In this issue: Recent Advances in Pediatric Laboratory Medicine

The Journal of the International Federation of Clinical Chemistry and Laboratory Medicine

eJIFCC

Pediatric reference intervals for biochemical markers: gaps and challenges, recent national initiatives and future perspectives

Houman Tahmasebi[1,2], Victoria Higgins[1,2], Angela W.S. Fung[2], Dorothy Truong[2], Nicole M.A. White-Al Habeeb[2], Khosrow Adeli[1,2]

[1] *CALIPER Program, Pediatric Laboratory Medicine, The Hospital for Sick Children, Toronto, ON, Canada*
[2] *Department of Laboratory Medicine & Pathobiology, Faculty of Medicine, University of Toronto, ON, Canada*

ARTICLE INFO

Corresponding author:
Khosrow Adeli, Ph.D., FCACB, DABCC
Clinical Biochemistry
Pediatric Laboratory Medicine
The Hospital for Sick Children
555 University Avenue
Toronto, ON, M5G 1X8
Canada
E-mail: khosrow.adeli@sickkids.ca

Key words:
pediatric, children, adolescent, reference interval, CALIPER

ABSTRACT

Reference intervals provide valuable information to medical practitioners in their interpretation of quantitative laboratory test results, and are critical in the assessment of patient health and in clinical decision-making. The reference interval serves as a health-associated benchmark with which to compare an individual test result. While the concept of reference intervals and their utility appear straightforward, the process of establishing accurate and reliable reference intervals is considerably complex and involved. Currently, many pediatric laboratory tests are inappropriately interpreted using reference intervals derived from either adult populations, hospitalized pediatric populations, or from outdated and/or inaccurate technology. Thus, many pediatric reference intervals used in diagnostic laboratories are incomplete and may be inappropriate for clinical use. The use of inappropriate reference intervals impacts clinical decision-making and has potential detrimental effects on the quality of patient healthcare including misdiagnosis, delayed diagnosis, inappropriate treatments, and patient risk. These are critical gaps in pediatric healthcare and it is imperative to update and establish appropriate reference intervals for pediatric populations based on specific age- and

sex-stratifications. In the present review, specific issues, challenges and deficiencies in pediatric reference intervals for biochemical markers will be discussed. Early studies using hospitalized patients will be examined, followed by a review of recent national and global initiatives on establishing reference intervals from healthy pediatric population. We will highlight the achievements and milestones of the Canadian CALIPER project, including the establishment of a comprehensive biobank and database which has addressed several of these critical gaps. CALIPER's mandate is to establish and provide comprehensive, up-to-date pediatric reference intervals to all biochemical markers of pediatric disease. CALIPER has also begun knowledge translation initiatives to disseminate its data via peer-reviewed publication, an online database, and a smartphone application to allow greater access to CALIPER pediatric reference interval data. Finally, limitations, future perspectives and harmonization of pediatric reference intervals to improve pediatric diagnostics in Canada and worldwide will be discussed.

❖❖❖❖❖

INTRODUCTION

The measurement of disease biomarkers in clinical laboratories are used to screen, diagnose, and monitor a wide range of medical conditions. To interpret these laboratory test results, physicians compare patient test results with a reference interval, defined as the typical values derived from a healthy population [1]. Statistically, reference intervals are defined as the limiting values denoting a specified percentage (typically central 95%) of values from an apparently healthy reference population with 90% confidence. In the central 95% distribution model, the reference limits are determined by calculating the 2.5^{th} and 97.5^{th} percentiles of test results [2]. In this case, a total of 5% of values may be interpreted as abnormal or higher risk of disease and require further follow-up and investigation. In other cases, the reference interval definition may be modified, where a different percentile may be used, or either the upper or lower limit may be used if only a one-sided distribution is clinically significant [3]. Reference intervals are therefore a fundamental tool in test result interpretation and serve as a benchmark for health status [4].

Accurately established reference intervals are critical to clinical decision-making as a lack or inappropriate use of reference intervals may lead to adverse consequences including misdiagnosis, patient risk, inappropriate treatment, and/or higher healthcare costs, all of which impact the overall quality of patient healthcare. Many current reference intervals were determined decades ago with older and less accurate laboratory instruments and testing methodologies. Furthermore, as instrumentation and reagents are upgraded, reference intervals are not always appropriately updated [4]. With rapid advances in technology, there is also a lack of data on novel, emerging disease biomarkers. The lack of standardization and harmonization of assay methodology further contributes to variations of established reference intervals for both adults and pediatrics. These variations create confusion in the interpretation of test results for the same patient whose specimen may be tested in different laboratories.

Additionally, there are specific challenges related to the determination of pediatric reference intervals. The same reference intervals are sometimes used to interpret test results for both adults and children. However, *children are not small adults*; children have significant differences in physiology and metabolic state, physical size, organ maturity, bodily fluid compartments, and immune and hormone responsiveness when compared to adults [4]. Most notably, dynamic physiological changes, growth

and development profoundly influence biomarker concentrations. For example, it is well known that sex hormones, growth hormones, and bone alkaline phosphatase vary with a child's age and development [4]. Establishing appropriate pediatric reference intervals involves overcoming specific challenges mostly related to volume and quantity of healthy pediatric samples. Complex physiological factors may also necessitate the separation of reference interval (called partitions) to be age- and sex-specific, requiring a greater number of reference samples. Moreover, children also suffer from, or are more susceptible to, diseases that differ from adults, requiring unique or new biomarker reference interval determinations. To complicate this further, some of these diseases may be genetically inherited and occur at a lower frequency, posing challenges in acquiring adequate sample size for statistical calculations. Gap analyses of pediatric reference intervals have identified four major critical areas in pediatrics including bone markers [5], cardiovascular disease and metabolic syndrome risk markers [6,7], hormones of thyroid and growth hormone axes [8], and inborn errors of metabolism [9]. Data from published reference interval studies often suffer from limitations in design, small sample sizes, and the use of hospitalized patients [3]. Currently pediatric clinicians and laboratorians depend on scattered information and incomplete data from published (scientific journals and textbooks) and unpublished (hospital, private, reference laboratories) sources in laboratory test result interpretation. There is an urgent need to establish and update reference intervals for all populations and particularly pediatric populations.

While the concept and utility of reference intervals appear straightforward, the process to accurately establish or verify reference intervals is quite complex. The Clinical Laboratory Standards Institute (CLSI) and International Federation of Clinical Chemistry and Laboratory Medicine (IFCC) provide a guideline (C28-A3) on how to define, establish, and verify reference intervals [10]. When judging the validity of reference intervals, several integral factors should be considered including the use of a large healthy population, appropriate inclusion/exclusion criteria, awareness of physiological, pre-analytical and analytical factors that affect analyte concentrations, careful outlier exclusion, and appropriate statistical analysis.

The quality of reference intervals depends on the recruitment of a large number of healthy reference individuals within the age group(s) of interest and well-defined inclusion/exclusion criteria. Laboratories commonly verify and adopt a reference interval provided by a manufacturer or transfer a pre-existing reference interval with 20 healthy samples. One problem with verification and transference of pre-existing reference intervals is the quality of the original reference interval. Reference intervals supplied by manufacturers often lack information regarding the study's sample size, age, sex, and ethnic distribution, and often do not include pediatric populations. The success of transference also depends on comparability of the reference populations and analytical methods between the donor and receiving laboratories. To establish robust reference intervals, a *de novo* reference interval study with at least 120 healthy subjects per partition is needed. This is typically done for novel biomarkers. For a pediatric reference interval study that is divided into 5 age groups, it would require 600 healthy subjects or 1200 healthy subjects if also sex stratified [4]. This undoubtedly requires an enormous amount of effort, resources, time and cost. There are many inter-related variables that contribute to the validity of reference intervals including ethnic composition, geographic factors (climate), diet and food preferences, and lifestyle factors. Laboratories that serve ethnically

diverse population should evaluate and determine whether a single reference interval is valid or whether ethnic-specific reference intervals are necessary [4]. Inclusion and exclusion criteria require documentation of covariates (e.g. age, sex, and ethnicity) and assessment of general health using surveys. Pre-analytical factors include standardization of subject preparation (fasting, diurnal variation), specimen collection (posture, sample volume, tourniquet time), and specimen-handling, transport, and storage conditions [3]. Analytical factors include a detailed description of methodology including calibration traceability, imprecision, limit of detection, linearity, analytical measuring range, interferences, and variability factors. Outliers should be excluded through robust statistical methods including Dixon's [11] or Tukey's method [12] for outlier exclusion, parametric or non-parametric analysis, and covariate analysis. Overall, major challenges in the determination of pediatric reference intervals include the recruitment of a sufficiently large healthy pediatric population, the need for parental consent, and difficulties with pre-analytical variables (i.e. sample collection, sample volume, and a specialized phlebotomist with pediatric experience) [3].

In this review, the important milestones in addressing these gaps and challenges in pediatric reference interval determination will be highlighted. The early approaches to establishment of pediatric reference intervals derived from hospitalized patients as well as several recent global initiatives to close the gaps in pediatric reference intervals based on recruitment of healthy children will be discussed. The significant progress that the Canadian Laboratory Initiative on Paediatric Reference Intervals (CALIPER) project has achieved through pilot studies, the establishment of age- and sex-specific pediatric reference intervals for more than 100 biomarkers, and various sub-studies will be reviewed. Lastly, limitations, future perspectives, and goals for pediatric reference interval distribution and harmonization will also be discussed.

INDIRECT APPROACHES TO REFERENCE INTERVAL ESTABLISHMENT

Ideally, reference methods are determined based on a healthy population using a direct *a priori* approach where individuals are specifically selected for the study. However, the recruitment of many healthy individuals can be quite challenging, particularly in the pediatric population, as it is very costly and time consuming. Additionally, it is difficult to define an individual as "normal" or "healthy" as the potential for existing subclinical issues is unknown. Thus, indirect methods, also known as data mining, can be quite useful. The indirect approach uses existing data to establish reference intervals by identifying an acceptable reference population retroactively. One example of using the indirect approach is to identify a group of healthy individuals from hospital in- and/or out-patient populations to calculate reference intervals. In addition to having data readily available through the laboratory information system (LIS), indirect methods remove the need to recruit healthy individuals. This method represents a strong alternative to cases where laboratory markers are measured at high volume in community outpatient clinics and therefore should include a relatively healthy population.

The Hoffmann method proposed by Robert G. Hoffmann in 1963 [13], uses an indirect *a posteriori* method to determine reference intervals using available test results from hospital-based data from in- and out-patients. This approach makes two assumptions: 1) values obtained for a specific analyte follows a Gaussian distribution; and 2) majority of measurements made in the hospital represent normal individuals. Reference intervals are determined by plotting

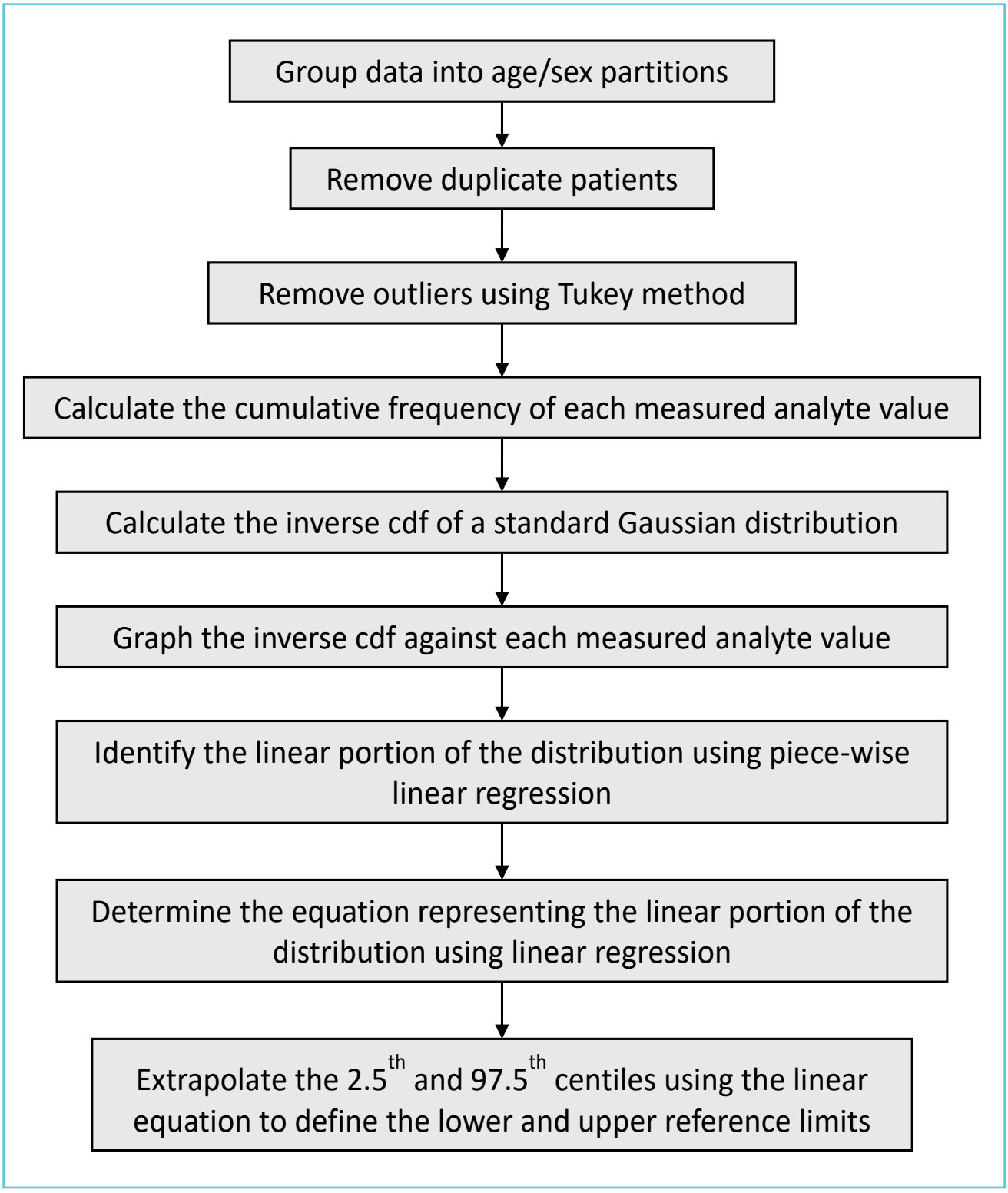
the cumulative frequency of a result against the analyte value. The linear portion of the resulting graph, centered on the 50th percentile, is chosen thereby giving these values the greatest weight. By extrapolating the linear portion of the graph, the 2.5th and 97.5th centiles are calculated, representing the normal range of values, if the assumptions are held true. Hoffmann used this approach with a relatively small number of patients (n=60) for glucose as a proof-of-concept. The application of this method using computer programs has eliminated much of the subjectivity of the Hoffmann approach and allows for the analysis of a very large number of samples.

Although the Hoffmann approach is highly cited, few authors have applied the Hoffmann method in their calculations. One notable exception is the study led by Steven Soldin at the Children's National Medical Center in Washington, DC, USA. Reference intervals were calculated using hospital-based data collected from patients from 1 day to 18 years old. This group reported reference intervals for several analytes including fertility hormones, thyroid hormones, adrenal hormones, and growth hormones [14]. As expected, age and sex-specific differences were identified for FSH, LH, estradiol, 17-hydroxyprogesterone, DHEA and testosterone. Sex-specific differences were also found for serum iron, homocysteine, IGF-1 and IgE levels. Lastly, age-related differences were found for free T4, TSH and 25-hydroxy vitamin D levels. Levels of 25-hydroxy vitamin D showed seasonal variation, with higher levels found in the summer months. This work, however, had some limitations. Much of the data was collected on a small number of Caucasian hospitalized patients and may not accurately reflect levels in a healthy multicultural population. Also, semimanual application of the Hoffman analysis of data, added subjectivity to the calculations.

To determine the accuracy of the Hoffmann method, Katayev *et al.* applied a computerized indirect Hoffmann approach to retrospectively determine reference intervals for hemoglobin, creatinine, calcium and thyroid stimulating hormone (TSH) and compared to previously published reference intervals [15]. Their method calculated cumulative frequencies of each test result and performed least squares analysis, applying a best-fit equation to the linear portion of the data. The statistical difference between the two methods was determined by calculating the reference change value (RCV), a factor representing within individual variation and analytical variation. They found that there was no statistically significant difference between the reference intervals calculated using the Hoffmann method and published reference intervals.

More recently, as part of the CALIPER initiative, Shaw *et al.* compared pediatric reference intervals calculated for 13 analytes (calcium, phosphate, iron, ALP, cholesterol, triglycerides, creatinine, direct bilirubin, total bilirubin, ALT, AST, albumin and magnesium) from hospital-based data using the Hoffmann approach (Figure 1) to reference intervals calculated by CALIPER [16]. They sought to determine; 1) whether the hospital-based reference intervals fell within the 90% confidence intervals calculated by CALIPER; 2) whether the hospital-based reference intervals fell within the RCV for each analyte; 3) by completing a reference interval validation study using reference samples from healthy children. None of the reference intervals calculated using the Hoffmann approach fell within the 90% confidence intervals calculated by CALIPER. When the RCV was used for comparison, only some of the Hoffmann calculated reference intervals fell within the RCV range calculated (creatinine 15 days–<1 year old and all phosphate partitions with the exception of 0–<14 days old). Given the wide biological variation in some analytes, it

Figure 1 Schematic of statistical approach to calculate reference intervals from hospital-based data based on the Hoffmann approach used in *Shaw et. al.* [16]



is not surprising that more of the Hoffmann-calculated reference intervals fell within the RCV than fell within the 90% confidence intervals calculated by CALIPER. Finally, all reference intervals were validated according to CLSI guidelines except for ALP (13–<15 year old male; 15–<17 year old male and female), creatinine (15–19 year old male), and iron (14–<19 year old female). Validation data for some age and/or sex partitions was not performed due to insufficient numbers. In this study, reference intervals calculated using the modified Hoffmann approach, were much wider than those calculated by CALIPER, suggesting that the use of hospital-based data may be limited in pediatrics, especially from a tertiary care center.

The main limitations of the Hoffmann approach are based on the assumptions made for the analysis. The method assumes that majority of measurements made in the hospital represent normal individuals, yet when using hospital-based data, the true proportion of unhealthy individuals included in the data set is unknown [17]. This is especially problematic when determining reference intervals for low volume tests that include a selective patient population, as the fewer available samples would likely skew the data due to the inclusion of unhealthy subjects [18]. Additionally, this approach assumes that the values obtained for a specific analyte follows a Gaussian distribution of healthy results which may not be true in all cases [13].

NATIONAL AND GLOBAL PEDIATRIC REFERENCE INTERVAL INITIATIVES BASED ON HEALTHY COHORTS

In contrast to a hospitalized patient population, the ideal reference population to establish population-based reference intervals is a group of well-defined individuals that are similar to the target patient in all respects other than the disease condition under investigation. Ultimately, the quality of a laboratory measurement depends on the quality of the reference interval that the value is compared with. In turn, the quality of reference intervals heavily depends on the selection and recruitment of a large number of appropriate reference individuals [2,19]. Although this is a challenging task, several national and international initiatives have recognized the critical gaps in pediatric reference intervals and the need to establish ranges that are robust and appropriate for the pediatric population (Table 1) [20,21].

The German Health Interview and Examination Survey for Children and Adolescents (KiGGS), performed by the Robert Koch Institute (RKI), is a national initiative aimed at providing information on several health aspects of German children and adolescents [22,23]. KiGGS has successfully determined reference intervals for numerous serum and urine laboratory biomarkers using healthy pediatric samples [20,23]. The baseline KiGGS study collected whole blood, serum and urine samples in addition to personal information (i.e. age, sex, socioeconomic status, geographical location, community population, and immigration status) from 17,641 pediatric subjects (aged 0-17 years) recruited from May 2003 to May 2006 [24]. Three main categories of analytes were examined depending on their relation to one of the following public health interests: nutrition, risk of non-communicable disease and immunization. 43 analytes were analyzed, for some of which median and 90% reference intervals were calculated, such as total, LDL and HDL cholesterol, triglycerides, and calcidiol, based on age, sex and other subjective information [24]. In an additional KiGGS study [25], the median, 25th and 75th percentile values for serum thyroid hormones and serum lipids were determined in 12,756 subjects ≥3 years of age. This study also examined biomarker relationships and identified a positive relationship between TSH and serum lipid biomarker

concentrations, except for HDL, even after adjusting for smoking status, age and sex [25].

The Nordic Reference Interval Project (NORIP) was established in 1998 to determine Nordic-specific reference intervals for common blood analytes [26,27]. Some of the general inclusion criteria included feeling healthy, aged ≥ 18 years, and individuals who are not pregnant, ill or hospitalized. Personal information obtained from subjects included age, sex, body mass index (BMI), cigarette and alcohol usage, ethnic origin and physical activity [28]. Serum, plasma and whole blood samples were obtained from 3,036 healthy adults (aged ≥ 18 years) in 102 Nordic laboratories in Denmark, Finland, Iceland, Norway and Sweden [21,26]. Reference intervals (central 95%) for 25 common serum biomarkers, including enzymes, were calculated based on IFCC recommendations, with 90% confidence intervals around upper and lower limits. Data were also partitioned based on age, sex and blood sample (plasma or serum). Interestingly, a 2013 study analyzed 21 blood biomarkers based on 1421 (596 males and 825 females) healthy Danish pediatric subjects between ages 5-20 who participated in The COPENHAGEN Puberty Study (from 2006-2008) [29], and the results were compared to the NORIP (adult) findings. This study used nonparametric statistics to calculate the central 95% RIs, and confidence intervals of upper and lower limits; the RIs were calculated for both sexes and 6 age groups. This paper compared the results from the oldest pediatric age group to the youngest NORIP adult results [29]. Although many of the results were similar, there were some differences between the two studies, some values being higher and others being lower in the pediatric study compared to the NORIP adult study. However, discrepancy in some of these analytes, such as alkaline phosphatase, lactate dehydrogenase and creatinine, was explained to be normal as these values are expected to increase or decrease by age (from adolescence to adulthood).

Children's Health Improvement through Laboratory Diagnostics (CHILDx) program has been establishing pediatric RIs based on a healthy cohort (6 months to 17 years) in Utah, United States since 2002 [30-32]. This group has determined RIs for a variety of analytes, such as vitamins [33], enzymes [32], hormones [34], coagulation tests [30] and bone markers [35]. A 2005 CHILDx paper describes reference intervals for seven common coagulation tests that were determined based on 902 healthy pediatric participants with ages ranging from 7 to 17 years [30]. Median, 95% reference intervals and 90% confidence intervals around upper and lower limits were determined for each parameter. The results were reported based on 3-year age groups. Another CHILDx study in 2011 focused on analyzing the serum levels of seven analytes, consisting of enzymes, prealbumin and uric acid, based on a healthy cohort of 1765 children and adolescents [32]. Participants were divided into 3-year age groups (except for the 6m-2y age group), and mean, median, and statistical difference between age and sex groups were determined. Subsequently, central 95% age-specific and sex-specific (sex differences for about one-third of analytes) RIs were determined using nonparametric statistics.

United States National Health and Nutrition Examination Survey (NHANES) evaluates the health status of pediatrics and adults in the U.S. population by receiving laboratory and interview details from their cohort, while recruiting thousands of additional participants each year [36-41]. The NHANES study examined the effects of age, sex, BMI, socioeconomic status and ethnicity on various health parameters, including biomarkers. For example, a study published in 2000, looked at the upper 95th percentile limit of C-reactive protein (CRP) concentrations in a sample size of more than 22,000 healthy

pediatric and adult individuals (from NHANES III) based on age, sex and ethnicity [42]. Females generally had a higher concentration compared to male counterparts. They also demonstrated that Caucasians and Hispanics have similar upper limit CRP values, compared to black adults. CRP levels were also higher in older adults compared to children. Furthermore, in 2004, NHANES III data were used to determine reference intervals for whole blood count based on approximately healthy 25,000 subjects ranging from pediatric to geriatric age groups (aged 10 to >75 years), partitioned based on age, sex and ethnicity (Mexican, white and black) [40]. Age-, sex- and/or ethnic-related differences were observed in some of the analytes. In 2012, NHANES data from 6062 healthy pediatric (ages 2-19 years) individuals were used to establish RIs for the same 3 ethnic groups (Mexican, white and black) [43]. In addition, 95% RIs, geometric means and statistical difference between the ethnic groups were determined for vitamins and lipids (total, LDL- and HDL-cholesterol). RIs were partitioned based on sex, age and other factors; this is interesting as it allowed for both genetic and environmental comparisons between the three ethnic groups. Overall, NHANES studies, along with other studies, highlight the importance of partitioning reference intervals based on sex, age and ethnicity since reference values are highly influenced by these factors [40,43,44].

The Lifestyle of Our Kids (LOOK) program is a longitudinal study that was initiated in Australia in order to study healthy children and adolescents, and the effects of physical activity on their health outcomes [45]. In one of the most significant studies done by this group, central 95% RIs and medians of 37 blood analytes were calculated from a sample of 852 healthy individuals [45]. Sex-specific as well as age-specific RIs were calculated for ages 8, 10 and 12 based on measurements in 2005, 2007 and 2009, respectively. Interestingly, they compared their results with those from other groups, including CALIPER RIs for ferritin, CRP, cholesterol, TSH and magnesium. In a 2012 paper, LOOK data from 854 pediatric individuals was used to calculate blood NT-proBNP RIs [46]. Median, 95% RIs and 90% confidence intervals were calculated based on sex (male, female and combined) and for the same 3 age groups (ages 8, 10 and 12). In addition, pairwise comparisons in concentrations between the sex groups and the 3 age groups were conducted with the use of Mann–Whitney U test. While no significant difference was observed between the sex groups of the same age, concentration differences were shown to be statistically significant between certain age groups.

Australasian Association of Clinical Biochemists (AACB) Committee for Common Reference Intervals and AACB Harmonisation Committee have been promoting the adaptation of common reference intervals, based on the values used in hospitals in Australia and New Zealand [47,48]. In a 2014 preliminary publication, reference intervals used by each laboratory, along with results of freshly obtained serum for various analytes, were reported by each of the 123 laboratories, partitioning analytes by sex when applicable [48]. This was used to identify differences in reference intervals used by each laboratory and analytical methods between the laboratories. Interestingly, it was demonstrated that for majority of analytes, reference interval variation was greater than analytical (i.e., measured sample) variation between the laboratories. Linear regression was also used to compare measured results against reported upper and lower reference limits reported by each laboratory [48]. Finally, the AACB group determined the location of the measured sample's value with regards to each laboratory's reference range, determining its relative location from upper and lower limits [48]; the

results were then compared between the laboratories. In another study published within the same year, harmonized 95% reference intervals were determined for 11 analytes in adults and 9 in pediatrics based on healthy subjects, while partitioning further based on age in pediatrics and sex in both pediatrics and adults (aged ≥ 18 years), where applicable [47]. Throughout the common reference interval selection process, many workshops and validation processes were held to identify analytical accuracy and biases, consider clinical importance, and compare the results to other studies (e.g. NORIP and Aussie Normals). Similar to other studies, this study highlights the importance of considering between-instrument/analytical method differences and the significance of local validation [44,49-51].

THE CANADIAN LABORATORY INITIATIVE ON PEDIATRIC REFERENCE INTERVALS (CALIPER) PROJECT

The Canadian Laboratory Initiative on Pediatric Reference Intervals (CALIPER), is aimed at creating a comprehensive database of pediatric reference intervals to disseminate for use by other pediatric centres nationally and globally. This initiative was launched through the Pediatric Focus Group of the Canadian Society of Clinical Chemists (CSCC) and is a collaborative effort amongst pediatric institutions across Canada. CALIPER is an *a priori* prospective study that has been recruiting thousands of healthy community children and adolescents and establishing age- and sex-specific reference intervals for many routine and specialized biochemical markers.

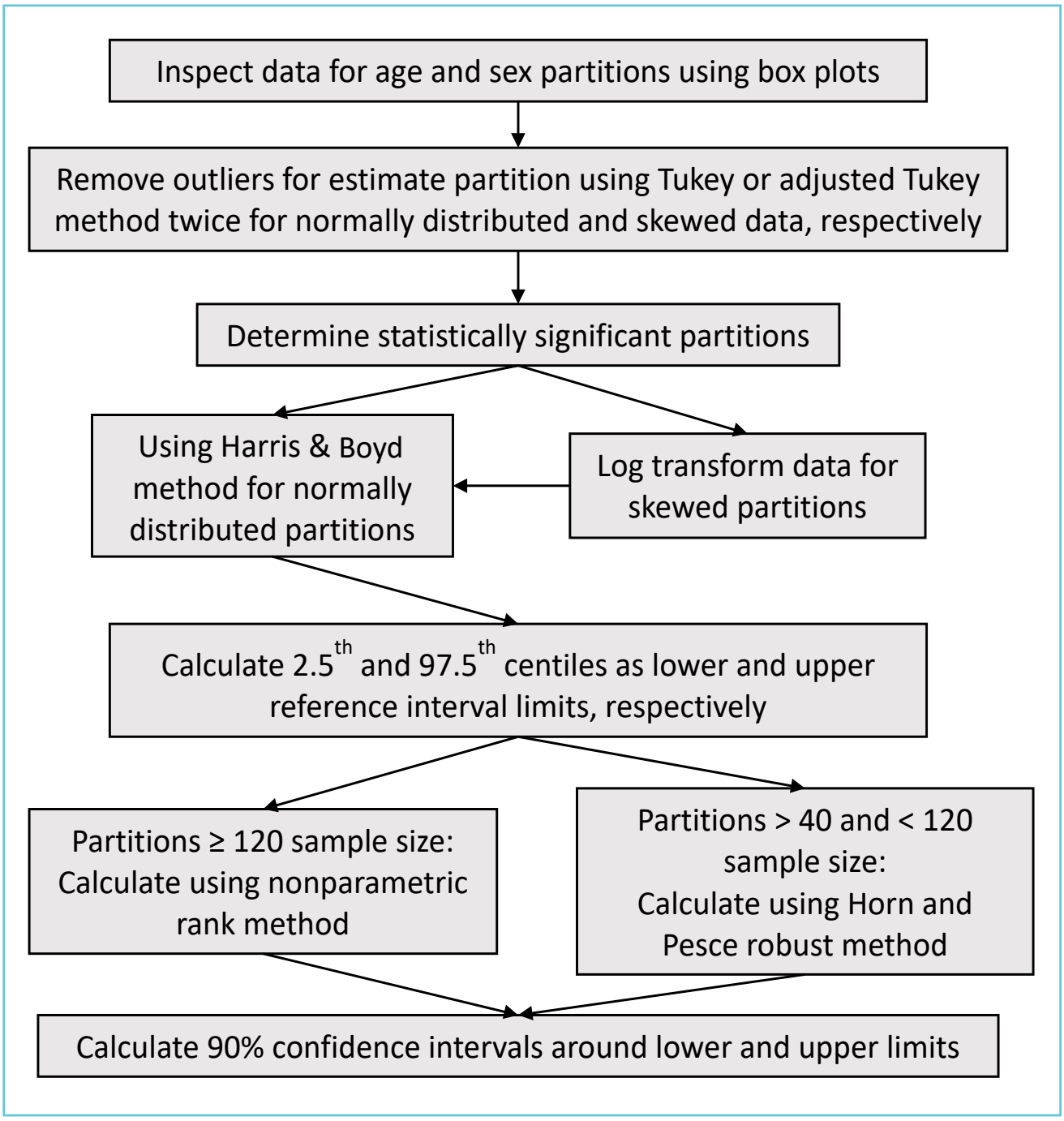
The early stages of CALIPER involved extensive planning to establish standardized procedures for pre-analytical, analytical and post-analytical aspects of the project. For example, blood collection, sample and statistical analysis were standardized to ensure consistency among different collection sites. A CALIPER team consisting of experienced research coordinators, project coordinators and volunteers was formed to help with promotion of the CALIPER campaign and recruitment of participants. Trained phlebotomists with expertise in pediatric sampling were also recruited to ensure ease and efficiency of collection.

Initially, preliminary pilot studies were conducted by CALIPER to refine project logistics and gain experience in sample collection, sample analysis, as well as statistical analysis and establishment of reference intervals. The initial CALIPER pilot studies included 2,809 serum and plasma specimens from apparently healthy and metabolically stable children from outpatient clinics. Overall, these initial studies analyzed over 50 chemistry and immunoassay biomarkers on the Abbott ARCHITECT ci8200 analyzer [52]. Using these data, CALIPER generated preliminary reference intervals according to CLSI and IFCC C28-A3 guidelines. Age and sex-specific reference intervals were established for five age groups. This was an important first step for CALIPER and formed the basis for more projects to come. However, as outlined by CLSI/IFCC C28-A3 guidelines, establishment of reference intervals should include recruitment of at least 120 healthy individuals per partition. Initial CALIPER pilot studies recruited apparently healthy children from outpatient clinics. According to CLSI guidelines, this was less than ideal since underlying disease in children from outpatient clinics could confound interpretation and establishment of reference intervals. Furthermore, a sufficient sample size of 120 patients per partition was not feasible in the initial CALIPER preliminary studies.

To address this limitation, subsequent CALIPER reference interval studies recruited healthy children and adolescents from Toronto and the Greater Toronto Area to establish a biobank

of healthy serum samples. With the help of the CALIPER team, over 9000 serum samples were collected from healthy children and adolescents allowing for appropriate sample size of 120 participants per partition, as per CLSI guidelines (Figure 2). Furthermore, to ensure adherence to CLSI guidelines and only inclusion of healthy participants, recruitment was limited to those without a history of chronic illness or acute illness within the previous month and without current

Figure 2 **Schematic of general statistical approach to calculate reference intervals in CALIPER study based on a healthy reference population [53,55]**

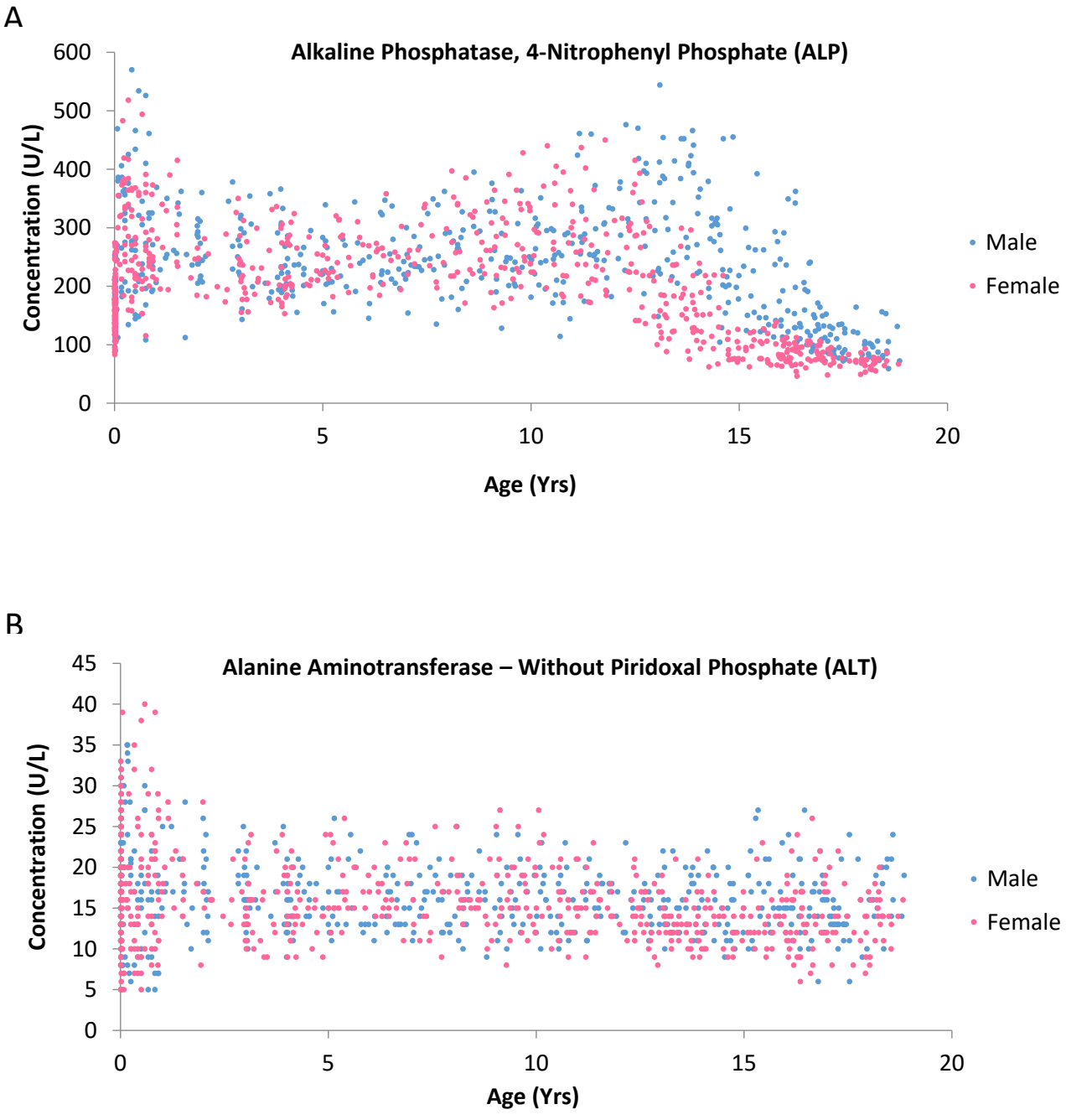


use of prescribed medication. Diverse collection sites, including community centres, day cares, churches and schools, further strengthened the CALIPER study and ensured that recruited participants accurately reflected the Canadian diversity. The first of these *a priori* studies, included the establishment of age-specific reference intervals for over 40 routine chemistry markers analyzed on the Abbott ARCHITECT c8000 system [53]. This hallmark study was the first of many CALIPER studies to begin to fill the gap in pediatric reference intervals for biomarkers routinely assessed in the pediatric population, including bone markers, markers of cardiovascular disease risk and metabolic syndrome. Findings from this study indicated that many analytes required multiple age partitions and the number of age partitions varied for each analyte. Oftentimes, developmental milestones are used to arbitrarily set age partitions. However, this study demonstrated that age partitions did not necessarily correlate with age developmental milestones, which was a paradigm shift from what was currently practiced. For example, alkaline phosphatase (ALP) required seven age partitions (0-14 days, 15 days - <1 yr, 1 - <10 yr, 10 - <13 yr, 13 - <15 yr, 15 - <17 yr, and 17 - <19 yr), as shown in Figure 3A, whereas alanine aminotransferase (ALT) required three age partitions (0 - < 1 yr, 1 - <13 yr, 13 - <19 yr), as shown in Figure 3B [53]. This study also performed preliminary analysis of differences between the major Canadian ethnic groups (i.e. Caucasian, South Asian, and East Asian). This analysis demonstrated biomarker concentration differences between ethnic groups for ALT, amylase, IgG, IgM, magnesium, total protein and transferrin [53]. While these three ethnic groups comprise a large proportion of Canada, future CALIPER studies will focus on expanding the analysis to other ethnicities.

As children grow and undergo pubertal development, there are significant changes in fertility hormone concentration. In 2013, CALIPER recruited healthy children and adolescents and measured 7 fertility hormones [54]. Age-specific reference intervals were required for all fertility hormones, and aside from prolactin, sex partitions were also required. This study also determined Tanner stage-specific reference intervals. Tanner staging is used to monitor progress of puberty in children. It is especially important to have Tanner stage-specific reference intervals for fertility hormones since every child enters puberty at various ages. Tanner staging is based on a 5-stage scale, with stage I correlating with pre-pubertal development and stage V correlating with adult development. Tanner staging was determined by providing participants with images of Tanner stages I to V and participants self-assessed their development relative to the diagram. Shortly following the examination of fertility hormones, pediatric reference intervals were established for additional endocrine and biochemical markers on the Abbott analyzer [55].

With the success of the initial CALIPER studies, the CALIPER initiative extended their studies to more specialized testing and moved beyond general automated analyzers. Age-specific steroid hormone reference intervals were completed on the AB SCIEX 4000 QTRAP mass spectrometer [56], HPLC analysis aided in establishment of vitamin A and vitamin E pediatric reference intervals [57], and 25-hydroxyvitamin D reference intervals were determined using LC-MS/MS analysis [58]. The 25-hydroxyvitamin D study also demonstrated elevated concentrations of 25-hydroxyvitamin D C3 epimer (C3-epi-25-OH-D_2 and C3-epi-25-OH-D_3) in neonates less than 1 year of age, which could interfere with 25-OH vitamin D measurements in this pediatric population [58,59]. Thus, caution should be exercised in measurement of 25-hydroxyvitamin D in neonates.

Figure 3 Scatter plot demonstrating the serum concentrations of (A) alkaline phosphatase and (B) alanine aminotransferase over the pediatric age range



*Adapted from: Colantonio et. al. [53].

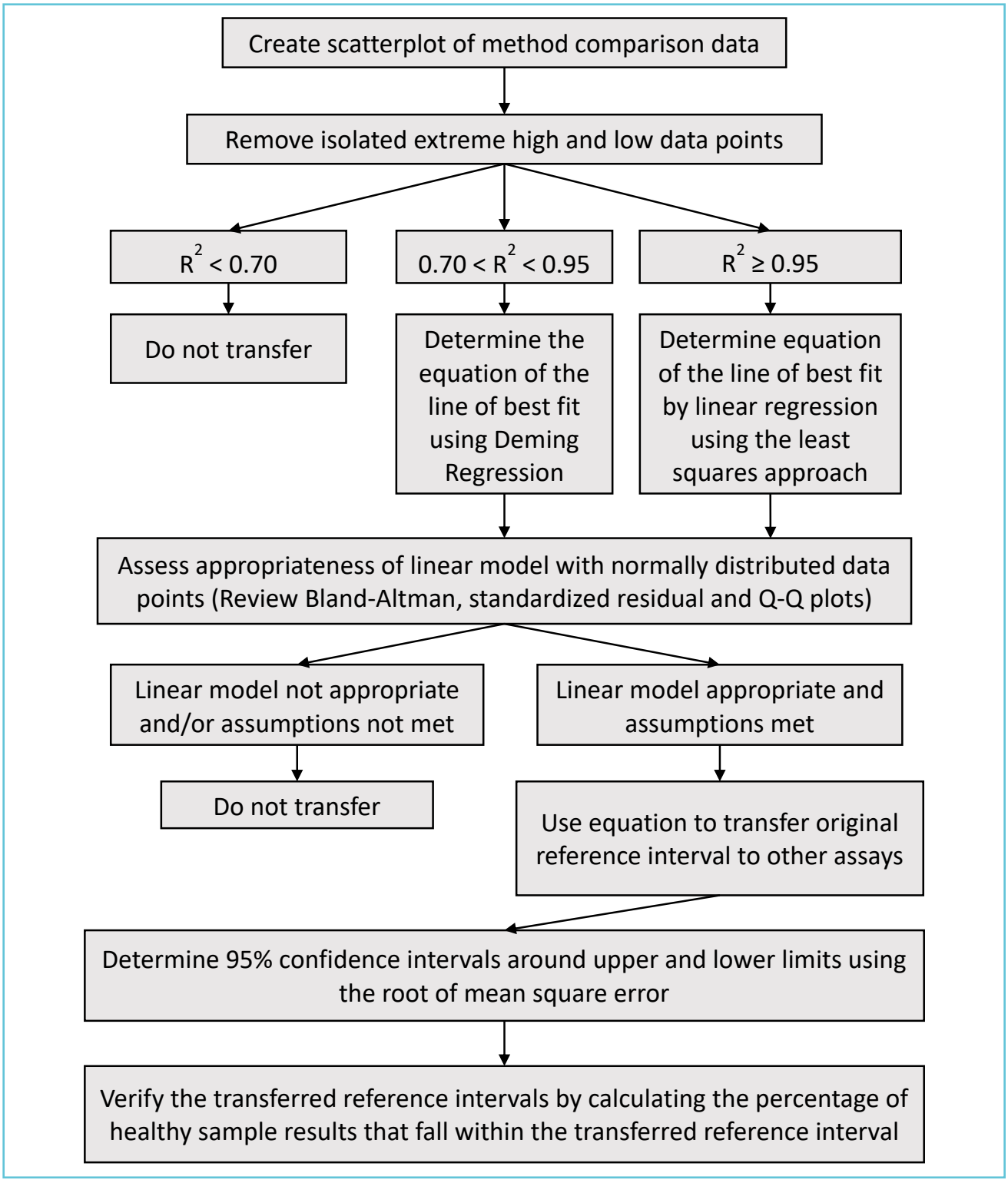
In addition to these large milestones, CALIPER allotted resources to smaller substudies. They have analyzed the effect of freezing conditions on samples and analyte stability [60], biological variation [61], and fasting on biomarker concentrations [62]. The stability of serum chemistry, protein and hormones were analyzed on three analyzers (Ortho Vitros Chemistry System, Roche Cobas Integra 400 Plus and Siemens Immulite 2500) using specimens that were frozen at -80°C [60]. The results of this study demonstrated that -80°C is a suitable storage method for serum samples since no significant deviations in analyte concentrations were observed. An understanding of biological variation is important for accurate laboratory test interpretation and for determining whether a change in biomarker concentration is clinically significant. This data is often lacking for many biomarkers, especially for the pediatric population, since it is often difficult to obtain sequential samples from one child. CALIPER analyzed the within and between-individual biological variation for the pediatric population. Four samples were obtained from 29 healthy participants, and over 30 analytes were analyzed to determine their biological variation [61].

To further expand the CALIPER database, recent studies have established age- and sex-specific reference intervals on the Abbott ARCHITECT immunoassays for cancer biomarkers [63], metabolic disease biomarkers [64], testosterone indices [65], and specialized biochemical markers [66]. To expand the CALIPER database to additional analyzers, a series of transference and verification studies have been performed to transfer reference intervals for chemistry analytes established on Abbott assays to assays on other analyzers including Beckman, Ortho, Roche, and Siemens [44,49-51]. The statistical algorithm CALIPER uses to transfer and verify reference intervals is summarized in Figure 4. To date, CALIPER has created a comprehensive and robust age- and sex-specific pediatric reference interval database for over 100 biomarkers. Through knowledge translation initiatives, CALIPER has also disseminated reference intervals through peer-reviewed publications, development of a free online database (www.caliperdatabase.com), and a smartphone application easily accessible for physicians.

LIMITATIONS OF PEDIATRIC REFERENCE INTERVALS AND FUTURE DIRECTIONS

Despite several advances in addressing the critical gaps in pediatric reference intervals, limitations of the completed and ongoing studies leave additional gaps to be addressed. Current initiatives have extensively examined how analyte concentrations vary with age and sex. However, the influence of additional covariates, such as ethnicity and body mass index (BMI), on analyte concentrations remain to be comprehensively analyzed. Several studies suggest that analyte concentrations may vary by ethnicity. For example, Gupta *et al.* discussed significant differences in serum prostate-specific antigen (PSA) concentration between various ethnicities, with relatively higher concentrations in African American subjects and lower concentrations in Asian subjects [67]. Therefore, ethnic specific PSA reference intervals should be considered in accurate interpretation of test results in cancer diagnosis. Another study by Troy *et al.* measured hematologic and immunologic reference intervals in healthy Zimbabwean infants and compared them to those established using mainly Caucasian subjects [68]. Interestingly, majority of Zimbabwean subjects were considered to have adverse events and immunodeficiency based on hemoglobin and CD4% reference intervals established using Caucasian subjects, respectively. Thus, ethnic-specific and locally-validated reference intervals are required for accurate laboratory test interpretation. The influence of ethnicity on the

Figure 4 Schematic of general statistical approach to transfer original reference intervals to other clinical chemistry assays



**Adapted from: Estey et. al. and Araujo et. al. [44,49].*

concentration of several analytes may be due to several factors including genetic differences, environmental factors, and dietary patterns. Several national pediatric reference interval initiatives have established reference intervals based on a reference population comprised of a single ethnicity [69-72]. As Canada is comprised of a multi-ethnic population, the reference population used in the CALIPER studies proportionally represents the main ethnic groups of the nation (i.e. Caucasian, South Asian, East Asian). CALIPER performed analysis on these ethnic groups, however, the relatively small sample size of ethnicities, other than Caucasian, limited the scope of this study to strictly preliminary analysis [53-55]. Large, comprehensive studies are warranted for a thorough understanding of how analyte concentrations differ between ethnic groups and where warranted the establishment of ethnic-specific reference intervals.

With pediatric obesity becoming an important public health concern, it is also important to understand how analyte concentrations change with BMI [73,74]. A reference population should be comprised of subjects who are representative of the local population, and therefore this definition becomes challenging when covariates such as BMI is constantly changing in the general population. As the average BMI in the general population increases, reference intervals for analytes that change with BMI may also be shifted, or a substantial subset of the local population must be excluded from the reference population. Therefore, it is important to understand which analytes are significantly influenced by BMI and if these changes are physiological and not of clinical significance, or if this change is clinically significant and may be indicative of subclinical progression of a metabolic disease. Understanding how analyte concentrations change with BMI are critical for laboratory specialists and physicians to interpret blood tests from overweight and obese pediatric patients. Some studies have been performed to examine the effect of BMI on analyte levels in a healthy population, however these studies only included adults [75] or were performed for a very limited number of analytes [76]. Further studies are needed to comprehensively examine the influence of BMI on analyte concentrations in the pediatric population.

The ultimate end goal of pediatric laboratory medicine is to achieve harmonization. Laboratory test interpretation, based on reference intervals and decision limits, remain highly variable and poorly harmonized across laboratories. This leads to great potential for inappropriate patient care when laboratory test results on the same sample can be interpreted differently depending on the reference interval reported by the laboratory. Several groups have launched initiatives to harmonize reference intervals including the NORIP [26], the UK Pathology Harmony project [77], the Australasian Harmonised Reference Intervals for Adults (AHRIA) and Australasian Reference Interval for Paediatrics (AHRIP) [47]. In Canada, a Working Group on reference interval harmonization has also been initiated to identify the variation in reference intervals being used in clinical practice and establish/recommend practice guidelines on the use of harmonized reference intervals in clinical laboratories across Canada.

Major challenges have been overcome and significant advances have been made in the establishment and wide-spread dissemination of accurately established pediatric reference intervals. However, with the continuously evolving technological and clinical advances, national and international research initiatives need to ensure pediatric reference intervals continuously improve and adapt to the changing environment.

Table 1 Major pediatric reference interval studies based on healthy children and adolescent populations

Study	Country	Age range (years)	Sex	Statistical method	Examples of groups of biomarkers studied	References
AACB	Australia and New Zealand	All age groups	Both	Central 95%	Common blood analytes (mostly ions and enzymes)	[47,48]
CALIPER	Canada	0-18	Both	Central 95%	Common biochemical markers Endocrine markers Tumor markers Vitamins Metabolic disease biomarkers Testosterone indices	[4,6,8, 18,44, 49-52, 54-57, 60-62, 66]
CHILDx	United States	0.5-17	Both	Median, mean and central 95%	Enzymes Coagulation tests Hormones Vitamins Bone markers	[30-35]
COPENHAGEN	Denmark	5-20	Both	Central 95%	Common blood analytes	[29]
KiGGS	Germany	0-18	Both	Median and central 90%	Nutrient deficiency markers Non-communicable diseases and lipids Immunology markers Thyroid hormones	[20, 22-25]
LOOK	Australia	8, 10 and 12	Both	Median and central 95%	Cardiac Biomarker Common blood analytes	[45,46]

NHANES	United States	All age groups	Both	2.5th, 25th, median, 75th, or 97.5th	Lipid profile Immunology and hematologic markers Vitamins Inflammatory markers	[36-43]
NORIP	Nordic Countries (Denmark, Finland, Iceland, Norway and Sweden)	≥ 18	Both	97.5 percentile or central 95%	Tumor markers Common blood analytes	[21, 26-28]

AACB = Australasian Association of Clinical Biochemists
CALIPER = Canadian Laboratory Initiative on Paediatric Reference Intervals
CHILDx = Children's Health Improvement through Laboratory Diagnostics
COPENHAGEN = The Copenhagen Puberty Study
KiGGS = German Health Interview and Examination Survey for Children and Adolescents
LOOK = Lifestyle of Our Kids
NHANES = National Health and Nutrition Examination Survey
NORIP = Nordic Reference Interval Project

REFERENCES

1. Horn PS PA. Reference Intervals. A User's Guide. Washington, DC: AACC Press; 2005.

2. Boyd JC. Defining laboratory reference values and decision limits: populations, intervals, and interpretations. Asian J Androl 2010;12:83-90.

3. Higgins V, Nieuwesteeg M, Adeli K. Reference Intervals: Theory and Practice. In: Clarke W, editor. Contemporary Practice in Clinical Chemistry, 3rd ed: AACC Press; 2016. p. 21-36.

4. Jung B, Adeli K. Clinical laboratory reference intervals in pediatrics: the CALIPER initiative. Clin Biochem 2009;42:1589-95.

5. Yang L, Grey V. Pediatric reference intervals for bone markers. Clin Biochem 2006;39:561-8.

6. Mansoub S, Chan MK, Adeli K. Gap analysis of pediatric reference intervals for risk biomarkers of cardiovascular disease and the metabolic syndrome. Clin Biochem 2006;39:569-87.

7. Davis GK, Bamforth F, Sarpal A, Dicke F, Rabi Y, Lyon ME. B-type natriuretic peptide in pediatrics. Clin Biochem 2006;39:600-5.

8. Delvin EE, Laxmi Grey V, Vergee Z, CALIPER Working Group. Gap analysis of pediatric reference intervals related to thyroid hormones and the growth hormone-insulin growth factor axis. Clin Biochem 2006;39:588-94.

9. Lepage N, Li D, Kavsak PA, Bamforth F, Callahan J, Dooley K, Potter M. Incomplete pediatric reference intervals for the management of patients with inborn errors of metabolism. Clin Biochem 2006;39:595-9.

10. Clinical and Laboratory Standards Institute (CLSI). Defining, establishing, and verifying reference intervals in the clinical laboratory; approved guideline - third edition CLSI document C28-A3. 2008.

11. Dixon W. Processing data for outliers. Biometrics 1953;9:74-89.

12. Tukey J. Expoloratory Data Analysis. Reading, MA: Addison-Wesley; 1977.

13. Hoffmann RG. Statistics in the Practice of Medicine. JAMA 1963;185:864-73.

14. Soldin SJ, Wong EC, Brugnara C, Soldin OP. Pediatric Reference Intervals, 7th ed. Washington, DC: AACC Press; 2011.

15. Katayev A, Balciza C, Seccombe DW. Establishing reference intervals for clinical laboratory test results: is there a better way? Am J Clin Pathol 2010;133:180-6.

16. Shaw JL, Cohen A, Konforte D, Binesh-Marvasti T, Colantonio DA, Adeli K. Validity of establishing pediatric reference intervals based on hospital patient data: a comparison of the modified Hoffmann approach to CALIPER reference intervals obtained in healthy children. Clin Biochem 2014;47:166-72.

17. Roberts WL, Rockwood AL, Bunker AM, Kushnir MM, Meikle AW. Limitations of the Hoffman approach to determine pediatric reference intervals for two steroids. Clin Biochem 2010;43:933,4; author reply 935.

18. Shaw JL, Binesh Marvasti T, Colantonio D, Adeli K. Pediatric reference intervals: challenges and recent initiatives. Crit Rev Clin Lab Sci 2013;50:37-50.

19. Sikaris K. Application of the stockholm hierarchy to defining the quality of reference intervals and clinical decision limits. Clin Biochem Rev 2012;33:141-8.

20. Kohse KP. KiGGS - the German survey on children's health as data base for reference intervals and beyond. Clin Biochem 2014;47:742-3.

21. Rustad P, Felding P, Lahti A, Hyltoft Petersen P. Descriptive analytical data and consequences for calculation of common reference intervals in the Nordic Reference Interval Project 2000. Scand J Clin Lab Invest 2004;64:343-70.

22. Kamtsiuris P, Lange M, Schaffrath Rosario A. The German Health Interview and Examination Survey for Children and Adolescents (KiGGS): sample design, response and nonresponse analysis. Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz 2007;50:547-56.

23. Kohse KP, Thamm M. KiGGS-the German survey on children's health as data base for reference intervals. Clin Biochem 2011;44:479.

24. Thierfelder W, Dortschy R, Hintzpeter B, Kahl H, Scheidt-Nave C. Biochemical measures in the German Health Interview and Examination Survey for Children and Adolescents (KiGGS). Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz 2007;50:757-70.

25. Witte T, Ittermann T, Thamm M, Riblet NB, Volzke H. Association between serum thyroid-stimulating hormone levels and serum lipids in children and adolescents: a population-based study of german youth. J Clin Endocrinol Metab 2015;100:2090-7.

26. Rustad P, Felding P, Franzson L, Kairisto V, Lahti A, Martensson A, Hyltoft Petersen P, Simonsson P, Steensland H, Uldall A. The Nordic Reference Interval Project 2000: recommended reference intervals for 25 common biochemical properties. Scand J Clin Lab Invest 2004;64:271-84.

27. Urdal P, Bolann B, Marstein S, Rustad P, Steensland H, Asberg A. Updated reference intervals for clinical chemical components. Tidsskr Nor Laegeforen 2004;124:1515-7.

28. Rustad P, Felding P, Lahti A, Nordic Reference Interval Project 2000. Proposal for guidelines to establish common biological reference intervals in large geographical areas for biochemical quantities measured frequently in serum and plasma. Clin Chem Lab Med 2004;42:783-91.

29. Hilsted L, Rustad P, Aksglaede L, Sorensen K, Juul A. Recommended Nordic paediatric reference intervals for 21 common biochemical properties. Scand J Clin Lab Invest 2013;73:1-9.

30. Flanders MM, Crist RA, Roberts WL, Rodgers GM. Pediatric reference intervals for seven common coagulation assays. Clin Chem 2005;51:1738-42.

31. Kushnir MM, Rockwood AL, Roberts WL, Pattison EG, Owen WE, Bunker AM, Meikle AW. Development and performance evaluation of a tandem mass spectrometry assay for 4 adrenal steroids. Clin Chem 2006;52:1559-67.

32. Clifford SM, Bunker AM, Jacobsen JR, Roberts WL. Age and gender specific pediatric reference intervals for aldolase, amylase, ceruloplasmin, creatine kinase, pancreatic amylase, prealbumin, and uric acid. Clin Chim Acta 2011;412:788-90.

33. Johnson-Davis KL, Moore SJ, Owen WE, Cutler JM, Frank EL. A rapid HPLC method used to establish pediatric reference intervals for vitamins A and E. Clin Chim Acta 2009;405:35-8.

34. Meikle AW, Kushnir MM, Rockwood AL, Pattison EG, Terry AH, Sandrock T, Bunker AM, Phanslkar AR, Owen WE, Roberts WL. Adrenal steroid concentrations in children seven to seventeen years of age. J Pediatr Endocrinol Metab 2007;20:1281-91.

35. Wyness SP, Roberts WL, Straseski JA. Pediatric reference intervals for four serum bone markers using two automated immunoassays. Clin Chim Acta 2013;415:169-72.

36. Mortensen ME, Caudill SP, Caldwell KL, Ward CD, Jones RL. Total and methyl mercury in whole blood measured for the first time in the U.S. population: NHANES 2011-2012. Environ Res 2014;134:257-64.

37. Kamycheva E, Goto T, Camargo CA,Jr. Celiac disease is associated with reduced bone mineral density and increased FRAX scores in the US National Health and Nutrition Examination Survey. Osteoporos Int 2016.

38. Breslow RA, Wideroff L, Graubard BI, Erwin D, Reichman ME, Ziegler RG, Ballard-Barbash R. Alcohol and prostate cancer in the NHANES I epidemiologic follow-up study. First National Health and Nutrition Examination Survey of the United States. Ann Epidemiol 1999;9:254-61.

39. Patel MA, Mener DJ, Garcia-Esquinas E, Navas-Acien A, Agrawal Y, Lin SY. Tobacco Smoke Exposure and Eustachian Tube Disorders in US Children and Adolescents. PLoS One 2016;11:e0163926.

40. Cheng CK, Chan J, Cembrowski GS, van Assendelft OW. Complete blood count reference interval diagrams derived from NHANES III: stratification by age, sex, and race. Lab Hematol 2004;10:42-53.

41. Hollowell JG, van Assendelft OW, Gunter EW, Lewis BG, Najjar M, Pfeiffer C, Centers for Disease Control and Prevention, National Center for Health Statistics. Hematological and iron-related analytes--reference data for persons aged 1 year and over: United States, 1988-94. Vital Health Stat 11 2005;(247):1-156.

42. Wener MH, Daum PR, McQuillan GM. The influence of age, sex, and race on the upper reference limit of serum C-reactive protein concentration. J Rheumatol 2000;27:2351-9.

43. Kant AK, Graubard BI. Race-ethnic, family income, and education differentials in nutritional and lipid biomarkers in US children and adolescents: NHANES 2003-2006. Am J Clin Nutr 2012;96:601-12.

44. Estey MP, Cohen AH, Colantonio DA, Chan MK, Marvasti TB, Randell E, Delvin E, Cousineau J, Grey V, Greenway D, Meng QH, Jung B, Bhuiyan J, Seccombe D, Adeli K. CLSI-based transference of the CALIPER database of pediatric reference intervals from Abbott to Beckman, Ortho, Roche and Siemens Clinical Chemistry Assays: direct validation using reference samples from the CALIPER cohort. Clin Biochem 2013;46:1197-219.

45. Southcott EK, Kerrigan JL, Potter JM, Telford RD, Waring P, Reynolds GJ, Lafferty AR, Hickman PE. Establishment of pediatric reference intervals on a large cohort of healthy children. Clin Chim Acta 2010;411:1421-7.

46. Koerbin G, Abhayaratna WP, Potter JM, Apostoloska S, Telford RD, Hickman PE. NTproBNP concentrations in healthy children. Clin Biochem 2012;45:1158-60.

47. Tate JR, Sikaris KA, Jones GR, Yen T, Koerbin G, Ryan J, Reed M, Gill J, Koumantakis G, Hickman P, Graham P. Harmonising adult and paediatric reference intervals in australia and new zealand: an evidence-based approach for establishing a first panel of chemistry analytes. Clin Biochem Rev 2014;35:213-35.

48. Jones GR, Koetsier SD. RCPAQAP First Combined Measurement and Reference Interval Survey. Clin Biochem Rev 2014;35:243-50.

49. Araujo PA, Thomas D, Sadeghieh T, Bevilacqua V, Chan MK, Chen Y, Randell E, Adeli K. CLSI-based transference of the CALIPER database of pediatric reference intervals to Beckman Coulter DxC biochemical assays. Clin Biochem 2015;48:870-80.

50. Abou El Hassan M, Stoianov A, Araujo PA, Sadeghieh T, Chan MK, Chen Y, Randell E, Nieuwesteeg M, Adeli K. CLSI-based transference of CALIPER pediatric reference intervals to Beckman Coulter AU biochemical assays. Clin Biochem 2015;48:1151-9.

51. Higgins V, Chan MK, Nieuwesteeg M, Hoffman BR, Bromberg IL, Gornall D, Randell E, Adeli K. Transference of CALIPER pediatric reference intervals to biochemical assays on the Roche cobas 6000 and the Roche Modular P. Clin Biochem 2016;49:139-49.

52. Chan MK, Seiden-Long I, Aytekin M, Quinn F, Ravalico T, Ambruster D, Adeli K. Canadian Laboratory Initiative on Pediatric Reference Interval Database (CALIPER): pediatric reference intervals for an integrated clinical chemistry and immunoassay analyzer, Abbott ARCHITECT ci8200. Clin Biochem 2009;42:885-91.

53. Colantonio DA, Kyriakopoulou L, Chan MK, Daly CH, Brinc D, Venner AA, Pasic MD, Armbruster D, Adeli K. Closing the gaps in pediatric laboratory reference intervals: a CALIPER database of 40 biochemical markers in a healthy and multiethnic population of children. Clin Chem 2012;58:854-68.

54. Konforte D, Shea JL, Kyriakopoulou L, Colantonio D, Cohen AH, Shaw J, Bailey D, Chan MK, Armbruster D, Adeli K. Complex biological pattern of fertility hormones in children and adolescents: a study of healthy children from the CALIPER cohort and establishment of pediatric reference intervals. Clin Chem 2013;59:1215-27.

55. Bailey D, Colantonio D, Kyriakopoulou L, Cohen AH, Chan MK, Armbruster D, Adeli K. Marked biological variance in endocrine and biochemical markers in childhood: establishment of pediatric reference intervals using healthy community children from the CALIPER cohort. Clin Chem 2013;59:1393-405.

56. Kyriakopoulou L, Yazdanpanah M, Colantonio DA, Chan MK, Daly CH, Adeli K. A sensitive and rapid mass spectrometric method for the simultaneous measurement of eight steroid hormones and CALIPER pediatric reference intervals. Clin Biochem 2013;46:642-51.

57. Raizman JE, Cohen AH, Teodoro-Morrison T, Wan B, Khun-Chen M, Wilkenson C, Bevilaqua V, Adeli K. Pediatric reference value distributions for vitamins A and E in the CALIPER cohort and establishment of age-stratified reference intervals. Clin Biochem 2014;47:812-5.

58. Yazdanpanah M, Bailey D, Walsh W, Wan B, Adeli K. Analytical measurement of serum 25-OH-vitamin D(3), 25-OH-vitamin D(2) and their C3-epimers by LC-MS/MS in infant and pediatric specimens. Clin Biochem 2013;46:1264-71.

59. Bailey D, Veljkovic K, Yazdanpanah M, Adeli K. Analytical measurement and clinical relevance of vitamin D(3) C3-epimer. Clin Biochem 2013;46:190-6.

60. Brinc D, Chan MK, Venner AA, Pasic MD, Colantonio D, Kyriakopolou L, Adeli K. Long-term stability of biochemical markers in pediatric serum specimens stored at -80 degrees C: a CALIPER Substudy. Clin Biochem 2012;45:816-26.

61. Bailey D, Bevilacqua V, Colantonio DA, Pasic MD, Perumal N, Chan MK, Adeli K. Pediatric within-day biological variation and quality specifications for 38 biochemical markers in the CALIPER cohort. Clin Chem 2014;60:518-29.

62. Pasic MD, Colantonio DA, Chan MK, Venner AA, Brinc D, Adeli K. Influence of fasting and sample collection time on 38 biochemical markers in healthy children: a CALIPER substudy. Clin Biochem 2012;45:1125-30.

63. Bevilacqua V, Chan MK, Chen Y, Armbruster D, Schodin B, Adeli K. Pediatric population reference value distributions for cancer biomarkers and covariate-stratified reference intervals in the CALIPER cohort. Clin Chem 2014;60:1532-42.

64. Teodoro-Morrison T, Kyriakopoulou L, Chen YK, Raizman JE, Bevilacqua V, Chan MK, Wan B, Yazdanpanah M, Schulze A, Adeli K. Dynamic biological changes in metabolic disease biomarkers in childhood and adolescence: A CALIPER study of healthy community children. Clin Biochem 2015;48:828-36.

65. Raizman JE, Quinn F, Armbruster DA, Adeli K. Pediatric reference intervals for calculated free testosterone, bioavailable testosterone and free androgen index in the CALIPER cohort. Clin Chem Lab Med 2015;53:e239-43.

66. Kelly J, Raizman JE, Bevilacqua V, Chan MK, Chen Y, Quinn F, Shodin B, Armbruster D, Adeli K. Complex reference value distributions and partitioned reference intervals across the pediatric age range for 14 specialized biochemical markers in the CALIPER cohort of healthy community children and adolescents. Clin Chim Acta 2015;450:196-202.

67. Gupta A, Gupta D, Raizada A, Gupta NP, Yadav R, Vinayak K, Tewari V. A hospital based study on reference range of serum prostate specific antigen levels. Indian J Med Res 2014;140:507-12.

68. Troy SB, Rowhani-Rahbar A, Dyner L, Musingwini G, Shetty AK, Woelk G, Stranix-Chibanda L, Nathoo K, Maldonado YA. Hematologic and immunologic parameters in Zimbabwean infants: a case for using local reference intervals to monitor toxicities in clinical trials. J Trop Pediatr 2012;58:59-62.

69. Yata N, Uemura O, Honda M, Matsuyama T, Ishikura K, Hataya H, Nagai T, Ikezumi Y, Fujita N, Ito S, Iijima K, Saito M, Keneko T, Kitagawa T. Reference ranges for serum cystatin C measurements in Japanese children by using 4 automated assays. Clin Exp Nephrol 2013;17:872-6.

70. Cho SM, Lee SG, Kim HS, Kim JH. Establishing pediatric reference intervals for 13 biochemical analytes derived from normal subjects in a pediatric endocrinology clinic in Korea. Clin Biochem 2014;47:268-71.

71. Abe Y, Okada T, Sugiura R, Yamauchi K, Murata M. Reference Ranges for the Non-High-Density Lipoprotein Cholesterol Levels in Japanese Children and Adolescents. J Atheroscler Thromb 2015;22:669-75.

72. Du Y, Sun H, Chen H, Wang C, Li Y, Wang X, Sun G. Reference Intervals for Routine Blood Tests in 468 Healthy Mongolian Children. Clin Lab 2015;61:1763-7.

73. Ogden CL, Yanovski SZ, Carroll MD, Flegal KM. The epidemiology of obesity. Gastroenterology 2007; 132:2087-102.

74. Ng M, Fleming T, Robinson M, Thomson B, Graetz N, Margono C, Mullany EC, Biryukov S, Abbafati C, Abera SF, Abraham JP, Abu-Rmeileh NM, Achoki T, AlBuhairan FS, Alemu ZA, Alfonso R, Ali MK, Ali R, Guzman NA, Ammar W, Anwari P, Banerjee A, Barquera S, Basu S, Bennett DA, Bhutta Z, Blore J, Cabral N, Nonato IC, Chang JC, Chowdhury R, Courville KJ, Criqui MH, Cundiff DK, Dabhadkar KC, Dandona L, Davis A, Dayama A, Dharmaratne SD, Ding EL, Durrani AM, Esteghamati A, Farzadfar F, Fay DF, Feigin VL, Flaxman A, Forouzanfar MH, Goto A, Green MA, Gupta R, Hafezi-Nejad N, Hankey GJ, Harewood HC, Havmoeller R, Hay S, Hernandez L, Husseini A, Idrisov BT, Ikeda N, Islami F, Jahangir E, Jassal SK, Jee SH, Jeffreys M, Jonas JB, Kabagambe EK, Khalifa SE, Kengne AP, Khader YS, Khang YH, Kim D, Kimokoti RW, Kinge JM, Kokubo Y, Kosen S, Kwan G, Lai T, Leinsalu M, Li Y, Liang X, Liu S, Logroscino G, Lotufo PA, Lu Y, Ma J, Mainoo NK, Mensah GA, Merriman TR, Mokdad AH, Moschandreas J, Naghavi M, Naheed A, Nand D, Narayan KM, Nelson EL, Neuhouser ML, Nisar MI, Ohkubo T, Oti SO, Pedroza A, Prabhakaran D, Roy N, Sampson U, Seo H, Sepanlou SG, Shibuya K, Shiri R, Shiue I, Singh GM, Singh JA, Skirbekk V, Stapelberg NJ, Sturua L, Sykes BL, Tobias M, Tran BX, Trasande L, Toyoshima H, van de Vijver S, Vasankari TJ, Veerman JL, Velasquez-Melendez G, Vlassov VV, Vollset SE, Vos T, Wang C, Wang X, Weiderpass E, Werdecker A, Wright JL, Yang YC, Yatsuya H, Yoon J, Yoon SJ, Zhao Y, Zhou M, Zhu S, Lopez AD, Murray CJ, Gakidou E. Global, regional, and national prevalence of overweight and obesity in children and adults during 1980-2013: a systematic analysis for the Global Burden of Disease Study 2013. Lancet 2014;384:766-81.

75. Deutschbein T, Mann K, Petersenn S. Total Testosterone and Calculated Estimates for Free and Bioavailable Testosterone: Influence of Age and Body Mass Index and Establishment of Sex-Specific Reference Ranges. Horm Metab Res 2015;47:846-54.

76. Wilasco MI, Goldani HA, Dornelles CT, Maurer RL, Kieling CO, Porowski M, Silveira TR. Ghrelin, leptin and insulin in healthy children: Relationship with anthropometry, gender, and age distribution. Regul Pept 2012;173:21-6.

77. Berg J, Lane V. Pathology Harmony; a pragmatic and scientific approach to unfounded variation in the clinical laboratory. Ann Clin Biochem 2011;48:195-7.