Functional Sensitivity of Seven Automated Thyroid Stimulating Hormone Immunoassays

William E. Owen, ARUP Institute for Clinical and Experimental Pathology, Salt Lake City, UT
Mary Lou Gantzer and Jeremy M. Lyons, Siemens Healthcare Diagnostics, Newark, DE
Alan L. Rockwood and William L. Roberts, Department of Pathology, University of Utah Health Sciences Center, Salt Lake City, UT
Abstract

Background
Serum thyroid stimulating hormone (TSH) measurements are useful for detecting clinical and subclinical primary hypothyroidism in ambulatory patients. For diagnosis of hypothyroidism, the functional sensitivity (FS) is an important performance criterion, and current guidelines recommend an FS of ≥ 0.02 m IU/l for "third" generation performance.

Methods
We evaluated TSH FS for the Access 2, ADVIA Centaur®, Architect i2000, Dimension® ExL, Modular Analytics E170, IMMULITE® 2000 and Dimension Vista® 1500 automated immunoassays using serum pools tested over a 6-week period using 2 reagent lots and 2 calibrations. FS was determined by fitting a power function to the imprecision data using KaleidaGraph software.

Results
The FS (m IU/l) for the Access 2, ADVIA Centaur, ARCHITECT i2000, Dimension ExL, Modular Analytics E170, IMMULITE 2000, and Dimension Vista 1500 systems were determined to be 0.039, 0.006, 0.007, 0.008, 0.003, and 0.003, respectively. The lowest and next to lowest pools had overall mean TSH concentrations of 0.005 to 0.022 m IU/l and 0.007 to 0.017 m IU/l, respectively.

Conclusions
All assays showed excellent performance in FS consistent with a "third" generation claim except for the Access 2 system. Further harmonization of TSH immunoassays is required, especially at lower concentrations.

Introduction

Most currently available automated thyroid stimulating hormone (TSH) immunoassays have a "third" generation claim, meaning a functional sensitivity (FS) of ≥ 0.02 m IU/l as recommended by the National Academy of Clinical Biochemistry. These assays, because of their sensitivity and specificity, are acceptable for use in the diagnosis of thyroid disease from overt hypothyroidism to overt hyperthyroidism. In hospitalized patients, "third" generation assays help to distinguish sick hyperthyroid patients with very low values of TSH (< 0.05 m IU/l) from patients with mild temporal TSH reduction (0.05 – 1.0 m IU/l) consistent with non-thyroidal illness. Clinical decisions may be guided or altered for patients with hyperthyroidism based on TSH results. Decisions about the dosing of exogenous thyroid hormone to patients with subnormal TSH can be aided by accurate quantification of TSH values (0.1 m IU/l). A study on the performance characteristics including FS of six TSH immunoassays having a "third" generation claim was published in 2005. Results of this study indicated the need for further efforts to harmonize TSH immunoassays including improved comparability at low TSH concentrations (0.2 m IU/l). Steps are being taken to achieve this goal. It was the purpose of our study to determine current FS performance of seven commercially available automated immunoassays and evaluate comparability at low TSH concentrations.

Materials and Methods

Seven automated TSH immunoassay analyzers were used in our study including Access 2 (Beckman Coulter), ADVIA Centaur (Siemens Healthcare Diagnostics), Architect i2000 (Abbott Diagnostics), Dimension ExL (Siemens), Modular Analytics E170 (Roche Diagnostics), IMMULITE 2000 and Vista 1500 (Siemens) systems. All immunoassays studied used chemiluminescence or electrochemiluminescent detection and are standardized to the World Health Organization Second International Reference Preparation (B10558). Testing on the Access 2, Architect i2000, Modular Analytics E170, and IMMULITE 2000 analyzers was performed at ARUP. The ADVIA Centaur, Dimension ExL, and Vista 1500 testing was performed at Siemens.

Our imprecision study met the requirements of the Clinical and Laboratory Standards Institute EP 17-A Protocol. Serum samples from individual subjects of 1–1.25 ml were obtained from a commercial supplier (vendor); a TSH value was supplied with each sample (tested primarily by a Modular Analytics E170 or an ADVIA Centaur system). Seven pools with the required concentrations (0.2 m IU/l) were prepared using serum from patients with subnormal TSH and were used to assess the performance characteristics of the seven immunoassays. These seven prepared pools were by combining 17–21 samples with comparable TSH concentrations (0.2 m IU/l). The lowest overall concentrations of 0.005 to 0.022 m IU/l and 0.007 to 0.017 m IU/l, respectively.

Table 1. Summary of functional sensitivity (FS) analyses.

<table>
<thead>
<tr>
<th>Instrument</th>
<th>Software</th>
<th>m IU/l (r)</th>
<th>m IU/l (r)</th>
<th>m IU/l (r)</th>
<th>m IU/l (r)</th>
<th>m IU/l (r)</th>
<th>m IU/l (r)</th>
<th>m IU/l (r)</th>
<th>m IU/l (r)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Access 2</td>
<td>Kaleida</td>
<td>0.031 (0.850)</td>
<td>0.039 (0.665)</td>
<td>0.037 (0.869)</td>
<td>0.049 (0.652)</td>
<td>0.039 (0.690)</td>
<td>0.039 (0.688)</td>
<td>0.037 (0.877)</td>
<td>0.049 (0.658)</td>
</tr>
<tr>
<td></td>
<td>Prism</td>
<td>0.006 (0.835)</td>
<td>0.009 (0.842)</td>
<td>0.006 (0.894)</td>
<td>0.006 (0.894)</td>
<td>0.010 (0.844)</td>
<td>0.010 (0.844)</td>
<td>0.006 (0.894)</td>
<td>0.006 (0.894)</td>
</tr>
<tr>
<td></td>
<td>Cricket</td>
<td>0.007 (0.914)</td>
<td>0.006 (0.914)</td>
<td>0.007 (0.958)</td>
<td>0.007 (0.958)</td>
<td>0.008 (0.958)</td>
<td>0.007 (0.958)</td>
<td>0.007 (0.958)</td>
<td>0.007 (0.958)</td>
</tr>
<tr>
<td></td>
<td>Excel</td>
<td>0.008 (0.982)</td>
<td>0.008 (0.944)</td>
<td>0.008 (0.978)</td>
<td>0.008 (0.978)</td>
<td>0.008 (0.981)</td>
<td>0.008 (0.981)</td>
<td>0.008 (0.981)</td>
<td>0.008 (0.981)</td>
</tr>
<tr>
<td></td>
<td>PSI-Plot (Natural Space)</td>
<td>0.006 (0.835)</td>
<td>0.009 (0.842)</td>
<td>0.006 (0.894)</td>
<td>0.006 (0.894)</td>
<td>0.010 (0.844)</td>
<td>0.010 (0.844)</td>
<td>0.006 (0.894)</td>
<td>0.006 (0.894)</td>
</tr>
<tr>
<td></td>
<td>PSI-Plot (Quasi Natural Space)</td>
<td>0.007 (0.914)</td>
<td>0.006 (0.914)</td>
<td>0.007 (0.958)</td>
<td>0.007 (0.958)</td>
<td>0.008 (0.958)</td>
<td>0.007 (0.958)</td>
<td>0.007 (0.958)</td>
<td>0.007 (0.958)</td>
</tr>
<tr>
<td></td>
<td>PSI-Plot (In ln Space)</td>
<td>0.007 (0.914)</td>
<td>0.006 (0.914)</td>
<td>0.007 (0.958)</td>
<td>0.007 (0.958)</td>
<td>0.008 (0.958)</td>
<td>0.007 (0.958)</td>
<td>0.007 (0.958)</td>
<td>0.007 (0.958)</td>
</tr>
</tbody>
</table>

a. PSI-Plot fits in "natural space" was performed using a user-defined true-nonlinear fit with original data.

b. PSI-Plot fit in "natural space" was performed using a predefined, built-in user space data fitting capability.

c. PSI-Plot fit in "ln space" was performed using a user-defined true-nonlinear fit with original data.

PSI-Plot fit in "ln space" was performed using a predefined, built-in user space data fitting capability.

e. Fit using all 7 pool results.

f. Fit excluding Pool 2 result.

g. PSI-Plot fit in "ln space" was performed using a predefined, built-in user space data fitting capability.
Results

The results for FS calculations by Kaleida, Prism, Cricket, and Excel for the seven immunoassays are summarized in Table 1. This table also includes results for FS calculations by PSI-PLOT determined using three different methods of analyzing the data. The fit for “natural space” was done using a user defined nonlinear equation (power law) to do a fit in natural variable space with the original data. The calculation using “quasi natural space” was carried out by using a pre-defined built in power law (y = a * x^b) fitting capability in the PSI-PLOT software using the original data. The “ln ln space” (i.e. natural log) fit was performed as follows: both axes were transformed by taking the natural log of the data, then the data were fit by ordinary linear regression, and finally the parameters from the fit were transformed back to ordinary variable space. Fig. 1 shows a comparison of power curve fitting by Kaleida and Prism of results for the Access 2 system. In panels A and C all seven pool results are shown. In panels B and D results for Pool 2 were excluded. Pool 2 (see arrow in panels A and C) could not be excluded as a statistical outlier but was considered a visual outlier. The data for the Access 2 system in Table 1 using Kaleida and Prism are summarized from the results of the power curve calculations (see statistics in Figure 1 caption).

The lowest pool (Pool 1) had mean TSH concentrations of 0.022, 0.016, 0.008, 0.005, 0.007, 0.015, and 0.008 m IU/l for Access 2, ADVIA Centaur, Architect i2000, Dimension ExL, Modular Analytics E170, IMMULITE 2000, and Vista 1500 systems, respectively. The next to lowest pool (Pool 2) had mean TSH concentrations for Access 2, ADVIA Centaur, Architect i2000, Dimension ExL, Modular Analytics E170, IMMULITE 2000, and Vista 1500 systems, respectively. The two lowest pools that were considered outliers were excluded from further analysis.

Discussion

All immunoassays we studied had an FS consistent with a “third” generation claim (0.02 m IU/l) except for the Access 2 system. Results summarized in Table 1 indicated that Access 2 had F50.02 m IU/l whether Pool 2 was included or excluded, and whether or not the claim was met did not depend on the software used for analysis. There are at least four variables that contribute to FS variability. These include the performance of the analyzer that is used if it has not been properly maintained, the reagent lots that are used, the characteristics of the patient pools that are tested, and software that is used to construct the power curve. In the current study, all instruments used were maintained according to manufacturer recommendations and the same sample pools and power curve software were also used across all instruments.

The results for Access 2 contrast with an earlier study where an FS of 0.020 m IU/l was obtained, while the Architect i2000, Modular Analytics E170, and IMMULITE 2000 gave comparable results.1 This study used Cricket for power curve analysis and determination of FS. The Access 2 method gave an FS of 0.028 m IU/l in a study conducted in 1997.2 The reason for differences in Access 2 performance in each of the three studies is unclear. The ADVIA Centaur FS result of 0.006 m IU/l contrasted with an earlier study where the FS was 0.039 mIU/l.3 The Centaur TSH assay in the previous study has since been reformulated to improve the FS; the reformulated assay was used in the current study, and as expected, a lower FS was obtained.

Figure 1. Comparison of Prism versus Kaleida analyses of Access 2 Results.
Our data indicate that harmonization of TSH immunoassays remains a challenge, though progress has been made. Differences seen in calculated FS based on the software used (Table 1) were an interesting and unexpected result from our study. Based on results for both "In In space" and "quasi-natural space" fit using the PSI-Plot software, we surmised that the "quasi-natural space" calculation followed the same log transform procedure as the "In In space" calculation we performed only it is done internally (hidden from the user). The results for "In In space" fit and "quasinatural space" fit are very similar to those calculated by Kaleida, Cricket, and Excel indicating that these software packages follow an internal procedure as seen with the "quasi-natural space." The Prism software results paralleled those of the "natural space" calculations where the original data with no log transformation were used for curve fit. In comparing the power curves plotted by Prism and Kaleida for the Access 2 study results (Figure 1), we observed that the Pool 2 result affected the Prism curve (Figure 1, panels A and B) substantially more than the Kaleida curve (Figure 1, panels C and D).

We speculated that the log transformation of the data makes the final calculated curve parameters more resistant to effects of divergent points. There were still uncertainties in how the different software packages performed the calculation for "r." However, it appeared that the calculation of "r" was done in the "natural space" for both Prism and the user programmed fit in PSI-Plot.

Our study indicated that the ADVIA Centaur, Architect i2000, Dimension ExL Modular Analytics E170, IMMULITE 2000, and Vista 1500 TSH immunoeasys have excellent performance with an FS consistent with a "third" generation claim. Further harmonization of TSH immunoassays is required, especially at lower TSH concentrations.

Acknowledgements

Siemens Healthcare Diagnostics provided reagents and support for this study. The Access 2, ARCHITECT i2000, and Modular Analytics E1700 analyzers at ARUP were provided by Beckman Coulter, Abbott Diagnostics, and Roche Diagnostics, respectively. The ARUP Institute for Clinical and Experimental Pathology provided support for this study.

References
11. Thyrotropin/Thyroid stimulating hormone (tsh) measurement. Thyroid. 2003;13:33–44.