Opinion Paper

Mary Kathryn Bohn and Khosrow Adeli*

Application of the TML method to big data analytics and reference interval harmonization

<https://doi.org/10.1515/labmed-2020-0133>Received November 12, 2020; accepted December 22, 2020;
published online January 22, 2021

Abstract: Significant variation in reported reference intervals across healthcare centers and networks for many well-standardized laboratory tests continues to exist, negatively impacting patient outcomes by increasing the risk of inappropriate and inconsistent test result interpretation. Reference interval harmonization has been limited by challenges associated with direct reference interval establishment as well as hesitations to apply currently available indirect methodologies. The Truncated Maximum Likelihood (TML) method for indirect reference interval establishment developed by the German Society of Clinical Chemistry and Laboratory Medicine (DGKL) presents unique clinical and statistical advantages compared to traditional indirect methods (Hoffmann and Bhattacharya), increasing the feasibility of developing indirect reference intervals that are comparable to those determined using a direct a priori approach based on healthy reference populations. Here, we review the application of indirect methods, particularly the TML method, to reference interval harmonization and discuss their associated advantages and disadvantages. We also describe the CSCC Reference Interval Harmonization Working Group's experience with the application of the TML method in harmonization of adult reference intervals in Canada.

Keywords: harmonization; reference interval; truncated maximum likelihood.

Introduction

Reference interval harmonization contributes to improved consistency in patient test result interpretation, thus

improving the standard of care through reduced variation in patient management [1]. Clinical laboratories provide the majority of objective medical data to patient charts, significantly influencing clinical decision-making [2]. However, the intricacies associated with laboratory testing from collection to reporting are often not well-recognized or appreciated clinically. Indeed, it is often assumed by patients, clinicians, and other healthcare professionals that the laboratory test results and associated reference intervals reported to patient charts will be identical independent of healthcare/laboratory center [3]. Unfortunately, this is not always the case. As part of the growing expectation for harmonized patient care across healthcare centers driven by network integration, adoption of electronic medical records, and increasing access for patients to their own medical laboratory data, harmonization of the total testing process remains a key priority in the field of laboratory medicine [4]. Indeed, several gains have been made toward the harmonization of pre-analytical and analytical phases of testing. Numerous initiatives have been developed to harmonize pre-analytical handling of specimens through the development of clear quality indicator goals as well as criteria for specimen rejection and acceptance [5, 6]. Increasing automation of laboratory testing lines has also facilitated the harmonization of pre-analytical specimen handling; however, there are still areas in need of improvement. In addition to pre-analytical harmonization, significant efforts by prominent societies (e.g., International Federation for Clinical Chemistry and Laboratory Medicine [IFCC] and American Association for Clinical Chemistry [AACC]) have focused on analytical standardization of key laboratory tests through the development of commutable reference standards and improved metrological traceability [7, 8]. These efforts have significantly improved result comparability for laboratory tests between analytical platforms and clinical laboratories.

Despite key advances in the harmonization of pre-analytical and analytical laboratory processes, test result interpretation continues to vary across healthcare centers. Appropriate and consistent test interpretation relies on reference intervals (RIs) and decision limits [9]. RIs can be defined as the 2.5th and 97.5th percentiles of test result

*Corresponding author: Khosrow Adeli, Molecular Medicine and Clinical Biochemistry, The Hospital for Sick Children, University of Toronto, Toronto, ON, Canada, Phone: +1 416 813 8682, E-mail: khosrow.adeli@sickkids.ca

Mary Kathryn Bohn, Molecular Medicine and Clinical Biochemistry, The Hospital for Sick Children, University of Toronto, Toronto, ON, Canada

values obtained from a healthy population [9]. These health-associated decision-making tools are provided in laboratory test reports to flag abnormal values and alert clinicians of the potential need for follow-up and/or treatment [9, 10]. Clinical laboratories are often hesitant to adopt harmonized or common RIs due to assumed analytical and/or population differences. This is a valid concern as several laboratory tests are not well standardized across manufacturer platforms and thus intra-laboratory RIs should be utilized to compensate for any observed bias. Global population differences, including ethnicity and environmental factors, may also impact laboratory test result values, potentially requiring distinct RIs [11]. However, several national surveys have reported wide variation in RIs across healthcare centers and networks in certain regions, even those using the same analytical platform for test measurement [12, 13]. In some cases, the variation reported in RIs has been shown to be higher than the allowable analytical error [12] or differences between analytical platforms. This suggests the lack of RI harmonization globally cannot be solely explained by analytical or population factors and is likely resultant from historic practices. There is a high risk of inappropriate test result interpretation when RIs are not appropriately harmonized, potentially leading to unnecessary interventions, including repeat testing, as well as erroneous or missed diagnosis. It is a key missing piece in laboratory medicine harmonization efforts.

Barriers to reference interval harmonization

There are several barriers that have delayed the development and implementation of harmonized RIs. Primarily, as discussed, RI harmonization is not appropriate for assays that are not analytically standardized. When significant analytical bias exists between instruments for a given assay, the use of a harmonized RI could lead to clinical errors such as inappropriate flagging and reduced clinical sensitivity. An additional barrier to the development and implementation of harmonized RIs is ensuring they are representative of the patient population for which they will serve, including age, sex, ethnicity, and other population factors. Several countries have undertaken large harmonization initiatives to both develop and implement harmonized RIs into clinical practice. These primarily include: The Nordic Reference Interval Project (NORIP) [14], Australasian Harmonised Reference Intervals for Adults (AHRIA) and Pediatrics (AHRIP) [15], UK Pathology Harmony [16], and the Canadian Society for Clinical Chemists Reference Interval Harmonization Working Group (CSCC hRI) [12] (Table 1). Each initiative differs in their sample collection, sample analysis, and statistical analysis procedures and can be grouped into three main categories: consensus, direct, and indirect. As undertaken by the UK Pathology Harmony group, consensus-based RI

Table 1: Summary of active reference interval harmonization initiatives and their respective approaches.

Harmonization initiative	Population	Main laboratory tests of interest	Overall approach	Main publications
Australasian Harmonized Reference Intervals for Adults (AHRIA) and Pediatrics (AHRIP)	Pediatric & adult	Clinical chemistry: electrolytes, proteins, enzymes.	<i>Combined.</i> Stakeholder discussions based on bias studies, direct Australian reference interval studies (AUSSIE Normals), indirect Bhattacharya RI analysis, and comparison of flagging rates.	[15, 40, 41]
Canadian Society for Clinical Chemistry (CSCC) Reference Interval Harmonization Working Group	Pediatric & adult	Clinical chemistry: electrolytes, proteins, lipids, enzymes, etc. Immunoassay: thyroid hormones, etc.	<i>Combined.</i> Adult: Outpatient data extraction from community reference laboratories, robust data cleaning, indirect RI estimation through TML method and comparison to direct data on healthy Canadians to support recommendations. Pediatric: harmonization through implementation of CALIPER RIs across Canada.	[3, 10, 12]
Nordic Reference Interval Project (NORIP)	Pediatric & adult	Clinical chemistry: electrolytes, proteins, enzymes, etc.	<i>Direct.</i> Recruitment of healthy individuals, sample analysis at multiple laboratories, and nonparametric RI calculations.	[14, 42]
UK Pathology Harmony	Adult	Clinical chemistry: electrolytes, proteins, enzymes.	<i>Consensus.</i> Consideration of variation reported across laboratories and literature regarding known bias.	[16]

harmonization involves evaluating RIs currently used for a laboratory test interpretation across healthcare centers and harmonizing around those values. This approach relies on a priori knowledge and can be easily influenced by scientific opinion. In contrast, some initiatives have recruited large representative cohorts of healthy individuals to directly establish evidence-based harmonized RIs. This direct approach to RI establishment is recommended by Clinical and Laboratory Standards Institute (CLSI) guidelines, but has many challenges in the context of harmonization [9]. Specifically, as harmonized RIs must take diverse populations and several analytical platforms into consideration, a large and representative cohort of healthy individuals is needed. For example, if a RI is to be harmonized for a given analyte, this would require recruitment of individuals from representative geographic regions (e.g., provinces or states) and sample analysis on the analytical platforms used in the laboratories aiming to adopt such RIs. These complexities present major challenges to RI harmonization efforts, including the extensive resources required for recruitment and outreach. Additionally, sampling bias can occur due to insufficient sample size or inappropriate sampling design. Pre-analytical considerations are also less likely to reflect routine operating procedures at a given laboratory, providing an idealized and sometimes non-realistic estimate.

With the emergence of the “big data era” and implementation of electronic health records, many statistical methods have been recently developed to derive robust RIs based on outpatient or inpatient data [17–19]. This indirect approach harnesses the power of stored laboratory data to establish RIs, eliminating the need for recruitment of healthy volunteers and relying on statistical techniques to isolate the “healthy” component of a mixed population dataset (diseased and non-diseased subjects). This approach also offers a more accurate representation of the pre-analytical and analytical procedures at a given laboratory and much lower permissible uncertainty in comparison to direct approaches. While some have expressed concern over the accuracy of these models, the IFCC Committee on Reference Intervals and Decision Limits (C-RIDL) recently stated that indirect approaches are not only a useful adjunct to traditional direct methods but have a number of significant benefits and advantages [20]. In the context of harmonization, indirect methods can be particularly beneficial as, in theory, laboratory data from multiple sites across the region to be harmonized could be extracted for RI derivation, representing the demographics, pre-analytical, and analytical procedures of key stakeholders. They also require minimal resources, lead to faster estimations, and may be considered more

ethical by limiting unnecessary blood draw. One disadvantage to indirect methods is that their outcome relies solely on statistical algorithms. However, as indirect techniques continue to improve due to enhanced computational power and robust statistical algorithms, their application to RI harmonization should be considered.

Indirect statistical techniques and their application to harmonization

Several indirect statistical techniques for RI establishment have been reported in the literature dating as far back as the 1960s, but few have been applied to harmonization. One of the most commonly used and reported indirect techniques is the Hoffmann method [21]. This method assumes that healthy test result values in a mixed population dataset can be represented by the Gaussian portion of the data. Developed in the pre-computer era, this method calculates the cumulative frequency of the data distribution. By plotting the data on normal probability axes, the Gaussian component is theoretically represented by a straight line, which can then be used to extrapolate the 2.5th and 97.5th percentiles [18]. While this method has been extensively published and even used in clinical practice, the RIs derived often are heavily influenced by diseased subpopulations [22]. It is also important to note that some have reported widespread incorrect implementation of the Hoffmann method, by plotting on a linear scale [23, 24]. When applying this method incorrectly, reference limits may be underestimated or overestimated, depending on the specifics of the proportions, means, and variances of the dataset [23, 24]. An arguable improvement of the Hoffmann method is the Bhattacharya method [25]. The Bhattacharya method was also developed in the pre-computer era, using a mathematical differential technique to straighten the Gaussian component of the data, representing healthy test result values. The slope and intercept of this Gaussian component are then used to derive the 2.5th and 97.5th percentiles [18]. Both techniques have been criticized due to their subjective nature and propensity to be heavily influenced by outlying points/subgroups. Additionally, the Bhattacharya method relies on bin size selection and can thus lack reproducibility. Most importantly, the key caveat to these models is the inherent assumption of normality. Both the C-RIDL committee as well as CLSI have criticized these methodologies due to this assumption, especially as laboratory data is likely to be skewed even after Box Cox transformation [9, 26]. The limitations of these models have rightly reduced the use of indirect methods for RI establishment clinically, including

harmonization efforts. However, major advancements have recently been made in computer automation and algorithm modelling that can be harnessed in several areas in healthcare, including precision medicine, machine learning, and RI establishment and harmonization.

Truncated Maximum Likelihood method

In 2007, Arzideh and colleagues proposed a novel method for indirect RI establishment entitled the Truncated Maximum Likelihood (TML) method [26, 27]. This modern and non-graphical computational approach harnesses the statistical power of maximum likelihood estimation to resolve competing distributions (i.e., healthy, non-healthy) in a patient dataset. Specifically, it applies a smoothed kernel density function to estimate the distribution of the entire dataset. The central portion of the dataset is then assumed to represent the “healthy” population and is modeled by a truncated power normal distribution family. The parameters of this distribution are estimated using maximum likelihood techniques and a goodness of fit test, similar to the Kolmogorov Statistic, is used to optimize these parameters. The 2.5th and 97.5th percentiles are then derived based on this distribution [26]. The major advantages of this technique include its computational power and ability to analyze millions of data points simultaneously as well as its reduced subjectivity, using likelihood and fitting techniques to mathematically select for the “healthy” population rather than visual or manual assessment. Additionally, it also makes no assumptions regarding the distribution of pathological values and only transforms the central component of the data, when necessary [26]. While the statistical algorithms applied in this method are quite complex, the German Society of Clinical Chemistry and Laboratory Medicine (DGKL) has made an easily accessible excel-based application to facilitate application globally [28]. The development of this method has led to an extensive list of publications, applying this approach to both pediatric and adult populations [26, 27, 29–34]. In adults, this technique has been applied to laboratory data (e.g., electrolytes, creatinine, TSH) from multiple laboratory centers, providing evidence of its potential use in RI harmonization [32, 34]. In pediatrics, this technique has been applied by Zierk and colleagues to develop continuous RI curves through cubic spline smoothing. This has been completed for many chemistry and hematology parameters, demonstrating excellent concordance to direct studies such as the Canadian Laboratory Initiative on Pediatric Reference Intervals (CALIPER) despite being performed in different

populations and on discrepant analytical platforms [29, 31, 35]. Given these encouraging data, the TML method has prompted discussion among laboratories regarding potential clinical adoption. Additionally, given the challenges associated with using the direct approach in the context of RI harmonization and the ability of the TML method to analyze millions of data points robustly, it presents unique advantages in the field of RI harmonization.

Reference interval harmonization in Canada: experience with TML method

The CSCC hRI working group was established in 2015 to develop and assist in the implementation of evidence-based harmonized RIs across Canada for select analytes. After robust review of the literature and emerging indirect technologies, this group established a novel approach to RI harmonization in adults involving: (1) extraction of data from community reference laboratories across Canada, (2) assessment and removal of outliers and monthly instability, (3) statistical evaluation of age, sex, and center-specific differences, (4) derivation of preliminary harmonized RIs using the TML method, and (5) comparison of established harmonized RIs to direct a priori data in the healthy Canadian population. Thus far, this approach has led to the development of preliminary harmonized RIs for 17 biochemical and immunochemical markers. It is important to note that the TML method is the final calculation step in this approach; however, many factors need to be considered prior to its application to optimize performance. First, indirect data sources need to be selected carefully. Ideally, data sources should serve the population of interest, covering the geographic regions that aim to be harmonized. Data sources should ideally be community reference laboratories that serve outpatients as opposed to tertiary hospitals with higher incidence of patients with severe disease [17]. Inpatient populations can reduce estimate accuracy due to the higher likelihood of medical intervention, including drug treatment, intravenous fluid, and surgery. Repeat measurements should also be removed to ensure the same patient is not included in the dataset more than once, keeping the first test result in calculations. An alternative approach would be to remove any individual with more than one test result over the data extraction period in an effort to exclude individuals with chronic conditions. However, this often reduces sample size significantly. Statistical and manual outlier removal should also be explored, depending on the analyte and population of interest. Second, the period of data

extraction should be selected to provide adequate sample size and minimize significant changes to test result measurement (e.g., transition to another analytical platform). The minimum number of test results recommended for indirect statistical techniques varies between methodologies. In the TML method, a minimum of 4,000 test results per partition is recommended for accurate and robust RI determination [26]. The period of data extraction needed to achieve this sample size may vary depending on test ordering patterns as well as age/sex-specific patterns for each analyte of interest. As the length of the data extraction period increases, there is increased potential for analytical drift to occur. Several groups have monitored this by graphing monthly medians and/or monitoring quality control and external quality assurance data retrospectively [20]. Finally, selecting appropriate age/sex partitioning as well as assessing differences between data sources can pose significant challenges. However, combining visual and statistical assessment, such as the Harris & Boyd method [36], with prior clinical knowledge can facilitate the selection of an appropriate age- and sex-specific partitioning strategy. Additionally, while there is currently no recommended approach to assess statistical and/or clinical differences between laboratory data from different data sources/laboratories, one could consider using validated statistical models such as the Harris & Boyd method, if sources served similar populations (e.g., outpatient). In addition, establishing RIs for each data source individually can provide support regarding the feasibility of their combination and ultimately harmonization. Combining statistical outcomes with existent clinical knowledge can also facilitate the comparison of center-specific RIs and the decision to harmonize. Analyte-specific measurement uncertainty and associated allowable performance limits can also be helpful in this regard.

Taking these factors into consideration, the CSCC hRI working group obtained very large datasets from four community reference laboratories in three Canadian provinces consisting of up to 14 million unique results per laboratory test over a two-year period. They then used robust data cleaning measures followed by the TML method to derive preliminary harmonized RIs. To strengthen their confidence in the preliminary estimates, they took advantage of the work completed and published by RI harmonization groups globally to evaluate discrepancies in recommendations implemented in other countries. They also compared estimates to direct data from the Canadian Health Measures Survey (CHMS) [37]. This study recruited 12,000 healthy children and adults to establish RIs for multiple chemistry, hematology, and immunoassay parameters [37–39]. This combined approach using indirect RI establishment and comparison to published direct RIs established using data

from Canada and other countries provides more rigor and support for the recommended harmonized RIs. Indeed, they observed encouraging concordance between derived harmonized RIs and those established in other cohorts based on consensus, direct, and indirect approaches [14–16, 37, 38]. In addition, by considering these data in conjunction with existent clinical knowledge, few analytes demonstrated differences in reference limits that could be considered clinically significant. These preliminary findings underscore the ability of the TML method to handle big data and effectively resolve competing distributions. It is important to note that ethnicity was not considered in data analyses. However, to support the implementation of derived harmonized RIs across Canada and evaluate their performance clinically, the CSCC hRI working group plans to conduct a verification program wherein healthy adult volunteers that reflect the ethnic distribution of Canada are recruited and serum and plasma samples are collected. Samples will then be sent to laboratories across Canada using different analytical platforms as part of a RI verification program as per CLSI guidelines [9]. Following the completion of this RI verification program, the CSCC hRI working group will focus on the implementation of recommendations across Canada. Implementation is a key and challenging step toward RI harmonization, requiring consistent advocacy and monitoring. Current strategies to optimize success include offering technical support, information packages (e.g., letters to clinicians, verification statistical program etc.) as well as virtual and in-person training/discussions. Long-term adoption monitoring schemes will also be explored.

Summary and outlook

With the increasing expectation for standardized patient care, appropriate RI harmonization is urgently needed. Electronic medical records have enabled the storage of mass amounts of laboratory data. The development of robust and automated statistical RI methods, such as the TML method, has the power to effectively use these data toward RI harmonization, eliminating the need for recruitment of healthy individuals and allowing for achievement of a representative population sample. However, its application should be combined with stringent data cleaning procedures as well as comparison to health-associated data to ensure confidence in the final RI calculations [37]. Finally, the use of retrospective clinical laboratory data is not limited to indirect RI establishment and harmonization. With the emergence of the big data era, we suspect the clinical laboratory will play a critical role in supporting research regarding the application of big data analytics to healthcare,

including machine learning algorithms in precision medicine and laboratory utilization strategies, among others.

Research funding: This work was supported by a Foundation Grant from the Canadian Institutes of Health Research (CIHR) (grant no. 353989). M.K.B. was supported by a CIHR Doctoral Scholarship.

Author contributions: All authors have accepted responsibility for the entire content of this manuscript and approved its submission.

Competing interests: Authors state no conflict of interest.

References

1. Plebani M. Harmonization in laboratory medicine: the complete picture. *Clin Chem Lab Med* 2013;51:741–51.
2. Pennestrì F, Banfi G. Value-based healthcare: the role of laboratory medicine. *Clin Chem Lab Med* 2019;57:798–801.
3. Parker ML, Adeli K. Pediatric and adult reference interval harmonization in Canada: an update. *Clin Chem Lab Med* 2019; 57:57–60.
4. Plebani M. Harmonization in laboratory medicine: requests, samples, measurements and reports. *Crit Rev Clin Lab Sci* 2016; 53:184–96.
5. Lippi G, Baird GS, Banfi G, Bölenius K, Cadamuro J, Church S, et al. Improving quality in the preanalytical phase through innovation, on behalf of the European Federation for Clinical Chemistry and Laboratory Medicine (EFLM) Working Group for Preanalytical Phase (WG-PRE). *Clin Chem Lab Med* 2017;55:489–500.
6. Lippi G, Banfi G, Church S, Cornes M, De Carli G, Grankvist K, et al. Preanalytical quality improvement. In pursuit of harmony, on behalf of European Federation for Clinical Chemistry and Laboratory Medicine (EFLM) Working group for Preanalytical Phase (WG-PRE). *Clin Chem Lab Med* 2015;53:357–70.
7. Stenman U-H. Immunoassay standardization: is it possible, who is responsible, who is capable? *Clin Chem* 2001;47:815–20.
8. Gillery P, Young IS. Standardization: recent achievements of the IFCC scientific division. *Clin Chim Acta* 2017;467:1–3.
9. Horowitz GL, Altaie S, Boyd JC. Defining, establishing, and verifying reference intervals in the clinical laboratory; approved guideline, 3rd ed. Wayne, PA, U.S.A: CLSI; 2010.
10. Adeli K, Higgins V, Trajcevski K, White-Al Habeeb N. The Canadian laboratory initiative on pediatric reference intervals: a CALIPER white paper. *Crit Rev Clin Lab Sci* 2017;54:358–413.
11. Tahmasebi H, Asgari S, Hall A, Higgins V, Chowdhury A, Thompson R, et al. Influence of ethnicity on biochemical markers of health and disease in the CALIPER cohort of healthy children and adolescents. *Clin Chem Lab Med* 2020;58:605–17.
12. Adeli K, Higgins V, Seccombe D, Collier CP, Balion CM, Cembrowski G, et al. National survey of adult and pediatric reference intervals in clinical laboratories across Canada: a report of the CSCC working group on reference interval harmonization. *Clin Biochem* 2017;50:925–35.
13. Jones GRD, Barker A, Tate J, Lim C-F, Robertson K. The case for common reference intervals. *Clin Biochem Rev* 2004;25:99–104.
14. Hilsted L, Rustad P, Aksglæde L, Sorensen K, Juul A. Recommended Nordic paediatric reference intervals for 21 common biochemical properties. *Scand J Clin Lab Invest* 2013;73: 1–9.
15. Tate JR, Sikaris KA, Jones GRD, Yen T, Koerbin G, Ryan J, 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* 2014;35:213–35.
16. Berg J, Lane V. Pathology harmony; a pragmatic and scientific approach to unfounded variation in the clinical laboratory. *Ann Clin Biochem* 2011;48:195–7.
17. Martinez-Sanchez L, Marques-Garcia F, Ozarda Y, Blanco A, Brouwer N, Canalias F, et al. Big data and reference intervals: rationale, current practices, harmonization and standardization prerequisites and future perspectives of indirect determination of reference intervals using routine data. *Adv Lab Med* 2020: 20200034.
18. Farrell C-JL, Nguyen L. Indirect reference intervals: harnessing the power of stored laboratory data. *Clin Biochem Rev* 2019;40: 99–111.
19. Giavarina D, Dorizzi RM, Soffiati G. Indirect methods for reference intervals based on current data. *Clin Chem* 2006;52:335–7.
20. Jones GRD, Haeckel R, Loh TP, Sikaris K, Streichert T, Katayev A, et al. Indirect methods for reference interval determination – review and recommendations. *Clin Chem Lab Med* 2019;57:20–9.
21. Hoffmann G, Lichtinghagen R, Wosniok W. Simple estimation of reference intervals from routine laboratory data. *J Lab Med* 2016; 39:000010151520150104. DOI: <https://doi.org/10.1515/labmed-2015-0104>.
22. Shaw JLV, 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.
23. Holmes DT, Buhr KA. Widespread incorrect implementation of the hoffmann method, the correct approach, and modern alternatives. *Am J Clin Pathol* 2019;151:328–36.
24. Katayev A, Fleming JK. Widespread implementation of the Hoffmann method: a second opinion. *Am J Clin Pathol* 2019;152: 116–7.
25. Bhattacharya CG. A simple method of resolution of a distribution into Gaussian components. *Biometrics* 1967;23:115–35.
26. Arzideh F, Wosniok W, Gurr E, Hinsch W, Schumann G, Weinstock N, et al. A plea for intra-laboratory reference limits. Part 2. A bimodal retrospective concept for determining reference limits from intra-laboratory databases demonstrated by catalytic activity concentrations of enzymes. *Clin Chem Lab Med* 2007;45: 1043–57.
27. Haeckel R, Wosniok W, Arzideh F. A plea for intra-laboratory reference limits. Part 1. General considerations and concepts for determination. *Clin Chem Lab Med* 2007;45:1033–42.
28. RLE Estimator. DGKL Working Group; n.d. [Accessed 22 Oct 2020].
29. Zierk J, Hirschmann J, Toddenroth D, Arzideh F, Haeckel R, Bertram A, et al. Next-generation reference intervals for pediatric hematology. *Clin Chem Lab Med* 2019;57:1595–607.
30. Zierk J, Arzideh F, Haeckel R, Rauh M, Metzler M, Ganslandt T, et al. Indirect determination of hematology reference intervals in adult patients on Beckman Coulter UniCell DxH 800 and Abbott CELL-DYN Sapphire devices. *Clin Chem Lab Med* 2019;57:730–9.

31. Zierk J, Arzideh F, Rechenauer T, Haeckel R, Rascher W, Metzler M, et al. Age- and sex-specific dynamics in 22 hematologic and biochemical analytes from birth to adolescence. *Clin Chem* 2015; 61:964–73.
32. Arzideh F, Wosniok W, Haeckel R. Indirect reference intervals of plasma and serum thyrotropin (TSH) concentrations from intra-laboratory data bases from several German and Italian medical centres. *Clin Chem Lab Med* 2011;49:659–64.
33. Zierk J, Arzideh F, Kapsner LA, Prokosch HU, Metzler M, Rauh M. Reference interval estimation from mixed distributions using truncation points and the Kolmogorov-Smirnov distance (kosmic). *Sci Rep* 2020;10:1704.
34. Arzideh F, Wosniok W, Haeckel R. Reference limits of plasma and serum creatinine concentrations from intra-laboratory data bases of several German and Italian medical centres. Comparison between direct and indirect procedures. *Clin Chim Acta* 2010;411:215–21.
35. Colantonio DA, Kyriakopoulou L, Chan MK, Daly CH, Brinc D, Venner AA, 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* 2012;58:854–68.
36. Harris EK, Boyd JC. On dividing reference data into subgroups to produce separate reference ranges. *Clin Chem* 1990;36:265–70.
37. Adeli K, Higgins V, Nieuwesteeg M, Raizman JE, Chen Y, Wong SL, 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* 2015;61:1049–62.
38. Adeli K, Higgins V, Nieuwesteeg M, Raizman JE, Chen Y, Wong SL, et al. Complex reference values for endocrine and special chemistry biomarkers 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* 2015;61:1063–74.
39. Adeli K, Raizman JE, Chen Y, Higgins V, Nieuwesteeg M, Abdelhaleem M, 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* 2015;61:1075–86.
40. Koerbin G, Tate JR. Harmonising adult reference intervals in Australia and New Zealand – the continuing story. *Clin Biochem Rev* 2016;37:121–9.
41. Jones GR, Koetsier S. Uptake of recommended common reference intervals for chemical pathology in Australia. *Ann Clin Biochem* 2017;54:395–7.
42. Rustad P, Felding P, Franzson L, Kairisto V, Lahti A, Mårtensson A, et al. The Nordic Reference Interval Project 2000: recommended reference intervals for 25 common biochemical properties. *Scand J Clin Lab Invest* 2004;64:271–84.