Data supplement accompanying:

A Progress report of the IFCC Committee for Standardization of Thyroid Function Tests

Linda M. Thienpont, Katleen Van Uytfanghe, Sofie Van Houcke, Barnali Das, James D. Faix, Finlay MacKenzie, Frank A. Quinn, Michael Rottmann, Annick Van den Brueel for the IFCC Committee for Standardization of Thyroid Function Tests (C-STFT)

Content

- Rationale for reporting of data with IVD manufacturers coded
- Calculation of the APTM
- Figure 1 – FT4 bias plots of data before and after recalibration by the IVD manufacturers.
- Figure 2 – TSH bias plots of data before and after recalibration by the IVD manufacturers.
Rationale for reporting of data with IVD manufacturers coded

The charter of the C-STFT is not to demonstrate the need for improved TFT standardization, as this has already been amply shown in the literature, and is indeed the reason for the existence of the C-STFT. Instead, the goal of the C-STFT is to conduct research that will ultimately lead to improved standardization of TFT through the identification and implementation of validated reference measurement systems.

This process requires several cycles of proof-of-concept studies, and to be most meaningful, must include many different TFT manufacturers. Early in the design of these studies, a consensus decision was made by the entire IFCC Working Group for STFT (the predecessor of the C-STFT) that data from these proof-of-concept studies would be reported with the assay manufacturer de-identified (each manufacturer would only know data for their own assays). The Working Group was composed of representatives from academia, professional societies, and IVD manufacturers. At first glance, this may seem contrary to principles of full disclosure, however, reporting these preliminary study data with IVD manufacturers coded actually protects the integrity and long-term objectives of the C-STFT in several ways.

First, at the start of the project, the key parts of the FT4 and TSH reference measurement system, i.e., the cRMP and APTM, were to be developed and validated. Also different studies were designed to explore options for assay standardization, and hence not primarily to be “method comparison” studies. Publishing the data from these studies established the current “state of the art”, and also showed the current direction and concepts being explored for standardization of TFT in the future. However, there is risk that if manufacturer-specific data are published, laboratories and clinicians may make premature decisions regarding
utilization of TFT based on preliminary data. This could cause confusion and potentially adversely impact patient care.

Second, reporting the data in coded form removes the temptation for any assay manufacturer to use the work of the C-STFT for commercial purposes. Ideally, as the work of the C-STFT progresses, each manufacturer would use the data to inform their assay standardization decisions, and begin to make plans for re-calibration, if required. Assay standardization is a complex task requiring the commitment of significant time and resources. There are also global regulatory and product registration implications.

Third, reporting of coded data is unlikely to have any short-term impact on patient care. Most clinical laboratories already have proficiency testing schemes in place that allow them to compare their results to peer-groups and other manufacturer's assays. In addition, laboratories are using method-specific reference intervals when reporting TFT test results.

Finally, once reference measurement systems for free thyroid hormones and TSH are validated for use in standardization efforts, it is the intention of the C-STFT to publish final data with manufacturers identified. Changes in assay standardization can have significant impact on patients, healthcare providers, the clinical laboratory, and assay manufacturers. Only when fully validated data are available can all stakeholders engage in a meaningful discussion about how best to advance the standardization of TFT.
Calculation of the APTM

For TSH the APTM was calculated with the adapted assay-specific outliers (see Main article, ‘data analysis’). The process was done iteratively because adaptation of outliers changes the APTM. Consequently, all assays were investigated for any particular features or influence on the “raw” APTM.

To come to the final APTM, the following decisions were taken. Assay B did not report results for the 2 lowest and the 2 highest samples of the “reduced range”, indicating that B had a smaller dynamic range than the other assays. Therefore, assay B was excluded from the APTM. This decision was additionally justified by the fact that the company participated with other assay(s) in the study. Furthermore, assays F, G, and I showed a strongly negative deviation from the other assays in the low concentration range. Assay I, additionally, deviated most from the other assays and showed issues with the %-differences between the duplicates (i.e., observation of a trend instead of random distribution around zero). Therefore, assay I was excluded from the APTM. Again, this decision was justified by the fact that the company participated with other assay(s) in the study. Then, assays F and G were calibrated to the APTM in the concentration range <1.1 mIU/L by adding a constant factor (F: 0.038 mIU/L; G: 0.042 mIU/L). This greatly improved the comparability of the assays to the APTM. For assay C, a suitable fit versus the APTM could be found only after multiplying the results in the range 2.5-40 mIU/L with the factor 0.87.

Finally, assays from the same company were averaged and their average was used for the calculation of the APTM. This was done to give each manufacturer the same weight and to fairly balance sample-related effects. In total, 8 data sets were averaged to obtain the final APTM. The TSH concentrations ranged from 0.042 mIU/L (sample 99155) to 80 mIU/L (P #006).
Data supplement accompanying: A Progress report of the IFCC C-STFT

Figure S1 – FT4 bias plots of data before and after recalibration by the IVD manufacturers.

%-Difference plots for FT4 before (grey symbols) and after recalibration (colored/black symbols) for each assay versus the ED ID-LC/tandem MS cRMP. Note that for assays L and M the symbols used in the preceding Phase III report were changed. The red broken lines are total error limits (8.6% (>5 pmol/L) and 0.43 pmol/L (≤5 pmol/L)). Note that 0.43 pmol/L equals 8.6% at 5 pmol/L.

A difference (%)

B difference (%)

Only 1st replicate used

D difference (%)

E difference (%)

After recalibration, P#007: 101%

F difference (%)

Free T4 ED-ID-MS (pmol/L)
Figure S1 ctd – FT4 bias plots of data before and after recalibration by the IVD manufacturers.

Data supplement accompanying: A Progress report of the IFCC C-STFT
Data supplement accompanying: A Progress report of the IFCC C-STFT

Figure S2 – TSH bias plots of data before and after recalibration by the IVD manufacturers.

%-Difference plots for TSH before (grey symbols) and after recalibration (colored/black symbols) for each assay versus the APTM. Note that for assays L, M and N the symbols in the preceding Phase III report were changed. The red broken lines are total error limits of 19.1%.
Figure S2 ctd – TSH bias plots of data before and after recalibration by the IVD manufacturers.