Establishing Pediatric Reference Intervals: A Challenging Task

Ferruccio Ceriotti*

Defining pediatric reference intervals (RIs) is one of the most difficult tasks in our profession. The continuously changing physiology of growing children makes their laboratory values a moving target. To measure specific analytes in the pediatric age range, one has to consider the interindividual variation of not only the homeostatic set point but also the developmental stage of each individual. In this sense, the first year of life and puberty are the most critical periods. In addition, ethnic and behavioral differences may also cause variations. The only way to deal with such complexity is to gather large populations of reference individuals; however, the ethical problems involved in drawing blood from a healthy newborn or child further complicate this already demanding effort. Therefore, although one of the seminal reports on RIs was published in 1960 by pediatrician Albert John Schneider (1), it is not surprising that very few data on pediatric RIs exist.

It is against this rather dark backdrop that the work of Colantonio and coworkers (2) described in this issue of the Journal sheds new light; however, a careful evaluation of this study is necessary to fully understand its relevance. Ceriotti et al. (3) have listed 4 criteria that must be considered in the appraisal of any RI study: (a) the selection of reference individuals; (b) the study design; (c) the analytical quality of the data; and (d) the statistical treatment of the data.

The selection of reference individuals represents one of the major strengths of the study of Colantonio et al. These investigators applied an “a priori” selection scheme by enrolling volunteers or using leftover samples from well-characterized healthy newborns or from select outpatient clinics (for dentistry, bone fractures, plastic surgery). Such an approach is particularly relevant because, as indicated by the IFCC and by CLSI approved guideline C28-A3 (4, 5), it permits the characteristics of the reference individuals to be clearly defined. Owing to the difficulties of enrolling a sufficient number of individuals, many authors in the pediatric field have advocated the use of the so-called data-mining or indirect-sampling approaches. These techniques, initially popularized by Hoffmann (6), are based on the assumption (subsequently confirmed by observation) that the majority of laboratory results are “normal.” Statistical calculations based on an underlying presumed distribution of the laboratory results are used to extract the healthy individuals from the study population; however, such approaches are fraught with problems. First, there is no certainty that the assumed underlying distribution is correct, especially when one considers that it is skewed in the majority of cases. Second, these methods provide no control of preanalytical variables (fasting status, type of tubes, and timing before centrifugation and/or analysis) and no control of analytical variables (calibration status, quality control, and so forth). Finally, the results of studies that have attempted to validate such methods have been disappointing (7, 8). These methods cannot, therefore, be endorsed to establish new RIs.

With regard to study design, Colantonio and coworkers (2) applied uniform criteria for sample procurement. They went to great lengths to recruit a sufficient number of individuals (>2000 children evenly divided between males and females), particularly in the younger age groups <1 year of age, for which there were about 1000 individuals.

With respect to the analytical quality of their data, Colantonio et al. reported all information regarding the analytical quality of the methods in their Supplemental Table 1. Although the characteristics and performance of the analyzer used are clearly described, the information on metrologic traceability is only indirect. They relied completely on the manufacturer’s indications and did not provide any direct confirmation of the traceability of the obtained data. It is important to stress the relevance of analytical quality, because the lack of standardization adds an additional layer to the already complex situation of RIs that vary according to age, sex, and ethnic group. Two aspects need to be considered: the metrologic traceability and the endorsement of non–state-of-the-art methods. Demonstrating traceability via comparison with higher-order methods or materials (9, 10) gives universal credibility and value to laboratory test results, independent of the analyzer or methods used. Any laboratory using a traceable method can adopt the intervals developed else-
where for a similar traceable method as long as there are no biological differences in the underlying populations that would lead to different RIs. Colantonio and coworkers plan to carry out transferability experiments to provide pediatric RIs for other analytical systems; however, this activity would not be necessary if they could have demonstrated the accuracy of their results by comparing them with commutable reference materials or with reference methods. In 1994, Norbert Tietz published an opinion article entitled, “Accuracy in Clinical Chemistry—Does Anybody Care?” (11). His plea seems to have been ignored if, 18 years later, we are still proposing method-dependent RIs for common clinical chemistry analytes. Standardization and harmonization are not just the fixation of some foolish metrologists. They represent ethical issues, as Bossuyt et al. have pointed out (12), and are very necessary to improve patient safety and clinical effectiveness (13).

A related problem of the Colantonio et al. study, in my opinion, is its seeming endorsement of non–state-of-the-art methods. To make progress in our discipline, we should promote only the best analytical principles. It is well known that changing habits and traditions takes time, and if we allow old-fashioned methods to survive (and developing a well-defined RI is certainly a good way of doing that), it will be difficult to convince people to abandon them. This consideration applies in particular to 3 analytes: creatinine, albumin, and transaminases. For creatinine, a recent report (14) showed that enzymatic methods, although not perfect, are clearly superior to Jaffe methods. The superiority of enzymatic methods is particularly true for pediatric values, for which nonspecific interferences may account for >50% of the signal, as is shown for the younger age groups (15 days to 4 years) in Table 1 of the Colantonio et al. report (2). For this reason, the IFCC Committee on Reference Intervals and Decision Limits decided to endorse only creatinine RIs obtained with a traceable enzymatic method (3). Not surprisingly, the enzymatic creatinine results of Colantonio et al. (2) are very similar to those proposed by the IFCC committee. This similarity in the results supports the concept that with proper analytical standardization—and in the absence of ethnicity-related modification—common RIs can be applied. For albumin, the data presented show large differences between RIs obtained via the 2 method principles (brom cresol green and brom cresol purple). Both methods are well known to be poorly specific (15), and the claimed traceability of these methods to the certified reference material ERM-DA470 is useless, because these methods do not measure albumin specifically, thus causing a break in the traceability chain. Albumin is a relevant parameter for various clinical presentations, ranging from nephrotic syndrome to nutritional assessment, especially in pediatrics, but it is also widely used to recalculate results for total calcium. Endorsement of such nonspecific measurement methods when more specific immunochromatographic approaches exist is, in my opinion, inappropriate. A third note concerns transaminases. Here, too, the IFCC is struggling to promote standardization, and nontraceable methods, such as those without pyridoxal phosphate, should be abandoned. Colantonio et al. (2) provided additional evidence for this fact in that lower alanine aminotransferase RIs for East Asian and Southeast Asian people occurred only with assays without added pyridoxal phosphate, thus suggesting that the lower values might be related to lower vitamin B₆ concentrations in these populations. Such data further support the use of optimized methods only.

In their statistical treatment of the data, Colantonio et al. applied the correct statistical approaches and conducted a careful statistical analysis of their data. Some may question their proposal of basing their definition of sex-related RIs only on a statistical criterion. For instance, the investigators used this criterion in proposing a sex-related RI for direct bilirubin (age group, 13–19 years), but the width of the 90% CI of the upper reference limit for girls is approximately 40% of the entire interval. According to CLSI guideline C28-A3 (5), an uncertainty this high (a CI >20%) points to the need for a larger reference population. An evaluation of the clinical relevance of the differences between the groups might also have reduced the number of analytes recommended for partitioning when such partitioning was based only on a statistical criterion. For analytes in which the RI changes continuously with age, such as creatinine, inorganic phosphate, and alkaline phosphatase, some of the recommendations regarding partitioning also appear debatable.

Colantonio and coworkers are to be commended, however, for making their entire database available online to readers, thus giving everyone an opportunity to analyze the data differently or to use more sophisticated statistical techniques. In particular, I am considering the use of fractional polynomials and exponential transformation (3, 16) to derive age-specific RIs for all of the analytes that show continuous age-related changes.

Setting aside all the issues related to assay standardization, the study of Colantonio and coworkers (2) is a milestone that represents a quantum leap forward in helping close gaps in pediatric RIs, especially for the first year of life, and even more so for the first few days of life. Given the large size and careful execution of this study, it surely represents a landmark achievement in pediatric clinical chemistry and provides results that will serve the field well into the future.
Author Contributions: All authors confirmed they have contributed to the intellectual content of this paper and have met the following 3 requirements: (a) significant contributions to the conception and design, acquisition of data, or analysis and interpretation of data; (b) drafting or revising the article for intellectual content; and (c) final approval of the published article.

Authors’ Disclosures or Potential Conflicts of Interest: Upon manuscript submission, all authors completed the Disclosures of Potential Conflict of Interest form. Potential conflicts of interest:

Employment or Leadership: None declared.

Consultant or Advisory Role: F. Ceriotti, Siemens Italia.

Stock Ownership: None declared.

Honoraria: None declared.

Research Funding: None declared.

Expert Testimony: None declared.

References