The Free Thyroxine Reference Measurement System

Session 1: “Impact of Reference Measurement Systems on Clinical Evidence”

Linda Thienpont
Linda.thienpont@ugent.be
**Background**

**Thyroid disease spectrum**
Subclinical to overt

**High prevalence**
Overt (0.8%); subclinical (5%)
Silent epidemic
High morbidity; chronic course

**Diagnosis**
Clinical diagnosis by symptoms
Laboratory data

→ **Worldwide, 60 million of FT4 tests yearly performed**

Background

FT4 testing

High burden on the healthcare system
Diagnosed patients need follow-up, and/or move from one hospital/clinician to another
Repeat measurements
Patients expect that measurement results are comparable over time, location and laboratory assay
Commercial immunoassays from different IVD-manufacturers; differently designed

→ Standardization

Where does FT4 testing stand in terms of standardization?
Considerable variation of FT4 results generated by immunoassays: means range from 10 to 17 pmol/L

→ Use of assay-specific reference intervals/decision limits (+ in publications)
Impact on Clinical Evidence

Laboratory data from non-standardized assays fail to address clinical and public health needs

• Definition of generally accepted clinical decision limits

• Development of evidence-based clinical practice guidelines for application of consistent standards of medical care

• Research translation into patient care and disease prevention activities (comparison of laboratory data across studies)

• Introduction of electronic patient records

→ Standardization urgently needed
Reference Measurement System*

- Measurand/SI
- Unit realization
  - Certified crystalline material
- Procedure (assign values)
  - Assay xyz
    - master procedure
  - Reference method
  - Assay xyz
    - end user’s procedure
  - Gravimetry

Primary calibration solution

- Working calibrator
  - Human serum panel

Product calibrator
  - (n calibrators)

Routine sample

- Patient xyz
  - *** nmol/L

Material (calibrate)

Metrological traceability

Uncertainty

*ISO 17511:2003

Thienpont - JCTLM Stakeholders' Meeting 2013 - Paris
2005: Working Group for Standardization of Thyroid Function Tests (WG-STFT) – Chair: L. Thienpont

2012: Committee (C-STFT) (same chair)

Terms of reference

“Develop reference measurement systems for thyroid hormones, i.e., TT4 & TT3, FT4 & FT3, TSH”

→ FT4 Reference Measurement System
Development of the FT4 RMS

Measurand
Quantity intended to be measured#

#Vocabulaire International de Métrologie – Concepts Fondamentaux et Généraux et Termes Associés (VIM) (3rd Ed.).

Measurand, full description

System – component (= analyte); kind-of-quantity

FT4 measurand, not really defined

FT4 Measurand

Component
Thyroxine that is not bound to proteins
Name: “Thyroxine(free)”; abbreviation: FT4

Kind-of quantity; unit
Amount-of-substance concentration; pmol/L

System
Plasma or serum under physiological conditions (pH 7.4, temperature 37°C).
Note: no demand to collect specimens under anaerobic conditions

IUPAC/IFCC format:
“Plasma/Serum – Thyroxine(free); amount-of-substance concentration” (pmol/L)
Development of the FT4 RMS

Thyroxine Primary Calibrator – IRMM-468#

<table>
<thead>
<tr>
<th>Substance</th>
<th>Certified mass fraction 1) [%]</th>
<th>Uncertainty 2) [%]</th>
</tr>
</thead>
<tbody>
<tr>
<td>thyroxine</td>
<td>98.6</td>
<td>0.7</td>
</tr>
</tbody>
</table>

1) The certified value is the purity after taking into consideration inorganic residues, water, ethanol and organic impurities detectable by HPLC-UV and HPLC-MS. The certified value is traceable to the International System of Units (SI).

2) The certified uncertainty is the expanded uncertainty estimated in accordance with the Guide to the Expression of Uncertainty in Measurement (GUM) [1]. It is expressed with a coverage factor $k = 2$, corresponding to a level of confidence of about 95%.

Development of the FT4 RMS

Reference measurement procedure
To measure the measurand as defined; should ideally be trueness-based

Possible for FT4?
FT4 Reference Measurement Procedure

Direct measurement in the system? NO

Indirect measurement in separated serum water? YES
Generation of Serum Water

Technically feasible
Ultrafiltration
Equilibrium dialysis

Potential pitfalls
Membrane type and cut-off (risk for protein leakage, adsorption)
Buffer type, pH, control
Dilution
Temperature, control
Donnan- and osmotic effects
Time of process (risk for free fatty acid generation, protein degeneration, etc.)

…
Generation of Serum Water

T4 in water separated from the serum sample vs FT4 in serum (in equilibrium with protein bound T4)

Identical?
Difficult/impossible to prove!

⇒ ED/UF separation may break the traceability chain. Trueness?
Reference Measurement Procedure

C-STFT opted for an international conventional reference measurement procedure based on

- Equilibrium dialysis
- Quantification of thyroxine in the dialysate with a trueness-based reference measurement procedure

→ ED ID-LC/tandem MS

Note

The measurand is operationally defined as

“Thyroxine in the dialysate from equilibrium dialysis of serum prepared under defined conditions”
Reference Measurement Procedure

Equilibrium dialysis
Strictly adhere to the defined equilibrium dialysis procedure (cf. CLSI C45-A#)


ID-LC/tandem MS RMP
Variants permissible
Calibration with IRMM-468

Reference Measurement Procedure


Validation of the ED part of the RMP

Performance of ED
Compliance with the predefined requirements

Sufficient robustness against relevant variables during dialysis
pH
Temperature
Dialysis time necessary to reach equilibrium
Membrane cut-off and brand
Generation of non-esterified fatty acids during dialysis

...
## Validation of Performance Attributes

<table>
<thead>
<tr>
<th></th>
<th>Imprecision (%)</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Within-run</td>
<td>Between-run</td>
</tr>
<tr>
<td>Inclusive ED [1]</td>
<td>2.8</td>
<td>2.4</td>
</tr>
<tr>
<td>MS-measurement [2]</td>
<td>1.7</td>
<td>1.0</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Trueness (Deviation from target, %)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Inclusive ED [1]</td>
<td>-0.2</td>
</tr>
<tr>
<td>MS-measurement [2]</td>
<td>+0.03</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Expanded (k = 2) uncertainty (%) for measurement protocol n = 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Incl. ED [1]</td>
<td>7.6 [3]</td>
</tr>
</tbody>
</table>

[1] n = 61  
[2] n = 66  
[3] Constant over the measurement range 1.8 – 79.8 pmol/L
Transferability of the RMP to a 2nd lab

Reference Material Institute for Clinical Chemistry Standards (ReCCS, Kawasaki, Japan)

15 samples; 4 replicates

Mean total CVs:
- UGent: 3.1% (1.3 - 4.5%)
- ReCCS: 4.2% (1.0 - 8.9%)

\[ y = 0.9489x + 0.9049 \]
\[ r = 0.9926 \]

- Listed in the JCTLM Database
FT4 RMS – Milestones

Three method comparisons
Phase I – III (2008 - 2012)

Objectives
Assess/investigate the standardization status of current FT4 assays

Assays’ quality of performance

Feasibility of standardization by method comparison with the FT4 conventional RMP on a clinical panel

Impact of standardization/recalibration
FT4 RMS – Milestones

Step-up approach

**Phase I:** Method comparison with high-volume sera from volunteers; mathematical recalibration\(^1\)

**Phase II:** Proof-of-concept but with inclusion of master calibrators and recalibration by IVD-manufacturers\(^2\)

**Phase III:** Method comparison with a clinically relevant panel (again with inclusion of master calibrators and and recalibration by IVD-manufacturers)\(^3\)

\(^3\)Submitted to *Eur Thyroid J* (revision)
Assessment of Standardization Status

**Phase III**

**Biases to ED ID-MS**

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

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

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

⇒ All assays strongly negatively biased
⇒ Physiological studies flawed

Concentration range of panel: 3 to 77 pmol/L
Assessment of Quality of Performance

No standardization without assessment of quality

Total error: difference plot after recalculation of data with regression equation: best and worst (Phase I)

Other performance attributes: imprecision, correlation, stability (within-run, between-), deviation from IQC target
Proof-of-Concept

Stability of relationship RMP – Routine immunoassays#

Relationship stable within the typical batch to batch variation of current assays

Recalibration# removes assay-specific biases

#Phase I: mathematical
#Phase II: master calibrator-based
Bias to the ED ID-MS RMP removed.
Residual dispersion nearly entirely due to within-assay effects
Effect of Standardization/Recalibration

Phase III

- Between-assay CV decreases from 9.7% (mid concentration range, before recalibration) to 3.4% (after –)
Most pronounced effect in the eu- & hyperthyroid range

FT4 concentrations will increase in general by 30 – 50%

△ Reference intervals
Way Forward?
Way Forward

Phase IV (timeline 2014-’15)

**Technically prepare standardization** (without direct implementation though); use a clinically relevant panel

**Establish an infrastructure for procurement of serum panels**#

**Set-up a network of reference laboratories**
Currently: UGent (L. Thienpont) & ReCCS (M. Umemoto); potential other candidates: Stanford University (J. Faix) & CDC (H. Vesper)

**Liaise with regulatory authorities**

#Van Houcke SK, Thienpont LM. “Good samples make good assays” - The torturous way to sourcing clinical samples for the thyroid standardization project. *Clin Chem Lab Med* 2013;51:967-72.
Way Forward

Liaise with key stakeholders

Establish reference intervals with standardized assays
In cooperation with IFCC Committee on Reference Intervals and Decision Limits (C-RIDL)

Educate manufacturers, clinicians and patients
In collaboration with IFCC Education and Management Committees

Coordinate implementation of standardization
All manufacturers/assays at the same time

Timeline: 2018?
Linda Thienpont
linda.thienpont@ugent.be