IFCC Committee for Standardization of Thyroid Function Tests (C-STFT)
Meeting at AACC 2013, Houston, Tx, USA, Monday July 29th (9:00 - 11:30 am)

PARTICIPANTS

The meeting attendance list is attached in annex 1.

OPENING OF THE MEETING

The chair (LT) welcomed the meeting attendees, presented the agenda and proposed to make a roll call. She expressed her gratitude that all IVD manufacturers (IVD MFs), but one, were present. She conveyed excuses from Dr. G. Baudino, the representative for BioMérieux.

1. Progress in laboratory testing of thyroid disease by standardization (FT4) and harmonization (TSH)

a. Milestones of C-STFT

LT presented the milestones of the C-STFT in a slide show (see annex 2). After the presentation, she reported that she did several efforts to get a written statement of the FDA regarding the need for renewal or not of the clearance of an assay after standardization/harmonization, however, she did not get it. She also regretted that no one of the FDA was present in the C-STFT meeting, in spite of what Dr. A. Gutierrez had promised in the last telephone call. *[Editorial note: after closure of the meeting, Yung W. Chan, FDA Chemistry Branch Chief, still showed up; apparently she misunderstood the time/location of the meeting].*

The need for a clear statement by the FDA was reiterated by the IVD MFs and thoroughly discussed. Obviously, different opinions exist on the FDA requirements for a cleared assay that is subject to standardization/harmonization, e.g.:

- As reported by M. Rottmann, a new 510(k) clearance will be needed, which includes a study for reference intervals (RIs).
- D. Clark thought that the FDA would investigate the needs case by case, and that they would not require renewal of clearance if the shift is within 10%, based on the comparison with the reference measurement procedure (RMP). Hence, he had the impression that the FDA considers now that demonstration of the match of a new assay with a RMP (instead of with the predicate assay) is acceptable. He also reported that from a recent contact with FDA, they are really open to meet with IVD MFs.
- B Cook stressed again that a written statement of the FDA was necessary. He added that, actually, there was a historical precedent, i.e., restandardization of PSA measurements, due to switching from the Hybritech to the WHO standard. In this case, FDA allowed to use the original numbers and a simple mathematical recalibration of the results. Thus no new FDA submission was needed, only addition of a supplement. He thought that from the PSA precedent, it appeared that FDA might be open to consider the approach of simple recalibration, without the need of new RIs.

The summary of the discussion that followed the aforementioned statements was that:

- harmonization of TSH will be a case study for many other analytes to come, as the AACC harmonization initiative progresses.

- The position statement of the FDA can be a benchmark for other regulatory bodies around the world.
- The FDA should be approached by industry, represented by ADVAMED, together with the AACC and the IFCC. It would also be good to include representatives from the clinical community and statisticians. Maybe it would also be worth to include representatives from other groups/committees, because all will face the same problem in the near future.
- It should also be tried to find out what the consequences of standardization/harmonization will be in the eyes of other regulatory bodies, e.g., in China, Brazil, Europe, etc. LT reported that for what concerns Europe, she had searched contact in Brussels with an EU-officer, who worked on the revision of the Directive. However, this visit was disappointing because she got the strong impression that the officer did not sufficiently understand the concept of metrological traceability as it is applied in clinical chemistry. S. Marivoet mentioned that the ISO-document on traceability is also under revision and that the concept 'comparability' will be added, in case no reference material nor RMP is available. Finally, the group agreed that the first contact to make regarding regulatory consequences of standardization/harmonization was with the FDA.

In conjunction with the above, 2 other topics were discussed:

1. **The need to establish new RIs.**
   Most IVD manufacturers would support the idea to incorporate a study to establish RIs in Phase IV of the C-STFT. However, it is clear that more than one study will be required, e.g., the recent study for RIs for common serum analytes in China.
   -G. Beastall reported that the centile figures in the Chinese RI study were quite similar to those used in the rest of the world. He also questioned, so to speak as the devil's advocate, whether one should continue to put so much emphasis on RI? Indeed, they may become more and more obsolete, as health care evolves to personalized medicine. Hence, he thought the focus rather should be on getting the assays aligned.
   -H. Vesper proposed to re-measure samples from previous NHANES studies as contribution to the establishment of a new RI after standardization/harmonization.
   -I. Young suggested that a close cooperation between C-STFT and the IFCC committee for Reference Intervals and Decision Limits (C-RIDL) could be beneficial.

2. **The need for harmonization/standardization – or (a strong) rationale to do so.**
   -Arguments con/pro harmonization: some IVD manufacturers saw the absence of complaints about TSH testing as an argument against the need for harmonization. Likewise, they found it prohibitive to justify spending millions of dollars. LT disagreed with this financial argument by reiterating that only 3-5 TSH assays really would be affected by harmonization. In this regard G. Beastall wondered whether the (<10%) effect of harmonization for most of the assays could not smoothly be introduced as a sort of 'lot change', which might indeed cause a shift in the aforementioned magnitude. Surprisingly, the IVD manufacturers argued that even a shift of 10% needed a regulatory justification. On the other hand, the fact that end-users complain about the discrepancy among methods and the non-concordance between FT4 and TSH-results (not necessarily due to the FT4 assays), was seen as an argument pro

harmonization. Also the ongoing debate on lowering the TSH upper range could be seen as a pro argument, although it must be said that several IVD representatives stated that the proposal is not supported anymore. They see a drop of the upper limit but as low as 2.5 mIU/L (maybe to ±4.5 mIU/L). G. Beastall stated that using a TSH RI is maybe too simplistic, as thyroid disease develops as a continuum. Clinicians do not know how to interpret a TSH of 2-5 mIU/L. Nevertheless, it puts a pressure on getting the RI correct. LT added that also the log/lin relationship between FT4 and TSH is under discussion. The relationship is considered much more complex. Another argument pro is the existence (although the use is under discussion) of common practice guidelines, whose recommendations can only be valid provided assays are comparable.

To close the pro/con discussion, LT repeated that TSH harmonization could be done by using a master equation, while sustaining the current traceability to the WHO-standard. The used process should than be mentioned in the assay inserts. Several IVD representatives noted that this has to be approved by the FDA.

b. International developments related to standardization/harmonization of thyroid hormone testing

- *Partnership for Accurate Testing of Hormones (PATH)*

H. Vesper briefly introduced this partnership. It joins 12 primarily clinical and medical organizations such as the AACE, AACC, … which understood and felt the need for standardization of hormone assays. Initially the focus was on standardization of testosterone assays, however, the scope was widened as the PATH also felt the need for standardization of estradiol, and since recently of thyroid hormones. The PATH has the intention to ask manufacturers to join the partnership.

- *UK consensus meeting on thyroid RIs – March 2013*

Attendees from C-STFT: F. MacKenzie, M. Rottmann, F. Quinn and G. Beastall
Rapporteurs: F. Quinn, G. Beastall and M. Rottmann

[Editorial note: the below is copied from the minutes from the C-STFT closed meeting at the IFCC Euromedlab 2013, where F. MacKenzie and M. Rottmann acted as rapporteur].

This meeting on initiative of J Barth was related to the 'Pathology Harmony' group in the UK. This group aims at harmonization of RIs for TFT, and has done the same already on a fairly arbitrary basis for several common chemistry measurands. The driver to have comparable test results in the UK is the implementation of the ‘electronical medical record’, in which all results, regardless their origin, are compiled. In the meeting, F MacKenzie gave an outline of the problem of insufficient standardization/harmonization of thyroid function tests based on the UK NEQAS data, while G. Beastall took the opportunity to present the work from C-STFT.

In general, the audience in the meeting was happy to see that the analytical aspects of the standardization/harmonization issue are tackled by the IFCC C-STFT. They showed very much confidence about the quality of these activities.

The representative of the British Thyroid Association (BTA) in the meeting mentioned that clinicians do not like the idea of common RIs, but want to treat their patients as individuals. In addition, they seem to have adapted to the non-standardized situation, and can live with it.

One is also concerned about the practicality of the implementation of standardization. It was suggested to rather harmonize (instead of standardize) the FT4 assays to the all

methods trimmed mean (AMTM), however, with knowledge of the relationship AMTM – conventional RMP. LT mentioned that she discussed this option, which is not new, already before with the chair of the SD (Prof. I. Young) and found him very much reluctant. He stated that the IFCC would never agree to use this approach internationally, but will always adhere to traceability to the conventional RMP.

Instead of using an analytical basis for common RIs, in the UK a novel approach was tested, i.e., thousands of data points (only gender and age were known) were combined and a RI was proposed. Then the RI was broken down according to the used methods. For TSH no difference was found, for FT4 the differences were not as big as expected. LT argued that the fact that no big differences were seen between the data obtained by different methods most probably depends on which data were used, how they were pooled and interpreted.

Editorial note: according to a previous personal report by G. Beastall to LT, the representative of a thyroid patients' association had made a firm statement in the meeting in favor of standardization/harmonization to reassure patients and GPs.

Finally, the meeting did not result in a particular outcome.

Addition made in the current meeting: the quest for common RIs does not restrict to the UK, but is worldwide at the order, e.g., in the USA medical reports sometimes include results from different laboratories, so that people struggle with the fact that each results needs an extra line for its accompanying RI.

c. Plan to establish a Network of FT4 reference laboratories

The University of Gent (UGent) has identified 3 possible partners, who were meanwhile invited to form, together with the reference lab of UGent, a Network of FT4 reference laboratories. Currently 2 laboratories are already able to provide FT4 RMP services, i.e., UGent and the Reference Material Institute for Clinical Chemistry Standards (ReCCS, Japan) (note: only UGent is JCTLM listed). The laboratories from CDC (H. Vesper) and Stanford University (J. Faix) committed to also develop the FT4 conventional RMP, and, hence, will be included. UGent initiated already the collection of a panel of 20 frozen sera from healthy subjects. Once available, the samples will be assigned FT4 concentrations by UGent and made available to the partners within the Network for validation of their performance.

2. Discussion on "Go"-decision for standardization (FT4) and harmonization (TSH)

- *Preparation of Final Phase IV method comparison*
- *Preparation/organization of stakeholders meeting*

LT proposed the IVD industry 2 options:

- Either put Phase IV (final Phase) on hold until the meeting with FDA has taken place *[editorial note: this had been recommended in the closed meeting at the IFCC Euromedlab 2013],*
  or
- Prepare technically for harmonization/standardization, however, without implementing until the stakeholders are well informed and prepared.
- Afterwards, she did a roll call and asked the IVD representatives which option they would support. For FT4, she reminded the requirements of the EU-Directive, i.e., IVD medical devices must be made traceable to a reference measurement system or RMP, if available. Note that the following statements by IVD representatives were made in their own name.

- F. Quinn was in favor of pursuing the scientific aspects of the C-STFT project, i.e., option 2, because this will show what is needed and can be achieved in terms of standardization/harmonization. For TSH each manufacturer can then decide what to do. A paper describing the final status, whit disclosure of results will be needed, so that each manufacturer can use it as a reference.
- J. Backus was also in favor of proceeding, however, predicted it would be challenging to get approval for additional funds without FDA-approval.
- LT argued that the cost for the Phase IV samples should not necessarily be that dramatic, provided all project participants contribute in soliciting clinicians for sample collections from their patients, rather than sourcing from a commercial vendor.
- P. Sibley agreed on pursuing the scientific goals of standardizing/harmonizing, however, was concerned about the expected shift in FT4 results after alignment of the assays to the conventional RMP. He recommended intensive education, but questioned how best to cope with this.
- LT posed in this regard the question, whether the HbA1c standardization case could be helpful?
- I. Young agreed that education will be a challenge. Although he did not consider all aspects of the HbA1c case a real example to follow, he mentioned that in the UK the education on the changes following HbA1c standardization was conducted in an exemplary fashion. The only requirement is that sufficient time is given for the process. He continued that, in contrast to HbA1c, the TSH case would not be that difficult, since harmonization would not require dealing with new units. He concluded that, if the community does not succeed in harmonizing TSH assays, it will never be able to harmonize any other measurand.
- With regard to the change in numbers for FT4 after standardization, J. Faix proposed to change the units from pg/mL to pmol/L to hide, so to speak, the change.
- M. Rottmann reiterated that he wanted a clear statement from regulatory bodies on the use of a master equation for TSH harmonization, as well as a clear demand from physicians. He saw the FT4 case differently in view of the availability of a conventional RMP. This means a real chance to make FT4 measurements comparable on a metrological traceability basis. He proposed to add in Phase IV a normal cohort (according to the NACB guidelines, 120 samples) next to a clinical cohort of 30-40 samples. This would allow to derive statistically valid 2.5%/97.5% centiles and demonstrate the real shift of the RI. Moreover, the data could be used in the educational process. This should be done via good publications, supported by everybody. F. Quinn added that education of clinicians should not be forgotten (might require publication in other journals). M. Rottmann added in this regard that it was utmost regrettable that the ATA is not aware of the activities of the C-STFT.
- The representatives from TOSOH confirmed that, although the regulatory aspect is important, they would want to move forward.
- J. Wassenberg allied to the previous statement. He added that future guidelines should incorporate the change that will have to be made upon standardization/harmonization.
- D. Clark was also in favor of going forward, however he repeated the importance of having multiple labs joining the reference measurement Network for FT4.
- B. Cook, finally, stated that although he finds that we do not need to proceed, it would be fine good to finish the C-STFT mission statement.

In conclusion of the above roll call, LT concluded that:

- **The group will prepare to be ready for FT4 standardization/TSH harmonization from a technical point of view, and hence proceed with Phase IV.**

- **A cohort of 120 "normal" samples will be included in the Phase IV, so that it becomes possible to investigate/confirm in parallel the new RI. It will be necessary to obtain enough sample volume so that the RI for FT4 can be established with the conventional RMP too. Ideally, the FT4 sample volume should be sufficient to also measure the TSH concentration with a least one immunoassay giving values close to the harmonization setpoint or all-procedure trimmed mean (APTM).**

- **A meeting with AACC/IFCC/ADVAMED representatives and the FDA will be organized[a].**

- **Implementation of standardization/harmonization will only be done, after sufficient evidence that all stakeholders are ready. Before, the C-STFT has set the date on 2018, but, if necessary, it can be postponed[b].**

Editorial note:
[a]LT met Dr. A. Gutierrez after the meeting. He confirmed that the FDA is willing to meet industry and all relevant partners at any time. In addition Dr. G. Myers promised to LT that he would try to facilitate the meeting (in view of the already existing good relationship between FDA-Advamed-AACC).
[b]With respect to the stakeholders meeting, H. Vesper stated that the CDC – PATH would be willing to cooperate in organizing this meeting.

3. Publications related to C-STFT activities

-Faix JD, Thienpont LM.Thyroid-Stimulating Hormone – Why efforts to harmonize testing are critical to patient care. Clin Labor News 2013;39 no. 5, 8-10.

-Thienpont LM, Van Uytfanghe K, Poppe K, Velkeniers B. Determination of free thyroid hormones. Best Pract Res Clin Endoc Metab (2013) (E-pub)

-Van Houcke SK, Van Aelst S, Van Uytfanghe K, Thienpont LM. Harmonization of immunoassays to the all-procedure trimmed mean - proof of concept by use of data from the insulin standardization project. Clin Chem Lab Med 2012;12:1-3.

4. Status of manuscripts in preparation

- *-Progress report of the IFCC Committee for Standardization of Thyroid Function Tests (authors on behalf of C-STFT: Thienpont LM, Van Uytfanghe K, Van Houcke S, Das B, Faix JD, MacKenzie F, Quinn F, Rottmann M)*
  The Phase III study will be submitted to the European Thyroid Journal (ETJ), because in a personal meeting of LT with its editor, Prof. W. Wiersinga, LT got the promise for acceptance, provided the manuscript would be accompanied by a well-sounding rationale for non-disclosure of the identity of results.
  We would not like to lose this engagement of Prof. W. Wiersinga and, therefore, our aim is to have the manuscript ready for final approval by the end of August. Hence, all remarks should be sent in by August 15th latest. If internal approval within a company is needed, please start the process now as it seems that the changes will be minor.

-Van Uytfanghe K, De Grande LA, Thienpont LM. A "Step-Up" approach for harmonization. Clin Chim Acta (Special issue entitled ""Harmonisation of Laboratory Testing - a global activity"; for publication in early 2014)
The manuscript is near to acceptance, provided some (feasible) revision.

-A statistical basis for harmonization of thyroid stimulating hormone immunoassays using a robust factor analysis model (authors: Stöckl D, Van Uytfanghe K, Van Aelst S, Thienpont LM)
LT explained that the manuscript was declined by Clin Chem because it does not disclose the identity of the IVD manufacturers, despite the fact that she tried to convince the Editor (N. Rifai) and Associate Editor (G. Miller) that doing so was of no relevance to the subject of the manuscript. LT stated that she will stick to the agreement made before with the C-STFT industry partners.
Editorial note: meanwhile we submitted the manuscript to Clin Chem Lab Med, and wait for the decision of the Editor.

5. Progress with regard to statistical estimation of the "All-Procedure Trimmed Mean" (APTM) – Factor Analysis Model

LT presented in short the statistical procedure in a slide show (see annex 3).

6. Financing of scientific secretariat at Ghent University

Until now, 6 (out of 8) manufacturers kindly committed to financing. Those who declined claimed that this was due to compliance issues and/or economy measures. In the closed meeting at IFCC Euromedlab 2013, the question was raised whether the non-contributing IVD manufacturers can simply continue to participate in the future activities of C-STFT, or after changing the conditions for participation? It was suggested to let them contribute more for the Phase IV sample collection.

CLOSURE OF MEETING

The chair thanked the attendees for their constructive contribution to the meeting.

As a result of the above discussions, the following "actions items" (2013-Bx) were defined for the project partners:

From now on		Responsibility	Timelines
2013-B1*	Organize a meeting with the FDA and obtain a written statement	UGent	As convenient
2013-B2	Review and send comments on the 1st draft of the Phase III manuscript. Finish 1st draft of the Phase III manuscript and send out for final approval by the C-STFT members and study participants	C-STFT members, participants from IVD manufacturers UGent, C-STFT members, participants from IVD manufacturers	August 15th 2013 September 2013
2013-B3	Develop a Network of FT4 reference laboratories	UGent	Initiated in summer and to be continued in fall 2013
2013-B4	Prepare Phase IV sample sourcing	UGent	Initiate ASAP – to be completed end of 2014
2013-B5	Prepare stakeholders meeting	UGent.in cooperation with CDC-PATH	To be decided
2013-B6	Prepare Phase IV measurements	UGent	End of 2014

* Note for the members of the C-STFT: to contrast with the action items decided to in the closed meeting at the IFCC Euromedlab 2013 and labeled A1-A6, we use here the symbols B1 through B6. For obvious reasons, actions A2 and A3 became obsolete.

Minutes made by:
Dr. Katleen Van Uytfanghe, on behalf of Prof. Dr. Linda Thienpont, chair of the IFCC WG-STFT
Laboratory for Analytical Chemistry, Faculty of Pharmaceutical Sciences, UGent
Harelbekestraat 72, B-9000 GENT, Belgium
Tel. +32 9 264 81 04
e-mail: linda.thienpont@ugent.be; katleen.vanuytfanghe@ugent.be

Annex 1

Name	Affiliation	e-mail address
Backus, John	Ortho Clinical Diagnostics	jbackus@its.jnj.com
Beastall, Graham	IFCC president and representative of the BTA	gbeastall@googlemail.com
Botelho, Juli	Centers of disease control and prevention	jbotelho@cdc.gov
Clark, Douglas	Siemens Healthcare	douglas.p.clark@siemens.com
Cook, Bernard	Beckman Coulter	bccook@beckman.com
Corocher, Nadia	Diasorin	Nadia.Corocher@Diasorin.it
Faix, Jim	Member of C-STFT (AACC; Stanford University)	jim.faix@stanford.edu
Gillery, Philippe	Member of the IFCC SD	pgillery@chu-reims.fr
Hishinuma, Akira	Corresponding member (Dokkyo Medical Laboratory)	a-hishi@dokkyomed.ac.jp
Horita, Hideki	Tosoh	hideki-horita-tv@tosoh.co.jp
Marivoet, Stefaan	Tosoh Bioscience, Inc.	Stefaan.Marivoet@tosoh.com
Narayanan, Shanti	Tosoh Bioscience, Inc.	Shanti.Narayanan@tosoh.com
Ogden, Judy	Tosoh Bioscience, Inc.	Judy.Ogden@tosoh.com
Quinn, Frank	Member of C-STFT (Abbott)	frank.quinn@abbott.com
Raneva, Violeta	ReCCS	v-raneva@reccs.net
Rottmann, Michael	Member of C-STFT (Roche)	michael.rottmann@roche.com
Shintani, Kouji	Tosoh	koji-shintani-sm@tosoh.co.jp
Sibley, Paul	Corresponding member (Siemens Medical Diagnostics)	paul.sibley@siemens.com
Thienpont, Linda	Chair IFCC C-STFT	linda.thienpont@ugent.be
Tsukamoto, Hisao	Tosoh	hisao-tsukamoto-rn@tosoh.co.jp
Tsuura, Masashi	Tosoh	masashi-tsuura-yv@tosoh.co.jp
Van Uytfanghe, Katleen	Scientific secretary IFCC C-STFT (University of Ghent, Belgium)	Katleen.VanUytfanghe@UGent.be
Vesper, Hubert	Centers of disease control and prevention	hvesper@cdc.gov
Wassenberg, James	Diasorin	James.wassenberg@diasoring.com
Young, Ian	Member of the IFCC SD	I.Young@qub.ac.uk

Excused

Name	Affiliation	e-mail address
Baudino, Gérard	BioMérieux	gerard.baudino@eu.biomerieux.com
Das, Barnali	Member of C-STFT (ACBI; Kokilaben Dhirubhai Ambani Hospital and Medical Research Institute at Mumbai, India)	barnali.das@relianceada.com
Mackenzie, Finlay	Member of C-STFT (ACB; UKNEQAS)	Finlay.Mackenzie@uhb.nhs.uk

Annex 2 – Milestones of C-STFT

IFCC Committee for Standardization of Thyroid Function Tests (C-STFT)

Annual meeting in conjunction with the AACC 2013 Conference

Chair
Linda Thienpont
Linda.thienpont@ugent.be

Scientific Secretary
Katleen Van Uytfanghe
Katleen.vanuytfanghe@ugent.be

Progress in Laboratory Testing of Thyroid Disease by Standardization (FT4) and Harmonization (TSH)

Milestones

Phase I – III

Milestones

3 Method comparison studies for FT4 and TSH

Showed good quality of performance, however, some room for improvement

Confirmed the need for standardization (FT4) and/or harmonization (TSH)

Demonstrated the feasibility by recalibration using a "targeted" panel (clinically relevant conc. range)

Targets set by:

- FT4 conventional RMP (ED ID-MS) (LoQ: 1.3 pmol/L)
- Statistically derived "All-Procedure Trimmed Mean" (APTM)

Phase III

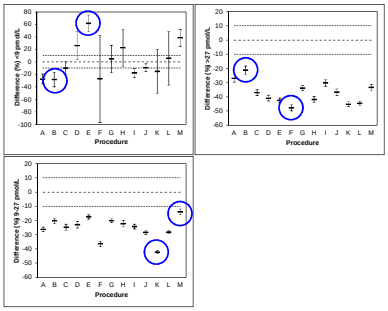
Status of standardization – FT4

Biases to ED ID-MS

9–27 pmol/L:
-25% (mean)
Range: -14% to -42%

<9 pmol/L:
2% (mean)
Range: -28% to 62%

>27 pmol/L:
-37% (mean)
Range: -21% to -48%



Conc. range of panel: : 3 to 77 pmol/L

Phase III

Status of standardization – FT4

Assay bias (%, average) vs ED ID-MS (sorted by bias in the range 9 – 27 pmol/L)

Assay	<9 pmol/L	9-27 pmol/L	>27 pmol/L
M	38.4	-14.0	-33.3
E	61.6	-17.5	-42.5
G	4.6	-20.2	-33.8
→ B	-28.3	-20.4	-21.2
H	22.5	-22.4	-42.0
D	26.0	-23.0	-40.9
I	-17.8	-24.3	-30.3
C	-10.2	-24.8	-37.1
→ A	-27.3	-26.3	-26.9
L	5.7	-28.2	-44.5
J	-9.3	-28.5	-36.8
F	-27.1	-36.7	-47.7
K	-15.2	-42.4	-45.1

M, E, G, H, D, L: Tend to pos. biases in low conc. range

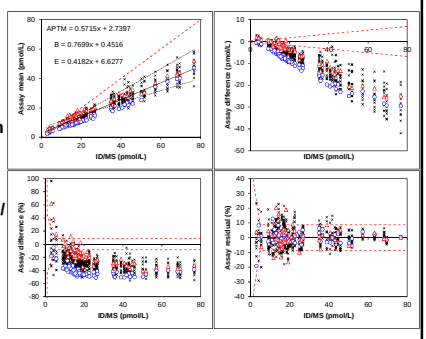
Phase III

Status of standardization – FT4

Summary
APTM = 0.57x + 2.74

M & K: most extreme Δ's: 34% in range 9 – 27 pmol/L; only 12% in range >50 pmol/L

B & E most extreme combinations of slope/ intercept: B = 0.77x + 0.45; E = 0.42x + 6.63 (conc.-dependent biases)

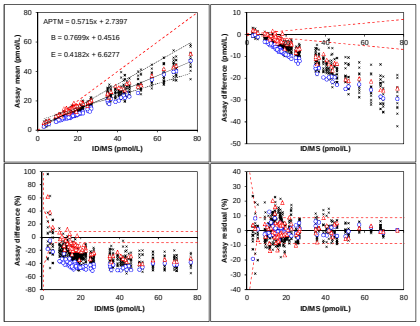


Phase III

Status of standardization – FT4

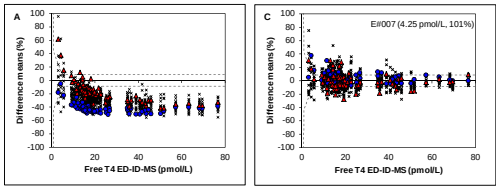
Summary
%-Residuals:
expected distribution after optimal recalibration: ±10%

Note: complies with the biological TE-limit of 9.6%



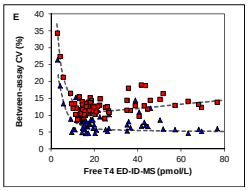
Impact of recalibration of FT4 assays

Recalibration eliminates biases to the ED ID-MS RMP



Impact of recalibration of FT4 assays

Between-assay CV considerably improved!

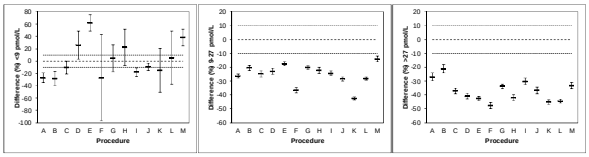


CV decreases from 9.7% (mid concentration range, before recalibration) to 3.4% (after –)

Impact of recalibration of FT4 assays

Huge changes on the overall market to expect!

Most impact in the eu- and hyperthyroid range
FT4 concentrations would increase in general by 30 – 50%

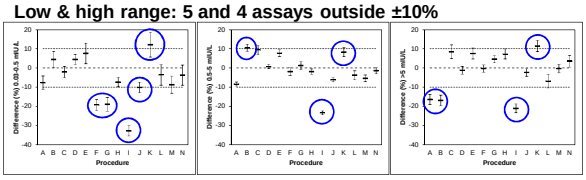


Phase III

Status of standardization – TSH

Biases to APTM

I and K most extreme Δ's (whole conc. range): ~33%
Normal range: good comparability, only 2 assays outside ±10%;
Low & high range: 5 and 4 assays outside ±10%



Concentration range covered by the panel: 0.04 – 80 mIU/L
➔Harmonization would benefit comparability most in the pathophysiological ranges

Phase III

Status of standardization – TSH

Assay bias (%, average) vs APTM (sorted by bias in the range 0.5 – 5 mIU/L)			
Assay	0.03-0.5	0.5-5	>5
I	-32.8	-23.4	-21.1
A	-7.7	-8.4	-16.5
J	-10.2	-6.2	-2.3
M	-8.7	-5.4	-0.3
L	-3.6	-3.8	-6.8
➔ F	-19.3	-1.9	-0.3
H	-7.4	-1.9	7.4
N	-3.7	-1.4	3.5
D	4.4	0.7	-1.2
➔ G	-19.0	1.2	4.6
E	7.6	7.7	7.6
K	12.1	8.3	11.5
C	-2.1	9.4	8.4
➔ B	4.4	10.5	-16.9

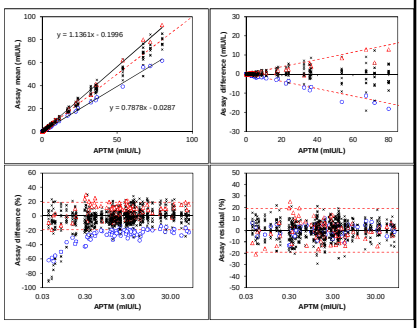
Phase III

Status of standardization – TSH

Summary
I & K most deviating:
0.03 – 0.5 mIU/L: 45%
0.5 – 5.0 mIU/L: 32%
>5.0 mIU/L: 33%

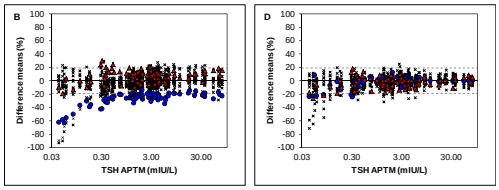
Slope & intercept:
I = 0.79x – 0.03
K = 1.14x – 0.2

%-Residuals:
expected distribution after optimal recalibration: ±20%
(biol. TE-limit 23.8%)



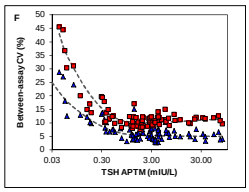
Impact of recalibration of TSH assays

Recalibration nicely centers the distribution of the assay differences around zero



Impact of recalibration of TSH assays

Between-assay CV considerably improved!

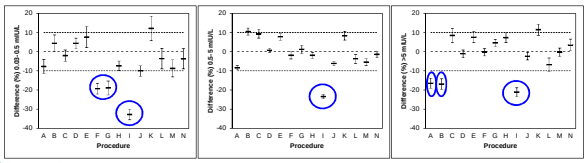


CV decreases from 9.1% (mid concentration range, before recalibration) to 5.9% (after –)

Impact of recalibration of TSH assays

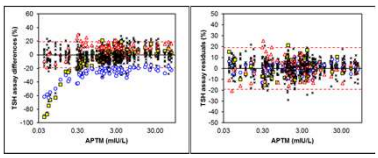
No dramatic effects on the overall market to expect!

Only 5 (out of 14) assays affected (conc. dependently)
I: overall; afterwards, harmonization status in the normal range quite impressive
F & G: low range adaptation (by constant factor)
Maybe most drastic because it may affect the assays' sensitivity claim
A: high range adaptation
B: Limited dynamic range, reformulation?



Impact of recalibration of TSH assays

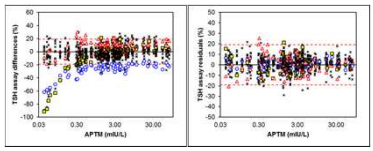
The numbers!!!



Assay 1 – Δ (most extreme positive deviation): to eliminate the bias, recalibration by ~12% needed

Assay 2 – o (most extreme negative deviation, very pronounced <1.1 mIU/L): to eliminate the bias in the conc. range >1.1 mIU/L, recalibration by ~30% is needed. The bias at the low end can be compensated by adding a constant factor of ~0.02 mIU/L. This factor most probably might affect the assay's limit of quantitation

Impact of recalibration of TSH assays



Assay 3 – □ (pronounced negative bias but only at the low conc. end): to eliminate that bias <1.1 mIU/L, the assay should be recalibrated by adding a constant factor of ~0.04 mIU/L. Above this range, calibration is all right. Again, the constant factor might affect the assay's limit of quantitation.

For all other assays, no recalibration is needed since the bias they show, is typically within the lot-to-lot changes of assays (~10%)

Path forward?

Timelines overview

2012

10 Phase III Final Report

10 Project Charter & Management concept

2013

01 Milestone Feasibility

02 "GO"-decision: Technical Part

03 Define design Phase IV; start sample procurement

04 Plan Stakeholder Meeting

2014

02 Phase IV Measurements

03 1st Stakeholder Meeting

IFCC C-STFT - Milestones - July 2013 20

Path forward?

Timelines overview

2015

02 2nd Stakeholder Meeting

03 Milestone Sustainability

04 "GO"-decision: Implementation

2016

02 Stakeholder Feedback Report

2017

01 Implement FT4 Standardization

02 Implement TSH Harmonization

11 Final Stakeholder Feedback Report

2018

03 Final Project Report – Project finished

IFCC C-STFT - Milestones - July 2013 21

Annex 3 – Factor Analysis Model

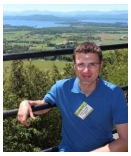
Progress with regard to statistical estimation of the "All-Procedure Trimmed Mean" (APTM)

Factor Analysis Model

proportionally

Factor Analysis Model

Stefan Van Aelst
Professor in Statistics
Ghent University
Faculty of Sciences
Department of Applied Mathematics and Computer Science



All-Procedure Trimmed Mean (APTM)

Statistical method developed

Factor Analysis (FA) model

Principal Component Analysis (PCA) is the standard procedure for the FA model

PCA estimates the so-called "composite reference values" (= APTM) as the (suitably centered) scores of the first principal component

All-Procedure Trimmed Mean (APTM)

PCA – Proof of concept

Van Houcke SK, Van Aelst S, Van Uytfanghe K, Thienpont LM. Harmonization of immunoassays to the all-procedure trimmed mean - proof of concept by use of data from the insulin standardization project. Clin Chem Lab Med 2012;12:1-3

"In conclusion, we reiterated the great potential of the APTM derived by PCA to contribute to the harmonization of laboratory measurement procedures. Our study confirmed not only the validity of this statistical approach, but also showed that it results in an equivalent quality of calibration as the reference measurement procedure approach. Naturally, the more mature measurement procedures contribute to the APTM, the better the harmonization will be."

All-Procedure Trimmed Mean (APTM)

FA model applied to the Phase III TSH dataset

The standard procedure PCA could not be used for data treatment, because of 2 limitations:

- **missing values and**
- **outliers**

Robust Alternating Regressions (RAR) as alternative

RAR after standardization of data (assumption of a homogeneous sample from an elliptically symmetric distribution not fulfilled)

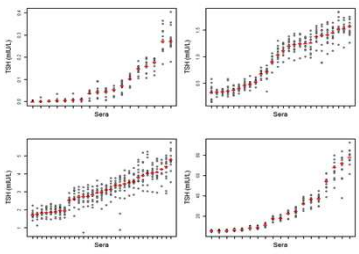
Standardization of the data implies that the estimated composite reference values have to be mapped back to the original data scale

All-Procedure Trimmed Mean by RAR

Statistical validity of the APTM

Dotplots for the data for each sample by all methods

APTM in the center of the data (–), without being affected by the outlying measurements or heterogeneity in distributions



All-Procedure Trimmed Mean by RAR

Statistical validity of the APTM

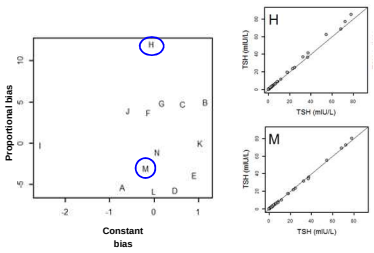
Robust R-squared (RR^2): gives the fraction of the heterogeneity in the measurements that can be explained by the FA model

RR^2 = 0.183 or 18.3% of the variability between the methods can be explained by proportional deviations of the methods from the estimated composite reference values

All-Procedure Trimmed Mean by RAR

Statistical validity of the APTM

Plot with the constant and proportional part of the systematic bias in all methods estimated by RAR



Scatterplot of method H confirms correctness of the proportional bias estimated by RAR (versus no ct -): see most measurement data lying above the first bisector, with a proportionally increasing distance to the line; for method M it confirms the small biases estimated by RAR

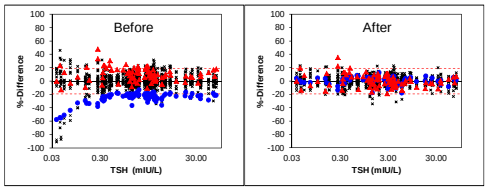
All-Procedure Trimmed Mean by RAR

Analytical validity of the APTM as recalibration basis

Differences of the results to the APTM before / after

Outcome

Before: 86% of the differences within ± 19.1%
After: 98% (even 95% were within ± 14%)



All-Procedure Trimmed Mean by RAR

Analytical validity of the APTM as recalibration basis

On the basis of the estimated APTM, we did for the results by all immunoasays (IAs):

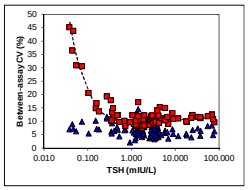
- **correlation analysis**
- **regression analysis**
- **$SD_{\%\text{-residuals}}$**

Outcome

- **Weighted linear regression analysis for 11 IAs**
- **Power function for only 3 IAs**
- **Range for r: from 0.9946 to 0.9996**
- **Range for $SD_{\%\text{-residuals}}$ from 3.4 to 9.7**

All-Procedure Trimmed Mean by RAR

Analytical validity of the APTM



Below 0.15 mIU/L, the between-assay CV decreases in average from 30% to ~8%, while for samples with a higher concentration from 11% to ~6%

All-Procedure Trimmed Mean (APTM)

Details: see poster session 07/Endocrinology/Hormones; 7/30/2013 9:30:00 AM

D. Stöckl, S. Van Houcke, S Van Aelst, K Van Uytfanghe, L. Thienpont. A statistical basis for harmonization of thyroid stimulating hormone assays using a robust factor analysis model.

Manuscript prepared

Stöckl D, Van Uytfanghe K, Van Aelst S, Thienpont LM. A statistical basis for harmonization of thyroid stimulating hormone immunoassays using a robust factor analysis model (submitted to Clin Chem but rejected because of non-identified data)