Clinical Chemistry 63:6
1141–1150 (2017)

Biological Variation Estimates Obtained from 91 Healthy Study Participants for 9 Enzymes in Serum

Anna Carobene,[1,10*] Thomas Røraas,[2] Una Ørvim Sølvik,[3] Marit Sverresdotter Sylte,[4] Sverre Sandberg,[2,3,4,10] Elena Guerra,[1] Irene Marino,[1] Niels Jonker,[5,10] Gerhard Barla,[5] William A. Bartlett,[6,10] Pilar Fernandez-Calle,[7,10] Jorge Díaz-Garzón,[7] Francesca Tosato,[8] Mario Plebani,[8] Abdurrahman Coşkun,[9,10] Mustafa Serteser,[9] Ibrahim Unsal,[9] and Ferruccio Ceriotti[1] on behalf of the European Biological Variation Study of the EFLM Working Group on Biological Variation

BACKGROUND: We sought to develop estimates of biological variation (BV) for 9 enzymes in blood serum as part of the European Biological Variation Study.

METHODS: Ninety-one healthy study participants (38 male and 53 female, 21–69 years old) were phlebotomized in each of 10 consecutive weeks at 6 European laboratories. The same preanalytical sample-handling protocol was followed at each center before transport to San Raffaele Hospital, Milan, Italy, for analysis. Sera were stored at −80 °C before analysis in duplicate within a single run on an ADVIA 2400 Clinical Chemistry System (Siemens Healthcare) following a protocol designed to minimize analytical imprecision. Assay traceability was established using frozen sera with target values assigned by reference methods. The results were subjected to outlier analysis before CV-ANOVA to deliver valid BV estimates. Results for 9 enzymes were subsequently partitioned for graphical display allowing visual assessment of the effects of country of origin, sex, and age on BV estimates.

RESULTS: We found no effect of country upon the observed variation, but overall sex-related differences were evident for alanine amino transferase (ALT), γ-glutamyl transferase (GGT), and creatine kinase (CK). The following estimates for within-subject BV (CV_I) and between-subject BV (CV_G), respectively, were obtained: ALT: 9.3%, 28.2%; aspartate aminotransferase: 9.5%, 20.3%; GGT: 8.9%, 41.7%; alkaline phosphatase : 5.3%, 24.9%; lactate dehydrogenase: 5.2%, 12.6%; CK: 14.5%, 31.5%; amylase: 6.8%, 30.4%; pancreatic α-amylase: 6.3%, 24.9%; and lipase (LIP): 7.7%, 23.8%.

CONCLUSIONS: All CV_I and some CV_G estimates were lower than those reported in the online BV 2014 updated database. Analytical performance specifications derived from BV can be applied internationally.

The measurements of the catalytic activities of enzymes are among the most frequently requested tests in clinical laboratory medicine. The IFCC continues to focus on the standardization of the methods for determination of catalytic activity of enzymes *(1)*, but as yet clear performance specifications for their measurement within clinical laboratories remain to be defined for contemporary methods. Availability of reliable biological variation (BV)[11] data will be important for the definition of analytical performance specifications (APS) for these enzyme measurements. In addition, well-characterized BV data have many other applications that include setting of reference change values (RCV) to assess the significance of change in serial results from patients and the assessment of the utility of population based reference intervals *(2)*. The first European Federation of Clinical Chemistry and Laboratory Medicine (EFLM) Strategic

[1] Servizio Medicina di Laboratorio, Ospedale San Raffaele, Milan, Italy; [2] Norwegian Quality Improvement of Primary Health Care Laboratories (Noklus), Haraldsplass, Hospital, Bergen, Norway; [3] Department of Global Public Health and Primary Care, University of Bergen, Bergen, Norway; [4] Laboratory of Clinical Biochemistry, Haukeland University Hospital, Bergen, Norway; [5] Certe, Wilhelmina Ziekenhuis Assen, Assen, the Netherlands; [6] Blood Sciences, Ninewells Hospital & Medical School, Scotland, UK; [7] Hospital Universitario La Paz, Madrid, Spain, and Quality Analytical Commission of the Spanish Society of Clinical Chemistry, Barcelona, Spain; [8] Department of Laboratory Medicine, University Hospital, Padua, Italy; [9] Acibadem University, School of Medicine, Atasehir, Istanbul, Turkey; [10] Biological Variation Working Group, European Federation of Clinical Chemistry and Laboratory Medicine, Milan, Italy (http://efcclm.eu/science/wg-biological-variation, www.biologicalvariation.com).

* Address correspondence to this author at: San Raphael Hospital, via Olgettina 60, 20132 Milan, MI, Italy. Fax 039-02-26434178; e-mail carobene.anna@hsr.it.

Received December 14, 2016; accepted January 24, 2017.
Previously published online at DOI: 10.1373/clinchem.2016.269811

[11] Nonstandard abbreviations: BV, biological variation; APS, analytical performance specification; RCV, reference change values; EFLM, European Federation of Clinical Chemistry and Laboratory Medicine; WG-BV, EFLM Working Group on BV; EuBIVAS, European Biological Variation Study; ALT, alanine amino transferase; ALP, alkaline phosphatase; AMY, α-amylase; AST, aspartate aminotransferase; CK, creatine kinase; GGT, γ-glutamyl transferase; LDH, lactate dehydrogenase; LIP, lipase; PAMY, pancreatic α-amylase; OSR, San Raffaele Hospital, Milan; RM, reference materials; IRMM, Institute for Reference Materials and Measurements; CV_I, within-subject biological variation; CV_G, between-subject biological variation; CV_{APS}, analytical performance specification for imprecision; B_{APS}, analytical performance specification for bias; TE_{ASP}, analytical performance specification for total error.

Table 1. Number, sex, age, and BMI of enrolled subjects by each laboratory.

	Men age 20–60 years	Age, years, median (range)	BMI, kg/m^2, median (range)	Women age 20–50 years	Age, years, median (range)	BMI, kg/m^2, median (range)	Women age 50–70 years	Age, years, median (range)	BMI, kg/m^2, median (range)
Italy-Milan (19 subjects)	9	38 (24–59)	25.2 (20.8–30.0)	7	34 (24–48)	22.7 (17.6–23.9)	3	58 (55–59)	22.8 (19.4–27.5)
Italy-Padua (14 subjects)	5	32 (27–35)	22.5 (19.0–23.5)	8	33 (27–49)	19.8 (18.7–23.2)	1	69	18.6
Norway (15 subjects)	7	37 (28–42)	24.3 (18.7–26.9)	6	43 (32–48)	21.7 (18.7–24.4)	2	63 (63–63)	24.6 (23.7–25.5)
The Netherlands (12 subjects)	4	36 (23–45)	24.0 (18.1–26.3)	6	39 (29–49)	21.7 (20.9–24.2)	2	60 (59–60)	23.0 (20.7–25.3)
Spain (16 subjects)	7	34 (26–54)	25.1 (19.5–32.5)	7	26 (24–48)	21.7 (17.9–23.1)	2	60 (60–60)	21.3 (21.2–21.4)
Turkey (15 subjects)	6	27 (22–35)	27.5 (22.2–29.9)	9	33 (21–38)	21.2 (18.3–27.3)	–	–	–
Total (91 subjects)	38	35 (22–59)	24.4 (18.1–32.5)	43	34 (21–49)	21.3 (17.6–27.3)	10	60 (55–69)	22.1 (18.6–27.5)

Conference Consensus Statement identified 3 models for assignation of APS based on knowledge of clinical outcomes, BV, and state of the art of measurement *(3)*. A task and finish group has been set up by the EFLM to allocate tests to different models *(4)*. While at this point in time clinical outcome based data are not available to enable APS setting for enzymes, BV data are available for APS setting and the other applications mentioned above in a database made available on the World Wide Web *(5)*.

The online database was last updated in 2014 by the Spanish Society of Laboratory Medicine *(5)*. The data applying to enzymes in that database, however, are problematic since they are mainly drawn from old publications. A systematic review of those publications revealed that they do not meet all of the requirements of the recently published EFLM checklist for BV studies *(6)* and that the assays used predate the introduction of IFCC standardized enzyme methods of measurement *(2, 7)*. For these reasons, the EFLM working group on biological variation (WG-BV) have derived a new set of BV data utilizing samples collected in a biobank created by the European Biological Variation Study (EuBIVAS) *(8)* employing contemporary methods to measure enzymes that form part of the core repertoire of the clinical biochemistry laboratory service. The study delivers BV data for alanine aminotransferase (ALT) 2.6.1.2, alkaline phosphatase (ALP) 3.1.3.1, α-amylase (AMY) 3.2.1.1, aspartate aminotransferase (AST) 2.6.1.1, creatine kinase (CK) 2.7.3.2, γ-glutamyl transferase (GGT) 2.3.2.2, lactate dehydrogenase (LDH) 1.1.1.27, pancreatic lipase (LIP) 3.1.1.3, and pancreatic α-amylase (PAMY) 3.2.1.1 using the large and well- defined set of samples collected by the WG-BV *(8)*. In addition, APS based on BV for these enzymes are proposed based on these new data.

Materials and Methods

Characteristics of enrolled study participants, inclusion and exclusion criteria of the health status of the participants, and the protocol used by all laboratories to collect and store serum samples have been previously described *(8)*. The EuBIVAS protocol was approved by the Institutional Ethical Review board of the coordinating central laboratory in agreement with the World Medical Association Declaration of Helsinki and by the Ethical Board/Regional Ethics Committee of each center. All participants gave informed consent.

SAMPLE COLLECTION AND HANDLING

Briefly, the project involved 6 European laboratories (Milan, Italy; Bergen, Norway; Madrid, Spain; Padua, Italy; Istanbul, Turkey; and Assen, The Netherlands). Blood samples were collected from 91 volunteers, 38 men and 53 women, age 21–69 years (Table 1) and allowed to clot followed by separation of serum by centrifugation at 3000*g* for 10 min at room temperature.

The separated serum was aliquoted in Nalgene cryovials and stored at −80 °C.

Initial study inclusion required that participants complete an enrollment questionnaire to verify their health status. Principal characteristics of the lifestyle of

the enrolled participants are shown in Table 1 in the Data Supplement that accompanies the online version of this article at http://www.clinchem.org/content/vol63/issue6. The volunteers provided blood samples once per week for 10 consecutive weeks.

The collection started in all centers between the 13th and the 16th week of 2015. For each participant, blood samples were drawn on a specific day (Tuesday to Friday) and at the same hour between 8 and 10 AM at each weekly visit, mainly by the same phlebotomists. At the end of the sample collection phase, the samples from all of the study participants at each center were sent as a single delivery, frozen on dry ice, to the coordinating central laboratory in San Raffaele Hospital in Milan (OSR) and stored in a dedicated freezer at −80 °C *(8)*.

ANALYTICAL METHODS

Measurements of enzyme activities were performed on ADVIA 2400 Clinical Chemistry System (Siemens Healthcare) in OSR, using the following Siemens reagents: ALT/GPT, code 03036926, analytical principle IFCC plus pyridoxal phosphate; AST/GOT, code 03039631, IFCC, plus pyridoxal phosphate; GGT, code 02011954; CK NAC, code 02096577; LDH lactate-pyruvate, code 03035814, ALP/AMP, code 03035814; AMY, code 03031177, PAMY code 01410820, Lipase code 01984894. These contemporary methods for measurement of enzyme activity are all standardized to IFCC reference measurement procedures with the exception of pancreatic amylase and lipase for which IFCC optimized methods are not available.

To verify traceability to the reference measurement systems, reference materials (RM) were analyzed in triplicate on 3 different days before starting the samples measurements. Specifically, for ALT, AST, GGT, ALP, LDH, and CK, fresh frozen serum pools at 2 different concentrations were analyzed. The target values of the pools were assigned by the IFCC reference measurement procedures performed manually by the Standardization Laboratory in OSR (Laboratory number 16 of the RELA—external quality control for reference laboratories) *(9)*. For AMY and PAMY, Institute for Reference Materials and Measurements (IRMM)/IFCC 456 material was analyzed as RM.

All samples from the same study participant were analyzed in duplicate within a single run. As internal QCs, Liquid Assayed Multiqual Level 1 (code 694) and Level 2 (code 695) (Bio-Rad Laboratories, Segrate, Italy) were analyzed in duplicate in each single run.

DATA ANALYSIS

Data was analyzed using CV-ANOVA, where data were first transformed using the CV-transformation *(10)*. Then, to assure homogeneity, outlier identification and removal was performed on replicates and samples on the transformed data. Homogeneity of analytical CV (between-replicates) was verified using Bartlett tests *(11)* and homogeneity of within-subject biological variation (CV_I) using Cochran tests *(12)*.

The Shapiro–Wilk test *(13)* was used to verify the normality of the residuals. For between-subject biological variation (CV_G) estimation, data were natural log transformed. The Shapiro–Wilk test was again used to verify the normality. The Dixon–Reed criterion *(14)* was used to detect outliers in between-subject mean values. The steady-state situation was verified by a visual check for a systematic common change in each enzyme's activity during the whole length of the study, while larger individual systematic changes would be detected by the homogeneity test of the within-subject BV.

The data sets from males and females in each group initially were analyzed separately. The statistical significance of the differences in the CV_I and CV_G between subgroups was determined by the overlapping of their 95% CI. The differences between mean values in male and female groups were tested using Student *t* tests. When no significant differences between male and female mean values were found, CV_I was set equal to the all-subjects CV_I and CV_G to the all-subjects CV_G. When mean values of males were significantly different from females (ALT, GGT, CK), the lowest of the 2 sex-related CV_Gs was used to calculate APS.

To evaluate differences among subgroups from different countries, visual inspection of the various participants' values ordered according to the country was conducted (Fig. 1A and 2A; also see online Supplemental Figs. 1–7A). In Figs. 1B and 2B (and online Supplemental Figs. 1–7B) participants' values are ordered by sex and age. Differences between the 2 female groups (younger than 50 years and menopause age) and sexes were examined by visual inspection. When doubt arose, CV-ANOVA of subgroups was performed.

BV estimates were used to calculate APS for analytical imprecision (CV_{APS}), for analytical bias (B_{APS}) and for total allowable error (TE_{APS}). The following criteria were applied: CV_{APS} = ½ all participants CV_I; $B_{APS} = 0.25(CV_I^2 + CV_G^2)^{0.5}$ and $TE_{APS} = (CV_{APS} \times 1.65) + B_{APS}$ *(15)*.

Data analyses were performed using Excel 2010 and XLSTAT, statistical software for Excel *(16)*.

Results

Number, sex, age, and body mass index of the selected study population, after exclusions of six participants based on the laboratory measurements made at each visit *(8)*, are shown in Table 1.

Online Supplemental Table 1 shows the characteristics of the participants. Of those enrolled, about 3% of participants were regular smokers, there was low to moderate alco-

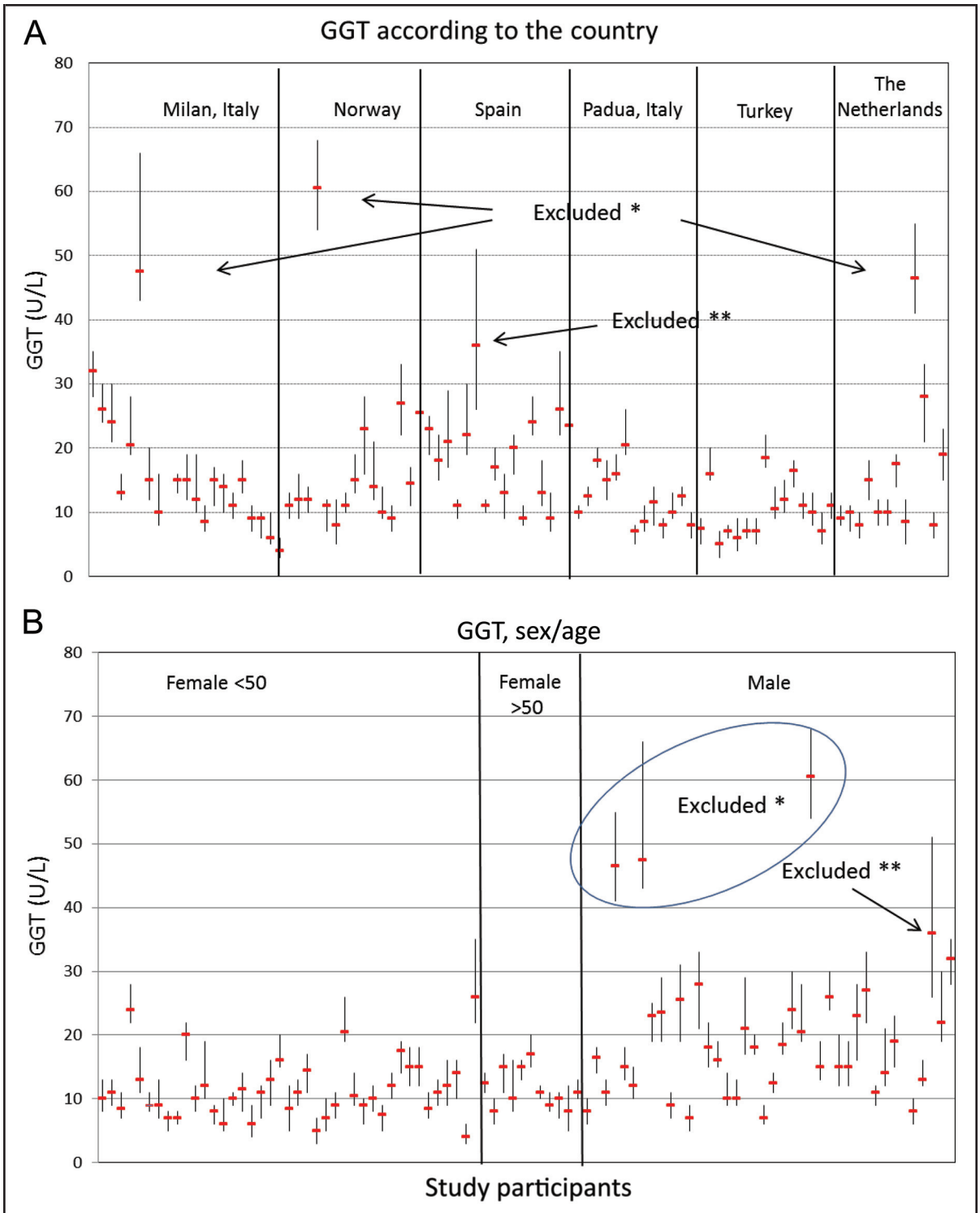

Fig. 1. Biological Variation of GGT.

(A), median (horizontal bar) (minimum – maximum) for each participant ordered by country. The arrows indicate excluded participants because of Reed criterion (*) or Bartlett test (**). (B), range minimum – maximum and median value (red hyphen) for each participant ordered by sex and age. The arrows indicate study participants excluded because of Reed criterion (*) or Bartlett test (**).

Table 2. Numbers of excluded samples and the reasons for exclusion.

	Data recognized as outliers						
	Homogeneity (Bartlett and Cochran tests)			Reed and Dixon	Numbers of results used to estimate CV_I data		
	Replicate (analytical homogeneity)	Samples (within homogeneity)	Subjects (within homogeneity)	Subjects (between)	Results	Subjects	Total % of outliers
ALT all subjects	5	39	3	0	1566	88	8.4%
ALT male	2	12	2	0	643	36	9.1%
ALT female	3	18	1	0	944	52	5.8%
AST all subjects	4	10	1	0	1655	90	2.6%
AST male	1	8	0	0	686	38	2.4%
AST female	3	4	1	0	970	52	3.0%
GGT all subjects	8	22	1	3	1586	87	4.2%
GGT male	2	10	1	3	608	34	5.9%
GGT female	6	10	4	0	908	49	10.3%
ALP all subjects	22	24	4	2	1532	85	9.0%
ALP male	15	5	3	2	587	33	12.5%
ALP female	7	22	0	0	943	53	5.2%
LDH all subjects	1	12	1	1	1638	89	2.6%
LDH male	1	3	1	1	656	36	3.8%
LDH female	0	6	0	0	988	53	1.2%
CK all subjects	0	53	7	0	1485	84	14.2%
CK male	0	17	3	0	624	35	13.1%
CK female	0	27	4	0	883	49	13.2%
AMY all subjects	1	25	2	0	1612	89	5.3%
AMY male	1	11	2	0	639	36	9.0%
AMY female	0	11	0	0	979	53	2.2%
PAMY all subjects	0	16	5	2	1552	84	7.7%
PAMY male	0	8	2	0	654	36	7.9%
PAMY female	0	8	3	2	896	48	7.8%
LIP all subjects	7	19	4	0	1584	87	7.3%
LIP male	4	11	0	0	678	38	3.7%
LIP female	3	6	3	0	930	50	7.7%

hol intake, practically no therapeutic drugs were used, and the participants had a rather high level of physical exercise. Details are shown in online Supplemental Table 1.

An overview of the numbers of excluded samples and the different outlier tests for exclusion are given in Table 2. In total, about 8% of the obtained data were excluded.

In online Supplemental Table 2, the results obtained when measuring reference materials are shown. All the results fall within the expanded uncertainty limits of the reference materials, except for the low concentration for ALT [result obtained: 42.7 U/L; target value: 39.6 (2) U/L]. A reference measurement system for pancreatic lipase was not available.

For each enzyme, a figure partitioning results for the country of origin (A) and for sex ordered by age, was drawn. Examples are presented for GGT (Fig. 1) and for LDH (Fig. 2). For the other enzymes, see online Supplemental Figs. 1–7.

As shown in Figs. 1 and 2 and in online Supplemental Figs. 1–7, no country-specific related differences were visually observed, while sex-related differences regarding the mean values were evident for some enzymes (ALT, GGT, and CK) (Table 3).

Table 3 shows the BV data for "all subjects" and for males and females. BV estimates are compared to those currently reported in online 2014 updated BV database *(5)*.

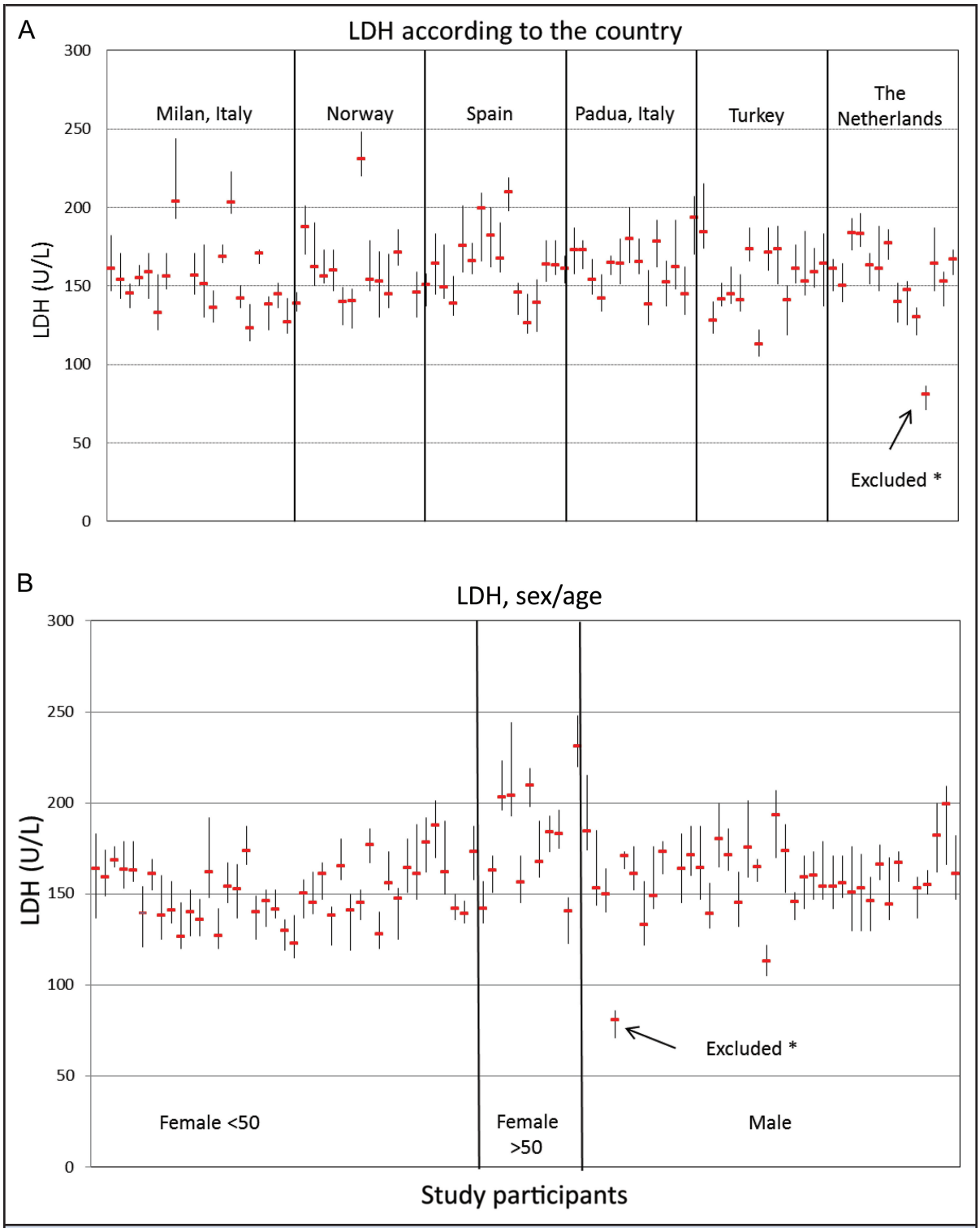

Fig. 2. Biological Variation of LDH.

(A), median (horizontal bar) (minimum – maximum) for each participant ordered by country. The arrows indicate excluded participants because of Reed criterion (*) or Bartlett test (**). (B), range minimum – maximum and median value (red hyphen) for each participant ordered by sex and age. The arrows indicate study participants excluded because of Reed criterion (*).

Table 3. BV data of 9 enzymes partitioned between males and females.[a]

	Number of subjects	Total number of results	Mean number of samples/subject	Mean number of replicates/sample	Mean value (95% CI), U/L	CV_A (95% CI), %	CV_I (95% CI), %	CV_G (95% CI), %	Online BV database[c] CV_I %	Online BV database[c] CV_G %
ALT all subjects	88	1566	8.99	1.96	22.3 (21.8–22.7)		**9.3 (8.7–10.0)**	28.0 (24.7–33.9)		
ALT males[b]	36	643	9.00	1.96	25.8 (25.1–26.5)	6.7 (6.4–7.1)	10.1 (9.2–11.2)	28.2 (22.6–37.4)	19.4	41.6
ALT females[b]	52	944	9.15	1.96	20.3 (19.9–20.8)		9.6 (8.8–10.5)	**25.2 (21.3–32.1)**		
AST all subjects	90	1655	9.30	1.96	20.9 (20.6–21.2)		**9.5 (9.0–10.2)**	**20.3 (17.7–24.2)**		
AST males	38	686	9.13	1.96	22.1 (21.8–22.5)	5.5 (5.2–5.8)	10.3 (9.5–11.3)	14.7 (12.4–20.3)	12.3	23.1
AST females	52	970	9.42	1.96	20.1 (19.7–20.4)		8.9 (8.3–9.8)	22.4 (18.7–28.3)		
GGT all subjects	87	1586	9.22	1.96	13.8 (13.4–14.1)		**8.9 (8.1–9.7)**	45.1 (38.9–54.2)		
GGT males[b]	34	608	9.09	1.94	17.5 (16.9–18.1)	8.7 (8.39.2)	8.3 (7.1–9.5)	41.7 (33.5–57.4)	13.4	42.15
GGT females[b]	49	908	9.33	1.97	11.9 (11.6–12.3)		8.3 (7.4–9.5)	**34.2 (28.1–43.4)**		
ALP all subjects	85	1532	9.47	1.98	61.2 (60.3–62.0)		**5.3 (5.0–5.7)**	**24.9 (21.4–29.3)**	6.45	26.1
ALP males	33	587	9.24	1.86	63.0 (61.6–64.3)	2.8 (2.6–2.9)	5.0 (4.65.5)	25.4 (20.3–34.1)		
ALP females	53	943	8.98	1.96	59.8 (58.7–60.8)		5.4 (5.15.9)	24.2 (20.6–30.9)		
LDH all subjects	89	1638	9.30	1.96	159.3 (158.1–160.6)		**5.2 (5.0–5.5)**	**12.6 (10.8–14.7)**		
LDH males	36	656	9.25	1.94	159.9 (158.3–161.5)	1.2 (1.1–1.2)	5.5 (5.1–6.0)	10.2 (8.1–13.3)	8.6	14.7
LDH females	53	988	9.40	1.97	158.8 (157.0–160.5)		5.2 (4.9–5.6)	14.0 (11.6–17.3)		
CK all subjects	84	1485	8.87	1.99	105.9 (103.4–108.3)		**14.5 (13.8–15.4)**	37.9 (32.8–45.8)		
CK males[b]	35	624	8.97	1.97	134.5 (130.4–138.6)	0.9 (0.9–1.0)	16.0 (14.8–17.5)	31.5 (25.2–42.5)	22.8	40
CK female[b]	49	883	9.02	2.00	87.5 (85.2–89.8)		15.7 (14.6–16.8)	**30.5 (24.8–38.4)**		
AMY all subjects	89	1612	9.18	1.95	69.9 (68.8–71.1)		**6.8 (6.5–7.2)**	**30.4 (26.5–36.3)**		
AMY males	36	639	9.06	1.92	69.0 (67.3–70.7)	1.0 (1.0–1.1)	6.3 (5.9–6.9)	30.0 (23.7–39.1)	8.7	28.3
AMY females	53	979	9.32	1.96	71.0 (69.5–72.6)		7.1 (6.7–7.6)	31.4 (26.6–40.0)		
PAMY all subjects	84	1552	9.29	1.98	32.5 (32.0–33.0)		**6.3 (6.0–6.7)**	**24.9 (21.9–30.1)**		
PAMY males	36	654	9.17	1.96	31.1 (30.4–31.7)	1.8 (1.8–1.9)	5.9 (5.5–6.5)	23.3 (18.8–30.8)	11.7	29.9
PAMY females	48	896	9.35	1.99	33.9 (33.2–34.6)		6.8 (6.3–7.3)	26.0 (22.3–34.2)		
LIP all subjects	87	1584	9.17	1.97	37.2 (36.7–37.8)		**7.7 (7.2–8.3)**	**23.8 (20.6–28.2)**		
LIP males	38	678	9.03	1.95	34.7 (33.9–35.4)	6.4 (6.1–6.7)	6.4 (5.5–7.2)	22.2 (18.1–29.2)	32.2	31.8
LIP females	50	930	9.38	1.97	39.0 (38.3–39.8)		8.9 (8.3–9.8)	23.9 (19.9–30.2)		

[a] Bold text indicates CV_G used to calculate analytical performance specifications.
[b] Significant difference between mean values in males and females (Student's *t*-test).
[c] Minchinela et al. *(5)*.

Table 4. APS for imprecision, bias and total error derived by the BV data reported in Table 3, compared with the desirable specifications reported in online database.[a]

	APS derived from new BV data			APS reported in online database		
	Imprecision %	Bias %	Allowable error %	Imprecision %	Bias %	Allowable error %
	CV_{APS}	B_{APS}	TE_{APS}	CV_{APS}	B_{APS}	TE_{APS}
ALT	4.7	6.7	14.4	9.7	11.48	27.48
AST	4.8	5.6	13.4	6.15	6.54	16.69
GGT	4.5	8.4	15.7	6.7	11.06	22.11
ALP	2.7	6.4	10.7	3.23	6.72	12.04
LDH	2.6	3.4	7.7	4.3	4.3	11.4
CK	7.3	8.4	20.4	11.4	11.5	30.3
AMY	3.4	8.4	13.4	4.4	7.4	14.6
PAMY	3.2	7.8	11.6	5.9	8.0	17.7
LIP	3.9	6.3	12.6	16.1	11.31	37.88

[a] Minchinela et al. *(5)*. The following criteria were applied: imprecision = ½ all subjects CV_I; bias = 0.25 $(CV_I^2 + CV_G^2)^{0.5}$ where CV_I = all subjects CV_I and CV_G = all subjects CV_G; allowable error = imprecision × 1.65 + bias.

Table 4 reports the APS derived from the BV data reported in Table 3.

Discussion

Serum enzyme activity measurements are frequently ordered from clinical laboratories. They represent key biomarkers for the diagnosis and monitoring of diseases of the liver, pancreas, skeletal muscle, and many other systems *(17)*. Characterization of the BV of measures of enzyme activity provides important knowledge that enables assessment of suitability of methods for practice and enables understanding of the clinical significance of the result. Previous studies of BV of enzyme activities have been published, but the validity of those data to practice today needs to be questioned due to changes to the methods through time and concern about the quality of the published studies from which they were derived *(7)*. The current study addresses both of these issues and delivers data from contemporary assays traceable in most cases to IFCC standard methodologies.

The main findings in the present study are that all the CV_I estimates and most of the CV_G estimates are lower than those reported previously (Table 3). The reason for these differences is probably multifactorial, but the current study has characteristics that deliver high confidence regarding these new data. The study addresses uncertainty by delivering a highly powered design that follows a rigorous protocol taking into account the critical list of factors to be considered when performing studies on BV *(6)*. The study not only employed a standardized preexamination- and examination process *(8, 18)* in a very large group of study participants, but also used an improved method for calculation of the CV_I and CV_G *(10)* and a sound approach to the analytical phase of the experiment. The general approach taken within this study therefore delivers more robust, and better characterized, estimates of BV (CV_I and CV_G) data than previously published *(10)*.

The transferability of the data from a single study into clinical applications in a mixed diseased and nondiseased population may appear problematic, but there should be fewer problems with their transferability in terms of performance specification setting. For the latter circumstance, providing that available methods meet the specifications set by use of these new data, we can be assured that the methods meet the most stringent requirements and are fit for application in most, if not all clinical situations. However, the ideal RCVs derived from these data may be a little oversensitive when applied in some clinical applications where the patient sampling conditions in field use are not as well controlled as they were in this study or where there is a specific impact of disease on the CV_I of the measurand. In such situations, alternate studies of BV may be required in particular reference populations. Users of the data may need to apply a pragmatic approach depending upon the application. The data presented in this paper are well defined and are accompanied by sufficient metadata to enable informed decisions regarding their applicability in the user's clinical practice.

The largest differences in CV_I estimates observed in comparison of current data with those reported in the online database are seen for ALT, CK, PAMY, and LIP.

In each case, they are lower and will have a large impact on the APS and other indices utilizing these data (e.g., much lower RCV) (Table 3).

There are some other interesting differences in that the CV_I for the contemporary ALT and AST assays delivered in this study are almost identical while CV_I of ALT is much higher than the CV_I of AST in the online BV database *(5)* although the numbers are not given with an uncertainty estimate (19.4% vs 12.3%, respectively). This might be explained by both the different numbers of papers used to obtain ALT value in comparison to those used to obtain AST value (9 and 13, respectively) in the online database updated 2014, and most probably by the reliability of the methods applied in papers used to obtain those data *(7)*. The values in the database are also the median values of the studies identified *(19)*.

The lower estimates in CV_I found in our study for CK compared with those seen in the online database updated 2014 *(5)* might be explained by the exclusion of 7 subjects detected as outliers in our study. They were excluded on the basis of identification of strenuous exercise in the presampling questionnaire (online Supplemental Table 1). Failure to exclude outliers arising from such causes would deliver higher CV_I values. Outlier removal is an absolute requirement, and it is essential that reports of BV data identify that this is undertaken as part of the data processing *(6)*.

For PAMY and LIP, the differences in CV_Is reported here and those in the online database may be explained by the fact the values reported in the online database are based, respectively, on 2 papers published more than 35 years ago *(20, 21)* and on 3 papers published more than 30 years ago *(20, 22, 23)*. Methodological changes are such that contemporary methods may have very different analytical performance characteristics and specificities.

The Milan conference proposed 3 models to define APS: outcome based, BV based, and state-of-the-art based *(3)*. For enzymes, the situation is particularly complex, and for this reason, they were not taken into consideration in the recent EFLM publication allocating measurands to different models *(4)*. In fact, owing to the very different clinical applications (screening, diagnosis, and monitoring) and to the various clinical conditions in which enzymatic analyses are used, it is difficult to define outcome-based APS. Since the enzymes studied in the present investigation demonstrated rather stable concentrations in healthy individuals for at least 10 weeks, we think that it is rational to use biological variation model to calculate APS for these enzymes. This way of calculating APS is based on the idea to minimize the signal/noise ratio where the analytical uncertainty only accounts for a smaller fraction of the biological variation. The APSs are shown in Table 4 and are stricter compared to what has been published earlier. We acknowledge that the calculation of TE is likely too high *(24)*, but we have included this information to be able to compare with earlier results.

Conclusions

The BV data and derived APS presented here for commonly assayed enzyme activities were derived using a best practice approach to the preanalytical, analytical, and postanalytical (statistical evaluation of the data) phases of the study. The results obtained confirm that the catalytic activities in serum of the 9 enzymes studied present a relatively high biological variability. Having a large data set (91 study participants, 10 samples per participant) meant that a certain number of data points could be removed, for legitimate clinical or statistical reasons (2.6% as AST and LDH to 14.2% of CK), without compromising the power of the study. The study design and delivery enables description of what appears to be the "most typical" within participant BV for each enzyme with attached confidence limits. The combination of careful statistical evaluation of the data together with the use of a well-controlled analytical methodology based on standardized measurement systems may explain why the BV estimates obtained are significantly lower than those reported in the online BV database *(5)*. The use of modern contemporary assays means that the data delivered here are more relevant to current clinical practice. In addition, the multicenter approach delivers a large multinational group of participants with an age distribution and sex balance sufficient to generate data that enable good estimation of the between participant variability and are reliable and powerful enough to deliver estimates of variability on the stratified data sets. Finally, the absence of clear differences in enzyme activity between groups from Turkey, Norway, The Netherlands, Spain, and Italy confirms that the obtained data are widely applicable across healthcare systems and that they can be used to deliver APS for systems to be used internationally.

Author Contributions: *All authors confirmed they have contributed to the intellectual content of this paper and have met the following 3 requirements: (a) significant contributions to the conception and design, acquisition of data, or analysis and interpretation of data; (b) drafting or revising the article for intellectual content; and (c) final approval of the published article.*

Authors' Disclosures or Potential Conflicts of Interest: *Upon manuscript submission, all authors completed the author disclosure form. Disclosures and/or potential conflicts of interest:*

Employment or Leadership: None declared.
Consultant or Advisory Role: None declared.
Stock Ownership: None declared.
Honoraria: None declared.
Research Funding: Siemens Healthcare donated all materials used for the measurements.
Expert Testimony: None declared.
Patents: None declared.

Role of Sponsor: The funding organizations played no role in the design of study, choice of enrolled patients, review and interpretation of data, and final approval of manuscript.

Acknowledgments: We thank Siemens Healthcare for allowing us to work on a dedicated Siemens ADVIA 2400 and for donating to us all materials used for the measurements.

References

1. Siekmann L, Bonora R, Burtis CA, Ceriotti F, Clerc-Renaud P, Ferard G, et al. IFCC primary reference procedures for the measurement of catalytic activity concentrations of enzymes at 37 degrees C. Part 1. The concept of reference procedures for the measurement of catalytic activity concentrations of enzymes. Clin Chem Lab Med 2002;40:631–4.
2. Carobene A. Reliability of biological variation data available in an online database: need for improvement. Clin Chem Lab Med 2015;53:871–7.
3. Sandberg S, Fraser FG, Horvath AR, Jansen R, Jones G, Oosterhuis W, et al. Defining analytical performance specifications: Consensus Statement from the 1st Strategic Conference of the European Federation of Clinical Chemistry and Laboratory Medicine. Clin Chem Lab Med 2015;53:833–5.
4. Ceriotti F, Fernandez-Calle P, Klee GG, Nordin G, Sandberg S, Streichert T, et al. Criteria for assigning laboratory measurands to models for analytical performance specifications defined in the 1st EFLM Strategic Conference. Clin Chem Lab Med 2017;55:189–94.
5. Minchinela J, Ricós C, Perich C, Fernández-Calle P, Álvarez V, Doménech MV, et al. Biological variation database and quality specifications for imprecision, bias and total error (desirable and minimum). The 2014 update. http://www.westgard.com/biodatabase-2014-update.htm (Accessed July 2016).
6. Bartlett WA, Braga F, Carobene A, Coşkun A, Prusa R, Fernandez-Calle P, et al. A checklist for critical appraisal of studies of biological variation. Clin Chem Lab Med 2015;53:879–85.
7. Carobene A, Braga F, Roraas T, Sandberg S, Bartlett WA. A systematic review of data on biological variation for alanine aminotransferase, aspartate aminotransferase and γ-glutamyl transferase. Clin Chem Lab Med 2013;51:1997–2007.
8. Carobene A, Strollo M, Jonker N, Barla G, Bartlett WA, Sandberg S, et al. Sample collections from healthy volunteers for biological variation estimates' update: a new project undertaken by the Working Group on Biological Variation established by the European Federation of Clinical Chemistry and Laboratory Medicine. Clin Chem Lab Med 2016;54:1599–608.
9. IFCC. RELA homepage. http://www.dgkl-rfb.de:81/4Daction/get_login (Accessed October 2016).
10. Røraas T, Støve B, Petersen PH, Sandberg S. Biological variation: the effect of different distributions on estimated within-person variation and reference change values. Clin Chem 2016;62:725–36.
11. Snedecor GW, Cochran WG. Statistical methods, 8th ed. Ames: Iowa State University Press; 1989.
12. Cochran WG. The distribution of the largest of a set of estimated variances as a fraction of their total. Ann Hum Genet 1941;11:47–52.
13. Shapiro SS, Wilk MB. An analysis of variance test for normality (complete samples). Biometrika 1965;52:591–611.
14. Dixon WJ. Processing data for outliers. Biometrics 1953;9:74–89.
15. Fraser CG, Petersen PH. Quality goals in external quality assessment are best based on biology. Scand J Clin Lab Invest Suppl 1993;212:8–9.
16. XLSTAT. Solutions. https://www.xlstat.com (Accessed September 2016).
17. Panteghini M, Bais R. Serum enzymes. In: Rifai N, Horvat AR, C Wittwer C, editors. Tietz textbook of clinical chemistry and molecular diagnostics, 6th ed. St. Louis: Elsevier Saunders; 2016.
18. Fraser CG, Harris EK. Generation and application of data on biological variation in clinical chemistry. Crit Rev Clin Lab Sci 1989;27:409–37.
19. Perich C, Minchinela J, Ricos C, Fernández-Calle P, Doménech MV, Simóne M, et al. Biological variation database: structure and criteria used for generation and update. Clin Chem Lab Med 2015;53:299–305.
20. Cummings ST, Fraser CG. Total amylase and pancreatic isoamylase in serum and urina: considerations from data on biological variation. Ann Clin Biochem 1989;26:335–40.
21. Huguet J, Fuentes-Arderiu X. Biological variation in the catalytic concentration of pancreatic alfa-amylase and triacylglycerol lipase in serum. Scand J Clin Lab Invest 1991;51:735–38.
22. Fraser CG, Cummings T, Wilkinson SP, Neville RG, Knox JDE, Ho O, MacWalter RS. Biological variability of 26 clinical chemistry analytes in elderly people. Clin Chem 1989;35:783–6.
23. Juan-Pereira L. Variabilitat biologica intraindividual de les magnitudes bioquimiques. Aplicacions cliniques [Doctoral Thesis]. [Barcelona]: Barcelona University; 1989.
24. Oosterhuis WP. Gross overestimation of total allowable error based on biological variation. Clin Chem 2011;57:1334–6.