

# Assessment of thyroid function tests and harmonization: opinion on thyroid hormone harmonization

Katherine A. Araque, Joanna Klubo-Gwiezdzinska, Lynnette K. Nieman, Kerry Welsh and Steven J. Soldin 

Received: 17 September 2019; revised manuscript accepted: 2 December 2019.

As noted in a 1973 *Clinical Chemistry* editorial, automation has resulted in precision displacing accuracy as the predominant standard for acceptable laboratory results; the author argues that 'meaningful measurement' must also address accuracy.<sup>1</sup> Because measurement of thyroid stimulating hormone (TSH) in conjunction with total triiodothyronine (TT3), total thyroxine (TT4), free T3 (FT3), and free T4 (FT4) is used for the diagnosis and management of thyroid disorders, it is important for all these assays to achieve optimal accuracy. A College of American Pathologists survey reported significant biases, so that more than 50% of results were unacceptably inaccurate in 13–60% of assays evaluated.<sup>2</sup> This inaccuracy, coupled with dissimilar reference intervals, complicates test interpretation and impacts the creation of clinical guidelines.

The International Federation of Clinical Chemistry (IFCC) Working Committee cofunded a project to harmonize thyroid hormone immunoassays (IAs), with the goal of achieving similar results with each IA method. To do this, clinical serum samples were run in multiple assays. All IAs were recalibrated to the statistically inferred targets with a robust mass spectrometry reference method.<sup>3</sup> Here, we describe some of the current limitations of thyroid hormone measurements by IAs that cannot be resolved by assay harmonization.

## Limitations of IAs

The reliability of TSH as an index of thyroid function has been questioned due to findings of significant intra-individual effects, which include aging, pregnancy, comorbidities leading to euthyroid sick syndrome, drug-related effects on hypothalamus-pituitary-thyroid axis, and diurnal and seasonal variations.<sup>4,5</sup> An investigation of 102

healthy volunteers using paired samples found that the mean afternoon (p.m.) TSH concentration is 10% higher for males (reference interval: 6–8 a.m.: 0.7–3.7 uIU/ml; 6–8 p.m.: 0.7–4.7 uIU/ml) and 17% higher for females (reference interval: 6–8 a.m. 0.5–4.3 uIU/ml; 6–8 p.m. 0.5–6.1 uIU/ml) compared with the mean morning (a.m.) concentration.<sup>5</sup>

TSH reference intervals should be age adjusted, particularly in individuals over 70 years old, to avoid diagnostic misclassification. The TEARS study revealed that the median TSH values in 153,127 adult participants without autoimmunity increased significantly with age, from 1.58 mU/liter at 31–40 years old to 1.86 mU/liter at >90 years old ( $p < 0.001$ ). In addition, the 2.5th percentile decreased with age and the 97.5th percentile increased with age.<sup>6</sup>

Medications can cause clinically significant variations in TSH levels (Table 1). Thus, the utility of a single reference interval for TSH as a standalone marker for hyper- and hypothyroidism is questionable.<sup>5</sup>

The diagnostic strategy of measuring both free thyroid hormone and TSH requires use of assays that correlate with the clinical presentation. Most immunoassays of free thyroid hormone are one-step direct analog IAs that are impacted by protein-binding variations. As free thyroid hormone concentrations measured by IA depend on serum binding proteins,<sup>2</sup> thyroxine binding globulin (TBG) deficiency or excess, abnormal transthyretin or albumin levels, and pregnancy have been found to alter free hormone levels. Similarly, medications can disrupt T3 and T4 binding to serum proteins and affect thyroid measurements (Table 1). Medical conditions such as cardiac surgery,

*Ther Adv Endocrinol Metab*

2019, Vol. 10: 1–4

DOI: 10.1177/

2042018819897049

© The Author(s), 2019.  
Article reuse guidelines:  
sagepub.com/journals-  
permissions

Correspondence to:

**Steven J. Soldin**

Department of Laboratory  
Medicine, National  
Institutes of Health,  
Clinical Center, 10 Center  
Drive, Building 10, Office:  
2C429, Bethesda, MD  
20892, USA  
[steven.soldin@nih.gov](mailto:steven.soldin@nih.gov)

**Katherine A. Araque**

Adult Endocrinology  
department, National  
Institutes of Health,  
National Institutes of  
Diabetes and Digestive and  
Kidney Diseases, NIDDK,  
Bethesda, MD, USA

Endocrinology Department  
Pacific Neuroscience  
Institute, John Wayne  
Cancer Institute, Santa  
Monica, CA, USA

**Joanna Klubo-  
Gwiezdzinska**

**Lynnette K. Nieman**  
Adult Endocrinology  
department, National  
Institutes of Health,  
National Institutes of  
Diabetes and Digestive and  
Kidney Diseases, NIDDK,  
Bethesda, MD, USA

**Kerry Welsh**

Food and Drug  
Administration, MD, USA

**Table 1.** Factors that alter thyroid hormone measurements by immunoassays in the absence of thyroidal illness.

<b>Assay-related factors</b>
Inaccurate measurements in end-stage renal disease, critical illness, cardiac surgery
Inconsistency of free thyroid hormone immunoassays due to different methodologies
Positive analytic bias for FT3 and FT4
Poor correlation coefficient between the log-transformed TSH and FT4
Poor sensitivity for thyroid hormones at low concentrations (e.g. fetus and newborns)
Biotin supplementation interference with assays using biotin-streptavidin detection techniques, antiruthenium antibodies
Heterophile antibodies interference
Thyroid hormone autoantibodies interference
Macro-TSH interference
<b>Clinical factors</b>
TSH physiology: age, sex, exercise, individual thyroid set points
Inter intra-individual variations: seasonal and diurnal fluctuations, biologically inactive forms of TSH with different degrees of sialylation and sulfation
Changes in concentration or modification of TBG, transthyretin and albumin
<b>Medication related factors</b>
Disruption of TBG levels and binding properties: heparin, furosemide, salicylates, certain NSAIDs, anti-seizure medications, oral estrogens, androgens, raloxifene, mitotane, fluorouracil, heroin, methadone, glucocorticoids, clofibrate, nicotinic acid, L-asparaginase.
<b>Assay interference: Biotin</b>
Fluctuations in thyroid hormone secretion: iodine, amiodarone, lithium, TKIs, immune modulators, thionamides, expectorants (guaifenesin), dopamine agonist, potassium iodine solutions, glucocorticoids, somatostatin analogs, retinoids, dobutamine, metformin
FT3, free triiodothyronine; FT4, free thyroxine; NSAID, nonsteroidal anti-inflammatory drugs; TGB, thyroid binding globulin; TKI, thyroid kinase inhibitor; TSH, thyroid stimulating hormone.

renal disease, and critical illness affects both free and total thyroid hormone levels. Assay harmonization will not resolve the misleading IA results for free and total thyroid hormone measurements. Lastly, antibody-binding assays for stimulating and blocking TSH receptor antibody measurements require methodological optimizations before they can be incorporated in clinical practice.<sup>7</sup>

### Comparisons of IA and liquid chromatography-tandem mass spectrometry results

Numerous studies have used paired samples to compare thyroid hormone IAs and liquid chromatography-tandem mass spectrometry (LC-MS/

MS) results. These correlations are usually acceptable in the euthyroid range. However, up to 50% of IA measurements in patients with low T3 states provide a discordantly higher result (assay bias) than those obtained by LC-MS/MS.<sup>8</sup>

Despite excellent precision, IAs often provide precisely the wrong value. In a population study, IAs were characterized by a good inverse linear relationship between FT4 and log-transformed TSH in a pediatric population (correlation coefficient  $-0.82$ ), while in adults the correlation coefficient of  $-0.48$  was suboptimal. In contrast, the FT4 measured by LCMS/MS was characterized by a good correlation coefficient of  $-0.90$  in children and  $-0.77$  in adults.<sup>9</sup> In a subsequent

study, a comparison of the thyroid hormone assay performance of LCMS/MS *versus* IA in a diverse group of 100 patients of any age with any medical diagnosis in a mixed healthcare setting (inpatient *versus* outpatient), demonstrated that the TT4 and TT3 values determined by the two different assays had a good correlation coefficient ( $r$ : 0.91–0.95). Conversely, the FT4 and FT3 correlation coefficient was suboptimal ( $r$ : 0.75 and 0.50, respectively). A better correlation was found for FT3 and TT3 than with FT4 and TT4. The IAs demonstrate discrepant values at the low and high end of the established reference range, with only a moderate correlation coefficient ( $r$ : 0.51–0.75).<sup>2</sup> More recently, Hannah-Shmouni and Soldin highlighted the importance of LCMS/MS in free thyroxine measurements.<sup>10</sup> This technique can avoid diagnostic misclassification of older adults with subclinical *versus* overt hypothyroidism.

Another study of 109 individuals were assigned to three equal-sized groups clinically characterized as hypothyroid, euthyroid, or hyperthyroid. For the entire group, the correlation coefficient of TSH with FT4 performed on the Siemens Immulite 2500 analyzer, was moderate [0.45, 95% confidence interval (CI) 0.29–0.59].<sup>11</sup> Analysis performed by LCMS/MS demonstrated a better correlation (coefficient of 0.84, 95% CI 0.77–0.88). Importantly, when the euthyroid group was removed from the analysis, the correlation coefficient for IA dropped to 0.2, while LCMS/MS was 0.72. In summary, the inverse log-linear correlation between TSH and FT4 was significantly improved when FT4 was assayed by LC-MS/MS compared with IA, indicating that FT4 results measured by LC-MS/MS agreed better with those obtained by TSH and the patient's clinical condition.<sup>12</sup>

A study of 40 patients classified as subclinical hypothyroidism using a FT4 IA found potential diagnostic misclassification: 65% of these patients had FT3 or FT4 values below the reference interval when measured by LC-MS/MS. The mass spectrometry findings agreed best with the clinical picture in this study and others.<sup>8,13</sup>

In addition to improved clinical correlation, especially at low hormone concentrations, LC-MS/MS methods have the advantage of superior analytical sensitivity and specificity.<sup>14</sup> The assay performance is paramount and permits appropriate clinical decision-making, particularly in the lower or higher ends of a given reference range, and

may change medical intervention from the watch-and-wait strategy to immediate implementation of targeted therapy. Unfortunately, the pitfalls of the IAs mentioned are not addressed by the current harmonization approach.

In summary, new data show that isolated TSH measurements by IA should no longer be regarded as the most reliable test of thyroid function assessment. Similarly, FT4 measured by IA can be affected by TBG concentrations, heterophilic nonspecific antibodies, steroids, and various medications (Table 1). Measurements of plasma FT3/TT3 with accurate methods complement the clinical workup due to its biological activity.<sup>15</sup> This phenomenon is frequently observed in Graves' disease T3-toxicosis, characterized by the elevation of FT3/TT3, rather than increment in FT4/TT4. IAs for TT3/FT3 and FT4 frequently give falsely normal results in individuals with hypothyroidism, suggesting subclinical rather than overt hypothyroidism.<sup>16</sup>

The direct measurement of thyroid hormones *via* LC-MS/MS is highly sensitive, specific, and precise, and these results correlate well with the patient's clinical presentation. Measurement of TSH by IA may need to be accompanied by measurements of FT4 and TT3. Optimal measurements of FT4 and FT3 should include removal of TBG by ultrafiltration or equilibrium dialysis followed by measurement with either LCMS/MS or IA, as this approach enables the most accurate assessment of the pituitary-thyroid axis. Therefore, this approach, as opposed to measurement of TSH levels alone or together with IA FT4,<sup>3</sup> is preferable, and is the recommended method of screening for thyroid abnormalities. In support of this, the high accuracy of LC-MS/MS has been recognized by the American Thyroid Association guidelines for management of thyroid disorders during pregnancy, as a gold standard for measurement of thyroid hormones by ultrafiltration or equilibrium dialysis LCMS/MS.<sup>17</sup>

To remove IAs interference requires use of higher economic cost LCMS/MS techniques.

From 2006 to the present time, all FT4/FT3 analyses at Children's National Medical Center were shifted from IA to mass spectrometry, under the leadership of both the Endocrinology Department and Dr Soldin. This resulted in improved diagnosis and patient management,

while the LCMS/MS unit generated a profit of approximately 2 million dollars per annum. An inaccurate IA result can lead to misdiagnosis, inappropriate treatments, and deleterious consequences for human health. Even if interference of a given IA is recognized by clinicians, additional testing, with its inherent financial costs, must be pursued. New information published over the last 8 years cast further doubt on the concept of IA thyroid hormone harmonization. The evidence accumulated over the past several years points out significant shortcomings in diagnostic accuracy of FT4, FT3, TT4, TT3, and TSH measurements performed by IA.<sup>18,19</sup>

### Funding

The author(s) received no financial support for the research, authorship, and publication of this article.

### Conflict of interest statement

The author(s) declare that there is no conflict of interest.

### ORCID iD

Steven J Soldin  <https://orcid.org/0000-0002-2098-2027>

### References

1. Cali JP. An idea whose time has come. *Clin Chem* 1973; 19: 291–293.
2. Steele, BW, *et al.* Analytic bias of thyroid function tests: analysis of a College of American Pathologists fresh frozen serum pool by 3900 clinical laboratories. *Arch Pathol Lab Med* 2005; 129(3): 310–317.
3. Thienpont LM, Van Uytendaele K, De Grande LAC, *et al.* Harmonization of serum thyroid-stimulating hormone measurements paves the way for the adoption of a more uniform reference interval. *Clin Chem* 2017; 63: 1248–1260.
4. Ling C, Sun Q, Khang J, *et al.* Does TSH reliably detect hypothyroid patients? *Ann Thyroid Res* 2018; 4: 122–125.
5. Sheikh SI, Parikh TP, Kushchayeva Y, *et al.* TSH should not be used as a single marker of thyroid function. *Ann Thyroid Res* 2018; 4: 151–154.
6. Vadiveloo T, Donnan PT, Murphy MJ, *et al.* Age- and gender-specific TSH reference intervals in people with no obvious thyroid disease in Tayside, Scotland: the thyroid epidemiology, audit, and research study (TEARS). *J Clin Endocrinol Metab* 2013; 98: 1147–1153.
7. Kahaly GJ. Bioassays for TSH receptor antibodies: quo vadis? *Eur Thyroid J* 2015; 4: 3–5.
8. Welsh KJ and Soldin SJ. Diagnosis of endocrine disease: how reliable are free thyroid and total T3 hormone assays? *Eur J Endocrinol* 2016; 175: R255–R263.
9. Jonklaas J, Kahric-Janicic N, Soldin OP, *et al.* Correlations of free thyroid hormones measured by tandem mass spectrometry and immunoassay with thyroid-stimulating hormone across 4 patient populations. *Clin Chem* 2009; 55: 1380–1388.
10. Weickert MO and Kyrou I. Thyroid hormone therapy for older adults with subclinical hypothyroidism. *N Engl J Med* 2017; 377: e20.
11. van Deventer HE and Soldin SJ. The expanding role of tandem mass spectrometry in optimizing diagnosis and treatment of thyroid disease. *Adv Clin Chem* 2013; 61: 127–152.
12. van Deventer HE, Mendu DR, Remaley AT, *et al.* Inverse log-linear relationship between thyroid-stimulating hormone and free thyroxine measured by direct analog immunoassay and tandem mass spectrometry. *Clin Chem* 2011; 57: 122–127.
13. Masika LS, Zhao Z and Soldin SJ. Is measurement of TT3 by immunoassay reliable at low concentrations? A comparison of the Roche Cobas 6000 vs. LC-MSMS. *Clin Biochem* 2016; 49: 846–849.
14. Wu AH and French D. Implementation of liquid chromatography/mass spectrometry into the clinical laboratory. *Clin Chim Acta* 2013; 420: 4–10.
15. Abdalla SM and Bianco AC. Defending plasma T3 is a biological priority. *Clin Endocrinol (Oxf)* 2014; 81: 633–641.
16. Biondi B and Cooper DS. The clinical significance of subclinical thyroid dysfunction. *Endocr Rev* 2008; 29: 76–131.
17. Anagnostis P, Lefkou E and Goulis DG. Re: “guidelines of the american thyroid association for the diagnosis and management of thyroid disease during pregnancy and the postpartum” by Alexander *et al.* (*Thyroid* 2017;27:315–389). *Thyroid* 2017; 27: 1209–1210.
18. Carvalho VM. The coming of age of liquid chromatography coupled to tandem mass spectrometry in the endocrinology laboratory. *J Chromatogr B Analyt Technol Biomed Life Sci* 2012; 883–884: 50–58.
19. Favresse J, Burlacu MC, Maiter D, *et al.* Interferences with thyroid function immunoassays: clinical implications and detection algorithm. *Endocr Rev* 2018; 39: 830–850.