

correspondence

Indirect Methods for TSH Reference Interval: At Last Fit for Purpose?

DOI: 10.1309/AJCPX4PISKJFDT8L

To the Editor

We read with great interest the article by Katayev et al¹ and the companion editorial by Horowitz² that discuss the production of reliable reference intervals for common tests such as calcium, creatinine, mean corpuscular volume, and thyroid-stimulating hormone (TSH). Katayev et al¹ used a computerized version of the technique proposed almost 50 years ago by Hoffmann³ in a large number of results stored in the laboratory information system serving 6 laboratories. According to Katayev et al,¹ “The computerized Hoffmann method for the indirect determination of RIs [reference intervals] produced intervals that were remarkably similar to peer-reviewed RIs.” Horowitz² disagrees and states “... the reference intervals generated in this way are strikingly different from the reference intervals in use....”

The articles disagree because TSH is being measured more and more precisely, but different assays show a relevant bias; manufacturers, laboratory professionals, and clinicians rarely appreciate the effect of this bias on the cutoffs quoted in guidelines. In some cases, we agree that TSH could indeed be unfit for purpose.⁴

Katayev et al¹ reported upper reference limits in 2 large sets of data using the ADVIA Centaur analyzer (Siemens Medical Solutions Diagnostics, Tarrytown, NY) of 3.05 and 3.19 mIU/L consistent with that proposed by the American Association of Clinical Endocrinologists (3 mIU/L).⁵

Since 2000, we have been using the same analyzer and the program GraphROC, which implements the “indirect” method proposed by Kairisto and Poola⁶ based on the Hoffmann method.⁷ In synthesis, the distribution is split and the mode (rather than the mean) of the hypothesized health-related distribution is forced to be the same as the mode in the original distribution. The health-related distribution consisted of 2 halves of 2 different gaussian distributions, with the same mode and mode frequency but different standard deviations. By using this procedure, we calculated the health-related limits using the 7,926 results obtained using the Advia Centaur analyzer (Siemens Healthcare Diagnostics, Deerfield, IL) between January 2002 and June 2007 in subjects older than 18 years among

the overall 21,862 TSH results downloaded from the laboratory information system of the laboratory of the Hospital Bambino Gesù, Rome, Italy. **Figure 1** shows the distribution of the results; the 2.5th and the 97.5th percentiles are 0.16 and 3.28 mIU/L, respectively, but a more accurate observation of the distribution allows detection of more effective limits, 0.5 and 3.05 mIU/L, respectively, which are very similar to those reported by Katayev et al.¹ Our results are consistent with American Association of Clinical Endocrinologists recommendation to lower the TSH upper reference limit that commonly ranges between 4 and 5 mIU/L. We disagree with the opinion of Horowitz² that the proposal by Katayev et al¹ is not a “better way” to establish reference intervals and should be simply considered a theoretical tool without practical applications.

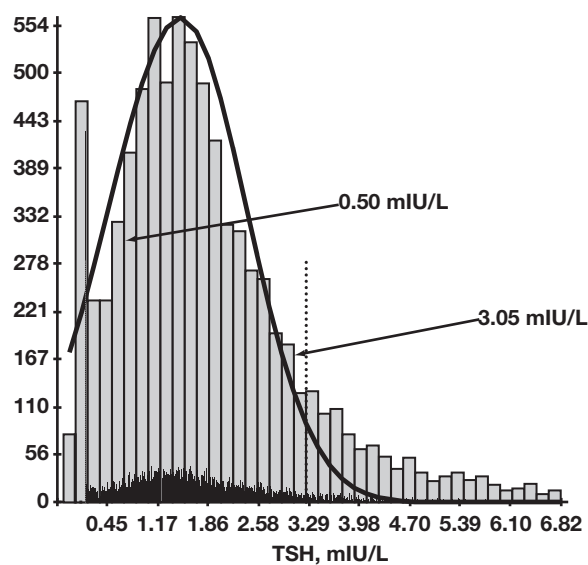


Figure 1 Health-related limits calculated using GraphROC from 7,926 thyroid-stimulating hormone results obtained in subjects older than 18 years retrieved from the laboratory information system of the Laboratory of the Hospital Bambino Gesù, Rome, Italy.

We think that the ability to calculate and validate the reference interval of such an important test as TSH in adults, pregnant women, children, and elderly people will help laboratory professionals to promote more effective interaction with clinicians. In our opinion, indirect methods yield reference limits for TSH not only consistent with those obtained using direct methods but also better than those proposed by the manufacturers.

References

1. Katayev A, Balciza C, Secombe DW. Establishing reference intervals for clinical laboratory test results: is there a better way? *Am J Clin Pathol*. 2010;133:180-186.
2. Horowitz GL. Estimating reference intervals [editorial]. *Am J Clin Pathol*. 2010;133:175-177.
3. Hoffmann RG. Statistics in the practice of medicine. *JAMA*. 1963;185:864-873.
4. Beckett G, MacKenzie F. Thyroid guidelines: are thyroid-stimulating hormone assays fit for purpose? *Ann Clin Biochem*. 2007;44:101-105.
5. Surks MI, Boucai L. Age- and race-based serum thyrotropin reference limits. *J Clin Endocrinol Metab*. 2010;95:496-502.
6. Kairisto V, Poola A. Software for illustrative presentation of basic clinical characteristics of laboratory tests: GraphROC for Windows. *Scand J Clin Lab Invest*. 1995;55(suppl 222):43-60.
7. Giavarina D, Dorizzi RM, Soffiati G. Indirect methods for reference intervals based on current data. *Clin Chem*. 2006;52:335-337.

The Author's Reply

I greatly appreciate the interesting discussion and support from our Italian colleagues regarding our article in the February 2010 issue of the *Journal*.¹ In our article, we pointed out that the indirect method for reference interval estimation has many more advantages than disadvantages and that the conventional direct sampling methods are prone to a number of limitations and false assumptions. I am pleased that Dorizzi et al reached the same conclusion and are successfully using the Hoffmann approach in their clinical and laboratory practice. It is not surprising that the TSH reference intervals calculated by us and Dorizzi et al are in perfect agreement. There is one very important factor in successfully performing the accurate and reproducible reference interval calculations by an indirect a posteriori study: the statistical method should be designed for the indirect method of laboratory test data analysis, and commonly used nonparametric and transformed parametric techniques may not be suitable in this case.

Since our article was written, we created the second version of the program that makes the calculations even more accurate by accounting for the biologic variation of analytes. After that improvement, we did many other reference interval calculations that proved over and over again the accuracy and viability of the computerized Hoffmann method.

Another important advantage that Dorizzi et al pointed out in their letter is that the indirect technique provides the

Romolo M. Dorizzi, MD

Corelab

Laboratorio Unico della Romagna

Cesena, Italy

Germana Giannone, BSc

Laboratorio

Paola Cambiaso, MD

Marco Cappa, MD

UOC Endocrinologia

Maurizio Muraca, MD

Laboratorio

Ospedale Bambino Gesù

Rome, Italy

unique opportunity to relatively easily establish reference intervals for difficult-to-sample populations such as pediatric and geriatric populations. For analytes like TSH, it is nearly impossible to find a completely disease-free population of reference subjects owing to a high prevalence of subclinical hypothyroidism. That is why most of the direct sampling studies for TSH reference intervals are lacking accuracy and tend to overestimate the upper reference limit.

In my opinion, providing the accurate population-based reference intervals is only one part of the clinically important information that laboratories provide for their clients. The other important part is help in the interpretative assessment of the test result value. Currently, many physicians are making clinical decisions based on the numeric value of the test result as if it is an absolute number (eg, if it is higher than the upper limit of the reference interval, it is abnormal). Laboratories need to educate their clients that any numeric test result may be affected by analytic and biologic variations of the given analyte. A "normal" result does not always exclude pathologic conditions, and an "abnormal" result does not necessarily mean that a condition is present. Providing the reference change value for the analyte together with a test result and a corresponding reference interval on the report may help in better clinical interpretation of laboratory data.²

It is time to think "out of the box" and accept what has been proved so many times: indirect a posteriori methods

work very well when the statistical technique is appropriate. Clinical laboratory medicine will broadly benefit from the simple and inexpensive reference interval calculation method that is accurate and reliable.

References

1. Katayev A, Balciza C, Seccombe DW. Establishing reference intervals for clinical laboratory test results: is there a better way? *Am J Clin Pathol*. 2010;133:180-186.

The Author's Reply

I commend Dorizzi and colleagues for their letter, in which they provide a detailed analysis of TSH data from their institution. It is heartening that the original article by Katayev and colleagues,¹ and, perhaps, my editorial² stimulated them to do this work.

Despite quoting me entirely out of context, Dorizzi and colleagues have done exactly what I recommended. What I wrote in my editorial was: "*If* [italics added] the reference intervals generated in this way are strikingly different from the reference intervals in use..., the laboratory needs to do some troubleshooting."

One of the points I made was that, since the incidence of subclinical hypothyroidism is higher in women than in men and increases with age, it would be interesting to look at reference intervals partitioned by sex and age. Dorizzi and his colleagues, like Katayev and colleagues,¹ have more than sufficient data to perform this analysis, as the original Hoffman article³ used just 500 points. Their data could tell us whether a TSH level of 3.5 mIU/L should be treated differently depending on whether the patient is a 55-year-old woman or a 40-year-old man.

Another point I addressed was that, even if the true 97.5th percentile for any partition is 3.2 mIU/L, it is not clear that laboratories should use that value as the upper reference limit. To the extent that TSH is commonly used as

References

1. Katayev A, Balciza C, Seccombe D. Establishing reference intervals for clinical laboratory test results: is there a better way? *Am J Clin Pathol*. 2010;133:180-186.
2. Horowitz GL. Estimating reference intervals [editorial]. *Am J Clin Pathol*. 2010;133:175-177.
3. Hoffmann RG. Statistics in the practice of medicine. *JAMA*. 1963;185:864-873.
4. Surks MI, Boucai L. Age- and race-based serum thyrotropin reference limits. *J Clin Endocrinol Metab*. 2010;95:496-502.
5. Baskin HJ, Cobin RH, Duick DS, et al. American Association of Clinical Endocrinologists medical guidelines for clinical practice for the evaluation of treatment of hyperthyroidism and hypothyroidism. *Endocr Pract*. 2002;8:457-469.

Alex Katayev, MD

Department of Science and Technology
Laboratory Corporation of America (LabCorp)
Elon, NC

2. Fraser CG. *Biological Variation: From Principles to Practice*. Washington, DC: AACC Press; 2001.

a screening test (ie, in asymptomatic people), using this value means that 2.5% of "normal" people will have "abnormal" values. Is that what we really want?

In their letter, Dorizzi and colleagues also suggest that a recent article⁴ endorses the recommendation made in 2002 by the American Association of Clinical Endocrinologists to lower the TSH upper reference limit to 3.0 mIU/L.⁵ In fact, the article is quite critical of that recommendation, echoing concerns published earlier.^{6,7} It recommends, as I suggest, that TSH reference intervals be partitioned because of known differences in distributions. It also emphasizes the fact that thyroid disease in asymptomatic people with TSH levels between 3.0 and 5.0 mIU/L is uncommon.

Quoting from my editorial: "As laboratory professionals, we need to think beyond simply providing central 95% reference intervals and more about what we want clinicians to do with the information we provide." I stand by that statement. I agree with Dorizzi and colleagues that indirect methods of validating reference intervals can be informative; where we disagree is on the manner in which we choose to use that information to enhance clinical practice.

Gary L. Horowitz, MD

Department of Pathology
Beth Israel Deaconess Medical Center
Boston, MA

6. Fatourechi V, Klee GG, Grebe SK, et al. Effects of reducing the upper limit of normal TSH values [letter]. *JAMA*. 2003;290:3195-3196.
7. Surks MI, Goswami G, Daniels GH. Controversy in clinical endocrinology: the thyrotropin reference range should remain unchanged. *J Clin Endocrinol Metab*. 2005;90:5489-5496.