Clinical Biochemistry xxx (2017) xxx-xxx

Contents lists available at ScienceDirect

Clinical Biochemistry

journal homepage: www.elsevier.com

National Survey of Adult and Pediatric Reference Intervals in Clinical Laboratories across Canada: A Report of the CSCC Working Group on Reference Interval Harmonization

Khosrow Adeli[a,*], Victoria Higgins[a], David Seccombe[b], Christine P. Collier[c], Cynthia Balion[d], George Cembrowski[e], Allison A. Venner[f], Julie Shaw[g]on behalf of the CSCC Reference Interval Harmonization (hRI) Working Group

[a] Pediatric Laboratory Medicine, The Hospital for Sick Children and Laboratory Medicine and Pathobiology, University of Toronto, Toronto, ON, Canada
[b] CEQAL, Vancouver, BC, Canada
[c] Kingston General Hospital and Department of Pathology and Molecular Medicine, Queen's University, Kingston, ON, Canada
[d] Department of Pathology and Molecular Medicine, McMaster University, Hamilton, ON, Canada
[e] Department of Laboratory Medicine and Pathology, University of Alberta, Edmonton, AB, Canada
[f] Calgary Laboratory Services and Department of Pathology and Laboratory Medicine, Cumming School of Medicine, University of Calgary, Calgary, AB, Canada
[g] Department of Pathology and Laboratory Medicine, The Ottawa Hospital, Eastern Ontario Regional Laboratories Association and University of Ottawa, Ottawa, ON, Canada

ARTICLE INFO

Keywords:
Reference interval
Reference range
Normal range
Harmonization
Result interpretation

ABSTRACT

Objective: Reference intervals are widely used decision-making tools in laboratory medicine, serving as health-associated standards to interpret laboratory test results. Numerous studies have shown wide variation in reference intervals, even between laboratories using assays from the same manufacturer. Lack of consistency in either sample measurement or reference intervals across laboratories challenges the expectation of standardized patient care regardless of testing location. Here, we present data from a national survey conducted by the Canadian Society of Clinical Chemistry (CSCC) Reference Interval Harmonization (hRI) Working Group that examines the variation in laboratory reference sample measurements, as well as pediatric and adult reference intervals currently used in clinical practice across Canada.

Design and Methods: Data on reference intervals currently used by 37 laboratories were collected through a national survey to examine the variation in reference intervals for seven common laboratory tests. Additionally, 40 clinical laboratories participated in a baseline assessment by measuring six analytes in a reference sample.

Results: Of the seven analytes examined, alanine aminotransferase (ALT), alkaline phosphatase (ALP), and creatinine reference intervals were most variable. As expected, reference interval variation was more substantial in the pediatric population and varied between laboratories using the same manufacturer. Reference sample results differed between laboratories, particularly for ALT and free thyroxine (FT4). Reference interval variation was greater than test result variation for the majority of analytes.

Conclusion: It is evident that there is a critical lack of harmonization in laboratory reference intervals, particularly for the pediatric population. Furthermore, the observed variation in reference intervals across instruments cannot be explained by the bias between the results obtained on instruments by different manufacturers.

1. Introduction

Harmonization in laboratory medicine is fundamental to ensure data obtained from different laboratories are comparable, and to improve the accuracy and consistency of results and their interpretation to optimize and standardize patient care [1]. Harmonization can be achieved by addressing several aspects of the total testing process from pre-analytical and analytical phases, through post-analytical phases.

* Corresponding author.
Email address: khosrow.adeli@sickkids.ca (K. Adeli)

http://dx.doi.org/10.1016/j.clinbiochem.2017.06.006
Received 26 March 2017; Received in revised form 6 June 2017; Accepted 20 June 2017
Available online xxx
0009-9120/

This can include everything from test requests, sample collection, handling and transportation, analytical method and calibration, reporting terminology and units, as well as reference intervals and result interpretation. Generally, patients and physicians assume they will receive the same, or at least comparable, results from any laboratory, and that their results will be interpreted consistently by all laboratories and physicians [2]. Unfortunately, result interpretation and for some methods, the result itself, may vary more than expected. Those requesting laboratory tests, receiving laboratory reports, developing information systems, and even laboratory professionals may be unaware of these differences, especially when results are obtained from different laboratories. Therefore, this study aims to examine the current variation of reference intervals in Canada and highlight the need for harmonized reference intervals with the main drivers being harmonized test result interpretation as well as improved and standardized patient care and safety.

The concept of common reference intervals was supported at an international meeting in 2003 [3], and since then several worldwide initiatives have been successful in harmonizing reference intervals. One of the first initiatives, reported in 2004, was the Nordic Reference Interval Project (NORIP), which established common reference intervals for 25 clinical chemistry analytes [4]. The UK Pathology Harmony project supported by the Association of Clinical Biochemistry (ACB) and the Institute of Biomedical Science and Royal College of Pathologists, reported harmonized reference intervals in 2011 for 12 and 35 clinical chemistry assays for pediatrics and adults, respectively, and subsequently expanded their program to include hematology and immunology [5]. The Australasian Association of Clinical Biochemists (AACB) and the Royal College of Pathologists of Australasia (RCPA) recommended a panel of 10 and 12 common reference intervals in 2014 for pediatrics and adults, respectively (Australasian Harmonized Reference Intervals for Paediatrics (AHRIP) and Australasian Harmonized Reference Intervals for Adults (AHRIA)) [6]. Additionally, the Japanese Association of Medical Technologists (JAMT) established common reference intervals in 2013 for 27 serum analytes for which certified reference materials are available and 9 routinely measured analytes through collaboration with core laboratories certified for metrological traceability nationwide [7].

The Reference Interval Harmonization (hRI) Working Group of the Canadian Society of Clinical Chemistry (CSCC) was formed in 2015 to address the critical need for harmonized laboratory test interpretation in Canada. The ultimate goal of the hRI Working Group is to establish evidence-based harmonized reference intervals and support their implementation in laboratories across the country. The main objectives of this working group are to a) review adult and pediatric reference intervals currently in use in clinical laboratories across Canada, b) assess the available evidence on reference values obtained in a priori studies of healthy populations, and c) develop appropriate recommendations and guidelines on the use of harmonized reference intervals across Canada. A national survey was conducted in April/May 2016 to collect data on the current reference intervals in use by laboratories for seven analytes. In addition, a reference sample was analyzed for six common analytes by participating laboratories to demonstrate the current comparability of results across laboratories and the relationship between measurement bias and the reference intervals in use by each laboratory. In the present report, we review the survey results demonstrating a considerable variation in both adult and pediatric reference intervals currently being used by clinical laboratories across the country, and in the baseline test results themselves. We also compare the currently used reference intervals in Canada with recently recommended reference intervals by Canadian Laboratory Initiative on Pediatric Reference Intervals (CALIPER) [8,9], Canadian Health Measures Survey (CHMS) [10,11], and the UK [5] and Australasian [6] harmonization initiatives.

2. Design and methods

The reference interval survey and commutable reference sample measurement relied on voluntary participation of clinical laboratories across Canada. Participants in the reference interval survey reported their reference intervals for any of the seven analytes which they measured in their laboratories. If the laboratory agreed to participate in the commutable reference sample measurement, they provided their address to receive the reference samples and submit their measurement results.

2.1. National Reference Interval Survey

2.1.1. Survey dissemination

The hRI Working Group created a survey consisting of stating your agreement or disagreement with three general statements regarding awareness of reference interval variation and a table to record reference interval information for seven analytes: aminotransferase (ALT), alkaline phosphatase (ALP), calcium, creatinine, free thyroxine (FT4), hemoglobin and sodium. The three general statements were: [1] There are significant gaps and inconsistencies in adult/geriatric reference intervals and decision limits currently used in clinical laboratories in Canada, [2] There are significant gaps and inconsistencies in pediatric reference intervals and decision limits currently used in clinical laboratories in Canada, and [3] There is a need for harmonized reference intervals and decision limits in clinical laboratories across Canada." In response to the three general statements, laboratories could respond "strongly agree", "agree", "no opinion", "disagree", or "strongly disagree". When providing reference interval information, laboratories were asked to provide units, instrument manufacturer, age range, sex, lower reference limit (LRL), upper reference limit (URL), and the reference interval source for each analyte. The survey was circulated electronically to the CSCC listserv by e-mail.

2.1.2. Statistical analysis

Results submitted by participating laboratories were sorted first by analyte and then by manufacturer. Results could not be further sorted by specific manufacturer instrument model due to a limited number of participating laboratories. Not all laboratories provided reference interval information on all seven analytes. Representative reference intervals for a child (2 year old male), adolescent (14 year old female), and adult (50 year old male) from each participating laboratory were plotted for each analyte and colour-coded by manufacturer. Reference intervals established from CALIPER [8,9], CHMS [10,11], UK Pathology Harmony [5], and Australasian Harmonised Reference Intervals [6] were shown on each plot (where available) for comparison. When the reference interval was reported as "<" a specific value, the URL is recorded as this value and the LRL is recorded as zero for graphical and calculation purposes. The range, mean, and percent variation (%V) for each LRL and URL were determined for each analyte and for each manufacturer.

2.2. Commutable reference sample measurement

2.2.1. Reference sample provided

The reference sample was prepared according to the Clinical Laboratory Standards Institute C37-A guidelines [12]. The reference pooled human serum sample was created by CEQAL with sera from 17 individual healthy donors (9 females (30–63 years) and 8 males (48–66 years)) who were not taking any prescription medication. The pooled serum was pre-filtered with an Acropak 1.2 μm filter, followed by sterile filtration with a 0.8/0.2 μm Acropak filter. 2 mL sterile

aliquots were dispensed into sterile screw-capped polypropylene cryovials. The samples were shipped on gel pack by overnight courier. Participating laboratories measured the following six analytes in singlicate: ALT, ALP, calcium, creatinine, FT4, and sodium. In addition to analyte results, each laboratory provided the following information: Laboratory ID and name, and instrument manufacturer.

2.2.2. Statistical analysis

Test results were sorted first by analyte and then by manufacturer. Again, results could not be further sorted by specific manufacturer instrument model due to limited participating laboratories. Not all laboratories performed all analyses. Total data submitted and data submitted within each manufacturer grouping was evaluated and the mean, between-laboratory coefficient of variation (CV_{BL}), and percent bias to target or all results median (ARM) was calculated. ALT, ALP, creatinine, and sodium results were evaluated against a reference "target" value established using CEQAL reference value assigned human serum samples. However, as reference value assigned human serum samples were not available for calcium and FT4 to establish "target" values, these results were evaluated against the ARM across all laboratories. Uncertainties of target value assignment are shown in Supplemental file 1.

2.3. Comparison between current laboratory reference intervals and reference sample measurement

The CV (%V and CV_{BL}) and percent bias were calculated for reported reference intervals and reference sample measurements, grouped by manufacturer. These values were compared to each other and the CV_{BL} was compared to the desirable total error allowable (%TE) determined by Ricos et al. and available on Westgard's website [13].

3. Results

Sixty-four laboratories (63 for question 1) responded to the survey statements, 37 laboratories provided reference interval information, and 40 laboratories measured the reference sample. Supplemental Fig. 1 shows the distribution of responses from laboratories to the survey statements. 87% and 91% of laboratories either agreed or strongly agreed that there are significant gaps and inconsistencies in adult/geriatric and pediatric reference intervals, respectively, across Canada. Furthermore, 92% of laboratories either agreed or strongly agreed that there is a need for harmonized reference intervals and decision limits in clinical laboratories. Figs. 1–3, Supplemental Figs. 2–5 and Table 1 depict the variation in reference intervals currently used by laboratories across Canada. Fig. 4, Supplemental Figs. 6–7 and Supplemental Table 1 depict the variation in results obtained from laboratories measuring the same reference sample. A comparison between CV_{BL} and %V values and percent bias to target (or ARM) for reference intervals and reference sample measurements are shown in Table 2.

When laboratories were grouped by manufacturer, some assays still had significant variation in reference intervals between laboratories using instruments from the same manufacturer. The same analytes which had significant variation in reference intervals (%V) also had significant variation in reference sample measurements (CV_{BL}). For example, reference intervals for ALT, ALP, creatinine, and FT4 varied the most, while calcium, and sodium reference intervals were less variable. Likewise, the reported sodium concentration varied the least across all laboratories (CV_{BL}: 1.1%), while the reported ALT concentration varied the most across all laboratories (CV_{BL}: 25%). Profound variation in ref-

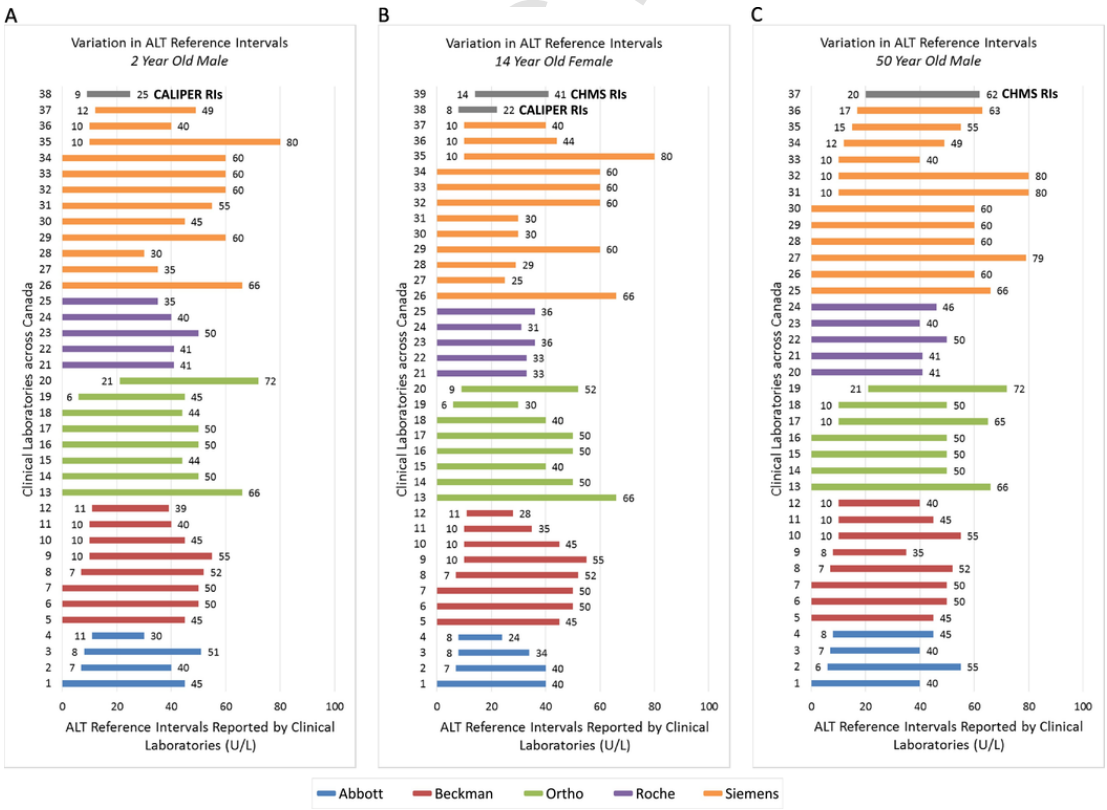

Fig. 1. Reference intervals for alanine aminotransferase (ALT) used in laboratories across Canada, grouped by instrument manufacturer for (A) a child (2 year old male), (B) an adolescent (14 years old female), and (C) an adult (50 year old male). Canadian Laboratory Initiative on Pediatric Reference Intervals (CALIPER) and Canadian Healthy Measures Survey (CHMS) reference intervals (RIs) are shown for reference.

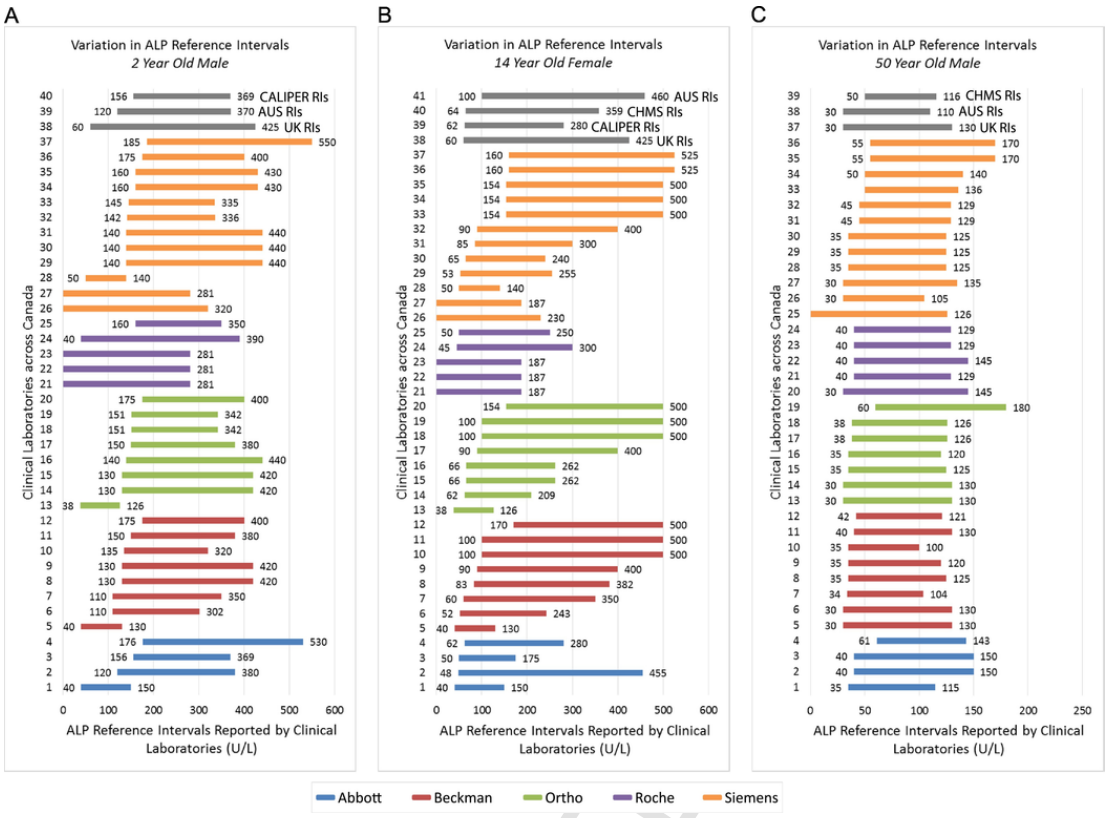

Fig. 2. Reference intervals for alkaline phosphatase (ALP) used in laboratories across Canada, grouped by instrument manufacturer for (A) a child (2 year old male), (B) an adolescent (14 years old female), and (C) an adult (50 year old male). Canadian Laboratory Initiative on Pediatric Reference Intervals (CALIPER), Canadian Healthy Measures Survey (CHMS), United Kingdom Pathology Harmony Project (UK) and Australasian Harmonized (AUS) reference intervals (RIs) are shown for reference.

erence sample measurements were also observed across laboratories using assays from the same manufacturer. In the majority of cases where the variation was compared between reference intervals and reference sample measurements, the variation in reference intervals was greater than the variation in reference sample results. Additionally, the difference in reference intervals between laboratories could not be explained by the percent bias in reference sample results obtained on that instrument/assay.

3.1. Alanine aminotransferase

ALT reference intervals were highly variable across laboratories for all age groups (Fig. 1, Table 1). The %V for ALT URLs were 23%, 30%, and 22%, for children, adolescents, and adults, respectively. The %V for most ALT LRLs are not shown due to LRLs of zero giving exaggerated differences. However, reported LRLs varied from 0 to 21 U/L for children and adults, and 0–11 U/L for adolescents. Reference intervals reported by laboratories using instruments from the same manufacturer remained highly variable.

The CV_{BL} for ALT across all laboratories was 25%, with Siemens assay users having the highest CV_{BL} of 20% (Table 2). Abbott, Beckman, and Roche assay users were all biased low compared to the assigned target value, while Ortho and Siemens assay users were biased high (Fig. 4A).

The %V was greater than the CV_{BL} for ALT in all cases (Table 2). This was true across all assay manufacturers analyzed together and for each instrument/assay user separately, except Roche. The percent bias for the reference sample measurement and reported reference intervals were similar for laboratories using assays from the same manufacturer, with the exception of Beckman users. Beckman assays were biased low (− 13%) compared to the reference sample target value, but their reported URLs were biased high (13%) compared to the ARM.

3.2. Alkaline phosphatase

ALP reference intervals varied substantially across laboratories for all age groups (Fig. 2, Table 1). The %V for ALP URLs were 28%, 42%, and 13%, for children, adolescents, and adults, respectively. Again, the %V for several LRLs were not shown due to LRLs of zero giving exaggerated differences. Nevertheless, the range of reported LRLs was substantial, with LRLs varying from 0 to 185 U/L, 0–170 U/L, and 0–61 U/L for children, adolescents, and adults, respectively. The %V for ALP across laboratories using assays from the same manufacturer also remained highly variable.

The CV_{BL} across all laboratories was 6.6%, with Beckman assay users having the highest CV_{BL} of 5.2% (Table 2). The average reported ALP concentrations from each manufacturer group was biased low compared to the target value (Fig. 4B).

The %V (for those calculated) was greater than the CV_{BL} in all cases for ALP (Table 2). This was the case across all assay manufacturers analyzed together and for each instrument/assay user separately. For example, the %V for URLs and LRLs for Beckman assay users was 36% and 47%, respectively, while the CV_{BL} was only 5.2%. Three of the five instrument/assay user groups had opposite bias in their reported reference intervals compared to their reported measurements. For example, the percent bias to ARM for Ortho users was 15% and 30% for URLs and LRLs, respectively. However, the measurements obtained by Ortho users were biased low (− 7.5%) compared to the target value.

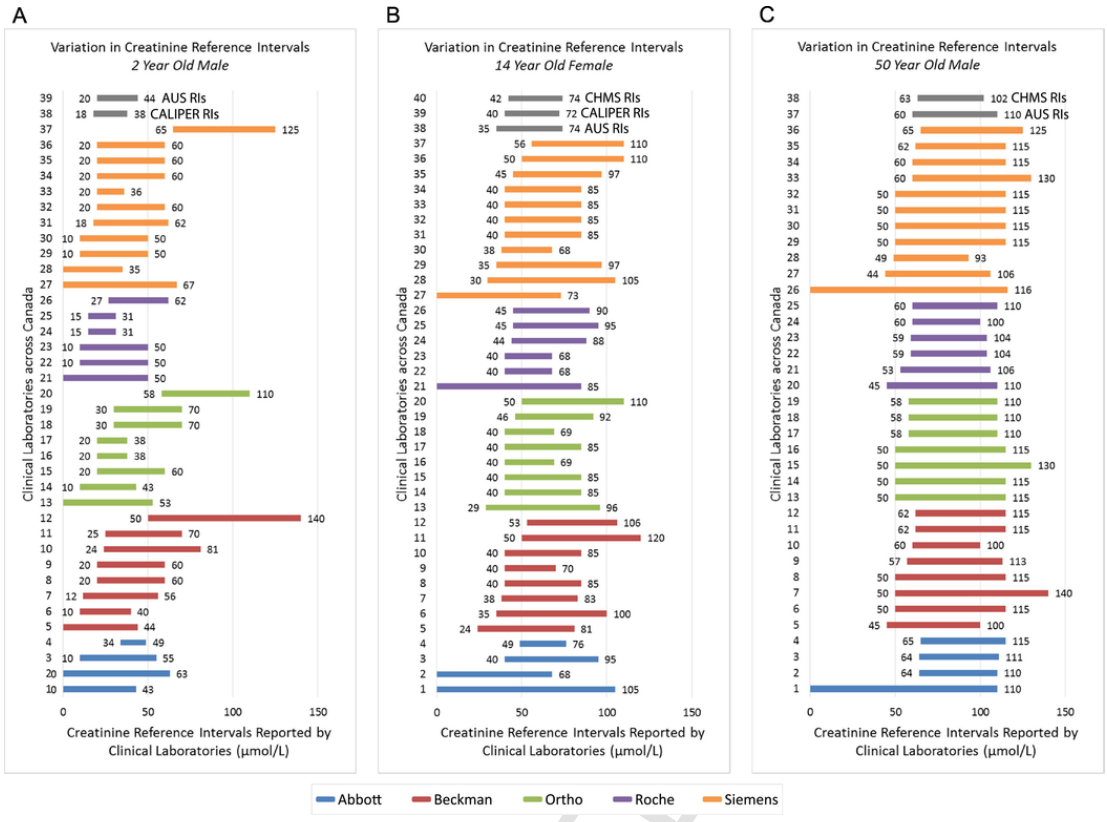

Fig. 3. Reference intervals for creatinine used in laboratories across Canada, grouped by instrument manufacturer for (A) a child (2 year old male), (B) an adolescent (14 years old female), and (C) an adult (50 year old male). Canadian Laboratory Initiative on Pediatric Reference Intervals (CALIPER), Canadian Healthy Measures Survey (CHMS) and Australasian Harmonized (AUS) reference intervals (RIs) are shown for reference.

3.3. Calcium

Calcium reference intervals had little variation across laboratories for all age groups and instrument users, with the highest %V being 4.5% for Siemens assay users reporting the LRL for a 14 year old female (Supplemental Fig. 2, Table 1). The %V for LRLs for children, adolescents and adults were 2.7%, 3.3%, and 2.3%, respectively.

The CV_{BL} across all laboratories was 2.6%, with Siemens assay users having the highest CV_{BL} of 2.8% (Table 2). The average reported calcium concentrations from each manufacturer group were compared to the ARM as no target value was available. Beckman and Ortho assay users were biased slightly high compared to the ARM (1.4% and 2.5%, respectively), while Abbott, Roche, and Siemens users were biased slightly low (Supplemental Fig. 6A).

For calcium, the %V was greater than the CV_{BL} in the majority of cases (Table 2). This was true across all assay manufacturers analyzed together and for each instrument/assay user group separately. The %V for URLs was only greater than the CV_{BL} for Abbott and Ortho assay users. Although a few instrument user groups had opposite bias in their reported reference intervals compared to their reported measurements, all percent biases were relatively low.

3.4. Creatinine

Creatinine reference intervals had large variation across laboratories, particularly for the pediatric age groups (Fig. 3, Table 1). The highest %V was 46% for Beckman assay users reporting the URL for a 2 year old male. Again, the %V for several LRLs were not shown due to LRLs of zero giving exaggerated differences.

The CV_{BL} across all laboratories was 5.0%, with Siemens assay users having the highest CV_{BL} of 6.9% (Table 2). Laboratories using instruments from all five manufacturers were biased high compared to the target value (Supplemental Fig. 6B).

The %V (for those calculated) was greater than the CV_{BL} for creatinine in all cases (Table 2). This was the case for all assay manufacturers analyzed together and for each instrument user separately. The majority of instrument user groups had fairly consistent bias between reported reference intervals and measured reference samples.

3.5. Free thyroxine

FT4 reference intervals were highly variable across laboratories for the majority of instrument/assay users, particularly for the pediatric age groups (Supplemental Fig. 3, Table 1). The %V for reported URLs decreased across the age range from 20%, to 17% to 13% for children, adolescents, and adults, respectively.

The CV_{BL} for FT4 across all laboratories was 14%, with both Beckman and Siemens assay users having the highest CV_{BL} of 7.5% (Table 2). Similar to calcium, the average reported FT4 concentrations from each manufacturer group were compared to the ARM as no target value was available. Abbott and Beckman assay users were biased low compared to the ARM, while Roche and Siemens assay users were biased high (Supplemental Fig. 7A).

The %V was greater than the CV_{BL} in the majority of cases for FT4 (Table 2). This was true for all assay manufacturers analyzed together and for each instrument/assay user separately, except for the average URL reported by Abbott assay users. The majority of instrument/assay user groups had fairly consistent bias between reported reference intervals and measured reference samples. However, Abbott assay users had

Table 1
Variation in reference intervals.

Analyte	Age/sex	Manufacturer	n	Lower reference limit (LRL)			Upper reference limit (URL)		
				Range	Mean	%V[a]	Range	Mean	%V
ALT (U/L)	2 years/male	All	37	0–21	4		30–80	49	22.7%
		Abbott	4	0–11	7		30–51	42	21.4%
		Beckman	8	0–11	6		39–55	47	12.1%
		Ortho	8	0–21	3		44–72	53	20.1%
		Roche	5	0–0	0		35–50	41	13.1%
		Siemens	12	0–12	3		30–80	53	26.5%
	14 years/female	All	37	0–11	3		24–80	44	30.2%
		Abbott	4	0–8	6		24–40	33	21.9%
		Beckman	8	0–11	6		28–55	45	20.4%
		Ortho	8	0–9	2		30–66	47	22.6%
		Roche	5	0–0	0		31–36	34	6.4%
		Siemens	12	0–10	3		25–80	49	36.8%
	50 years/male	All	37	0–21	5		35–80	54	22.2%
		Abbott	4	0–8	5		40–55	45	15.7%
		Beckman	8	0–10	6		35–55	47	14.2%
		Ortho	7	0–21	6		50–72	58	16.8%
		Roche	5	0–0	0		40–50	44	9.8%
		Siemens	12	0–17	6		40–80	63	19.6%
ALP (U/L)	2 years/male	All	37	0–185	113		126–550	355	27.9%
		Abbott	4	40–176	123	48.8%	150–530	357	43.8%
		Beckman	8	40–175	123	32.2%	130–420	340	28.1%
		Ortho	8	38–175	133	30.8%	126–440	359	28.1%
		Roche	5	0–160	40		281–390	317	16.0%
		Siemens	12	0–185	120		140–550	379	27.6%
	14 years/female	All	37	0–170	75		126–525	331	41.9%
		Abbott	4	40–62	50	18.2%	150–455	265	52.3%
		Beckman	8	40–170	87	46.5%	130–500	376	35.7%
		Ortho	8	38–154	85	41.7%	126–500	345	43.2%
		Roche	5	0–50	19		187–300	222	23.1%
		Siemens	12	0–160	94		140–525	359	41.1%
	50 years/male	All	37	0–61	39		100–180	131	12.9%
		Abbott	4	35–61	44	26.3%	115–150	140	11.9%
		Beckman	8	30–42	35	12.0%	100–130	120	9.9%
		Ortho	7	30–60	38	27.0%	120–180	134	15.4%
		Roche	5	30–40	38	11.8%	129–145	135	6.5%
		Siemens	12	0–55	39		105–170	135	13.9%
Calcium (mmol/L)	2 years/male	All	37	2.02–2.30	2.19	2.7%	2.45–2.85	2.64	3.4%
		Abbott	4	2.12–2.30	2.23	3.8%	2.62–2.85	2.68	4.2%
		Beckman	8	2.12–2.25	2.19	1.8%	2.55–2.70	2.63	2.3%
		Ortho	8	2.10–2.25	2.19	2.1%	2.51–2.74	2.63	3.6%
		Roche	5	2.10–2.20	2.18	2.0%	2.58–2.70	2.66	2.3%
		Siemens	12	2.02–2.30	2.19	3.5%	2.45–2.80	2.64	4.2%
	14 years/female	All	37	2.02–2.30	2.16	3.3%	2.51–2.80	2.62	2.1%
		Abbott	4	2.20–2.29	2.24	2.0%	2.60–2.80	2.66	3.5%
		Beckman	8	2.10–2.25	2.14	2.3%	2.58–2.70	2.61	1.5%
		Ortho	8	2.10–2.25	2.15	2.9%	2.51–2.74	2.61	2.7%
		Roche	5	2.10–2.20	2.14	2.2%	2.55–2.65	2.60	1.5%
		Siemens	12	2.02–2.30	2.17	4.5%	2.55–2.70	2.62	1.9%
	50 years/male	All	36	2.02–2.25	2.13	2.3%	2.50–2.80	2.59	2.0%
		Abbott	4	2.10–2.25	2.19	2.9%	2.55–2.80	2.64	4.1%
		Beckman	8	2.10–2.15	2.12	1.1%	2.58–2.70	2.61	1.5%
		Ortho	7	2.10–2.20	2.13	2.3%	2.52–2.60	2.57	1.3%
		Roche	5	2.10–2.20	2.15	1.8%	2.50–2.65	2.59	2.2%
		Siemens	12	2.02–2.18	2.11	2.3%	2.52–2.62	2.57	1.3%
Creatinine (μmol/L)	2 years/male	All	37	0–65	18		31–140	59	39.7%
		Abbott	4	0–34	11		43–63	53	16.3%
		Beckman	8	0–50	20		40–140	69	45.9%
		Ortho	8	0–58	24		38–110	60	39.7%
		Roche	6	0–27	13		31–62	46	26.9%
		Siemens	11	0–65	18		35–125	60	39.4%
	14 years/female	All	37	0–56	37		68–120	88	15.9%

Table 1 (*Continued*)

Analyte	Age/sex	Manufacturer	n	Lower reference limit (LRL)			Upper reference limit (URL)		
				Range	Mean	%V[a]	Range	Mean	%V
		Abbott	4	0–49	22		68–105	86	19.8%
		Beckman	8	24–53	40	22.2%	70–120	91	17.7%
		Ortho	8	29–50	41	14.8%	69–110	86	15.7%
		Roche	6	0–45	36		68–95	82	14.1%
		Siemens	11	0–56	38		68–110	91	15.5%
	50 years/male	All	36	0–65	52		93–140	113	7.9%
		Abbott	4	0–65	48		110–115	112	2.1%
		Beckman	8	45–62	55	12.0%	100–140	114	10.9%
		Ortho	7	50–58	53	8.0%	110–130	115	6.2%
		Roche	6	45–60	56	10.7%	100–110	106	3.7%
		Siemens	11	0–65	49		93–130	115	8.2%
Free T4 (pmol/L)	2 years/male	All	32	7–12	9	16.1%	15–37	20	19.7%
		Abbott	7	9–11	10	10.7%	18–21	19	6.0%
		Beckman	10	7–10	8	9.8%	15–23	19	15.0%
		Roche	6	8–12	10	18.6%	20–37	24	26.2%
		Siemens	9	7–11	9	13.4%	16–26	21	14.7%
	14 years/female	All	32	7–13	9	15.3%	14–28	20	16.6%
		Abbott	7	9–11	10	10.0%	17–19	18	4.8%
		Beckman	10	7–9	8	9.5%	14–23	19	19.8%
		Roche	6	8–13	10	17.1%	20–28	23	12.6%
		Siemens	9	7–11	9	12.7%	16–26	21	14.6%
	50 years/male	All	31	7–12	9	16.7%	15–25	21	12.9%
		Abbott	6	9–11	9	7.0%	19–20	19	1.4%
		Beckman	10	7–9	8	10.5%	15–23	20	17.3%
		Roche	6	9–12	11	11.9%	22–25	23	5.1%
		Siemens	9	7–12	10	14.6%	18–25	22	9.2%
Hemoglobin (g/L)	2 years/male	All	35	100–115	111	4.2%	127–150	138	3.7%
		Abbott	3	110–115	112	2.6%	135–140	138	2.1%
		Beckman	15	100–115	111	4.7%	130–150	138	3.8%
		Siemens	3	110–115	112	2.2%	135–140	137	2.1%
		Sysmex	14	102–115	110	4.5%	127–147	138	4.3%
	14 years/female	All	35	108–130	119	2.9%	133–165	156	3.8%
		Abbott	3	120–120	120	0.0%	153–160	158	2.6%
		Beckman	15	115–130	119	3.2%	149–165	157	2.8%
		Siemens	3	120–120	120	0.0%	153–160	158	2.6%
		Sysmex	14	108–125	118	3.3%	133–160	154	5.0%
	50 years/male	All	34	115–140	134	4.2%	150–180	173	3.7%
		Abbott	2	138–140	139	1.0%	170–180	175	4.0%
		Beckman	15	115–140	134	5.1%	150–180	173	4.4%
		Siemens	3	135–140	137	2.1%	175–180	177	1.6%
		Sysmex	14	120–137	133	3.7%	160–180	171	3.3%
Sodium (mmol/L)	2 years/male	All	37	132–138	135	1.1%	143–148	145	1.1%
		Abbott	4	133–138	135	1.5%	143–147	145	1.1%
		Beckman	8	132–135	134	0.9%	143–148	145	1.1%
		Ortho	8	133–137	135	1.2%	143–148	146	1.2%
		Roche	5	135–138	136	1.0%	145–147	146	0.6%
		Siemens	12	133–136	135	0.7%	143–148	145	1.2%
	14 years/female	All	37	133–137	135	0.9%	143–148	146	1.1%
		Abbott	4	133–135	135	0.7%	143–147	146	1.3%
		Beckman	8	133–135	134	0.8%	143–148	145	1.1%
		Ortho	8	133–137	135	1.2%	143–148	146	1.2%
		Roche	5	135–137	136	0.6%	145–147	146	0.6%
		Siemens	12	133–136	135	0.7%	143–148	146	1.2%
	50 years/male	All	37	133–138	135	0.9%	144–148	146	0.7%
		Abbott	4	133–135	135	0.7%	145–147	146	0.8%
		Beckman	8	133–138	135	1.2%	145–148	146	0.7%
		Ortho	8	133–137	135	1.1%	145–148	146	0.8%
		Roche	5	135–137	136	0.6%	145–147	146	0.6%
		Siemens	12	133–136	135	0.7%	144–148	146	0.9%

Table 1 (*Continued*)
The missing CV data for lower reference limits is due to several lower limits of zero giving exaggerated differences. CV = (SD / mean) * 100.
n = sample size.
[a] %V: percent variation for reference intervals.

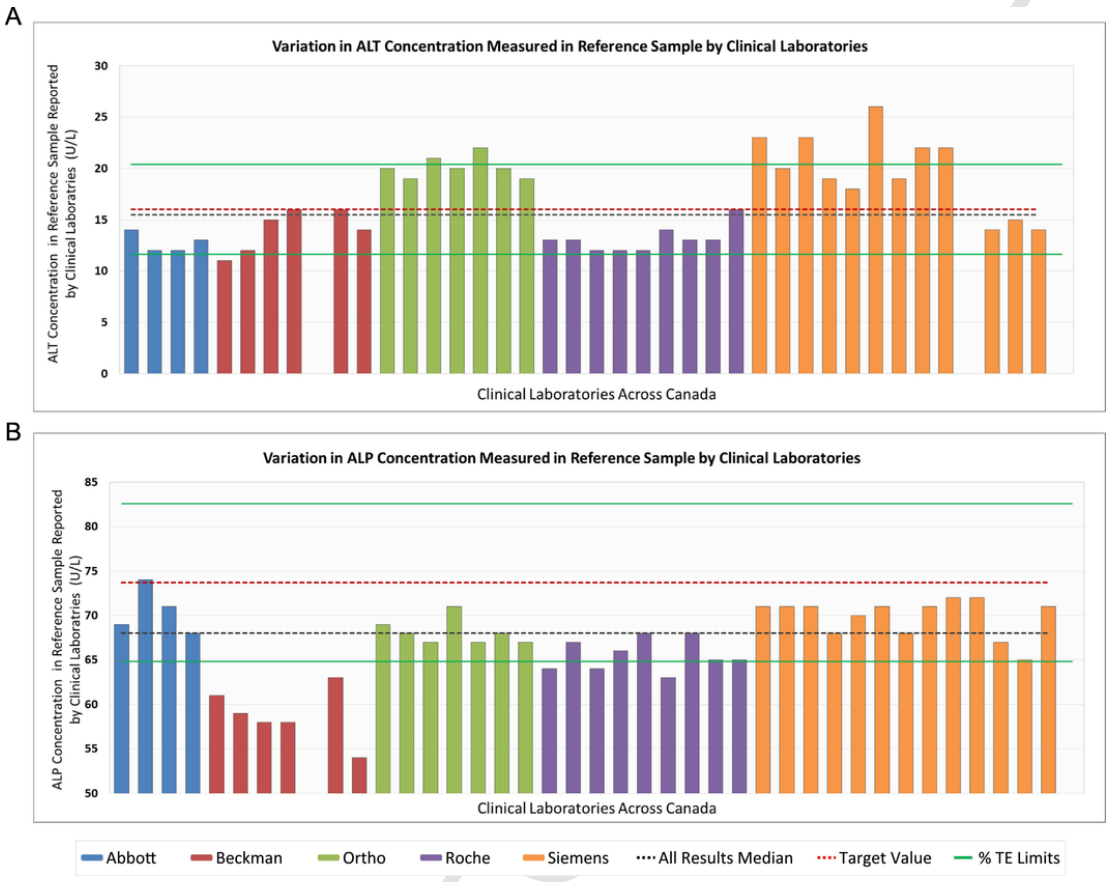

Fig. 4. (A) Alanine aminotransferase (ALT) and (B) alkaline phosphatase (ALP) measurements of a commutable reference sample obtained by clinical laboratories across Canada, grouped by instrument manufacturer. The All Results Median (ARM), target value, and total error allowable (%TE) limits of the target value are shown for reference (%TE = 28% and 12% for ALT and ALP, respectively).

a percent bias to ARM of 10% for the LRL, although the percent bias to target of the reference sample measurement was − 7.3%.

3.6. Hemoglobin

Reference intervals had little variation across laboratories for all instrument users and age ranges (Supplemental Fig. 4, Table 1). The %V across all instrument/assay users were very low, with the URL %V being 3.7%, 3.8%, and 3.7% for children, adolescents, and adults, respectively. Hemoglobin was only included in the reference interval survey, not the reference sample measurement. As variation in reported reference intervals and reference sample measurement cannot be compared, hemoglobin will not be discussed further.

3.7. Sodium

There was little variation in sodium reference intervals across laboratories for all instrument users and age groups (Supplemental Fig. 5, Table 1). Across all instrument users, %V for LRLs were 1.1%, 0.9%, and 0.9% for children, adolescents, and adults, respectively.

The CV_{BL} for sodium across all laboratories was only 1.1%, with Ortho assay users having the highest CV_{BL} of 1.2% (Table 2). All instrument users were slightly biased low compared to the target value (Supplemental Fig. 7B).

The %V and CV_{BL} for sodium were both very low (Table 2). This was the case for all laboratories analyzed together and for each instrument/assay user separately. The majority of instrument user groups had fairly consistent bias between reported reference intervals and measured reference samples. However, Beckman assay users had a percent bias to ARM of 7.5% for the URL, although the percent bias to target of the reference sample measurement was − 2.6%.

4. Discussion

Harmonized reference intervals are essential to achieve standardized patient care. Ensuring laboratory results are comparable across laboratories and are interpreted consistently will substantially reduce medical error and enhance the use of electronic medical records. Laboratories need to use evidence-based reference intervals obtained from reliable sources to accompany laboratory test results. The present survey assessed the current state of pediatric and adult reference intervals in Canada by obtaining information on reference intervals currently used for seven pre-selected analytes from 37 laboratories across the country. These laboratories were spread geographically across Canada, allowing a representative snapshot of the reference intervals in use. Reference samples were also sent to laboratories to assess the variation in results for six analytes. It is evident from this survey and measurement data that the state of reference intervals in Canada is alarmingly

Table 2
Comparing variation and bias between reference sample results and reference intervals.

Analyte	Instrument	CV_{BL}	%V (LRL)	%V (URL)	% bias to target or ARM (test sample measurement)[a]	% bias to ARM (LRL)	% bias to ARM (URL)
ALT, U/L	All	24.6%		**30.2%**			
	Abbott	7.5%		**21.9%**	− 20.3%		− 13.8%
	Beckman	15.0%		**20.4%**	− 12.5%		**12.5%**
	Ortho	5.3%		**22.6%**	25.9%		18.1%
	Roche	9.7%		6.4%	− 18.1%		− 15.5%
	Siemens	19.7%		**36.8%**	22.4%		21.7%
ALP, U/L	All	6.6%		**41.9%**			
	Abbott	3.8%	**18.2%**	**52.3%**	− 4.3%	− 23.1%	− 11.7%
	Beckman	5.2%	**46.5%**	**35.7%**	− 20.2%	**33.7%**	**25.2%**
	Ortho	2.1%	**41.7%**	**43.2%**	− 7.5%	**30.0%**	**15.0%**
	Roche	2.8%		**23.1%**	− 11.1%	− 70.8%	− 25.9%
	Siemens	3.1%		**41.1%**	− 5.2%	**44.2%**	**19.5%**
Calcium, mmol/L	All	2.6%	**3.3%**	2.1%			
	Abbott	1.2%	**2.0%**	**3.5%**	− 0.5%	**4.0%**	**2.4%**
	Beckman	2.2%	**2.3%**	1.5%	1.4%	**− 0.3%**	0.4%
	Ortho	2.2%	**2.9%**	**2.7%**	2.5%	0.1%	0.3%
	Roche	1.9%	**2.2%**	1.5%	− 0.6%	− 0.4%	**0.1%**
	Siemens	2.8%	**4.5%**	1.9%	− 1.4%	**0.9%**	**0.8%**
Creatinine, μmol/L	All	5.0%		**15.9%**			
	Abbott	3.7%		**19.8%**	1.1%	**− 44.4%**	1.2%
	Beckman	1.3%	**22.2%**	**17.7%**	1.8%	0.0%	7.4%
	Ortho	1.3%	**14.8%**	**15.7%**	3.7%	1.6%	1.6%
	Roche	4.1%		**14.1%**	2.9%	**− 10.8%**	**− 3.1%**
	Siemens	6.9%		**15.5%**	6.8%	**− 5.9%**	7.0%
FT4, pmol/L	All	14.3%	**15.3%**	**16.6%**			
	Abbott	6.0%	**10.0%**	4.8%	− 7.3%	**10.2%**	− 11.4%
	Beckman	7.5%	**9.5%**	**19.8%**	− 14.2%	− 10.7%	− 6.5%
	Roche	4.7%	**17.1%**	**12.6%**	19.4%	14.1%	13.0%
	Siemens	7.5%	**12.7%**	**14.6%**	4.5%	2.2%	2.1%
Sodium, mmol/L	All	1.1%	0.9%	1.1%			
	Abbott	0.7%	0.7%	**1.3%**	− 1.4%	− 0.4%	**0.3%**
	Beckman	0.8%	0.8%	**1.1%**	− 2.6%	− 0.7%	**7.5%**
	Ortho	1.2%	1.2%	1.2%	− 0.5%	**0.2%**	**0.7%**
	Roche	0.9%	0.6%	0.6%	− 1.4%	**0.6%**	**0.4%**
	Siemens	0.6%	**0.7%**	**1.2%**	− 0.9%	− 0.1%	**0.3%**

%V and % bias to target values based on reference limits for a 14 year old female. The missing %V data for LRLs is due to several LRLs of zero giving exaggerated differences. The missing % bias to ARM data for ALT LRLs is due to an ARM value of zero, rendering an invalid % bias to ARM calculation. Bolded values indicate those where %V is greater than CV_{BL}. Bolded % bias to ARM values indicate those where the URL or LRL bias is opposite to the test sample measurement bias. % bias to target or ARM are shown for individual manufacturer users, not across all laboratories. This is because the % bias varies substantially between different manufacturer users, and therefore the overall % bias has no real meaning. % bias to target (or ARM) = deviation / target(or ARM) × 100, CV = (SD/mean) ∗ 100.

ARM: all results median, CV_{BL}: between-laboratory variation for commutable reference sample measurement, %V: percent variation for reference intervals, LRL: lower reference limit, URL: upper reference limit.

[a] % Bias calculation for test sample measurements were calculated as % bias to target values for all analytes except calcium and FT4 which were calculated as % bias to ARM.

variable, even between laboratories using instruments from the same manufacturer. It is also evident that the degree of reference interval variability is not warranted based on differences in laboratory reference sample measurements. This is supported by the fact that the between-laboratory variation in reference intervals (%V), even between laboratories using the same instrument, is greater than the variation in reference sample results (CV_{BL}). The AACB Harmonisation Group in Australia showed similar findings among laboratories in Australia and we are now able to conclude that this is also the case for Canadian Laboratories [14,15].

The largest %V in the reported URLs for ALT was 30% for 14 year old females. The broad range in URLs across laboratories for ALT has been a vexing issue and is thought to be the result of poorly characterized reference populations used to establish URLs, which may even include persons with subclinical liver disease [16]. Although ALT is also highly variable when measured in a reference sample by different laboratories (CV_{BL} = 25%), the variation of reported URLs was greater. Furthermore, the overall ALT variability in reference sample measurements was lower than the %TE determined by Ricos et al. of 28% [13]. However, only 75% of participating laboratories' ALT result for the reference sample were within 28% of the target value. Assays measure the catalytic activity of ALT, rather than the actual amount of the enzyme. Therefore, if components of the enzymatic reaction system (e.g. pH and buffer, temperature, presence of activators and/or inhibitors, substrate nature and concentration) are changed, the magnitude of the measured activity will also change [17]. As a result, if two procedures measure the activity of the same enzyme, but under different analytical conditions, they may give different results.

Reported reference intervals for ALP also varied across laboratories, particularly for reference intervals reported for children (%V = 28%) and adolescents (%V = 42%). ALP concentrations are much higher in the pediatric population, decrease in adolescence and remain fairly constant throughout adulthood [10]. It is quite evident from Fig. 2A-B that some laboratories are using appropriate pediatric reference intervals to interpret ALP pediatric tests results, while other laboratories are using adult reference intervals to inappropriately interpret pediatric test results (i.e. report lower reference intervals). The variation in reference intervals was much greater than the variation in test results, with the URL %V for a 14 year old female of 42% compared to the

CV_{BL} of 6.6%. Furthermore, the overall CV_{BL} was lower than the %TE of 12% [13], but only 75% of laboratories' ALP result for the reference sample were within 12% of the target value. ALP is measured using an enzymatic assay in which ALP catalyzes the cleavage of phosphate from 4-nitrophenyl phosphate to form 4-nitrophenoxide. 4-nitrophenoxide then undergoes spontaneous rearrangement at alkaline pH to form a colourless compound, which is measured by absorbance [18]. As mentioned above, varying components of the enzymatic reaction system can lead to different results.

Reference intervals reported for creatinine also varied across laboratories, most significantly for children (%V = 40%). Creatinine concentrations are much lower in the pediatric population, increase in adolescence and remain fairly constant through adulthood [10]. It is similarly evident in Fig. 3A–B that some laboratories are using appropriate pediatric creatinine reference intervals, while some laboratories are using inappropriate adult reference intervals (i.e. report higher reference intervals). The CV_{BL} for creatinine measurements was relatively low (i.e. 5.0%) and was below the %TE of 8.9% [13] and 88% of laboratories' creatinine result for the reference sample were within 8.9% of the target value. Reliable serum creatinine measurements are required for accurate glomerular filtration rate (GFR) estimation, and therefore are also important to accurately diagnose and monitor patients with chronic kidney disease. Therefore, The National Kidney Disease Education Program (NKDEP) Laboratory Working Group, in collaboration with international professional organizations, developed a plan to enable standardization and improved accuracy of serum creatinine measurements in clinical laboratories worldwide [19].

The variation in reference intervals for FT4 were again higher than the variation in measured reference samples, with a %V of 17% for the URL of a 14 year old female, compared to a CV_{BL} of 14%. However, the CV_{BL} still exceeded the %TE of 8% [13] and only 38% of laboratories' FT4 result for the reference sample were within 8% of the ARM. The need for standardization of thyroid function tests, particularly FT4 tests, has been recognized, resulting in the IFCC forming a Working Group for Standardization of Thyroid Function Tests [20]. Among their goals is to standardize FT4 tests and subsequently allow the use of common reference intervals for result interpretation [20]. Free hormones often have greater discordance between methods than total-hormone methods, as their concentrations are lower and adequate separation of free from bound hormone must occur [21].

Reference intervals for calcium and sodium varied the least between laboratories and across manufacturer users. However, the CV_{BL} for calcium was just within the %TE of 2.6% and only 74% of laboratories' calcium results were within 2.6% of the ARM [13]. Furthermore, the CV_{BL} for sodium exceeded the %TE of 0.73% [13]and only 53% of laboratories' sodium results were within 0.73% of the target value. Both sodium and calcium are physiologically tightly regulated and therefore it is expected that their reference intervals be similar across laboratories. CALIPER pediatric reference intervals for calcium required only one age partition from 1–<19 years [8]. Reference intervals were established for both calcium and sodium using CHMS data for ages 3–79 years, which also showed that both URLs and LRLs changed very little across the age range [10].

Laboratories are often concerned with adopting the same reference interval as other laboratories because they claim their instrument/assay and/or local population is unique and therefore they need to establish their own set of reference intervals. If measurements of a particular analyte are biased high for a certain instrument or manufacturer, the reported reference intervals would be expected to be biased high to compensate. When we compared the percent bias between reported URLs and LRLs and reference sample results, it was evident that this was not the case; instead, there was a discordance in the bias between the average reference sample result and the average reported reference limit in the majority of cases. Therefore, this suggests that the variation in reference intervals across instruments cannot be explained by the bias between the results obtained on instruments by different manufacturers.

For some analytes or methods, it may be more practical to use instrument-specific reference intervals to accommodate for calibration and method differences until a higher level of agreement between manufacturers can be reached. Instrument-specific reference intervals will facilitate standardized laboratory test interpretation, while ensuring the potential instrument bias in test results is considered when adopting an appropriate reference interval. This would be a logical approach, especially for the majority of analytes that are not yet standardized based on primary reference methods and/or lack traceability to a primary or secondary reference material [22]. However, from our assessment of reference sample results obtained from laboratories across Canada, several laboratories using instruments from the same manufacturer still varied substantially. In addition to analytical variation across instruments posing a challenge to harmonized reference intervals, population-specific differences may also hinder the feasibility of harmonized reference intervals. For example, Ichihara et al. observed extensive regional differences across Asia in commonly measured analytes even after adjusting for age, sex, and lifestyle variables, suggesting a profound effect of genetic and environmental factors on observed regional differences [23]. Thus, in addition to being instrument-specific, harmonized reference intervals may also only be feasible across populations with minimal biological variation.

For harmonized reference intervals to be used across the country it is important that laboratory test results are first comparable between laboratories. It is important to ensure all laboratories using harmonized reference intervals are part of a standardized quality assurance program to ensure both laboratory test results and laboratory test result interpretation are consistent. However, involvement, commitment, and most importantly, support, from several stakeholders, including clinicians, accreditation programs, quality assurance programs, clinical chemists, and other laboratory professionals is crucial to the widespread implementation and overall success of harmonized reference intervals in Canada. The CSCC hRI Working Group plans to work towards establishing harmonized reference intervals using the most up-to-date evidence-based reference intervals available to ultimately improve patient care.

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.clinbiochem.2017.06.006.

Acknowledgements

This survey was supported by CSCC and CEQAL. CSCC helped support financially, while CEQAL supported by provided reference samples for the reference sample measurements. The authors thank all laboratories and clinical chemists across Canada who participated in the reference interval survey and/or reference sample measurement.

References

[1] A.K. Aarsand, S. Sandberg, How to achieve harmonisation of laboratory testing - the complete picture, Clin Chim Acta Int J Clin Chem. 432 (2014 May 15) 8–14.
[2] M. Panteghini, Implementation of standardization in clinical practice: not always an easy task, Clin. Chem. Lab. Med. 50 (7) (2012 Feb 29) 1237–1241.
[3] L.M. Cruz-Carlos, N. Monge-Azemar, X. Fuentes-Arderiu, Report of the II European symposium on clinical laboratory and in vitro diagnostics industry: "physiological reference values: a shared business?", Clin Chim Acta Int J Clin Chem. 338 (1–2) (2003 Dec) 165–169.
[4] P. Rustad, P. Felding, L. Franzson, V. Kairisto, A. Lahti, A. Mårtensson, et al., The Nordic Reference Interval Project 2000: recommended reference intervals for 25 common biochemical properties, Scand. J. Clin. Lab. Invest. 64 (4) (2004) 271–284.
[5] J. Berg, V. Lane, Pathology harmony; a pragmatic and scientific approach to unfounded variation in the clinical laboratory, Ann. Clin. Biochem. 48 (3) (2011 May 1) 195–197.

[6] J.R. Tate, K.A. Sikaris, G.R. Jones, T. Yen, G. Koerbin, J. Ryan, et al., 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. 35 (4) (2014 Nov) 213–235.
[7] Y. Yamamoto, S. Hosogaya, S. Osawa, K. Ichihara, T. Onuma, A. Saito, et al., Nationwide multicenter study aimed at the establishment of common reference intervals for standardized clinical laboratory tests in Japan, Clin. Chem. Lab. Med. 51 (8) (2013 Aug) 1663–1672.
[8] D.A. Colantonio, L. Kyriakopoulou, M.K. Chan, C.H. Daly, D. Brinc, A.A. Venner, et al., 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. 58 (5) (2012 May 1) 854–868.
[9] D. Bailey, D. Colantonio, L. Kyriakopoulou, A.H. Cohen, M.K. Chan, D. Armbruster, et al., 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. 59 (9) (2013 Sep) 1393–1405.
[10] K. Adeli, V. Higgins, M. Nieuwesteeg, J.E. Raizman, Y. Chen, S.L. Wong, et al., Biochemical marker reference values across pediatric, adult, and geriatric ages: establishment of robust pediatric and adult reference intervals on the basis of the Canadian Health Measures Survey, Clin. Chem. 61 (8) (2015 Aug) 1049–1062.
[11] K. Adeli, J.E. Raizman, Y. Chen, V. Higgins, M. Nieuwesteeg, M. Abdelhaleem, et al., Complex biological profile of hematologic markers across pediatric, adult, and geriatric ages: establishment of robust pediatric and adult reference intervals on the basis of the Canadian Health Measures Survey, Clin. Chem. 61 (8) (2015 Aug) 1075–1086.
[12] C37-A, Preparation and validation of commutable frozen human serum pools as secondary reference materials for cholesterol measurement procedures; approved guideline, In: Clinical Laboratory Standards Institute (CLSI), 1999.
[13] C. Ricos, V. Alvarez, F. Cava, J. Garcia-Lario, A. Hernandez, C. Jimenez, et al., Desirable specifications for total error, imprecision, and bias, derived from intra- and inter-individual biologic variation, Internet 2014. Available from: http://www.westgard.com/biodatabase1.htm.
[14] G.R.D. Jones, A. Barker, J. Tate, C.-F. Lim, K. Robertson, The case for common reference intervals, Clin. Biochem. Rev. 25 (2) (2004 May) 99–104.
[15] G.R. Jones, S.D. Koetsier, RCPAQAP First Combined Measurement and Reference Interval Survey, Clin. Biochem. Rev. 35 (4) (2014 Nov) 243–250.
[16] C.E. Ruhl, J.E. Everhart, Upper limits of normal for alanine aminotransferase activity in the United States population, Hepatology 55 (2) (2012 Feb) 447–454.
[17] I. Infusino, P. Mauro, Standardization in clinical enzymology, EJIFCC. 20 (3) (2009 Oct) 141–147.
[18] D.H. Vroon, Z. Israili, Alkaline phosphatase and gamma glutamyltransferase, in: H.K. Walker, W.D. Hall, J.W. Hurst (Eds.), Clinical Methods: the History, Physical, and Laboratory Examinations, 3rd ed., Butterworths, Boston, 1990. Internet. cited 2017 Jan 6. Available from: http://www.ncbi.nlm.nih.gov/books/NBK203/.
[19] G.L. Myers, Recommendations for improving serum creatinine measurement: a report from the laboratory working Group of the National Kidney Disease Education Program, Clin. Chem. 52 (1) (2006 Jan 1) 5–18.
[20] L.M. Thienpont, J.D. Faix, G. Beastall, Standardization of free T4 and harmonization of TSH measurements: a request for input from endocrinologists and other physicians, Eur Thyroid J. 4 (4) (2015 Oct 6) 271–272.
[21] G.G. Klee, Harmonization and standardization of thyroid function tests, Clin. Chem. 56 (6) (2010 Jun) 879–880.
[22] J.R. Tate, G. Koerbin, K. Adeli, Opinion paper: deriving harmonised reference intervals - global activities, EJIFCC. 27 (1) (2016 Feb) 48–65.
[23] K. Ichihara, Y. Itoh, C.W.K. Lam, P.M.K. Poon, J.-H. Kim, H. Kyono, et al., Sources of variation of commonly measured serum analytes in 6 Asian cities and consideration of common reference intervals, Clin. Chem. 54 (2) (2008 Feb) 356–365.