

ADVANCES IN PEDIATRIC REFERENCE INTERVALS FOR BIOCHEMICAL MARKERS: ESTABLISHMENT OF THE CALIPER DATABASE IN HEALTHY CHILDREN AND ADOLESCENTS

NAPREDAK U OBLASTI PEDIJATRIJSKIH REFERENTNIH INTERVALA ZA BIOHEMIJSKE MARKERE: IZRADA BAZE PODATAKA CALIPER KOD ZDRAVE DECE I ADOLESCENATA

Kimiya Karbasy^{1,2}, Petra Ariadne¹, Stephanie Gaglione^{1,2}, Michelle Nieuwesteeg¹, Khosrow Adeli^{1,2}

¹CALIPER program, Department of Pediatric Laboratory Medicine, The Hospital for Sick Children, University of Toronto, Toronto, Ontario, Canada

²Department of Laboratory Medicine and Pathobiology, University of Toronto, Toronto, Canada

Summary

Clinical laboratory reference intervals provide valuable information to medical practitioners in their interpretation of quantitative laboratory test results, and therefore 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. Unfortunately, critical gaps currently exist in accurate and up-to-date pediatric reference intervals for accurate interpretation of laboratory tests performed in children and adolescents. These critical gaps in the available laboratory reference intervals have the clear potential of contributing to erroneous diagnosis or misdiagnosis of many diseases. To address these important gaps, several initiatives have begun internationally by a number of bodies including the KiGGS initiative in Germany, the Aussie Normals in Australia, the AACC-National Children Study in USA, the NORICHILD Initiative in Scandinavia, and the CALIPER study in Canada. In the present article, we will review the gaps in pediatric reference intervals, challenges in establishing pediatric norms in healthy children and adolescents, and the major contributions of the CALIPER program to closing the gaps in this crucial area of pediatric laboratory medicine. We will also dis-

Kratak sadržaj

Klinički laboratorijski referentni intervali pružaju lekarima podatke koji su važni za tumačenje rezultata kvantitativnih laboratorijskih testova i stoga su od ključne vrednosti za procenu zdravstvenog stanja pacijenta i donošenje kliničkih odluka. Referentni interval služi kao reper u smislu zdravlja s kojim će se porediti rezultat pojedinačnog testa. Nažalost, trenutno postoje velike razlike u tačnim i savremenim pedijatrijskim referentnim intervalima za tačno tumačenje laboratorijskih testova koji se obavljaju kod dece i adolescenata. Ove velike razlike u dostupnim laboratorijskim referentnim intervalima očito lako mogu dovesti do dijagnostičke greške ili pogrešnog dijagnostikovanja mnogih bolesti. Nekoliko inicijativa pokrenuto je od strane više tela na međunarodnom nivou sa ciljem da se reši problem ovih razlika, među njima inicijativa KiGGS u Nemačkoj, Aussie Normals u Australiji, AACC – Nacionalna dečija studija u SAD, Inicijativa NORICHILD u Skandinaviji i studija CALIPER u Kanadi. U ovom članku daćemo pregled razlika u pedijatrijskim referentnim intervalima, izazova vezanih za utvrđivanje pedijatrijskih normi kod zdrave dece i adolescenata i glavnih doprinosa programa CALIPER u pogledu otklanjanja razlika u ovoj najvažnijoj oblasti pedijatrijske laboratorijske medicine.

Address for correspondence:

Khosrow Adeli
Clinical Biochemistry, The Hospital for Sick Children,
University of Toronto, Toronto, Ontario, M5G 1X8 Canada

List of abbreviations: CALIPER, Canadian Laboratory Initiative on Pediatric Reference Intervals; RIDL, Reference Intervals and Decision Limits; IFCC, International Federation of Clinical Chemistry and Laboratory Medicine; CLSI, Clinical Laboratory Standards Institute; GGT, Gamma-Glutamyl Transferase; SHGB, sex hormone-binding globulin; FSH, follicle-stimulating hormone; LH, luteinizing hormone; AFP, α -fetoprotein; TSH, thyroid-stimulating hormone; T4, total thyroxine; T3, total triiodothyronine; CRP, C-reactive protein.

cuss the recently published CALIPER reference interval database (www.caliperdatabase.com) developed to provide comprehensive age and gender specific pediatric reference intervals for a larger number of biochemical markers, based on a large and diverse healthy children cohort. The CALIPER database is based on a multiethnic population examining the influence of ethnicity on laboratory reference intervals. Thus the database has proved to be of global benefit and is being adopted by hospital laboratories worldwide.

Keywords: biochemical markers; pediatric; reference intervals; CALIPER; children; adolescents

The Necessity of Reference Intervals

Reference intervals are health-associated benchmarks essential for the interpretation of quantitative laboratory test results by medical practitioners. An interval is formally defined as a statistically derived range of values determined from a reference interval study encompassing the central 95% of values from a healthy reference population. Biomarker test results lying outside of the reference interval suggest an abnormal result and as such, establishing accurate reference intervals is crucial to informed clinical decision-making.

Reference Interval Verification

Clinical laboratories are individually responsible for assuring the validity of reference intervals used to interpret test results they report. Regulatory, accreditation, and licensing organizations require verification or establishment of reference intervals for all quantitative test methods offered by the laboratory, with the exception of tests that utilize decision cut-off limits. If a laboratory opts not to establish the reference interval, verification must be completed using well-defined, systematic methods with supporting documentation demonstrating adherence to international standards. The process of validating an externally determined reference interval is referred to as »transference«, and is required to prevent incorrect classification of test results as normal or abnormal. If a laboratory chooses to conduct a reference interval study, the Clinical Laboratory Standards Institute (CLSI) guidelines must be followed (C28-A3) (1).

Reference intervals are not necessarily transferable between labs due to differences in the reference populations of the donor and receiving laboratories, and differences in analytical methods. In particular, dissimilarities between reference populations can include variations in ethnic compositions, geographic factors, diet preferences, and lifestyle. Analytical methods can vary in measurement principle, calibration, reagent formulation, operating environment, and lot-to-lot variation in reagents and calibrators. To eliminate the potential for error resulting from these differences, a transference study can be completed to

Takođe će biti reči o nedavno objavljenoj bazi podataka o referentnim intervalima CALIPER (www.caliperdatabase.com) koja treba da pruži sveobuhvatne pedijatrijske referentne intervale specifične za uzrast i pol za veći broj biohemijskih markera, zasnovane na velikom i raznolikom skupu zdrave dece. Baza podataka CALIPER zasniva se na multietničkoj populaciji, zahvaljujući čemu se istražuje i uticaj etničke pripadnosti na laboratorijske referentne intervale. Ova baza podataka pokazala se globalno korisnom, zbog čega je usvajaju bolničke laboratorije širom sveta.

Ključne reči: biohemijski markeri, pedijatrija, referentni intervali, CALIPER, deca, adolescenti

assess a reference interval on the demographics of its reference individuals, pre-analytical and analytical details, and statistical methods. While transference of a reference study requires 20 reference individuals, a *de novo* reference interval study is far more laborious and expensive, and is usually reserved for emerging analytes or new analytical methods.

An alternative approach is to establish common reference intervals using a multicenter study design (2). Common reference intervals offer significant advantages including: 1) reduced costs in establishing the reference interval, 2) involvement of multiple laboratories, 3) implementation of a single universal reference interval to interpret a test result that does not require transference, and 4) wide adoption of the interval by laboratories with similar reference populations using the same metrologically traceable method. Nonetheless, several challenges currently impact the common reference interval approach. The lack of harmonization of methods by manufacturers and the limited availability of reference materials and methods for most analytes reduces the ability of multiple laboratories to collaborate. Presently, there are no official guidelines for setting and monitoring the minimum analytical quality goals that a laboratory must achieve to contribute to or use reference values, and reference intervals may not be robust when applied to ethnically diverse populations.

Challenges in Establishing Pediatric Reference Intervals

Necessity of Pediatric Reference Intervals

Critical attention is required to close several accuracy gaps with respect to establishing reference intervals for use in the pediatric population. The application of adult reference intervals in a pediatric setting is inappropriate due to differences in physical size, organ maturity, immune and hormonal responsiveness, nutrition, and metabolism, which collectively influence normal analyte concentrations in children. Additionally, unique biomarkers for children and neonates are often unaccounted for in adult reference intervals, and partitions (or separate reference intervals) are necessary for children and neonates of different age groups, genders, and ethnicities (1).

Gap Analysis of Various Biomarkers

Several studies exemplified the need for up-to-date pediatric reference intervals for biomarkers associated with cardiovascular, endocrine, and metabolic diseases systems (3–6). Although pediatric reference intervals have previously been calculated for various biomarkers, many of these were performed prior to the vast technological advancements in recent years. In addition, most published reference intervals have been determined in hospitalized inpatients or clinic outpatients, and in many cases appropriate partitioning for key covariates including age, gender, and ethnicity have not been made. These deficiencies in appropriate reference intervals pose a serious risk to pediatric care, potentially subjecting infants to further blood collection, pain, anxiety, infection risk, lengthier hospital stays, and unpleasant or invasive diagnostic procedures. The potential for incorrect or delayed diagnosis and the administration of inappropriate treatment is unacceptable. Thus, determining age- and gender-specific pediatric reference intervals is essential to patient safety (1).

Reference Interval Methodology

Procedures for establishing reference intervals have been recommended by the CLSI and by the RIDL (Reference Intervals and Decision Limits) Committee of the International Federation of Clinical Chemistry and Laboratory Medicine (IFCC). The most recent updated guideline, entitled *Defining, Establishing, and Verifying Reference Intervals in the Clinical Laboratory, Approved Guideline—Third Edition, Version C28-A3*, includes protocols to determine reference intervals for new or existing analytes, and to validate existing reference intervals developed in a different laboratory (1). These guidelines clearly indicate that reference individuals must be recruited from a healthy population according to a pre-described definition of healthy. Non-healthy individuals must be clearly distinguishable and recruitment criteria should be derived from known sources of biological variation for the analyte. An *a priori* approach is also advised to recruit healthy reference individuals. Indirect sampling may be employed when collecting sufficient numbers of reference samples is difficult, however, the CLSI guideline does not generally endorse the statistical analysis of data derived from a previously sampled population (indirect sampling) (1). Questionnaires must be filled out by reference individuals with informed consent and, noting that the intention of the samples is not for medical investigation, ethical approval for the study must be obtained. A minimum of 120 samples must be collected for all analytes in order to accurately calculate a 95% confidence interval. Sourcing a large number of healthy children can be difficult and requires parental permission. Potential sources include a campaign in elementary and secondary schools (with the school board's sup-

port), and leftover specimens from routine testing of healthy newborns. Key challenges associated with the collection of samples from children include small sample volumes, higher costs associated with the inability to analyze a large number of analytes from small samples, and difficulty collecting high volumes of samples for the many required age partitions.

The CALIPER Initiative to Close the Gaps in Pediatric Reference Intervals

The Canadian Laboratory Initiative on Pediatric Reference Intervals (CALIPER) project is a collaborative research initiative spearheaded by The Hospital for Sick Children in Toronto, Canada in collaboration with several pediatric clinical laboratories across Canada. The overall objective is to develop and maintain a comprehensive database of reference intervals for laboratory tests that span the pediatric age range (birth to 18 years), and include age-, gender-, and ethnicity-specific partitions. A comprehensive menu of both traditional and emerging biomarkers have been considered in developing the database.

CALIPER Pilot Studies

In the preliminary phases of the CALIPER project, 14 chemistry and 15 immunoassay analytes were analyzed using the Abbott ARCHITECT ci8200 system (7). Subsequently, an additional 14 chemistry analytes were tested using the same platform (8). Together, a total of 2809 serum samples from both studies were collected. Metabolically healthy pediatric outpatients from five age groups: 0–12 months, 1–5 years, 6–10 years, 11–14 years, and 15–20 years were deemed appropriate to participate in this study (7, 8). Following the CLSI/IFCC C28-P3 guidelines, 120 subjects were recruited for each age group, with the exception of the birth to 1 year age range, as the sample size proved difficult to achieve (7). After laboratory analysis, the 2.5th and 97.5th percentiles were established and the central 95th confidence interval was created using the Robust method (7, 8). Age and gender partitions were specified using the Harris-Boyd method. Problems with time of collection for some biomarkers, such as thyroid hormone, and lack of sufficient sample size in some age groups, resulted in weaknesses in the reference intervals for certain analytes (7). Another pilot study was performed using the Roche cobas® 6000 analyzer to further enhance the applicability of CALIPER-derived pediatric reference intervals (9). In this study, 28 chemistry and immunoassay analytes that were previously tested using the Abbott ARCHITECT system, were retested using 600 new healthy outpatients (9). Specifically, for this study, reference intervals were determined after eliminating samples found flagged for hemolysis, icterus, and lipemia, as well as outlier exclusion using the CLSI recommended Dixon test for

Table I Summary of CALIPER studies published to-date.

	Number and type of Analytes	Analytical platform	Reference
Pilot studies	24 chemistry assays 15 immunoassays	Abbott ARCHITECT ci8200	Chan et al. (2009) (7)
	14 chemistry assays	Abbott ARCHITECT ci8200	Chan et al. (2008) (8)
	11 chemistry assays 17 immunoassays	Roche cobas® 6000	Kulasingam et al. (2010) (9)
Full Reference Interval studies (in Healthy Community Children)	11 chemistry assays 12 enzymes 3 lipids/lipoproteins 14 protein markers	Abbott ARCHITECT ci8000	Colantonio et al. (2012) (10)
	7 fertility hormones	Abbott ARCHITECT ci2000	Konforte et al. (2013) (12)
	14 immunoassays	Abbott ARCHITECT ci2000	Bailey et al. (2013) (11)
	8 steroid hormones	Sciex 4000 QTRAP mass spectrometer	Kyriakopoulou et al. (2013) (13)
Transference studies	16–27 chemistry assays, enzymes, lipids/lipoproteins, protein markers	Transference from Abbott ARCHITECT ci8000 Assays To Assays on Beckman Coulter Dx800, Ortho Vitros 5600, Roche cobas® 6000, Siemens Vista 1500	Estey et al. (2013) (14)

non-Gaussian distributed populations (9). Gender was considered a more important variable than age, and therefore gender partitions were established first, if necessary (9). A summary of CALIPER pilot studies can be found in (Table I).

A Priori Studies using a CALIPER Cohort of Healthy Community Children

Over the past five years, CALIPER has been recruiting a cohort of healthy children and adolescents across the community at schools, community centers, daycares, and special community clinics. The CALIPER promotional campaign has been very successful and has resulted in recruitment of over 8500 children and adolescents into the study. Serum samples collected from this cohort have been used to establish a CALIPER biobank. The biobank specimens have subsequently been used to complete several reference interval studies. Since 2012, CALIPER has established pediatric reference intervals for over 60 analytes, including 40 biomarkers assessed using the Abbott ARCHITECT c8000 system (10), 21 using the Abbott ARCHITECT i2000SR (11, 12), and 8 using the Sciex 4000 QTRAP mass spectrometer (13) (Table I). These reference intervals have been published in scientific journals and can also be accessed on the CALIPER website, as well as on a recently developed mobile application.

In a study published in 2012, specimens collected from 2188 healthy participants (0 to 18 years of age) from a multiethnic population were analyzed in order to establish age- and sex- stratified reference

intervals for 40 serum biochemical markers using the Abbott ARCHITECT c8000 analyzer (10). The establishment of normative values was in conformity with CLSI C28-A3 statistical guidelines (1), and the assessed analytes included 11 serum chemistry assays, 12 enzymes, 3 lipids or lipoproteins, and 14 protein markers. The study also investigated differences between the 3 most prevalent ethnic groups in Canada (Caucasians, East Asians, and South Asians). Results confirmed that child growth and development influence the concentration of analytes in healthy children and adolescents, since all of the assays (except lipase) required partitioning by age (10). Partitioning was especially necessary within the 0 to 1 year age range, as levels of many analytes were remarkably different within the first 14 days, compared to the rest of the first year of life (10). For a number of biomarkers, sex partitioning within some age groups was also necessary. Although levels for seven assays displayed ethnic differences, ethnicity was not considered a major determinant of normative values (10). However, further studies assessing the impact of ethnicity on normal analyte concentrations in the pediatric population are necessary, since the number of samples obtained from ethnicities other than Caucasians in this study was considered small.

Considering that the reference intervals for the 40 biochemical markers assessed in 2012 were only applicable to hospitals/laboratories using the Abbott ARCHITECT platform, reference intervals for many analytes were transferred to other systems in order to make the CALIPER database more broadly applicable in Canada and worldwide (14) (Table I). The reference intervals were calculated for the Beckman Coulter

DxC800, Ortho Vitros 5600, Roche Cobas 6000 and Siemens Vista 1500 systems if deemed transferable. For a reference interval to be considered transferable, the relationship between values obtained using the Abbott and other systems must have met some required statistical criteria and assumptions. For example, only normative values for analytes that displayed proper correlation between the systems (R^2 value greater than 0.7) were transferred. In addition, the residuals distribution was normal and unbiased, and residuals could not cluster or follow any identifiable pattern when plotted against the corresponding Abbott value (14). These assumptions were visually assessed using normal Q-Q plots, Bland Altman plots, and residual versus fit plots. After the new reference intervals were determined, verification of the reference interval was necessary. In order for a reference interval to be considered valid, at least 80 to 90% of the levels obtained from reference individuals (approximately 100 healthy CALIPER participants) should fall within the lower and upper limits. In this study, many of the reference intervals previously determined using Abbott were transferable to the other systems (14). Carbon dioxide and magnesium were exceptions, since Abbott reference intervals for these assays were not transferable to any of the other systems due to poor correlation. Also, normative values for phosphate were not transferable to the Beckman Coulter assay (however, it could be transferred to all of the other systems examined) (14). In spite of the strong correlation, Gamma-Glutamyl Transferase (GGT) reference intervals were also deemed non-transferable because the assumptions in the three graphs assessed were not followed. A similar situation occurred with C-reactive protein, but only for the Roche Cobas system (therefore, the reference interval could be transferred to the other systems) (14). The other analytes were considered transferable for all of the four systems (14). In brief, transference from Abbott to the four other major clinical chemistry platforms broadens the scope of the CALIPER database and allows more children to benefit from the establishment of accurate sex and age partitioned pediatric reference intervals. It is important to note, however, that the comparisons performed between the Abbott and Beckman assays make no assumption as to assay accuracy or which is more correct/accurate, and the differences are largely due to differences in calibration of assays between the two platforms.

The project also established pediatric reference intervals for hormones of the hypothalamus-pituitary-gonadal axis (fertility hormones) (12) and for steroid hormones (13). In a study published in 2013, age-, sex- and Tanner stage-specific normative values for seven fertility hormones were established on the Abbott ARCHITECT i2000SR system: estradiol, testosterone (second generation), progesterone, sex hormone-binding globulin (SHBG), prolactin, follicle-

stimulating hormone (FSH), and luteinizing hormone (LH) (12). Complex changes in analyte concentration according to age were observed for all of the assessed hormones. There were also statistically significant gender differences for five out of the seven analytes (12). Partitioning within the first year of life was also observed for these assays, which further demonstrates the importance of establishing adequate age partitions for pediatric reference intervals. Pubertal development was assessed by self-reported Tanner stage for a subset of participants from 9 to 15 years of age, for whom data was available (12). All of the assessed fertility hormones displayed Tanner-stage specific differences, and the patterns were different when comparing boys to girls. There were modest differences among ethnic groups for FSH and SHBG, but no differences were observed for other analytes (12). However, due to the low number of participants from some ethnicities, ethnicity-specific reference intervals for some analytes could not be calculated. This is a limitation that the project is currently trying to surpass by recruiting participants from specific ethnicities (East Asian and South Asian) to better represent the Canadian population.

In another study also published in 2013, reference values for 8 steroid hormones (cortisol, corticosterone, 11-deoxycortisol, androstenedione, 21-hydroxyprogesterone, testosterone, 17-hydroxyprogesterone, and progesterone) were established after the development of a mass spectrometric method for measuring serum steroids (13). The developed assay was an accurate and sensitive method for simultaneously measuring these eight steroid hormones using the Sciex 4000 QTRAP mass spectrometer and a small volume of serum (200 μ L). This methodology is important because the ability to accurately and simultaneously measure steroid hormones in a small aliquot of serum is essential in a pediatric setting/facility due to the challenges and risks associated with obtaining a large volume of blood from children. CALIPER samples from 337 participants from 0 to 18 years of age were used to determine the normative values for the 8 steroid hormones, and age- and sex-specific reference intervals were established (13). In the case of cortisol, one extra variable was taken into account: the time of day when the sample was collected, and therefore reference intervals for this analyte were set for each period of the day. Age-partitioning, especially from birth to 14 days, was necessary for all analytes (13), which is in accordance with findings from a previous CALIPER study (10). The pattern for testosterone was especially interesting and worth noting since it illustrates the importance of partitioning pediatric reference intervals by age and sex: in newborns, serum concentration was significantly higher in boys than girls. From 1 to 13 years, the levels were very low and similar for both sexes. After 13 years, concentrations increased dramatically (20-fold) in males, peaking at 15 years. On the other hand, lev-

els were relatively constant for females from 13 to 16 years, after which time there was a modest increase (but 10-fold less than that observed in males).

Also in 2013, sex- and age- partitioned reference intervals were determined for an additional 14 analytes using the Abbott ARCHITECT i2000 system. These included α -fetoprotein (AFP), cobalamin (vitamin B12), folate, total homocysteine, ferritin, cortisol, troponin I, 25(OH)-vitamin D, intact parathyroid hormone, thyroid-stimulating hormone (TSH), total thyroxine (T4), total triiodothyronine (T3), free T4, and free T3 (11). Samples from 1482 healthy children were assessed in this study. Age- and sex- partitioning was necessary for all of the analytes (11). The effect of ethnicity was also assessed, and statistically significant differences in the levels of total T4, total T3, free T4, cobalamin, ferritin, intact parathyroid hormone, and 25(OH)-vitamin D across ethnic groups were observed. In the case of total T3 and free T4, the effect was small, but total T4 and ferritin were considerably higher in East Asians and lower in South Asians compared to Caucasians. Cobalamin was noticeably higher in East Asians, 25(OH)-vitamin D was lower in East Asians, and intact parathyroid hormone was higher in South Asians (11).

CALIPER Sub-studies

Parallel to its core studies, which establish and transfer reference intervals, the CALIPER project has also published a number of other articles, which investigate relevant questions and challenges involved with establishing normative values for a pediatric population.

For example, considering the high costs and intrinsic challenges associated with obtaining blood from healthy children in their communities (schools, day care centers, festivals etc), we compared CALIPER reference intervals established using the community-based approach with normative intervals obtained using a modified version of Hoffman's original method (15). The Hoffman method uses hospital in- and out- patient blood samples (which are easier and more convenient to obtain) to determine normative values. Reference intervals for 13 analytes were determined using the Hoffman method on the Vitros 5600 system (15). The obtained reference intervals were compared to the corresponding intervals determined by CALIPER (after transferring the reference intervals established in 2012 using the Abbott ARCHITECT system to the Vitros 5600 system). The same sex and age partitions established in the previous CALIPER study were used. In order to compare the two methods, the 90% confidence interval corresponding to each lower or upper limit previously identified by CALIPER was used. None of Hoffman limits fell within the corresponding 90% confidence intervals defined by the CALIPER methodology (15). This

result suggests that using the Hoffman approach might not be a suitable method for determining pediatric reference intervals, at least not in a tertiary care hospital such as the Hospital for Sick Children (where the patient samples were collected), since this method requires that the majority of assessed individuals be healthy children, which may not be the reality at an acute care hospital center. The study concluded that a more feasible approach may be to include only out-patient data, or use samples from a community hospital/health center (15).

Another important CALIPER study performed in parallel with the main project explored the between- and within- individual biological variation in levels of 38 analytes in a 1-day study involving 29 healthy participants from 4 to 18 years of age (16). In the clinical setting, it is essential to consider the effect of biological variation on analyte levels in order to properly interpret test results. Few studies have investigated the complex and dynamic biological variation present within pediatric samples, since most previous studies have focused on adult populations. Therefore, this CALIPER study was relevant because it established pediatric reference change values, which describe the variation that must be observed before a change in patient values should be considered clinically important. Reference change values might be more appropriate to use than reference intervals in some situations, particularly when there is significant within-subject variation in the levels of the analyte. In addition, the effect of time of sample collection on analyte concentration was also investigated (blood was drawn from each participant 4 times throughout the day). Acknowledgement of biological variation is very important in cases where analytes display cyclic changes throughout the day. The majority of analytes assessed in this study displayed significant individuality. In general, the within- and between- individual variation was consistent with data reported for the adult population, and only four analytes (C-reactive protein (CRP), GGT, ceruloplasmin, and glucose) showed marked differences in both within- and between- individual variation in the pediatric population when compared to adults (16). Nevertheless, there were essential lessons that illustrate the importance of specifically studying pediatric samples. For example, the biological variation of glucose was more pronounced in children than in adults (16). This is likely due to the fact that children tend to have lower glucose levels when fasting (17). The biological variation of iron in children was also interesting: there was less within-individual variation than in adult populations (16), which is in accordance with the observation that infants and young children lack diurnal variation of serum iron due to the absence of a sustained period of sleep (18). On the other hand, there was higher between-individual variation (16), which may be explained by the fact that iron deficiency is common in pediatric populations, especially in girls from 12 to 19 years of age (19).

Current CALIPER Studies

The CALIPER project has been continuing to establish accurate and accessible pediatric reference intervals by analyzing common biomarkers across various platforms, as demonstrated by a number of studies recently presented at the Canadian Society of Clinical Chemists Annual Meeting in Prince Edward Island (June 2014), and the American Association of Clinical Chemistry in Chicago, IL (July 2014). Using the Abbott ARCHITECT ci4100 platform, 14 special chemistry and endocrine markers (including free testosterone indexes) were tested using 400-700 samples from healthy children aged birth to 18 years.

To date, there has been a lack of pediatric reference intervals for biochemical markers that are key for prognosis and diagnosis of various childhood cancers. The CALIPER project aimed to establish new pediatric reference value distributions for tumor markers and determine the effects of important covariates like age, gender, and ethnic background. Using the Abbott ARCHITECT ci4100 system, 11 analytes were tested illustrating significant fluctuations in circulating levels for 10 of the 11 cancer biomarkers.

Additionally, testing for 29 immunoassays has been recently performed on the Beckman Coulter Dxl Immunoassay system to ensure broader application of the CALIPER database. The CLSI EP28-A3c guidelines were followed when statistically analyzing all samples. It was noted that concentrations for some assays were higher on the Beckman Coulter platform compared to the Abbott platform necessitating instrument-specific reference intervals.

To ensure a wider application of the CALIPER database, reference intervals for an additional 34 analytes were studied for potential transference from the Abbott ARCHITECT to the Beckman Coulter Synchron Unicel DxC800 platform. The methodology was essentially the same as described for the previous transference study, and included an assessment of the quality of the correlation between the two systems for each analyte, establishment of the transference equation, assessment of the distribution of residuals (which must be unbiased and normal), and verification of the transferred reference interval. Reference intervals for most Beckman assays were transferable and considered verified following analysis of healthy CALIPER specimens.

Concluding Remarks & Future Endeavors

Since the beginning of the project, CALIPER has taken huge steps toward providing up-to-date and comprehensive pediatric reference intervals that are accessible to health institutions in Canada and globally. The CALIPER outreach and recruitment efforts have led to collection of over 8500 blood specimens from healthy community children and adolescents, which has enabled the creation of a CALIPER biobank where specimens are stored at -80°C for later use. Over 70 common chemical biomarkers have been studied using the Abbott platforms and many of these have been successfully transferred to the four other analytical platforms used in clinical laboratories including Beckman Coulter DxC, Ortho Vitros 5,1, Roche Cobas, and Siemens Vista.

The long-term objective of the CALIPER program is to ensure that every child presenting at any clinic or hospital in Canada (and around the world) benefits from the CALIPER program and the availability of up-to-date and accurate pediatric reference intervals. In the short-term, current and future projects are focused on establishing a comprehensive pediatric reference interval database on all major clinical chemistry and immunoassay platforms. This is being achieved by both transference studies and *de novo* reference interval studies of biochemical and immunochemical assays.

A key objective of the CALIPER program has been to ensure extensive knowledge translation through peer-reviewed publications, development of an online and easy-to-use database listing reference intervals for over 70 analytes, accessible through the CALIPER website (www.caliperdatabase.ca), and most recently, a smartphone application for Apple and Android smartphones and tablets. Once downloaded, the new CALIPER app allows any physician or laboratorian worldwide to access the CALIPER database without the need for a network connection. Efforts are currently underway to continue full knowledge translation through new peer-reviewed publications, additions to the online database, and updates to the CALIPER application available through the Apple Store and Google Play.

Conflict of interest statement

The authors stated that they have no conflicts of interest regarding the publication of this article.

References

1. 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.
2. Ferré-Masferrer M, Fuentes-Arderiu X, Alvarez-Funes V, Güell-Miró R, Castiñeiras-Lacambra MJ. Multicentric reference values: shared reference limits. *Eur J Clin Chem Clin Biochem* 1997; 35: 715–8.
3. Yang L, Grey V. Review Pediatric reference intervals for bone markers. *Clin Biochem* 2006; 39: 561–8.
4. 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.
5. Delvin EE, Grey V, Vergee Z. Gap analysis of pediatric reference intervals related to thyroid hormones and the growth hormone–insulin growth factor axis. *Clin Biochem* 2006; 39: 588–94.
6. Lepage N, Li D, Kavsak PA, Bamforth F, Callahan J, Dooley K, et al. Incomplete pediatric reference intervals for the management of patients with inborn errors of metabolism. *Clin Biochem* 2006; 39: 595–9.
7. Chan MK, Seidon-Long I, Aytekin M, Quinn F, Ravalico T, Ambruster D, et al. 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.
8. Chan MK, Quinn F, Preston N, Ravalico T, Ambruster D, Adeli K. Paediatric reference intervals for 14 chemistries and immunoassays on the Abbott Architech ci8200 system—A CALIPER (Canadian Laboratory Initiative on Pediatric Reference Range Database.) pilot study. *Clin Biochem* 2008; 41: 1278.
9. Kulasingam V, Jung BP, Baradaran S, Chan MK, Aytekin M, et al. Pediatric reference intervals for 28 chemistries and immunoassays on the Roche cobas® 6000 analyzer—A CALIPER pilot study. *Clin Biochem* 2010; 42: 1045–50.
10. Colantonio DA, Kyriakopoulou L, Chan MK, Daly CH, Brinc D, Venner AA, Pasic MD, Ambruster 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.
11. Bailey D, Colantonio D, Kyriakopoulou L, Cohen AH, Chan MK, Ambruster 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.
12. Konforte D, Shea JL, Kyriakopoulou L, Colantonio D, Cohen AH, Shaw J, Bailey D, Chan MK, Ambruster 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.
13. 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.
14. 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.
15. 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.
16. 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.
17. Lamers KJ, Doesburg WH, Gabreels FJ, Romsom AC, Renier WO, Wevers RA, Lemmens WA. Reference values of blood components related to fuel metabolism in children after an overnight fast. *Clin Chim Acta* 1985; 145: 17–26.
18. Schwartz E, Baehner RL. Diurnal variation of serum iron in infants and children. *Acta Paediatr Scand* 1968; 57: 433–5.
19. Cooper M, Greene-Finestone L, Lowell H, Levesque J, Robinson S. Iron sufficiency of Canadians. *Health Rep* 2012; 23: 41–8.

Received: July 1, 2014

Accepted: July 2, 2014