Clinica Chimica Acta 467 (2017) 70–82

Contents lists available at ScienceDirect

Clinica Chimica Acta

journal homepage: www.elsevier.com/locate/clinchim

A global multicenter study on reference values: 1. Assessment of methods for derivation and comparison of reference intervals

Kiyoshi Ichihara [a,*], Yesim Ozarda [b], Julian H Barth [c], George Klee [d], Ling Qiu [e], Rajiv Erasmus [f], Anwar Borai [g], Svetlana Evgina [h], Tester Ashavaid [i], Dilshad Khan [j], Laura Schreier [k], Reynan Rolle [l], Yoshihisa Shimizu [m], Shogo Kimura [a], Reo Kawano [a,n], David Armbruster [o], Kazuo Mori [p], Binod K Yadav [q], on behalf of, Committee on Reference Intervals and Decision Limits, International Federation of Clinical Chemistry and Laboratory Medicine:

[a] Faculty of Health Sciences, Yamaguchi University Graduate School of Medicine, Ube, Japan
[b] Dept of Medical Biochemistry, Uludag University School of Medicine, Bursa, Turkey
[c] Blood Sciences, Leeds General Infirmary, Leeds, UK
[d] Laboratory Medicine and Pathology, Mayo Clinic College of Medicine, Rochester, MN, USA
[e] Dept of Clinical Laboratory, Peking Union Medical College, Beijing, China
[f] Chemical Pathology, University of Stellenbosch and National Health Laboratory Services, Tygerberg, South Africa
[g] King Abdullah International Medical Research Center, King Saud bin Abdulaziz University for Health Sciences, Pathology, King Abdulaziz Medical City, Jeddah, Saudi Arabia
[h] Beckman Coulter LLC, Moscow, Russia
[i] Dept of Laboratory Medicine, P. D. Hinduja National Hospital and Medical Research Centre, Mumbai, India
[j] Dept of Pathology, National University of Medical Sciences, Rawalpindi, Pakistan
[k] Dept of Clinical Biochemistry, Faculty of Pharmacy and Biochemistry, University of Buenos Aires, Argentina
[l] Newborn Screening Center Visayas, West Visayas State University Medical Center, Iloilo, Philippines
[m] Dept of Medical Life Science, Faculty of Medical Bioscience, Kyushu University of Health and Welfare, Nobeoka, Japan
[n] Clinical Research Center, Yamaguchi University Hospital, Ube, Japan
[o] Abbott Laboratories, Abbott Park, IL, USA
[p] Beckman Coulter Japan, Tokyo, Japan
[q] Dept of Biochemistry, Maharajgunj Medical Campus, Institute of Medicine, Kathmandu, Nepal

ARTICLE INFO

Article history:
Received 31 December 2015
Received in revised form 2 September 2016
Accepted 20 September 2016
Available online 22 September 2016

Keywords:
Reference intervals
Multicenter study
Latent abnormal values exclusion method
Harmonization
Standardization
Standard deviation ratio
Serum panel
Modified Box-Cox power transformation
Ethnic differences

ABSTRACT

Objectives: The IFCC Committee on Reference Intervals and Decision Limits coordinated a global multicenter study on reference values (RVs) to explore rational and harmonizable procedures for derivation of reference intervals (RIs) and investigate the feasibility of sharing RIs through evaluation of sources of variation of RVs on a global scale.

Methods: For the common protocol, rather lenient criteria for reference individuals were adopted to facilitate harmonized recruitment with planned use of the latent abnormal values exclusion (LAVE) method. As of July 2015, 12 countries had completed their study with total recruitment of 13,386 healthy adults. 25 analytes were measured chemically and 25 immunologically. A serum panel with assigned values was measured by all laboratories. RIs were derived by parametric and nonparametric methods.

Results: The effect of LAVE methods is prominent in analytes which reflect nutritional status, inflammation and muscular exertion, indicating that inappropriate results are frequent in any country. The validity of the parametric method was confirmed by the presence of analyte-specific distribution patterns and successful Gaussian transformation using the modified Box-Cox formula in all countries. After successful alignment of RVs based on the panel test results, nearly half the analytes showed variable degrees of between-country differences. This finding, however, requires confirmation after adjusting for BMI and other sources of variation. The results are reported in the second part of this paper.

Conclusion: The collaborative study enabled us to evaluate rational methods for deriving RIs and comparing the RVs based on real-world datasets obtained in a harmonized manner.

* Corresponding author at: Department of Clinical Laboratory Sciences, Faculty of Health Sciences, Yamaguchi University Graduate School of Medicine, Minamikogushi 1-1-1, Ube 755-8505, Japan.
E-mail address: ichihara@yamaguchi-u.ac.jp (K. Ichihara).

http://dx.doi.org/10.1016/j.cca.2016.09.016
0009-8981/

Non-standard abbreviations

3N-ANOVA	three-level-nested ANOVA
Alb	albumin
AFP	alpha-fetoprotein
ALP	alkaline phosphatase
ALT	alanine aminotransferase
AMY	amylase
ARG	Argentina
AST	aspartate aminotransferase
BMI	body mass index
Ca	calcium
CDL	clinical decision limit
CEA	carcinoembryonic antigen
CK	creatine kinase
CI	confidence interval
Cl	chloride
CLSI	Clinical and Laboratory Standards Institute
C3	complement component 3
C4	complement component 4
CA125	carcinoma antigen 125
CHN	China
Cre	creatinine
C-TLM	Committee on Traceability in Laboratory Medicine
CRM	certified reference materials
CRP	C-reactive protein
CV(b)	CV of the regression slope b
dDL	drugs for dyslipidemia
dHT	drugs for hypertension
dHU	drugs for hyperuricemia
DL	decision limit
DMS	data management system
Drk	drinking habit
E2	estradiol
Fe	iron
FSH	follicular stimulating hormone
GCA	general chemistry analytes
GGT	gamma-glutamyltransferase
Glu	glucose
GBR	Great Britain
GH	growth hormone
HDL-C	HDL-cholesterol
HBV	hepatitis B virus
HCV	hepatitis C virus
HIV	human immunodeficiency virus
HRT	hormone replacement therapy
IgA	immunoglobulin A
IgG	immunoglobulin G
IgM	immunoglobulin M
IND	India
IP	inorganic phosphate
JPN	Japan
K	potassium
LAVE	latent abnormal values exclusion
LDH	lactate dehydrogenase
LDL-C	LDL-cholesterol
LH	luteinizing hormone
LL	lower limit
MALRA	major-axis linear regression analysis
MRA	multiple regression analysis
Me	median
Mg	magnesium
Na	sodium
NP	non-parametric
OC	oral contraceptives
P	parametric
Pr	probability
PAK	Pakistan
PHL	Philippines
PSA	prostate specific antigen
PTH	parathyroid hormone
Prog	progesterone
PRL	prolactin
QC	quality control
RI	reference interval
RMP	reference measurement procedure
RT	reference tests used in the LAVE method
RUS	Russia
RV	reference value
r_p	partial correlation coefficient
SAU	Saudi Arabia
SD	standard deviation
SDR	standard deviation ratio
Sk	skewness
SV	sources of variation
TBil	total bilirubin
Testo	testosterone
TC	total cholesterol
Tf	transferrin
TG	triglycerides
TP	total protein
TUR	Turkey
TSH	thyroid stimulating hormone
UA	uric acid
UL	upper limit
ZAF	South Africa

1. Introduction

The reference interval (RI) is defined simply as the prediction interval which includes the central 95% of reference values (RVs), or test results from well-defined healthy individuals (reference individuals). Establishment of well-controlled, reliable RIs is an important mission for all clinical laboratories. In reality, it is very challenging, because it is not easy to recruit a sufficient number of reference individuals, to control pre-analytical variables, and to apply all statistical methods in appropriate manners. The international guideline entitled "Defining, establishing, and verifying reference intervals in the clinical laboratory" was first published as a possible solution in 1996 by collaboration between the Clinical and Laboratory Standards Institute (CLSI) and the International Federation of Clinical Chemistry and Laboratory Medicine (IFCC) (the latest 2010 version is designated as CLSI/IFCC EP28-A3c (formerly, C28-A3) [1]. However, the descriptions are generally theoretical in nature, and the rationales of the recommendations have not been well evaluated by implementing actual, down-to-earth RI studies. In fact, there have been controversies over its pragmatic aspects, such as the rationale for secondary exclusion, the use of parametric vs. nonparametric derivation, and determining how to judge the need for partitioning RVs [2–5].

In the init terim period, for the achievement of global standardization for major laboratory tests, the current consensus is to derive the RIs in a reproducible manner based on sufficient sample size by collaboration between multiple laboratories and to use the derived RIs in common or by transference [6–8]. The Scandinavian groups conducted a multicenter study for derivation of common RIs based on standardized test values and found virtually no between-country differences in the Nordic countries [9]. However, the appropriateness of the study protocol and the method used for evaluating between-laboratory differences need to be evaluated by use of newer statistical methods. In addition, the IFCC Committee on Plasma Proteins concurrently conducted two RI studies mainly aimed at deriving common RIs for major serum proteins in East and Southeast Asian countries in 2000 and 2004; it revealed apparent between-country differences in many of the analytes, especially those of inflammatory markers [10,11].

With this background and its mission of promoting proper implementation of multicenter RI studies, the Committee on Reference Intervals and Decision Limits (C-RIDL) was established by the IFCC in 2005. The primary project of the C-RIDL, as planned in early 2010, was to clarify between-country differences in RIs on a global scale and to seek the most practical and harmonizable methodologies for conducting the RI studies. The key strategy of the global study was to make RVs comparable among the countries through measurement of a common panel of serum samples. After conducting a feasibility study to confirm the validity of cross comparison of test results based on the panel test results [12] and the elaboration of the common protocol [13], the global multicenter study was launched at the end of 2011 on a trial basis. The more