Standardization of Thyroid Function Tests

Saturday, Dec 3rd 2011
39th meeting
Location: Nycomed Belgium
1080 Brussels

Linda Thienpont
Linda.thienpont@ugent.be
Thyroid dysfunction: clinical importance

Relatively high prevalence in adults

Significant clinical consequences

New perspectives resulting from research\textsuperscript{e.g.,1-3}

Subclinical thyroid dysfunction linked to various adverse clinical outcomes (all-cause mortality; pregnancy)


\textsuperscript{3}Negro R et al. Thyroid antibody positivity in the first trimester of pregnancy is associated with negative pregnancy outcomes. J Clin Endocrinol Metab 2011;96:E920-4.
Accuracy of clinical diagnosis limited

Clinical manifestations of thyroid disease vary in character and severity among patients

Symptoms often non-specific and slowly progressing

→ Adequate laboratory assessment is needed

→ Strategies for laboratory testing and interpretation are welcome

Note: yearly volume of requests for testing of TSH and free thyroxine (FT4) alone is estimated to be in the order of 180 and 60 million US$, respectively#

#Weinzierl C, Beckman-Coulter, personal communication (2009).
Lab testing and interpretation strategies

Guidelines


Lab testing and interpretation strategies

Guidelines (continued)


Lab testing and interpretation strategies

Leading journals


– Surks MI et al. Subclinical thyroid disease: scientific review and guidelines for diagnosis and management. *JAMA* 2004;291:228-238 [Review].
Lab testing and interpretation strategies

Not to forget:

Hot topics currently under debate in scientific literature

Among others: is screening for thyroid dysfunction in pregnancy worthwhile? e.g., 1,2

Do we need age-, gender- & ethnicity-specific TSH reference limits? e.g., 3

Should the TSH upper reference limit be lowered? e.g., 4,5


A closer look into the strategies

Guidelines sometimes quote **absolute** decision limits/cut-off values or reference intervals

Literature sometimes reports on clinical trials in terms of **absolute values** without reference to the laboratory test used

**Examples:**

... *the panel defined the reference range of normal TSH concentration as 0.45 – 4.5 mIU/L*

... *the panel recommends thyroid hormone therapy in individuals with elevated serum TSH concentrations whose FT4 concentration is below the reference range of 10.3 – 25.7 pmol/L*

... *a TSH of 2.5 mIU/L has been accepted as the upper limit of normal TSH in the first trimester of pregnancy*
Measurement paradigm

Prerequisite to work with absolute values

“Laboratory tests that claim the same measurand should give equivalent results within meaningful clinical constraints”
Measurement paradigm

Do current thyroid function tests give equivalent results?
Working Group for Standardization of Thyroid Function Tests (WG-STFT) – Chair: LM Thienpont (since 2005)

Mission statement:

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

*Standardize the tests*
Reference measurement system

Material [Calibration]

Procedure [Value assignment]

Definition of measurand & unit

Primary reference measurement procedure

Secondary reference measurement procedure

Manufacturer’s master procedure

End user’s routine measurement procedure

Traceability

Uncertainty

Primary calibrator

Manufacturer’s working calibrator

Manufacturer’s product calibrator

Routine sample

Result
WG-STFT – Mission statement

Develop
Reference materials
Reference measurement procedures
Reference laboratories

Reference measurement system

Before the mission statement “standardize”…

Assess the current standardization status and quality of performance
Method comparison with a panel of native sera

Investigate the feasibility of standardization

Implement (sustainable) standardization
WG-STFT – Mission statement

RMS for free thyroid hormones: accomplished

References


WG-STFT – Mission statement

RMS for total thyroid hormones: accomplished

References


WG-STFT – Mission statement

RMS for TSH: accomplished

References


WG-STFT – Mission statement

Assessment of the current standardization status and quality of performance: accomplished

References


Measurement paradigm

Do current thyroid function tests give equivalent results?
Standardization status

Free thyroxine (FT4)
IFCC WG-STFT – Project phase I (2008)

15 tests
Means varied from 10 to 17 pmol/L
Except 2 tests, all measured FT4 much lower than the RMP (up to 42%)

⇒ Hyperthyroid FT4 conc. with 1 test, still eu- with another
⇒ Standardization is needed
Standardization status

Free triiodothyronine (FT3)
IFCC WG-STFT – Project phase I (2008)

14 tests

Means varied from 3.7 to 6.5 pmol/L

Most tests were negatively biased vs the RMP (up to 30%)

1 test had a positive bias of 22%

Standardization is needed
Standardization status

TSH
IFCC WG-STFT – Project phase I (2008)

16 tests
Medians of 11 tests differed <10% from the surrogate RMP (overall median)
Using the respective regression equations and a TSH RI from 0.4 to 4 mIU/L, the most discrepant tests would give values ranging from 0.34 to 3.24 mIU/L, and 0.39 to 4.55 mIU/L

Lowering the upper reference limit is impossible without standardization
Standardization status

Total thyroxine (TT4)
IFCC WG-STFT – Project phase I (2008)

11 tests

Means varied from 75 to 102 nmol/L

Overall good agreement with the RMP: 7 tests gave means that differed less than 10%

➔ Only 4 tests need standardization
Standardization status

Total triiodothyronine (TT3)
IFCC WG-STFT – Project phase I (2008)

12 tests

Means varied from 1.45 to 1.89 pmol/L

Nearly all tests were positively biased vs the RMP: 7 deviated >10%; 2 of them even >20% (up to 32%)

⇒ Standardization is needed
Quality of performance

Assessment of quality of performance
Against quality specifications, e.g., TE limits.
After recalculation of results with regression equation, e.g., FT4 best and worst quality of performance
IFCC WG-STFT – Project phase I (2008)
Quality of performance

Assessment of quality of performance
IFCC WG-STFT – Project phase I (2008)

TSH

\[ r^2 = 0.998 \]
\[ r^2 = 0.992 \]
\[ r^2 = 0.972 \]
Free T4 in pregnancy

Comparison (trimester specific) of FT4 values by ED ID-LC/tandem MS and 3 immunoassays in non-pregnant controls (n = 26) and pregnant (n = 107)
Free T4 in pregnancy

First conclusion

Current thyroid function tests need standardization

- Diagnosis/follow-up against test specific reference intervals/clinical decision thresholds is still recommended

Standardization efforts must be accompanied by assessment of the quality of thyroid tests

- Communicate in a transparent way with clinicians on the influence of analytical quality on a test result

- Infer adequate quality specifications preferably in consultation with clinicians

- Improve the quality of some thyroid function tests
First conclusion

Testing of FT4 in pregnancy

Immunoassays are sensitive to binding protein alterations, but to a grossly different extent

Some show the “true” changes of “FT4” during pregnancy (as observed with the reference measurement procedure ED ID-LC/tandem MS)

⇒ FT4 in pregnancy can be tested with commercial tests, but results have to be interpreted test specific; the same test should be used for follow-up
Feasibility of standardization

Phase II - Proof of concept – FT4

**Objective**: demonstrate consistency in time of method comparisons against the reference measurement procedure (or surrogate –) (= sustainability)

![Graph showing ratio of assay mean to MS mean for Phase 1 and Phase 2.](image-url)
Feasibility of standardization

Phase II - Proof of concept – TSH

![Graph showing ratio assay-mean/trimmed-mean for Phase 1 and Phase 2. The x-axis represents different labels: D, O, M, F, B, H, J, A, K, G, C, L, N, E. The y-axis shows the ratio assay-mean/trimmed-mean ranging from 0.7 to 1.3. The graph includes data points for each phase and horizontal lines indicating the mean values.](image)
Feasibility of standardization

Standardization = recalibration on the basis of a method comparison

Outcome of standardization – Phase I & II

FT4 – Situation **before** and **after** recalibration
Feasibility of standardization

FT4 – Between test CV before (□) and after (Δ) recalibration
Feasibility of standardization

Outcome of standardization
Phase I & II

TSH – Situation **before** and **after** recalibration
Feasibility of standardization

TSH – Between test CV before (□) and after (Δ) recalibration

![Graph showing TSH Trimmed mean (mIU/L) vs. Between assay CV (%)]
Second conclusion

Standardization is feasible

Technical means are available

- Disseminate on activities/plans of the WG

- Implement (sustainable) standardization
Way forward
Way forward

Transformation of WG-STFT into Committee

Objective: involve a broader forum of stakeholders in preparation of the implementation phase

Stakeholders

(IVD manufacturers)
Physicians and their patients
Laboratory directors
Professional societies
Pharmaceutical companies
Regulatory agencies

Additional vehicles
Medical journals
Way forward

Up to now, feasibility of standardization shown from method comparison with samples from “apparently healthy” individuals

BUT

What about the performance of the tests on clinical samples?

→ Submit commercial tests to a method comparison with samples from different patient populations (hypo-, eu- and hyperthyroid patients) – Project Phase III (2012)

→ Clinical samples, repository panels are needed!
Phase III study

Establish a strong physician–laboratory interface

January 2010 – Call to clinicians for:
- Interest in joining the WG activities/discussions
- Help with procurement of clinical samples/repository panels

Several positive replies (Europe, Japan, USA), BUT: all restricted to interest in joining the WG activities

Hindrances with regard to procurement of clinical samples
- Approval by local ethical committees
- Needed sample volume (15 mL serum/~ 30 mL blood)

⇒ Route not further pursued
Phase III study

Clinical samples:
FT4 n = 90: TSH n = 100
15 mL of serum per sample

Commercial source for samples

Contact person: Dr. Jim Boushell
Address: 10 Commerce Way, Norton, MA 02766, USA
# Phase III study

<table>
<thead>
<tr>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Sample Type:</strong></td>
</tr>
<tr>
<td><strong>Matrix:</strong></td>
</tr>
<tr>
<td><strong>Supporting Data:</strong></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Inclusion Criteria:</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Informed Consent Obtained</td>
<td></td>
</tr>
<tr>
<td>2. Free T4 concentration within the suggested reference range corresponding to a diagnosis of:</td>
<td></td>
</tr>
<tr>
<td>a. Euthyroid:</td>
<td>0.8 – 1.7 ng/dL</td>
</tr>
<tr>
<td>b. Hyperthyroid:</td>
<td>&gt; 1.7 ng/dL</td>
</tr>
<tr>
<td>c. Hypothyroid:</td>
<td>&lt; 0.8 ng/dL</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Exclusion Criteria</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1. No Informed Consent</td>
<td></td>
</tr>
<tr>
<td>2. Current T4 or recombinant TSH therapy</td>
<td></td>
</tr>
<tr>
<td>3. Current diagnosis of any of the listed severe thyroidal conditions:</td>
<td></td>
</tr>
<tr>
<td>a. Chronic Renal Failure</td>
<td></td>
</tr>
<tr>
<td>b. Liver Cirrhosis</td>
<td></td>
</tr>
<tr>
<td>c. Advanced (active) malignancy</td>
<td></td>
</tr>
<tr>
<td>d. Sepsis</td>
<td></td>
</tr>
<tr>
<td>e. Trauma</td>
<td></td>
</tr>
<tr>
<td>f. Prolonged Fasting or Starvation</td>
<td></td>
</tr>
<tr>
<td>g. Heart Failure</td>
<td></td>
</tr>
<tr>
<td>h. Myocardial Infarction</td>
<td></td>
</tr>
</tbody>
</table>
## Phase III study

<table>
<thead>
<tr>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Sample Type:</strong></td>
</tr>
<tr>
<td><strong>Matrix:</strong></td>
</tr>
</tbody>
</table>

**Supporting Data:**
- Demographics (Age / Gender / Ethnicity)
- **TSH** concentration within corresponding medical diagnosis of:
  - Euthyroid
  - Hyperthyroid
  - Hypothyroid
- Current Medications & Treatment

<table>
<thead>
<tr>
<th>Inclusion Criteria:</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Informed Consent Obtained</td>
</tr>
<tr>
<td>2. TSH concentration within the suggested reference range corresponding to a diagnosis of:</td>
</tr>
</tbody>
</table>
  - Euthyroid: 0.3 – 3.0 uIU/mL |
  - Hyperthyroid: < 0.3 uIU/mL |
  - Hypothyroid: > 3.0 uIU/mL |

<table>
<thead>
<tr>
<th>Exclusion Criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. No Informed Consent</td>
</tr>
<tr>
<td>2. Current T4 or recombinant TSH therapy</td>
</tr>
<tr>
<td>3. Current diagnosis of any of the listed severe thyroidal conditions:</td>
</tr>
</tbody>
</table>
  - Chronic Renal Failure |
  - Liver Cirrhosis |
  - Advanced (active) malignancy |
  - Sepsis |
  - Trauma |
  - Prolonged Fasting or Starvation |
  - Heart Failure |
  - Myocardial Infarction. |
Phase III study

Criteria for enrolment of individuals
Preferably, not under treatment. If treated, information on the type of treatment and when it has been started should be captured.

Exclusion of patients diagnosed with a severe non-thyroidal illness (NTI) (state of dysregulation with abnormal levels of T3, T4, FT3 and/or FT4 although the thyroid gland does not appear to be dysfunctional).

In practice, exclusion of critically ill patients suffering from CKD, liver cirrhosis, sepsis, trauma, advanced (active) malignancy, prolonged fasting/starvation, heart failure, MI, psychiatric disorder.

Agreement on a fair cost price per donation (450 US$)
### Phase III study: status update (30/11/11)

<table>
<thead>
<tr>
<th>Population</th>
<th>Details</th>
<th>Target n</th>
<th>Enrollment n</th>
</tr>
</thead>
<tbody>
<tr>
<td>TSH A1 &lt;&lt;&lt;conc.</td>
<td>Hyper-thyroid</td>
<td>10</td>
<td>2</td>
</tr>
<tr>
<td>A2: 0.01- 0.1 mIU/L*</td>
<td></td>
<td>10</td>
<td>1</td>
</tr>
<tr>
<td>A3: 0.1-0.3 mIU/L</td>
<td></td>
<td>10</td>
<td>9</td>
</tr>
<tr>
<td>B: 0.3-3.0 mIU/L</td>
<td>Eu –</td>
<td>30</td>
<td>30</td>
</tr>
<tr>
<td>C1: 3.0-50 mIU/L</td>
<td>Hypo –</td>
<td>20</td>
<td>20</td>
</tr>
<tr>
<td>C2: &gt;50 mIU/L</td>
<td></td>
<td>20</td>
<td>2</td>
</tr>
<tr>
<td>FT4 D: &gt; 2.2 ng/dL</td>
<td>Hyper –</td>
<td>30</td>
<td>5</td>
</tr>
<tr>
<td>E: 0.8-2.2 ng/dL</td>
<td>Eu –</td>
<td>30</td>
<td>29</td>
</tr>
<tr>
<td>F: 0.2-0.8 ng/dL</td>
<td>Hypo –</td>
<td>30</td>
<td>9</td>
</tr>
</tbody>
</table>

*Test specific
Call of the Chair of WG-STFT

Would you be willing to support us by active involvement in patient (and sample) recruitment?
If yes, contact me!

Linda.thienpont@ugent.be
Tel. +32 9 264 81 04