Establishment of pediatric reference intervals on a large cohort of healthy children

Emma K. Southcott a, Jennifer L. Kerrigan a, Julia M. Potter a,b, Richard D. Telford b,c, Paul Waring d, Graham J. Reynolds b,e, Antony R.A. Lafferty b,c, Peter E. Hickman a,b,⁎

a ACT Pathology, The Canberra Hospital, Canberra, Australia
b Australian National University Medical School, Canberra, Australia
c Commonwealth Institute, Canberra, ACT, Australia
d Department of Chemistry, Australian National University, Canberra, Australia
e Paediatric and Child Health, Canberra Hospital, Canberra, Australia

Article info

Article history:
Received 6 May 2010
Received in revised form 17 June 2010
Accepted 17 June 2010
Available online 22 June 2010

Keywords:
Children
Pediatric
Reference intervals
Biomarkers

Abstract

Background: Reference intervals are essential in assessing the significance of laboratory results. There have been limited studies generating reference intervals from pediatric populations. We have studied a large cohort of healthy children on 3 separate occasions at 2 yearly intervals.

Methods: 852 healthy 8 year old children were enrolled in a community-based multidisciplinary longitudinal study investigating how early physical activity contributes to health. The same children came back for reassessment at ages 10 and 12 years. Blood samples were analyzed for a total of 37 different chemistries, immunoassays or derived values.

Results: Reference intervals were derived for all the analytes for males and females separately.

Conclusion: Whilst our results are largely in agreement with previously published work, we have shown that for a number of analytes, previously published work is distorted by subclinical disease.

© 2010 Elsevier B.V. All rights reserved.

1. Introduction

The availability of reference intervals is essential for the practice of modern medicine—only by having a feel for what is “normal” can a physician judge how “abnormal” a laboratory result is. Laboratory reference intervals are of particular importance in assessing pediatric patients where there may be marked changes in results at different ages due to physiological causes. The advent of puberty is a signal example.

Development of reference intervals is problematic. The procedure is expensive, time-consuming and fraught with difficulty in precisely defining the population(s) being studied. Further problems arise because particular analyzers used may have systematic biases for the estimation of individual analytes [1–3]. Once reference intervals have been reported, they may be used by other laboratories without appropriate consideration for the precise population studied and without accounting for subsequent changes in methodologies.

It is particularly difficult to obtain good reference intervals for pediatric populations. For many reasons including their relatively small blood volume and societal reluctance to allow blood to be taken from healthy children, there has been a relatively small number of studies performed looking at reference intervals in children. Of those published very few have been performed looking at healthy, community children [4–8] and most have been developed from pediatric hospital populations with statistical manipulation to remove outliers. Soldin has been the principle source of pediatric reference intervals [9–11] and more recently the CALIPER group [12] have provided useful information in this regard.

We have had a unique opportunity to collect longitudinal reference interval data on a cohort of unambiguously healthy children. These children have been participants in the LOOK (Lifestyle Of Our Kids) study which is a community-based multidisciplinary longitudinal study looking at the role of physical activity in preventive medicine [13]. As part of these studies, the same cohort of children was bled on three separate occasions, at ages 8, 10 and 12 years. We have performed laboratory measurements on a range of important and common analytes and in this paper we report these results.

2. Materials and methods

2.1. Ethics approval

This study was approved by several ethics committees including those of the Australian Institute of Sport and ACT Health.

2.2. Population studied

The LOOK study is a community-based multidisciplinary longitudinal study investigating how early physical activity contributes to health. A cohort of 852 (412F 440M) healthy eight year old children were recruited in 2005 and control and intervention groups were
assigned of approximately equal numbers and followed over the next four years. The suburban junior schools selected for the study were representative of Australian city dwellers, both for average income and racial mix [13]. The body mass indices for the children were [median (2.5–97.5 percentiles)] 16.59 (13.80–23.23) for the 8 year olds, 17.62 (14.36–25.38) for the 10 year olds and 18.97 (14.96–27.39) for the 12 year olds.

As part of this assessment, blood was collected at ages 8, 10 and 12 years and serum prepared for the analysis of a range of common analytes and analytes of particular interest in the development of cardiovascular disease [13].

A number of children had significant medical conditions at enrollment, including type 1 diabetes, hypothyroidism, asthma and epilepsy. Where children required medications for management of their condition, they were excluded from the reference interval study. During the course of the study several more clinical conditions were identified, including 5 cardiac abnormalities, 2 of which required surgery and one case of acute leukemia. The results of those children were subsequently removed from our analysis.

2.3. Blood collection and sample handling

Blood was collected from children at their school on three separate occasions in 2005, 2007 and 2009. The cohort of children fell progressively over the four years as some children initially enrolled in the program moved to other schools within the ACT or interstate where the LOOK program was not offered. Only 15 children formally withdrew from the program.

Children were requested to come to school fasting. A substantial breakfast was given to all participants afterwards and this assisted in giving a very high level of compliance with the request. However, a small number did not comply, and samples were rejected if:

- “fasting” blood glucose indicated impaired fasting glycaemia, but with glycated hemoglobin being less than 5.5% [14];
- A sample in one of the earlier collections had an isolated high HOMA-IR [15] which fell to normal at a later collection; and in addition
- Any subjects self-reporting that they were unwell on the day of collection or had eaten prior to collection.

It is likely that a small number of non-fasting samples have been included in our analysis. We felt that we could only exclude samples on absolutely objective grounds or risk possible significant artifactual bias in results. Because of the possibility of a small number of non-fasting results being included, we have presented the data for glucose and HOMA-IR as cumulative frequency plots as these best demonstrate the reality of the changes.

All samples were collected in springtime when early morning ambient air temperature was between 10 and 15 °C. After collection into gel separator tubes, samples were left at room temperature for 30 min to clot, and then centrifuged for 10.5 min at 1505 g (Spintron GT-25LL, Dandenong VIC, Australia). No more than 4 h passed between sample collection and placement on the laboratory autoanalyzer. A small number of analytes were analyzed retrospectively on serum stored at ~70 °C, and samples were subjected to a single freeze–thaw cycle.

Analyses were performed on all available samples for a particular age group, without selection. Where there is variation in the number of samples assayed for a particular analyte in a particular age group, it is because of insufficient sample being available or poor sample quality.

2.4. Instrumentation and assays

Serum insulin and homocysteine were measured on the Abbott AxSYM (Abbott Park, Illinois, USA) and glycated haemoglobin was analysed using the PRIMUS CLC 385 HPLC system (Kansas City, USA). All other analyses (Table 1) were performed on an Abbott C8200 analyser (Abbott Park, Illinois, USA) using routine methods for that analyser. Interference from hemolysis, icterus and lipemia was assessed on all samples using onboard automated procedures.

Glycated hemoglobin was measured in 2005 using the DCA2000. The subsequent collections in 2007 and 2009 had the analysis performed using the PRIMUS CLC 385 HPLC system. The change was made because of the significant positive bias associated with the DCA2000 results which is not observed with the PRIMUS system. This reason also precludes us from including the 2005 results in our analysis. Analytical performance on the PRIMUS had a CV of 1.9% at a glycated hemoglobin of 5.4%, and a CV of 2.2% at a glycated hemoglobin of 9.2%.

We deliberately elected not to measure hormones related to puberty (estradiol, testosterone, LH and FSH) because the age at which children develop puberty-related changes in these hormones is extremely variable and we felt the wide spread in results would make them of little value.

HOMA-IR (homeostatic model assessment) is a derived value used to assess pancreatic β cell function and insulin resistance, by measurement of fasting glucose and insulin concentration. HOMA-IR values tend to be higher (indicating insulin resistance) in persons who are overweight and/or unfit.

2.5. Statistical analysis

All subjects were coded whilst testing to prevent study results being recorded into patient routine results files. The results were electronically transmitted from the laboratory equipment into a laboratory information database and extracted to Microsoft EXCEL to reduce data entry errors. Any patient information added into the laboratory information database and EXCEL database which was initially checked twice and subsequently routinely checked with data processing.

At the end of each week of testing results were screened for missing results or anomalies. Results were removed if HIL (hemolysis, icterus and lipemia) indices implied interference with analytes. Subject’s results were sent to the chemical pathologist to screen for disease and confirm all subject were healthy.

At the end of each collection period, analytes were graphed using scatter plots and histograms to identify outliers. Clinical details were rechecked with subjects with outlier values. Outliers were removed when associated with known illness, such as diabetes or if any subjects self-reporting that they were unwell or had eaten prior to collection.

These scatter plots were also used to examine if transportation distance from the schools to the laboratory affected sample quality. Ten schools placed near and far from the laboratory were chosen to compare their results using robust and sensitive analytes. Descriptive statistics of albumin, bicarbonate, cholesterol, gamma GT, glucose, insulin, lactate dehydrogenase, phosphate, and potassium were used in the comparison and no significant difference was seen between these schools.

Data were separated into males and females and the median, average, minimum, maximum, central 50th percentile, variance, and skewness were calculated. Maximums, minimums and difference in the variance of male and female populations were used to identify analytes with outliers. The central 50th percentiles were plotted to observe trends in analytes due to growth.

Two way t-tests were used and a smaller p value was chosen due to the number of comparisons that were performed between gender and groups, hence P<0.01 was used for significance. When a difference was found between males and females within an age group, significance was displayed in the higher value. For differences within a gender group due to growth, significance was displayed on the later year. We have not displayed a small number of analyses that showed a male–
Table 1

<table>
<thead>
<tr>
<th>Analyte</th>
<th>2005 Adults</th>
<th>2007 Adults</th>
<th>2009 Adults</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cholesterol</td>
<td>273</td>
<td>274</td>
<td>275</td>
</tr>
<tr>
<td>HOMA</td>
<td>227</td>
<td>228</td>
<td>229</td>
</tr>
<tr>
<td>Homocystine</td>
<td>222</td>
<td>223</td>
<td>224</td>
</tr>
<tr>
<td>Potassium</td>
<td>251</td>
<td>252</td>
<td>253</td>
</tr>
<tr>
<td>LDH</td>
<td>214</td>
<td>215</td>
<td>216</td>
</tr>
<tr>
<td>ALT</td>
<td>221</td>
<td>222</td>
<td>223</td>
</tr>
<tr>
<td>AST</td>
<td>252</td>
<td>253</td>
<td>254</td>
</tr>
<tr>
<td>Bicarbonate</td>
<td>226</td>
<td>227</td>
<td>228</td>
</tr>
<tr>
<td>Bilirubin</td>
<td>221</td>
<td>222</td>
<td>223</td>
</tr>
<tr>
<td>Calcium</td>
<td>222</td>
<td>223</td>
<td>224</td>
</tr>
<tr>
<td>Calcium (corr)</td>
<td>222</td>
<td>223</td>
<td>224</td>
</tr>
<tr>
<td>Cholesterol</td>
<td>222</td>
<td>223</td>
<td>224</td>
</tr>
<tr>
<td>Creatinine</td>
<td>222</td>
<td>223</td>
<td>224</td>
</tr>
<tr>
<td>CRP</td>
<td>222</td>
<td>223</td>
<td>224</td>
</tr>
<tr>
<td>Ferritin</td>
<td>222</td>
<td>223</td>
<td>224</td>
</tr>
<tr>
<td>Glucose</td>
<td>222</td>
<td>223</td>
<td>224</td>
</tr>
<tr>
<td>HDL</td>
<td>222</td>
<td>223</td>
<td>224</td>
</tr>
<tr>
<td>HOMA</td>
<td>222</td>
<td>223</td>
<td>224</td>
</tr>
<tr>
<td>Homocystine</td>
<td>222</td>
<td>223</td>
<td>224</td>
</tr>
<tr>
<td>Insulin</td>
<td>222</td>
<td>223</td>
<td>224</td>
</tr>
<tr>
<td>Iron</td>
<td>222</td>
<td>223</td>
<td>224</td>
</tr>
<tr>
<td>Lipase</td>
<td>222</td>
<td>223</td>
<td>224</td>
</tr>
<tr>
<td>Magnesium</td>
<td>222</td>
<td>223</td>
<td>224</td>
</tr>
<tr>
<td>Phosphatase</td>
<td>222</td>
<td>223</td>
<td>224</td>
</tr>
<tr>
<td>Potassium</td>
<td>222</td>
<td>223</td>
<td>224</td>
</tr>
<tr>
<td>Pre-albumin</td>
<td>222</td>
<td>223</td>
<td>224</td>
</tr>
<tr>
<td>Sodium</td>
<td>222</td>
<td>223</td>
<td>224</td>
</tr>
<tr>
<td>Total protein</td>
<td>222</td>
<td>223</td>
<td>224</td>
</tr>
<tr>
<td>Triglyceride</td>
<td>222</td>
<td>223</td>
<td>224</td>
</tr>
<tr>
<td>TSH</td>
<td>222</td>
<td>223</td>
<td>224</td>
</tr>
</tbody>
</table>

(continued on next page)
female difference when the change was considered not to be of clinical value.

Reference intervals were derived by nonparametric methods and data are presented as median with 2.5–97.5 percentile ranges. Male and female data are shown separately for all analytes and age groups. Because of the large numbers of subjects available to us, we did not utilise analysis with the robust statistical methods as described by Horn et al. [16].

The calculations of CRP were influenced by the reporting of the ABBOTT hsCRP assay which has a lower limit of ~0.2 mg/L. Of the 748 children who were tested in 2005, 216 were reported to have results as less than 0.2 mg/L and 91 of 588 in 2007 and 166 of 490 in 2009. The 2.5th percentiles and medians for CRP results were affected by this detection limit.

3. Results

Table 1 shows the results for 37 different analytes and derived values for males and females on three separate occasions with samples collected two years apart.

While there were several analytes for which there were statistically significant differences either with age, or between sexes, only those which are of some clinical import are considered below.

Amongst the important differences observed were:

- for alkaline phosphatase (ALP), females were higher than males at ages 8 and 10 but the males had higher enzyme activity by age 12;
- males had higher CK activity than females at all ages, but we have not been able to control these results for amount of physical activity;
- cholesterol was higher in females than males at ages 10 and 12 whilst HDL was higher in the males at all ages. Triglyceride concentration was higher in females at all ages;
- urate was significantly higher for males at age 12;
- ferritin was slightly higher in males than females at age 12. At the time of blood collection, 50 of 256 females were menstruating and this may explain this difference;
- whilst glucose concentration at a particular age was nearly identical whilst glucose concentration at a particular age was nearly identical for both sexes.

Fig. 1a and b shows the cumulative frequency distribution of glucose concentration for males and females, whilst Fig. 2a and b shows the cumulative frequency distribution for both sexes for the derived HOMA-IR. HOMA-IR increased with age for both sexes but females were higher than males at all ages.

4. Discussion

Whilst there have been several pediatric studies that have recruited healthy children, many of these studies focused on specific health issues such as lipid levels and did not explore a broad menu of biochemical analytes [17–19]. These community-based studies were cross-sectional and captured the health of a specific age group and did not follow up the same subjects.

We have had a unique opportunity to collect reference interval data on unequivocally healthy normal children, as opposed to previous studies which have predominantly used hospital patients [9,10,12]. Our study was performed on fasting subjects and is further strengthened by our ability to perform repeat measurements on the same cohort of children on three occasions over four years, thus removing concerns about population variation as we move from age group to age group. We deliberately avoided making hormonal measurements related to puberty as we felt our population, particularly the 12 year olds would be heterogeneous in their pubertal development. We have presented our data for males and females separately so that readers have full information.

Whilst studies have been performed previously looking at healthy community children [4–8] these studies were performed many years ago, were on now historical instrumentation and were cross-sectional not longitudinal. Though these papers used different instrumentation they do show biochemical trends with growth for urea, ALP, LDH and ALT which we also observed. These biochemical trends are due to underlying physiological changes and need to be considered when developing reference intervals. Ideally reference ranges should be retested with each generation change of assays to assist in the identification of “abnormal” results.

Alternative collection methods have been used, and perhaps the best example is the CALIPER study from Canada where surplus blood was collected at hospital outpatient visits from children considered to be free from metabolic diseases [12]. One downside of collecting surplus blood is that there would be limited information of subject histories and the reason for blood collection.

The CALIPER study is of particular interest to us as it was conducted using identical instrumentation and testing methods (Abbott cib200), and so results could be directly compared without concerns of between-instrument biases. However, their subjects were not fasting as were ours, and another background consideration must be possible variations due to different ethnicities of participants in the studies. A further suboptimal element from the CALIPER study was the pooling of children over 5 year age ranges. Particularly around puberty, changes can occur rapidly and pooling data from children over large age ranges has the potential to mask important changes. An example is magnesium where their 6–10 age groups had a range of 0.8–1.5 mmol/L falling to 0.8–1.0 mmol/L in the 11–14 year old group. Our 8 year olds had a range of 0.8–1.0 mmol/L and this suggests that a physiological fall in magnesium concentration occurs before 10 years of age. With these provisos were there any major differences in reference intervals between the 2 studies and are there any lessons to be learned from this?

Whilst for many analytes the differences in results were small, supporting in principle the experimental design for the CALIPER study, we have demonstrated that for a small number of analytes, use of subjects with even a hint of disease, may result in reference intervals that are different to those seen in truly healthy children.

A very interesting difference was noted with ferritin. CALIPER quoted a range of 8.8–184.7 μg/L for 6–10 year olds. For our 8 year...
olds we found a narrower range of 12–75 μg/L (combined). Ferritin is particularly prone to artifactual increase with any inflammation and this is the most likely explanation for this major difference in results. In support of this is that CALIPER found substantially higher CRP in 3 of the 4 age groups they investigated.

Cholesterol is a negative acute phase analyte and concentration falls in persons who are stressed [20]. For this reason studies using hospital or clinic derived samples have the potential to show a negative bias. Cholesterol values in eight and ten year olds from this study were higher than those observed in the CALIPER study even though these were non-fasting results, supporting the thesis that some CALIPER subjects had subclinical disease.

Although our data cannot be directly compared to that from CALIPER (they did not provide data for their 6–10 year age group) we found the TSH range to be substantially lower than in the slightly older children. We have shown that TSH is a hormone which may increase acutely in response to stress [21] and this too may reflect some subclinical illness in the CALIPER children. We can compare our

---

**Fig. 1.** Fasting serum glucose (mmol/L) cumulative frequency histogram with cumulative percentage for a) boys and b) girls, at ages 8, 10 and 12 years.
data with Zurakowski et al. [22] who collated outpatient’s results from the Children Hospital in Boston over a three year period. Their TSH and free T4 median for 6–10 year olds were similar to ours but their range was broader. The use of outpatient information increases the chance of subjects having an underlying illness and explains the higher 97.5th values. For the 8 year old children (only age group studied) we found a narrower range of fT4 results than the CALIPER study, again supporting the thesis that their children were showing the effects of low level stress. However, we obtained nearly identical results to a recent study also using this assay [23]. Importantly, this assay showed a poor correlation with TSH concentration and the validity of the assay is in question.

Other results which are of interest and broadly in agreement with already published data are those for phosphate and prealbumin, emphasizing the differences between children and adults [4,8].

For some analytes there appeared to be significant differences with age, but on further investigation these apparent changes proved to be artifactual. Consider cholesterol as an example. The upper limit in
8 year old children was substantially higher than for the 10 or 12 year old cohorts. Similarly the upper limits for total calcium and albumin were higher for the 8 year olds than the older children. The younger children had interpreted the request to fast before the blood was taken, as an instruction to neither eat nor drink and several of the children were mildly dehydrated at presentation. Thus these seemingly significant results are an artefact as identified when the calcium concentration was corrected for the albumin concentration. This raises the interesting question as to whether we should include a correction factor for cholesterol with albumin in the same way that we correct total calcium!

Our data showing the increasing HOMA-IR with age and also being much more pronounced in females is of interest. Because we could not exclude a small number of children being non-fasting we have presented these data as a cumulative frequency distribution and this change is apparent at all points on the graph. At this stage we do not have any information as to whether the exercise intervention part of the LOOK study may have modiﬁed this change with age.

In summary, we have performed a longitudinal study looking at the same cohort of healthy children at ages 8, 10 and 12 years. We have measured a large number of common analytes and present reference intervals based upon our measurements. Whilst our proposed reference intervals are largely in agreement with already published work we have shown that for a number of analytes, previously published work is distorted by subclinical disease.

Acknowledgement

This study was supported by a grant from the Commonwealth Institute (Australia).

References