Cautions in the Adoption of Common Reference Intervals

Transferability of test results for a given analyte has long been a goal of the laboratory community. Achievement of transferability will require development of standardized methods that have the least possible measurement uncertainty and are traceable to reference standards. Once accomplished, transferability will allow comparison of laboratory results regardless of the individual performing laboratory and will provide long-term stability of results, enabling longitudinal results to be followed for individual patients. Transferability remains an elusive goal, however. Analytical hindrances include the lack of established reference measurement systems for many quantities, lack of traceability of field methods to the reference system, and lot-to-lot variability in reagents and calibrators.

When no population-specific differences exist, the use of common reference intervals for transferable tests can be envisioned. Enthusiasm for the use of common reference intervals has grown since publication of the findings of the Nordic Reference Interval Project (1, 2), in which the feasibility of this approach was demonstrated. However, when Ichihara and coworkers tried to implement this approach in Asia (3), they found significant differences in laboratory test results between similar reference populations in 6 Asian cities.

In this issue of Clinical Chemistry, Ichihara and coworkers (4) have conducted a follow-up study to explore possible causes of the between-city differences they previously noted. As in their earlier study, these investigators removed the between-laboratory component of variability by analyzing deep-frozen specimens from all 6 cities in a single central laboratory. Their current study confirmed the presence of large between-city differences for several quantities. Although such differences may seem surprising in view of the Nordic Reference Interval experience, their confirmation in 2 independent studies is compelling and requires careful consideration.

Although specimen-handling and population-selection differences could explain the observed between-city differences, Ichihara and coworkers appear to have carefully designed their study to control for these factors, and therefore such explanations seem unlikely. It is much more likely that the observed differences had a biological basis or were due to environmental factors. Using a multivariate model, Ichihara and coworkers (4) found the observed differences persisted after controlling for the effects of age, body mass index, sex, and lifestyle factors including smoking, alcohol intake, diet, and exercise, and these investigators hypothesized that the differences they observed likely reflect genetic or environmental effects.

Whether the differences observed by Ichihara and coworkers are large enough to merit separate reference intervals for each population is an important question that cannot be fully answered. Unfortunately, most methods used to decide whether to partition reference intervals do not work well when more than 2 populations are being considered (2). Attempting to avoid this problem, Ichihara and coworkers developed a method based on estimates of SDs derived from nested ANOVA. Adapting a measure originally posed by Callum Fraser in analytical goal-setting for assay bias, the Ichihara criterion compared the ratio of SD for the component of variation under consideration (e.g., between-city, between-sex, etc.) to the SD corresponding to the between-individual component of variation. An empirical ratio threshold of >0.4 was used to decide the need for partitioning. Although this criterion appears to produce reasonable suggestions for partitioning in the dataset of Ichihara and coworkers, to be accepted on a more general basis it needs further testing and validation in independent datasets and comparison with other methods such as the clustering method recently described by Gellerstedt and Peterson (5).

Additional studies similar to the one conducted by Ichihara and coworkers are needed to fully document the prevalence and magnitude of differences in test results across other populations and what factors account for those differences. Such studies will be key in deciding whether common reference intervals should be applied. Until it can be shown that no population-related differences exist for a given test, individual laboratories will need to continue making an assessment of whether the population they serve is similar to the population used in defining a proposed common reference interval before adopting that interval for use (6).

Grant/funding Support: None declared.
Financial Disclosures: The author has served as a consultant to Abbott Diagnostics and as a statistical consultant to Bayer Diagnostics and Sanofi-Aventis and has performed product evaluations for Abbott Diagnostics.
James C. Boyd  
Department of Pathology  
Division of Laboratory Medicine/Clinical Laboratories  
P.O. Box 800214  
Charlottesville, VA 22908-0214  

References  


DOI: 10.1373/clinchem.2007.098228