CALIPER database of paediatric reference intervals: key milestones and future directions

ABSTRACT

Accurately established reference intervals are essential to interpret laboratory test results and assess patient health. Poorly established reference intervals can lead to misdiagnosis, subjecting patients to anxiety, unnecessary testing, and/or infection risk. The clinical importance of reference intervals is well recognised. However, establishing robust reference intervals is a complex process, especially for the paediatric population. Therefore, available reference intervals are often incomplete, cover a limited paediatric age interval, and/or do not consider gender differences. CALIPER, a collaborative study among Canadian paediatric centres, is addressing these critical gaps by determining age- and sex-specific paediatric reference intervals for over 80 biomarkers using samples collected from over 8,500 children and adolescents. These reference intervals established on the Abbott ARCHITECT have been transferred to other major analytical platforms, broadening the utility of the CALIPER database. The effect of diurnal variation, post-prandial effects, biological variation, and storage temperature on analyte concentration has also been assessed. Knowledge translation initiatives, including peer-reviewed publications, an online database, and a smartphone application, allow physicians and laboratory technicians worldwide to easily access the CALIPER database. This project has made great progress in addressing critical knowledge gaps in paediatric reference intervals, ultimately benefiting paediatric healthcare across Canada and globally.

Key words: reference intervals, pediatrics, children, adolescents, partitioning

Introduction

Reference intervals established from a healthy population are critical to accurately interpret laboratory tests. Reference intervals or ‘normative values’ are defined as the central 95% of values from a healthy reference population. Test results that lie outside this range are considered abnormal, and often indicate that the patient requires additional testing and/or medical treatment. Therefore, poorly established or inappropriate reference intervals can lead to misdiagnosis and can subject patients to further anxiety, unnecessary testing, and/or infection risk. Although the importance of reference intervals in the clinical setting is well recognised, the process of establishing reliable reference intervals is quite complex. In addition, with the significant advances in technology and increases in newly identified biomarkers for disease, published reference intervals quickly become out-of-date. Therefore, there is an urgent need to update existing reference intervals, and also to establish robust reference intervals for newer tests.

In order to be accurate, reference intervals must be established from a large, healthy population. Additionally, appropriate inclusion/exclusion criteria must be applied to ensure that only healthy individuals are included. The normal levels of many biomarkers can also be influenced by physical size, organ maturity, rates of growth and development, immune and hormone responsiveness, nutrition, and metabolism [1, 2]. Therefore, these factors must be considered when establishing reference values, and adult reference intervals are not applicable in the paediatric setting. Specifically, the paediatric population poses additional challenges when establishing reference intervals. Children respond differently to infections compared to adults, and are also...
exposed to unique infections as a result of their ‘immunologically naïve’ state [3]. In addition, obtaining parental consent is necessary when recruiting children and adolescents, and attaining sufficient sample volume can be difficult, especially from neonates. In some cases, partitioning of the reference population may also be required when there are significant differences between subgroups of different ages, genders, or ethnicities. According to the Clinical Laboratory Standards Institute (CLSI) guideline C28-A3, 120 samples per partition are necessary to generate significantly robust reference intervals [4], and depending on the number of partitions needed, the resulting sample size required can be very large.

Based on current literature, the vast majority of previously established paediatric reference intervals have been incomplete, covered a limited paediatric age interval, and/or have not always considered both genders.

Our group has reviewed the literature and identified critical gaps in reference intervals for four paediatric sub-specialties, including bone markers [5], risk markers for cardiovascular disease and metabolic syndrome [6, 7], hormones of the thyroid and growth hormone axes [8], and markers of inborn errors in metabolism [9]. In summary, this research found that existing reference intervals have been established using outdated instruments and methodologies, and/or protocols that were not in accordance with currently recommended guidelines; thus, previously established paediatric reference intervals are generally inappropriate. Increased awareness that the use of unsuitable reference values can pose potential patient safety risks by subjecting children to unnecessary further blood collection, infection risk, and lengthier hospital stays, has led to the development of global initiatives to address the critical gaps in paediatric reference intervals. These initiatives include the Children’s Health Improvement through Laboratory Diagnostics (CHILDx) programme in the United States (www.childx.org), the German Health Interview and Examination Survey for Children and Adolescents (KiGGS) in Germany [10], and the Canadian Laboratory Initiative on Paediatric Reference Intervals (CALIPER) in Canada [1, 2, 11]. The CALIPER project was initiated by The Paediatric Focus Group of the Canadian Society of Clinical Chemists to address the critical gaps in paediatric reference intervals. This collaborative project between seven paediatric hospitals across Canada (in St. John’s, Montreal, Toronto, Ottawa, Hamilton, Saskatoon, and Vancouver) has made significant strides towards improving healthcare for the paediatric population in Canada, and globally.

The significant progress that CALIPER has achieved through initial pilot studies, large-scale reference interval studies for over 80 paediatric biomarkers, and various sub-studies will be reviewed here.

**Pilot studies**

In the initial pilot phases of the CALIPER project, studies were performed to analyze levels of 14 chemistries and immunoassays [12], and later, an additional 15 immunoassays and 24 chemistries [13], all using the Abbott ARCHITECT ci8200 analyser. Across these two studies, a total of 2,809 serum and plasma specimens were collected from apparently healthy and metabolically stable paediatric subjects attending outpatient clinics [13]. Age- and gender-specific reference intervals were established according to CLSI/IFCC C28-P3 guidelines, with the upper and lower limits defined as the 2.5th and 97.5th percentiles, and corresponding 95% confidence intervals were determined for each limit [4, 13]. Reference intervals were established for five age groups, chosen somewhat arbitrarily: birth–12 months, 1–5 years, 6–10 years, 11–14 years, and 15–18 years [13]. These studies were an important initial step for this project, but limitations were still apparent. Namely, samples were collected from outpatients rather than healthy community children, and the recommended number of 120 specimens for each subgroup was not always collected, especially in the birth–12 months age group [12, 13].

In 2010, 25 common analytes were studied in another pilot study using the OrthoVitros 5600 Integrated System. The results were partitioned based on sex and age, according to five arbitrarily chosen age groups, which were merged if deemed appropriate [14]. In this study, the Robust method of Horn and Pesce (Horn 2005) was used to calculate reference intervals, as the requirement to obtain a minimum sample size of 120 or greater per partition could not be met [14]. In the same year, another pilot study was performed for 11 chemistries and 17 immunoassays on the Roche Cobas 6000 analyzer, using 600 outpatient samples [15]. Data was stratified by gender and five arbitrarily chosen age groups, but due to the limited number of collected specimens it was necessary to combine some age partitions [15].

In 2011, the Roche Modular Analytics E 170 system was used to determine reference intervals for free thyroxine (FT4) and thyroid stimulating hormone (TSH), for two predetermined age groups: 1–6 and 7–17 years, but reference intervals were not stratified by gender (Henderson 2011). That same year, age- and gender-specific reference intervals were determined for five bone markers, and the effects of BMI-z score, Vitamin D₃, and parathyroid hormone (PTH) on the levels of these analytes were investigated (Huang 2011). Although these pilot studies were important in assessing the feasibility of, and establishing a foundation for, the CALIPER project, many limitations still needed to be addressed in order to generate robust and accurate
reference intervals that could be used to improve paediatric healthcare across Canada, and worldwide. At this stage, the aims for CALIPER were to obtain an extensive number of blood samples from a cohort of healthy community children and adolescents and to establish appropriately partitioned age and gender-specific paediatric reference intervals from the neonatal age to the end of adolescence.

A priori reference interval studies in healthy community children

Since 2009, CALIPER began its a priori studies, where healthy community children and adolescents have been recruited to establish a biobank of serum samples that can be used to establish representative reference intervals for a large number of clinically relevant paediatric tests. More than 8,500 children and adolescents have been recruited as a result of the CALIPER promotional campaign. Since 2012, CALIPER has established paediatric reference intervals for over 80 biochemical analytes, including 40 biomarkers assessed on the Abbott ARCHITECT C8000 system [16], 21 on the Abbott ARCHITECT i2000SR [17, 18], 11 using the Abbott ARCHITECT ci4100 [19], eight using the Scieq 4000 QTRAP mass spectrometer [20], and two using High Performance Liquid Chromatography (HPLC) [21]. All paediatric reference intervals have been published in high-impact scientific journals and are easily accessible through the CALIPER website (www.caliperdatabase.com), as well as the recently developed smartphone application.

In a study published in 2012, samples from 2,188 healthy children and adolescents (1–18 years) from the greater Toronto area were used to establish reference intervals for 40 biochemical markers using the Abbott ARCHITECT c8000 system [16]. However, for participants < 1 year of age, samples from selected outpatient clinics were used. The ethnic composition of participants was based on the 2006 Canadian census data for the province of Ontario, with the major ethnic groups appropriately represented, including Caucasians, East Asians, and South Asians [16]. Statistically significant differences were observed among the three major ethnic groups for seven of the 40 analytes, but the sample sizes for ethnicities other than Caucasian were too small to determine reliable reference intervals by ethnicity [16]. Results showed that child growth and development greatly affected normal analyte levels in the healthy paediatric population (especially during the first 14 days after birth), and further partitioning by gender was also necessary for a number of biomarkers. This was the first study to establish reference intervals based on a large number of samples collected from healthy community children of various ethnic backgrounds. In contrast, previous studies used samples from hospitalized patients [22], or determined reference intervals from small sample sizes [20, 23] or homogenous populations [24].

In 2013, CALIPER analyzed seven fertility hormones and eight steroid hormones using the Abbott ARCHITECT i2000SR system and the Scieq 4000 QTRAP mass spectrometer, respectively [18, 20]. Reference intervals for seven fertility hormones were stratified by age, sex, and Tanner stage, to further explore the effect of sexual development on endocrine marker concentrations [18]. Samples from 1,234 community children and adolescents (aged 0–18) were used, with the exception of children < 1 year of age, where again samples were collected from apparently healthy neonates and children in hospital maternity wards and selected outpatient clinics [18]. Similarly to the previous study, differences in analyte concentrations between ethnic groups were observed for some analytes; however, the small sample size for some ethnic groups prohibited calculation of ethnicity-specific reference intervals [18]. Many age partitions were required for all fertility hormones, with additional gender partitioning for a number of markers, further illustrating the importance of considering the effects of developmental changes when establishing paediatric reference intervals [18]. Age- and sex-specific reference intervals were also determined for eight steroid hormones using samples obtained from 337 healthy participants [20]. Age partitioning was needed, especially within the first 14 days, which was in agreement with previous CALIPER studies [16, 18]. In 2013, another 1,482 samples from the CALIPER biobank were analyzed on the Abbott ARCHITECT i2000SR to calculate age- and sex-specific reference intervals for an additional 14 endocrine and specialty biochemical markers [17].

Most recently, in 2014, additional studies determined age- and sex-specific reference intervals for Vitamins A and E using 342 blood samples [21], and for 11 cancer biomarkers using between 400 and 700 samples, depending on the analyte [25]. Vitamins A and E were analyzed using High Performance Liquid Chromatography (HPLC) [21] and cancer biomarkers were measured using the Abbott ARCHITECT ci4100 [19]. In these more recent studies, CALIPER expanded its collection sites to the Hamilton region as well as the greater Toronto area.

Transference studies

Most CALIPER reference intervals have been established using Abbott ARCHITECT assays, meaning that the database was initially only applicable to...
laboratories using the Abbott ARCHITECT platform, thereby limiting the information to only a subset of paediatric centres. Establishing de novo reference intervals for every common analytical platform can be very complex, challenging, and costly; therefore, CALIPER has used a method for transferring the reference intervals established in one laboratory (donor) to other (receiving) laboratories, based on the CLSI guidelines [4]. The first step in this process is to assess the degree of similarity between methods by running patient samples on both systems, in order to establish the comparability of the two analytical systems under study [4]. If there is sufficient correlation between assay results on the two systems, and bias compared to the old system is negligible, a mathematical relationship between the two analytical systems can be determined and used to calculate new reference intervals for the receiving laboratory [4]. Next, the receiving laboratory must validate the transferred reference intervals using a relatively small number of samples obtained from healthy individuals [4]. In 2013, CALIPER performed a series of studies to transfer and validate the reference intervals established for the Abbott ARCHITECT to assays performed on four other commonly used clinical chemistry analyzers [26]. Specifically, 28 basic chemistry assays were transferred to the Siemens Vista 1500 platform, 16 to the Beckman Coulter DxC800, 21 to Ortho Vitros 5600, and 19 to Roche Cobas 6000 [26]. The assay-specific, age- and sex-stratified paediatric reference intervals presented in this study greatly expanded the utility of the CALIPER reference interval database, and have increased the applicability of CALIPER reference intervals to paediatric centres across Canada, and worldwide [26].

CALIPER substudies

Additional CALIPER studies have focused on the examination of other factors that can alter analyte concentrations in the healthy population. These include diurnal variation [27], post-prandial effects [27], biological variation [28], and storage temperature [29]. Additionally, CALIPER has assessed the validity of establishing paediatric reference intervals based on data obtained from hospitalized patients, in contrast to our own database established from samples obtained from healthy children [30].

In 2012, CALIPER determined the stability of chemistry, protein, and hormone analytes stored at –80°C on three analytical instruments (i.e. Ortho Vitros Chemistry System, Roche Cobas Integra 400 Plus, and Siemens Immulite 2500) [29]. Samples from children (aged 0–18) were analyzed for 57 biochemical markers at monthly time points throughout the course of one day, and the samples collected included both fasted and fed states [27]. This study revealed that for certain analytes there were significant differences in concentration related to non-fasting state and/or time of day of blood draw [27]. It is critical for paediatricians to be aware of these types of variation and which analytes are affected, as an uninformed decision to order non-fasting, random, or serial samples can contribute to misdiagnosis or mismanagement of patients [27].

In 2014, CALIPER studied the short-term biological variation of 38 chemistry, lipid, enzyme, and protein analytes in the paediatric population using four samples from 29 healthy subjects (aged 4–18) at different time points throughout the course of one day [28]. This was the first study to analyze the components of biological variation in a healthy paediatric cohort, including within- and between-person biological variation, reference change values, and index of individuality. These findings generated a wealth of information, not only regarding the physiological changes that occur within and between individuals, but also regarding which quality specifications are appropriate for a given test, and how and when to use reference intervals [28].

Also in 2014, another sub-study was performed to compare paediatric reference intervals calculated from hospital-based patient data (using the Hoffman approach) to those calculated using samples collected from healthy community children and adolescents as part of the CALIPER study [30]. Reference intervals calculated from hospital-based data were generally wider than those calculated by CALIPER and none fell completely within the 90% confidence intervals calculated by CALIPER [30]. Although using hospital-based patient data has an obvious advantage in that it removes the need to recruit healthy individuals, this study showed that it is not an appropriate method for establishing accurate reference intervals [30].
**Future directions**

CALIPER has made significant advances in addressing the limitations in paediatric laboratory data by establishing up-to-date paediatric reference intervals that are accessible to healthcare centres across Canada, and worldwide.

To date, CALIPER has established robust age- and sex-specific paediatric reference intervals for over 80 biochemical analytes commonly used in clinical practice, and has also performed a number of sub-studies to further validate and support the clinical application of these reference intervals. In addition, CALIPER has begun expanding its database for application in paediatric centres across the country and globally, by performing transference studies for commonly used analytical platforms. These transference studies are the first step towards the development of ‘common reference intervals’, which can be used to interpret test results obtained from different laboratories using different analytical equipment. ‘Common reference intervals’ are determined through collaboration between national or international laboratories, which use analytical methods that are traceable to a common reference method. Although major obstacles exist, including the lack of harmonisation of methods by manufacturers as well as the paucity of reference materials and methods for the majority of analytes, the concept of common reference intervals has been suggested as representing at least in part ‘the way forward’ [31].

The CALIPER initiative serves as a key starting point towards achieving this goal, as a key objective of CALIPER is to transfer its paediatric reference interval database to all major analytical platforms to facilitate application of common reference intervals in paediatric centres across Canada.

Other CALIPER initiatives include expanding the database to include reference values for the adult and geriatric populations. This will aid improved healthcare for patients of all ages, and the same rigorous statistical approach will be employed to examine samples from healthy community subjects. CALIPER has already begun collaboration with Statistics Canada to obtain access to the vast amount of data and blood samples that have been collected through the Canadian Health Measures Survey (CHMS) from subjects aged 3–79 years, in order to establish age- and sex-specific reference intervals for the adult and geriatric populations. CALIPER will also use this data to update our paediatric reference intervals, in order to generate a comprehensive database that is truly representative of the Canadian population. In addition, CALIPER also plans to stay up to date by establishing new reference intervals for emerging biomarkers of disease that are relevant to clinical practice.

The CALIPER project has also been focused on knowledge translation initiatives, and has begun to disseminate reference intervals through peer-reviewed publications, the development of an accessible online database (www.caliperdatabase.ca), as well as a new smartphone CALIPER reference application (available on iTunes) that allows physicians and laboratory technicians worldwide to easily access the CALIPER database.

The CALIPER project has made extensive progress towards its goal of ensuring that every child presenting at any clinic or hospital in Canada or elsewhere benefits from the extensive, accurate, and available knowledge this project has provided. We will continue to pursue this goal through future endeavours.

**References**


