, and in the case of FT4, agrees better with the gold standard equilibrium dialysis assay [10]. Numerous pharmacological agents affect thyroid hormone measurements and may further confuse interpretation of results beyond the problems encountered with immunoassay measurements. Drugs may affect thyroxine binding globulin levels (for example estrogens), thyroid hormone binding (for example carbamazepine), TSH levels (for example glucocorticoids), conversion of T4 to T3 (for example amiodarone) and may even induce thyroid disease (for example lithium) [13–16].

2. Materials and methods

2.1. Sample collection

Outpatient serum samples received at the NIH Clinical Center (NIH-CC) for 4 months in 2013 with elevated TSH values and FT4 concentrations within the laboratory reference interval were selected for inclusion in the study. All samples were collected between 8 and 11 am. Samples were stored at −70 degrees Celsius until MS analysis. Patients' samples were excluded if they had known thyroid disease, were receiving thyroid

* Corresponding author at: Department of Laboratory Medicine, NIH Clinical Center, Building 10, Room 2C-249, Bethesda, MD 20892, USA. Tel.: +1 301 496 3453; fax: +1 301 402 1885.
E-mail address: Steven.soldin@nih.gov (S.J. Soldin).

0009-8981/$ – see front matter. Published by Elsevier B.V.
http://dx.doi.org/10.1016/j.cca.2013.12.034

hormone replacement therapy or other medications that are known to affect FT4 values or cause elevation in TSH values. Samples from patients older than sixty years and those that were positive for anti thyroid peroxidase antibodies were also excluded. The study was approved by the Institutional Review Board of the NIH (Clinical Protocol number 93-CC-0094).

2.2. Immunoassay measurements

Samples were analyzed at NIH-CC Department of Laboratory Medicine. TSH and FT4 samples were processed according to usual laboratory procedures. FT4, reference interval 9.8–19.4 pmol/l (0.76–1.50 ng/dl), was measured by a direct (analogue) immunoassay method on a Dimension Vista (Siemens Healthcare Diagnostics). TSH (reference interval, RI: 0.40–4.00 mIU/l) and free triiodothyronine (FT3, RI: 2.8–6.5 pmol/l or 1.8–4.20 pg/ml) were also measured on the Vista. Anti–thyroid peroxidase antibody (ATPOA) (RI: <35 IU/ml) testing was performed on Immulite XPI 2000 (Siemens Healthcare Diagnostics). Two or three levels of commercially available internal quality control material were analysed at the start of each run

2.3. LC–MS/MS measurement

The FT4, reference interval used was 17.4–30.9 pmol/l (1.35–2.40 ng/dl) and for FT3, reference interval 2.3–9.5 pmol/l (1.5–6.1 pg/ml). Samples were analysed in batches. Analyses were performed as per methods previously published [8,10,17]. Briefly four hundred microliters of human plasma/serum was filtered through a Centrifree YM-30 ultrafiltration device by centrifugation at 37 degrees Celsius, and 450 μl of deuterium labeled internal standard, T4-d_5 in methanol was then added to 150 μl of ultrafiltrate for deproteinization. After vortexing and centrifugation, 500 μl of supernatant was diluted with 400 μl of distilled de-ionized water and a 650 μl aliquot was injected onto a C-18 column. After washing, the switching valve was activated and the analytes were eluted from the column with a water/methanol gradient into the MS/MS system. Quantification by multiple reaction-monitoring (MRM) analysis was performed in the negative mode(ESI−). Three levels of internal quality control were analysed at the beginning and end of each run.

2.4. Statistical analysis

Non-normally distributed data (MS FT4 values) was normalised by log-transformation before analysis and back-transformed for data presentation. We used the Kolmogorov–Smirnov test to test for normality, and we used Pearson's correlation coefficient, Bland–Altman difference plots, and Passing Bablock regression analysis to evaluate the methods. Statistical analysis was performed on Medcalc Version 12 (MedCalc Software). A $pf < 0.05$ was considered statistically significant.

3. Results

Fifty seven samples with increased TSH and immunoassay FT4 within normal reference limits were collected. Following exclusion of patients ≥ 60 y and those positive for ATPOA, a total of 40 samples from patients aged 6 – 59 y were included in the study. TSH values ranged between 4.3 and 8.2 mIU/l. Analysis of variance (ANOVA) revealed no statistically significant difference for values between males and females. See Table 1 for summary of results.

The CVs for the immunoassay methods used were as follows: CV of TSH at a concentration of 5.8 mIU/l was 3–6%; FT4 at 12.4 pmol/l (0.96 ng/dl) was 4–6%; FT3 at 19.2 pmol/l (1.49 pg/ml) was 6.9–7.2%. The CVs for the LC–MS/MS assays were: FT4 4.1–6.6% at concentrations of 8.5 pmol/l(0.66 ng/dl) and 33.8 pmol/l (2.62 ng/dl); FT3 ≤9% at concentrations between 0.23 pmol/l (0.15 pg/ml) and 3.44 pmol/l (2.22 pg/ml).

Sixty-five percent ($n = 26/40$) of patients had mass spectrometry FT4 or FT3 or both FT4 and FT3 values below mass spectrometry reference limits. (Fig. 1 summarises findings). Sixteen of the 26 patients (62%) had only low MS FT4 results. Three of the 26 patients (12%) had only low MS FT3 results and 7 patients (27%) had both low MS FT4 and MS FT3 results. Fifty eight percent ($n = 23/40$) of patients that would be classified as subclinical hypothyroidism as per immunoassay FT4 measurements had LC–MS/MS FT4 values that were below the reference interval. Patients with LC–MS/MS FT4 results below the reference interval had on average values 16% below the lower limit of the LC–MS/MS specific reference interval. The majority had values that were greater than 10% below the lower limit. Thirteen patients ($n = 13/23$) had MS FT4 results that were >10% below the low reference limit. Nine patients (9 of 23) had MS FT4 values >15% below the low reference limit.

Pearson's correlation coefficient ($n = 40$) between IA FT4 and MS FT4 was 0.55 with a 95% confidence interval (CI) of 0.28–0.73 and between IA and MS for FT3 it was 0.30 (95% CI 0.14–0.45). Regression analysis in the population studied (Figs. 2 and 3) for mass spectrometry versus immunoassay showed poor correlations between the two methods for both FT4: Slope 2.85 (95% CI 1.80 to 6.29), intercept −1.40 (95% CI −4.56 to −0.40) and FT3: Slope 1.29 (95% CI 0.66 to 2.09), intercept −2.16 (95% CI −4.72 to −0.12).

There was significant bias between MS and immunoassay values, with FT4 immunoassay values being on average 48% lower than MS values and FT3 immunoassay values on average 36% higher than MS results. Bland–Altman percentage difference plots show poor agreement between IA and MS data (Figs. 4 and 5); the 95% limits of agreement for IA FT4 and MS FT4 are between 13.4% and −82.3%. The

Table 1
Summary of results.

Test	All patients ($n = 40$)	Female ($n = 24$)	Male ($n = 16$)	ANOVA *P* value
Age (years)	Mean: 39.7 (24.5–44.9) SD: 16.2	Mean: 39.3 (32.3–46.4) SD: 16.6	Mean: 40.3 (31.7–48.8) SD: 16.1	$P = 0.32$
TSH (mIU/l)	Mean: 5.8 (95% CI, 5.3–6.3) SD: 1.5	Mean: 6.0 (95% CI, 5.3–6.7) SD: 1.7	Mean: 5.5 (95% CI, 4.8–6.2) SD: 1.2	$P = 0.36$
FT4 (IA) (ng/dl)	Mean: 1.0 (95% CI, 0.9–1.1) SD: 0.2	Mean: 1.1 (95% CI,0.9–1.2) SD: 0.2	Mean: 0.9 (95% CI,0.9–1.0) SD: 0.1	$P = 0.07$
FT3 (IA) (pg/ml)	Mean: 3.0 (95% CI, 2.9–3.2) SD: 0.4	Mean: 3.0 (95% CI, 2.8–3.2) SD: 0.4	Mean: 3.0 (95% CI, 2.8–3.3) SD: 0.4	$P = 0.73$
FT4 (MS) (ng/dl)	Mean: 1.4 (95% CI, 1.3–1.6) SD: 0.6	Mean: 1.5 (95% CI,1.3–1.7) SD: 0.7	Mean: 1.3 (95% CI,1.2–1.5) SD: 0.3	$P = 0.39$
FT3 (MS) (pg/ml)	Mean: 1.9 (95% CI, 1.7–2.1) SD: 0.6	Mean: 2.0 (95% CI,1.8–2.2) SD: 0.6	Mean: 1.8 (95% CI,1.4–2.1) SD: 0.5	$P = 0.26$

[a]Abbreviations: SD (standard deviation); TSH (Thyroid stimulating hormone); FT4 (free thyroxine); FT3 (free tri-iodothyronine); IA (immunoassay); MS (mass spectrometry).