The European Biological Variation Study (EuBIVAS): Biological Variation Data for Coagulation Markers Estimated by a Bayesian Model

Aasne K. Aarsand,[a,b,*] Ann Helen Kristoffersen,[a,b] Sverre Sandberg,[a,b,c] Bård Støve,[d] Abdurrahman Coşkun,[e] Pilar Fernandez-Calle,[f,g] Jorge Díaz-Garzón,[f,g] Elena Guerra,[h] Ferruccio Ceriotti,[i] Niels Jonker,[j] Thomas Røraas,[b,†] and Anna Carobene,[h,†] on behalf of the European Federation of Clinical Chemistry and Laboratory Medicine Working Group on Biological Variation

BACKGROUND: For biological variation (BV) data to be safely used, data must be reliable and relevant to the population in which they are applied. We used samples from the European Biological Variation Study (EuBIVAS) to determine BV of coagulation markers by a Bayesian model robust to extreme observations and used the derived within-participant BV estimates [$CV_{P(i)}$] to assess the applicability of the BV estimates in clinical practice.

METHOD: Plasma samples were drawn from 92 healthy individuals for 10 consecutive weeks at 6 European laboratories and analyzed in duplicate for activated partial thromboplastin time (APTT), prothrombin time (PT), fibrinogen, D-dimer, antithrombin (AT), protein C, protein S free, and factor VIII (FVIII). A Bayesian model with Student *t* likelihoods for samples and replicates was applied to derive $CV_{P(i)}$ and predicted BV estimates with 95% credibility intervals.

RESULTS: For all markers except D-dimer, $CV_{P(i)}$ were homogeneously distributed in the overall study population or in subgroups. Mean within-subject estimates (CV_I) were <5% for APTT, PT, AT, and protein S free, <10% for protein C and FVIII, and <12% for fibrinogen. For APTT, protein C, and protein S free, estimates were significantly lower in men than in women ≤50 years.

CONCLUSION: For most coagulation markers, a common CV_I estimate for men and women is applicable, whereas for APTT, protein C, and protein S free, sex-specific reference change values should be applied. The use of a Bayesian model to deliver individual $CV_{P(i)}$ allows for improved interpretation and application of the data.

Introduction

Coagulation markers play a central role in a variety of clinical settings, such as evaluation of a patient presenting with suspected thromboembolism or increased bleeding tendency, monitoring anticoagulant therapies, evaluation of liver function, and as risk assessment markers. To ensure correct interpretation of coagulation markers in these and other contexts, data on biological variation (BV) are necessary. BV data are used to set analytical quality specifications (APS), to assess changes in a measurement series within an individual by the reference change value (RCV) (1), to examine the use of population-based reference intervals (2), and to derive personalized reference intervals (3). However, for these applications, BV estimates must be reliable and representative for the populations to which they are applied. BV components include the within-subject BV (CV_I), which describes the natural fluctuation of the concentration around a set point within an individual, usually reported as an average (mean) CV_I estimate for the study population, and the between-subject BV (CV_G), which describes the variation between the set points of different individuals (4).

BV studies are usually undertaken as prospective experimental studies in healthy volunteers, but different

[a]Department of Medical Biochemistry and Pharmacology, Haukeland University Hospital, Bergen, Norway; [b]Norwegian Organization for Quality Improvement of Laboratory Examinations (Noklus), Haraldsplass Deaconess Hospital, Bergen, Norway; [c]Department of Global Public Health and Primary Care, University of Bergen, Bergen, Norway; [d]Department of Mathematics, University of Bergen, Bergen, Norway; [e]Department of Medical Biochemistry, Acibadem Mehmet Ali Aydınlar University School of Medicine, Atasehir, Istanbul, Turkey; [f]Department of Laboratory Medicine, Hospital Universitario La Paz, Madrid, Spain; [g]Analytical Quality Commission of Spanish Society of Laboratory Medicine (SEQCML), Madrid, Spain; [h]Laboratory Medicine, IRCCS San Raffaele Scientific Institute, Milan, Italy; [i]Central Laboratory, Fondazione IRCCS Ca' Granda, Ospedale Maggiore Policlinico, Milan, Italy; [j]Certe, Wilhelmina Ziekenhuis Assen, Assen, the Netherlands.

*Address correspondence to this author at: Department of Medical Biochemistry and Pharmacology, Haukeland University Hospital Bergen, N-5020 Bergen, Norway. Fax (+47)55-973115; e-mail aasne.aarsand@helse-bergen.no.

[†]Shared last authorship.

Received February 5, 2021; accepted April 19, 2021.
DOI: 10.1093/clinchem/hvab100

statistical approaches are applied to deliver the BV estimates. The most frequently used method is that detailed by Fraser and Harris, in which duplicate analysis of samples is followed by analysis of variance (ANOVA) (5) or CV-ANOVA (6). However, these approaches depend on laborious data analysis including assessment of outliers at 3 levels, and variance homogeneity of both the analytical variation (CV_A) component and the CV_I to provide generalizable results. Furthermore, data normality is assumed to construct confidence intervals (CI) (6). As shown by systematic reviews, these criteria are fulfilled by only a small number of BV studies (7). To overcome these issues, we recently explored a Bayesian approach for estimating BV (8). Our work has shown that a Bayesian model applying an adaptive Student *t* distribution is robust to extreme observations and does not require variance homogeneity to provide relevant results (8). Furthermore, in the Bayesian analysis, each individual's personal CV_I, termed the within-participant CV_I [$CV_{P(i)}$], can be estimated (8). The distribution of the individual $CV_{P(i)}$ can be used to assess whether data are homogeneous, and, as such, if a central CV_I estimate is representative for the study population, if subgroup analysis is required or if the data are too heterogeneous for an average CV_I estimate to be of value. In this way, the Bayesian approach provides an opportunity to evaluate whether the derived BV data are fit for purpose. The aims of our study were to deliver BV estimates for the following coagulation markers; activated partial thromboplastin time (APTT), prothrombin time (PT), fibrinogen, D-dimer, antithrombin (AT), protein C, protein S free, and factor VIII (FVIII) derived from the European Biological Variation Study (EuBIVAS) population (9), and to use the distribution of the $CV_{P(i)}$ to assess the applicability of using mean CV_I estimates in clinical practice.

Materials and Methods

STUDY POPULATION AND SAMPLE COLLECTION

The EuBIVAS is a multicenter study involving 6 laboratories situated in Italy, Norway, Spain, the Netherlands, and Turkey (9). For this study on coagulation markers, 92 healthy participants; 39 men (21–59 years), 43 women (21–49 years), and 10 women (55–69 years) (Table 1 in the online Data Supplement) were included, after assessment of eligibility based on clinical information and laboratory analyses, as previously detailed (9). For each individual, fasting blood samples were drawn for 10 consecutive weeks (April to June 2015). The EuBIVAS protocol was approved by the Institutional Ethical Review Board of San Raffaele Hospital in agreement with the World Medical Association Declaration of Helsinki and by the Ethical Board/Regional Ethics Committee as relevant for each involved center. All participants signed informed consent.

PREANALYTICAL AND ANALYTICAL PHASE

The same protocol was applied by all sites (9). Sample materials for coagulation analyses were collected in BD Vacutainer® sodium citrate tubes (3.2%). Samples were centrifuged at 3000*g* for 10 minutes at room temperature within 1 hour of sampling. Platelet-poor plasma was removed without disturbing the sedimented cells and aliquoted in Nalgene cryovials, which were frozen rapidly at −80 °C by immersion in a bowl with methanol and dry ice. All samples from each laboratory were sent collectively, frozen on dry ice, to San Raffaele Hospital, where they were stored at −80 °C until analysis in 2017. The frozen samples were rapidly thawed in a 37 °C water bath for 5 minutes and thoroughly mixed prior to being analyzed in duplicate in the same analytical run on an ACL Top 750 CTS (Instrumentation Laboratory S.p.A., a Werfen company). Only 3 results were below the limit of detection, all of which were D-dimer (< 21 ng/mL). The following single results outside clinical decision limits were excluded; D-dimer of 951 (week 9) and 3730 ng/mL (week 10) in a woman 26 years of age and AT of 74.9% (week 8) in a man 31 years of age. Single lot reagents, calibrators, and internal quality controls were used during the study period (online Supplemental Table 2), and no systematic changes in the concentrations of controls were detected (online Supplemental Table 3).

DATA ANALYSIS

Before calculating the BV measures, data were adjusted to counteract the effect of a trend in the results, by multiplying data at each time point by the factor "overall median"/"time point median", providing a constant median over the sampling period. Trend adjustment was applied to the overall study population and separately to the following 4 subgroups; men, all women, and women below and above 50 years, before analysis. To deliver the BV estimates, a Bayesian model, previously described in detail (8), was applied. In this model, Student *t* distributions, instead of normal distributions as in classical ANOVA, are assumed for the analytical and within-subject effects to make the model more robust to extreme observations (8). In addition, previous information (priors) (10) is included in the model, to deliver more precise estimates when analyzing with Bayesian updating. In our model, we applied the following prior distributions and hyperparameters for the priors, as detailed in (8). $N_{truncated}$ indicates that only the positive part of the normal distribution is used in the estimation routine and SDs are defined as positive

- $SD_G \sim N_{truncated}\,(SD_G, 0.1 \times SD_G)$

Table 1. Estimates of within-subject (CV_I) and between-subject (CV_G) biological variation and analytical variation (CV_A) for all study groups.

Measurand[a]	Subgroup	No. of individuals	No. of results	Average no. of samples/ individual	Mean conc.	$\mu CV_{P(i)}$ (95% CrI)	Estimated Harris-Brown ratio[b]	$dCV_{P(i)}\_50$ (20%–80%)	Predicted Harris-Brown ratio[b]	CV_G (95% CrI)	CV_A
APTT (seconds)	All	92	1799	9.8	29.7	2.9 (2.5–3.3)	13.4	3.0 (2.4–3.8)	31.5	7.2 (6.4–8.1)	3.7
	Men	**39**	**758**	**9.7**	**30.1**	**1.9 (1.6–2.2)**	**15.3**	**2.0 (1.7–2.2)**	**21.5**	**6.8 (5.8–8.0)**	**4.8**
	Women	**53**	**1041**	**9.8**	**29.4**	**3.3 (2.8–4.0)**	**14.9**	**3.6 (2.7–4.7)**	**38.9**	**7.4 (6.5–8.5)**	**2.2**
	Women ≤50	43	841	9.8	29.2	3.1 (2.6–3.9)	19.0	3.4 (2.4–4.5)	40.8	7.5 (6.5–8.5)	2.2
	Women >50	10	200	10.0	30.3	2.7 (1.6–4.3)	32.3	3.2 (1.9–4.8)	65.6	7.1 (5.9–8.4)	3.1
PT (seconds)	**All**	**92**	**1800**	**9.8**	**11.8**	**2.6 (2.3–2.9)**	**9.4**	**2.6 (2.4–2.9)**	**13.1**	**5.1 (4.6–5.6)**	**1.7**
	Men	39	760	9.7	12.0	2.6 (2.2–3.0)	14.7	2.6 (2.3–3.0)	19.6	5.1 (4.5–5.7)	2.1
	Women	53	1040	9.8	11.7	2.7 (2.4–3.0)	6.7	2.7 (2.5–2.9)	11.0	4.5 (3.9–5.1)	1.7
	Women ≤50	43	840	9.8	11.8	2.7 (2.5–3.0)	6.0	2.8 (2.6–3.0)	10.4	4.4 (3.8–5.0)	1.7
	Women >50	10	200	10.0	11.5	2.5 (1.7–3.6)	17.8	2.8 (2.2–3.5)	37.2	4.3 (3.6–5.0)	2.2
Fibrinogen (mg/dL)	**All**	**92**	**1799**	**9.8**	**294.3**	**10.2 (8.9–11.7)**	**19.4**	**10.6 (7.9–13.3)**	**31.6**	**17.3 (15.3–19.4)**	**4.7**
	Men	39	759	9.7	281.1	8.8 (7.1–10.9)	26.6	9.4 (6.6–12.3)	38.4	17.1 (14.6–19.8)	6.3
	Women	53	1040	9.8	304.0	10.9 (9.3–12.9)	12.6	11.5 (9.5–13.7)	25.7	17.0 (14.8–19.5)	5.5
	Women ≤50	43	840	9.8	304.3	11.4 (9.6–13.6)	10.9	12.1 (10.3–14.2)	24.3	16.7 (14.2–19.3)	5.9
	Women >50	10	200	10.0	302.4	7.6 (5.2–11.3)	21.2	8.8 (6.6–11.8)	46.6	17.6 (14.7–20.7)	6.1
D-dimer (ng/mL FEU)	All	92	1790	9.7	271.0	29 (24.7–34.1)	30.6	30.1 (19.9–40.9)	42.4	35.6 (32.3–39.3)	33.2
	Men	39	756	9.7	258.3	26.7 (20.8–34.6)	22.1	29.6 (20.5–40.5)	49.3	35.3 (31.6–39.4)	32.4
	Women	53	1034	9.8	280.3	31.2 (26–36.6)	32.6	32.2 (21.8–43.1)	40.6	30.5 (26.6–34.6)	23.8
	Women ≤50	43	834	9.7	280.0	34.3 (28.5–40.4)	25.9	35.6 (26.4–45.2)	33.7	30.3 (26.3–34.5)	24.6
	Women >50	10	200	10.0	281.7	17.2 (10.4–24.8)	37.2	18.2 (11.7–25.1)	53.3	27.7 (23.0–32.6)	17.9
AT (%)	**All**	**92**	**1792**	**9.7**	**110.3**	**3.5 (3.1–4.0)**	**14.6**	**3.6 (3.0–4.2)**	**21.2**	**8.1 (7.2–9.1)**	**5.4**
	Men	39	753	9.7	112.5	3.1 (2.6–3.7)	5.9	3.2 (2.9–3.7)	20.2	8.0 (6.9–9.2)	4.7
	Women	53	1039	9.8	108.7	3.8 (3.2–4.4)	18.2	3.9 (3.2–4.6)	23.3	7.9 (6.8–9.0)	5.1
	Women ≤50	43	839	9.8	107.7	3.9 (3.3–4.6)	17.5	4.1 (3.4–4.8)	22.3	7.7 (6.6–8.9)	4.9
	Women >50	10	200	10.0	112.7	2.7 (1.8–4.0)	28.3	3.0 (2.2–3.9)	44.7	7.8 (6.4–9.2)	3.9

Continued

Table 1. (continued)

Measurand[a]	Subgroup	No. of individuals	No. of results	Average no. of samples/ individual	Mean conc.	$\mu CV_{P(i)}$ (95% CrI)	Estimated Harris-Brown ratio[b]	$dCV_{P(i)}\_50$ (20%–80%)	Predicted Harris-Brown ratio[b]	CV_G (95% CrI)	CV_A
Protein C (%)	All	92	1797	9.8	105.8	5.4 (4.9–6.0)	14.7	5.5 (4.8–6.3)	18.4	18.6 (16.4–21.0)	2.4
	Men	**39**	**756**	**9.7**	**105.7**	**4.6 (3.8–5.4)**	**10.2**	**4.8 (4.2–5.4)**	**18.6**	**18.6 (15.8–21.8)**	**3.5**
	Women	**53**	**1041**	**9.8**	**105.9**	**6.1 (5.5–6.8)**	**12.7**	**6.2 (5.5–7.0)**	**16.1**	**19.1 (16.5–22.0)**	**2.2**
	Women ≤50	43	841	9.8	102.3	6.4 (5.8–7.0)	9.9	6.5 (5.9–7.0)	13.5	18.5 (15.7–21.6)	2.2
	Women >50	10	200	10.0	120.8	5.2 (3.6–7.5)	23.0	5.8 (4.2–7.9)	48.8	19.4 (15.9–23.1)	2.2
Protein S free (%)	All	92	1795	9.8	96.3	4.0 (3.5–4.6)	25.0	4.1 (3.1–5.2)	30.9	16.2 (13.8–19.1)	2.0
	Men	**39**	**755**	**9.7**	**102.9**	**3.2 (2.6–3.9)**	**32.5**	**3.3 (2.3–4.4)**	**38.2**	**16.7 (12.8–21.5)**	**1.8**
	Women	53	1040	9.8	91.5	4.5 (3.8–5.2)	16.4	4.7 (3.9–5.5)	22.2	17.9 (14.6–21.9)	2.1
	Women ≤50	**43**	**840**	**9.8**	**89.8**	**4.7 (3.9–5.6)**	**15.7**	**5.0 (4.2–5.8)**	**21.1**	**18.7 (15.0–23.1)**	**2.2**
	Women >50	**10**	**200**	**10.0**	**98.8**	**3.0 (1.9–4.4)**	**25.6**	**3.4 (2.5–4.4)**	**45.3**	**23.3 (18.2–28.3)**	**1.8**
FVIII (%)	**All**	**91**	**1780**	**9.8**	**114.6**	**8.3 (7.5–9.2)**	**8.0**	**8.4 (7.7–9.3)**	**13.3**	**23.3 (20.5–26.5)**	**5.9**
	Men	38	738	9.7	111.5	7.6 (6.5–8.7)	21.4	7.7 (6.2–9.2)	25.4	23.1 (19.4–27.4)	6.7
	Women	53	1042	9.8	116.7	8.8 (7.7–10.0)	5.9	9.0 (8.2–10.0)	14.0	24.4 (21.0–28.2)	6.7
	Women ≤50	43	842	9.8	117.3	8.4 (7.3–9.5)	9.2	8.6 (7.8–9.4)	13.7	24.7 (21.1–28.7)	5.8
	Women >50	10	200	10.0	114.3	9.5 (6.6–13.7)	18.4	10.8 (7.9–15)	52.0	24.8 (20.2–29.7)	6.7

CV_I measures are reported both as the mean within-participant BV estimate [$\mu[CV_{P(i)}]$] with 95% credibility intervals (CrI) and the 50th [$dCV_{P(i)}\_50$] percentile (20th, 80th) of the predicted distributions. Results in bold are estimates recommended for use.

[a]APTT; activated partial thromboplastin time, PT; prothrombin time, AT; antithrombin, FVIII; factor VIII.

[b]The Harris–Brown heterogeneity ratio ($100\%\{\sigma[CV_{P(i)}]/\mu[CV_{P(i)}]\}$) is expected to be below 100%/sqrt(2*S*) for a homogeneous population with an average number *S* samples from each individual (13); i.e., below 22.4% if 10 samples are included.

- $SD_{P(i)} \sim N_{truncated}\ \{\mu[SD_{P(i)}], \sigma[SD_{P(i)}]\}$
- $\mu[SD_{P(i)}] \sim N_{truncated}\ (SD_I, 0.1 \times SD_I)$
- $\sigma[SD_{P(i)}] \sim N_{truncated}\ (0, 0.1)$
- $SD_A \sim N_{truncated}\ (1/2 SD_I, 0.1 \times SD_A)$

where μ and σ refer to the mean and SD, respectively. To estimate these priors, which were included in the analysis of the overall group as well as subgroups, previously published data from Kristoffersen (11) and de Maat (12) were used (online Supplemental Table 4).

Inference in the model results in posterior distributions, obtained computationally, which provide estimates of the mean $\{\mu[CV_{P(i)}]\}$ and SD $\{\sigma[CV_{P(i)}]\}$ of the $CV_{P(i)}$, and CV_G and CV_A with corresponding SDs. Based on the estimated parameters, randomly generated $CV_{P(i)}$ were thereafter used to deliver the predicted distributions $\{d[CV_{P(i)}]\}$, from which the 20th [$dCV_{P(i)}$_20], 50th [$dCV_{P(i)}$_50], and 80th [$dCV_{P(i)}$_80] percentiles and the SD were calculated. In the following, when the term CV_I is used, it refers to the estimated measure, i.e., the $\mu[CV_{P(i)}]$. The credible interval is defined as the central portion of the posterior distribution that contains 95% of the values. A lack of overlap between the 95% credible intervals of the $\mu[CV_{P(i)}]$ was used to indicate differences between subgroups. As an indicator of the degree of heterogeneity, the estimated Harris–Brown heterogeneity ratio; $100\%\{\sigma[CV_{P(i)}]/\mu[CV_{P(i)}]\}$ was calculated (13). For a homogeneous population with an average of S samples from each individual, this ratio should be below 100%/sqrt(2S) (13), i.e., for our study below 22.4% as 10 samples were included. In addition, we calculated the predicted Harris–Brown ratio by using the $dCV_{P(i)}$_50 and SD from the $d[CV_{P(i)}]$. The posterior distribution of the parameters in the model were also used to calculate individual $CV_{P(i)}$ with 95% credible intervals for all study participants. For selected measurands, we performed a regression analysis with the response being the $CV_{P(i)}$ and the predictor being the study participants' homeostatic set point.

SOFTWARE

We used R (v.4.0.1) with tidyverse (v.1.3.0) and rstan (v.2.19.3) running on a Manjaro Gnu/Linux box. Stan is a probabilistic programming language for implementation of Bayesian inference (14). The main codes are available at https://gitlab.com/thoror/bayesian_bv_paper (8).

Results

The $CV_{P(i)}$ for the study participants, with different symbols indicating men, women below and above 50 years, are displayed for all measurands in Figs. 1–4. $CV_{P(i)}$ appeared homogeneously distributed for all study participants for PT (Fig. 1), fibrinogen (Fig. 2), and AT and FVIII (Fig. 4), in line with the estimated Harris–Brown heterogeneity ratios being <22.4% (Table 1). For APTT (Fig. 1) and protein C (Fig. 3), the Harris–Brown ratio also indicated homogeneous data, but CV_I estimates were lower in men than in women below 50 years (Table 1, Fig. 5). For protein S free (Fig. 3), the Harris–Brown ratio was above the cutoff for the overall group, but below in females, with differences in subgroup estimates observed between men and women ≤50 years (Table 1, Fig. 5). For D-dimer, the CV_I estimate was lower in women >50 years compared to below, but the Harris–Brown heterogeneity ratios were high for both subgroups (Table 1, Fig. 2). No differences in $CV_{P(i)}$ were observed for participants from the different sites (online Supplemental Table 1), smokers ($n = 20$), or females using oral contraceptives ($n = 4$). The study participants' individual $CV_{P(i)}$ appeared related to the homeostatic set point for fibrinogen ($\beta = 0.012$; $P = 0.005$) and protein S free ($\beta = -0.047$; $P < 0.0001$) (online Supplemental Fig. 1). Based on visual inspection, there was no clear relationship between the individual $CV_{P(i)}$ for each study participant for the different coagulation markers, with the exception of fibrinogen and protein S free (online Supplemental Fig. 2).

Discussion

Many studies have reported BV estimates for commonly analyzed measurands such as coagulation markers. However, the methodological quality and power of these studies vary, with potential impact on the applicability of these estimates in clinical practice. We used samples from the highly powered EuBIVAS study to deliver BV estimates by a Bayesian method. The Bayesian approach offers several advantages compared to classical frequentist methods. These include not needing to fulfill formal requirements related to outlying data and variance homogeneity. Furthermore, our Bayesian model provides the study participants' $CV_{P(i)}$ which, among other applications, allow us to assess the distribution of the data used as basis for the BV estimates (Figs. 1–4). BV components are typically reported in the form of a mean estimate, and it is therefore essential that this is representative for the population from which it has been derived, for the estimates to be safely used.

APPLICABILITY OF COAGULATION MARKER BV ESTIMATES

In our study, we report both the estimated $\mu[CV_{P(i)}]$ and the 50th percentile of the predicted distribution; the [$dCV_{P(i)}$_50] (Table 1). The predicted values take into account that extreme values may occur more often

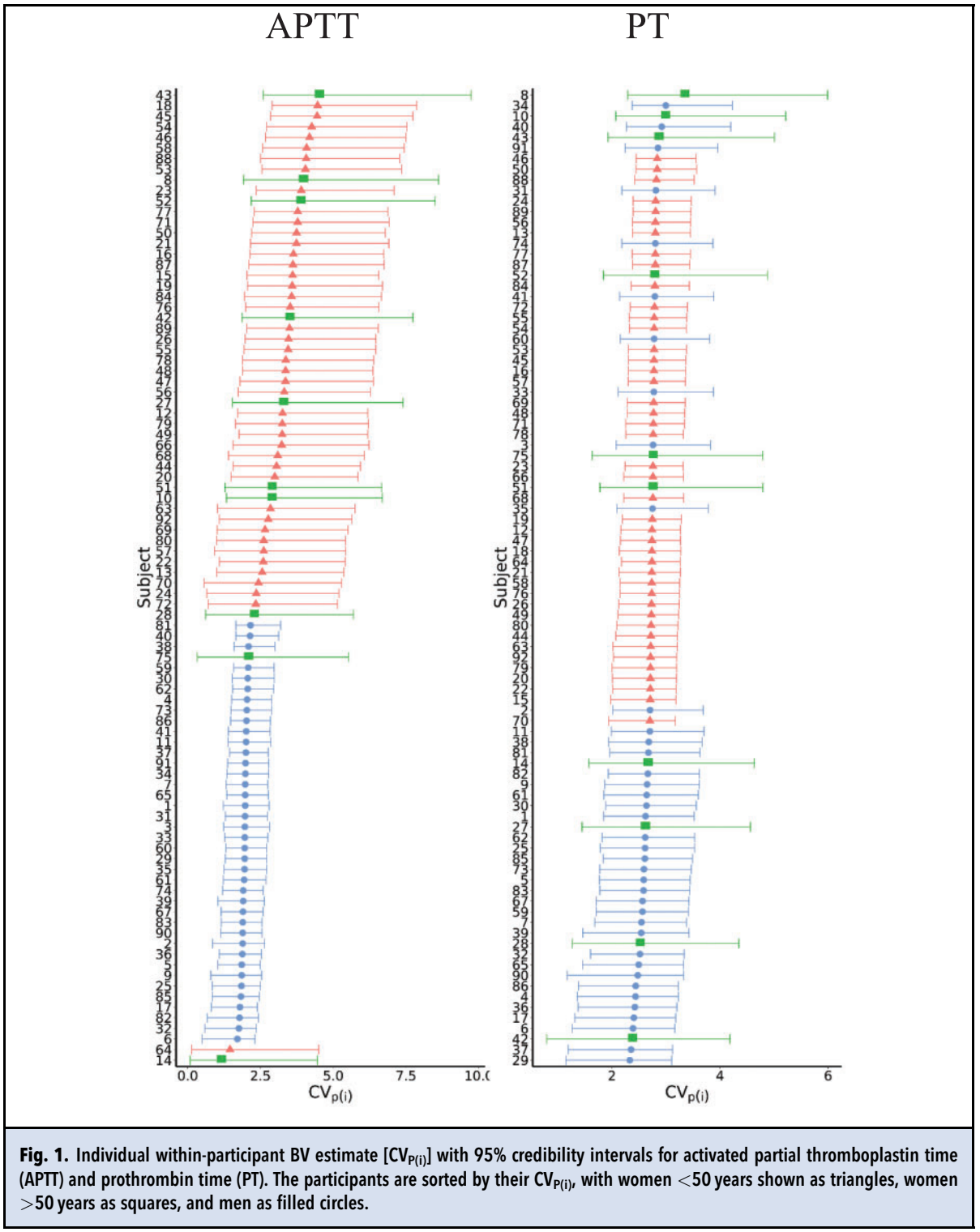

Fig. 1. Individual within-participant BV estimate [$CV_{P(i)}$] with 95% credibility intervals for activated partial thromboplastin time (APTT) and prothrombin time (PT). The participants are sorted by their $CV_{P(i)}$, with women <50 years shown as triangles, women >50 years as squares, and men as filled circles.

than the posterior distribution suggests and are thus associated with larger SD. Typically, the estimated and predicted point estimates are similar, but larger differences may be evident when the number of included participants is small. In our study, the largest difference was observed for FVIII in women above

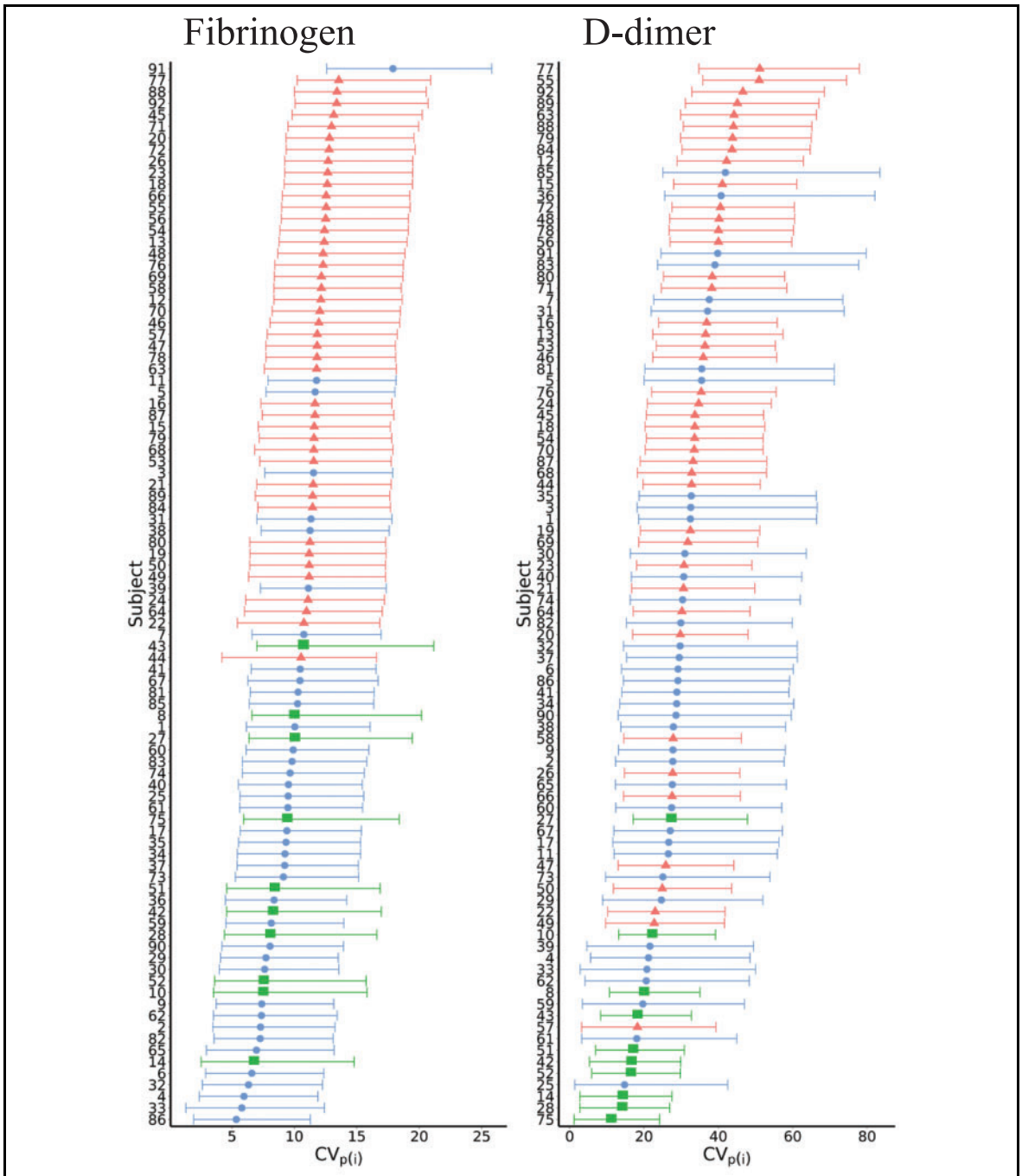

Fig. 2. Individual within-participant BV estimate [$CV_{P(i)}$] with 95% credibility intervals for fibrinogen and D-dimer. The participants are sorted by their $CV_{P(i)}$, with women <50 years shown as triangles, women >50 years as squares, and men as filled circles.

50 years, where the μ[$CV_{P(i)}$] was 9.5% and the $dCV_{P(i)}$_50 was 10.8%. To appraise the homogeneity of the data, the predicted and the estimated Harris–Brown heterogeneity ratios can be used (Table 1) (8). The estimated Harris–Brown heterogeneity ratio, calculated based on the estimated measures, reflects the actual

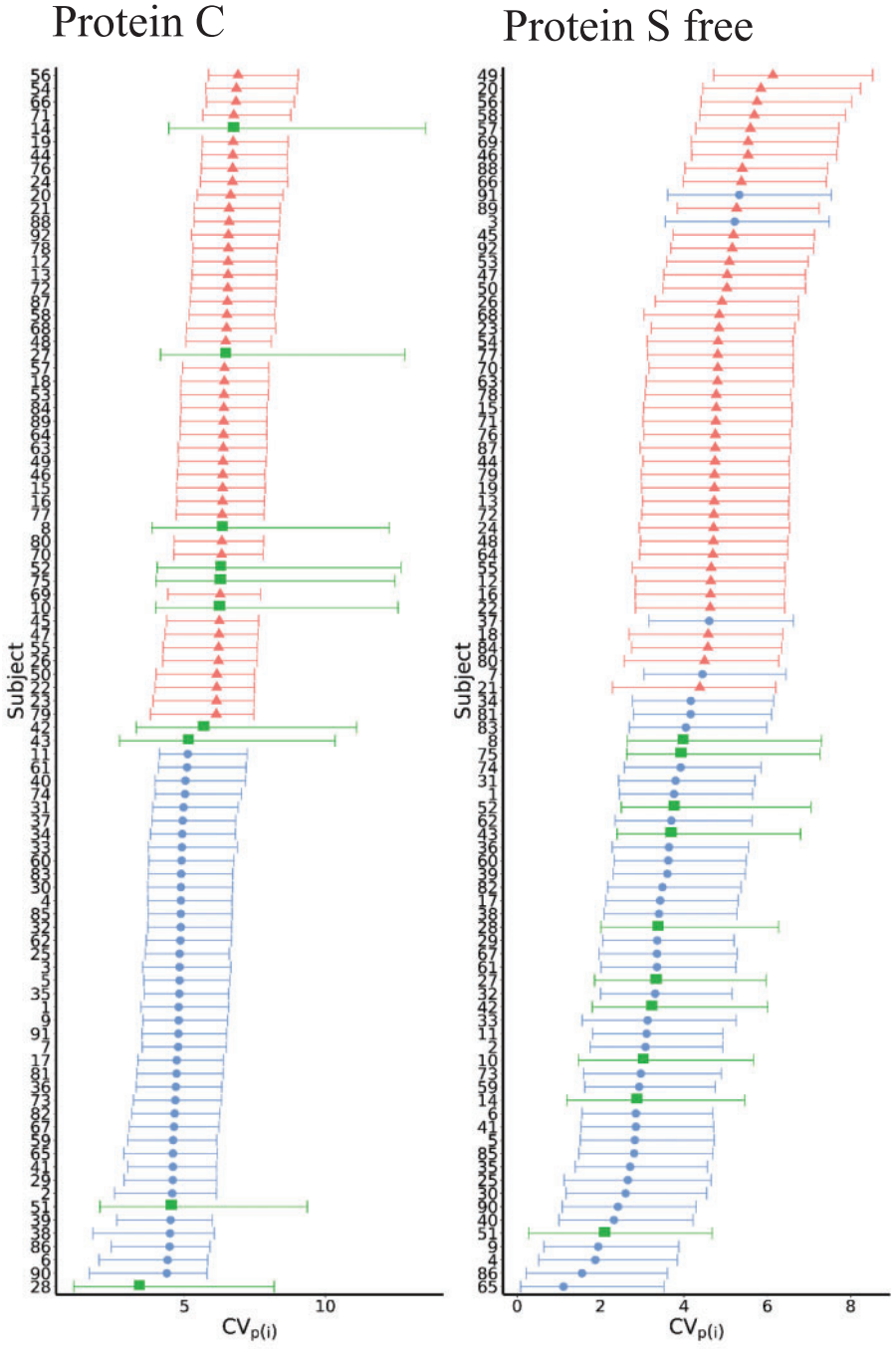

Fig. 3. Individual within-participant BV estimate [$CV_{P(i)}$] with 95% credibility intervals for protein C and protein S free. The participants are sorted by their $CV_{P(i)}$, with women <50 years shown as triangles, women >50 years as squares, and men as filled circles.

study data, and is therefore used to assess the distribution of the BV estimates in our study. This ratio is below the cutoff of 22.4% for APTT, PT, fibrinogen, AT, protein C, and FVIII for the overall study population (Table 1). This indicates that these data are homogenous, as also illustrated in Figs. 1–4, and that using a

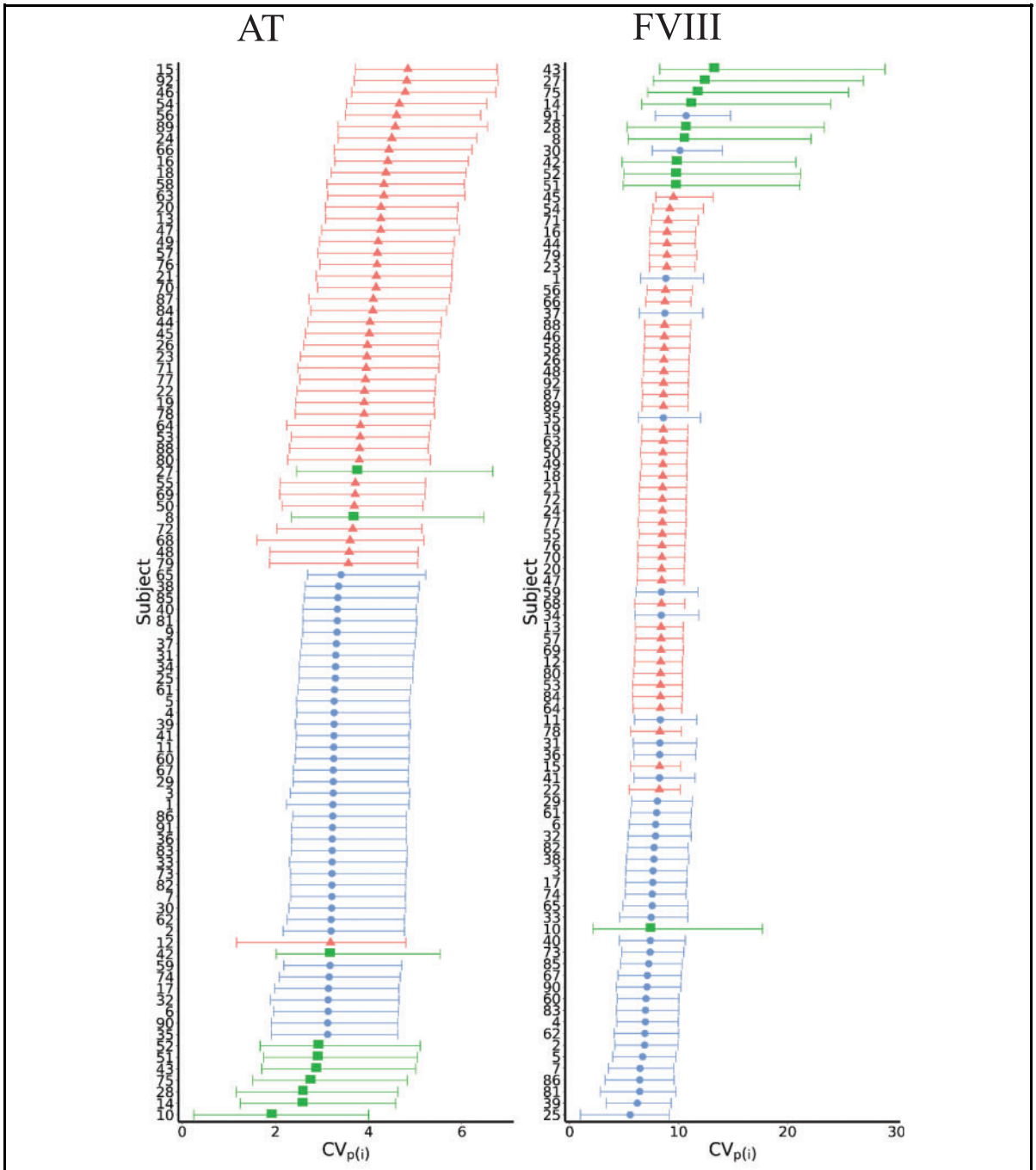

Fig. 4. Individual within-participant BV estimate [$CV_{P(i)}$] with 95% credibility intervals (CrI) for antithrombin (AT) and factor VIII (FVIII). The participants are sorted by their $CV_{P(i)}$, with women <50 years shown as triangles, women >50 years as squares, and men as filled circles.

common CV_I estimate is appropriate. For APTT and protein C, the CV_I estimates derived for men are lower than those for women ≤50 years (Table 1) as illustrated in Fig. 1 for APTT; where $CV_{P(i)}$ for men and women below 50 years are clearly separated, and for protein C in Fig. 3. This indicates that applying different CV_I

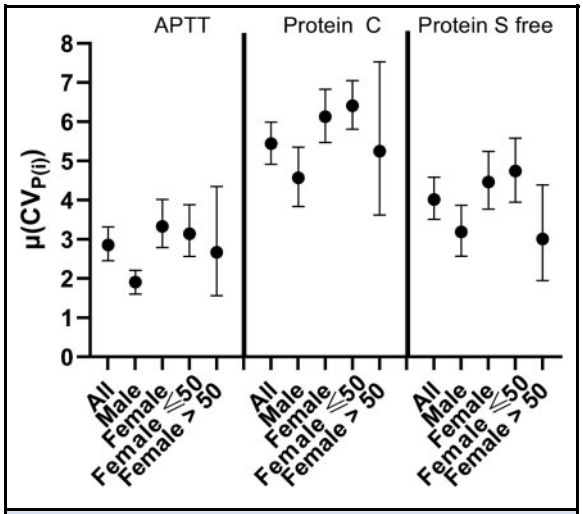

Fig. 5. Mean within-participant BV estimates $\{\mu[CV_{P(i)}]\}$ with 95% credibility intervals for activated partial thromboplastin time (APTT), protein C, and protein S free.

estimates for women and men may be relevant for these markers, depending on the purpose for which the data are being used. For protein S free, the Harris–Brown heterogeneity ratio was below the cutoff only in women ≤50 years, and differences in estimates were observed between subgroups (Table 1, Fig. 5). Thus, in this setting, applying subgroup estimates appears most appropriate, despite the somewhat heterogeneous data in men and women >50 years (Fig. 3). Generally, providing robust estimates for women >50 years is challenging due to the small number of participants in this group. This is particularly evident for FVIII (Fig. 4), where the broad individual 95% credibility intervals for these individuals reflect both the variation in the individual $CV_{P(i)}$ as well as the small number of participants. For D-dimer, the Harris–Brown heterogeneity ratios were high in the overall study population and in the subgroups, as illustrated in Fig. 2. Thus, the data are so heterogeneously distributed that to apply an average CV_I estimate for this measurand would generally not be recommended. The high amount of "noise" in the D-dimer data may be related to low concentrations of D-dimer in healthy individuals, with most study participants having concentrations around 200–300 ng/mL, associated with high CV_A estimates (Table 1).

In studies where classical methods such as ANOVA are applied to deliver BV estimates, outlying results have to be excluded to fulfill outlier and variance homogeneity criteria, but how much data must be excluded depends on the heterogeneity of the data (8). For some measurands, outlying results may make up a substantial amount of the data (15), which reduces the generalizability of the resulting BV estimates. In our study, we only excluded 3, clearly pathological, results from the data set. The low estimated degrees of freedom for the analytical component in our study indicate extreme observations in the data set between replicates (online Supplemental Table 5), but these were not necessary to exclude.

In a setting with homogeneous data, the classical frequentist and the Bayes approach will provide similar results (8). A disadvantage with applying a Bayesian model is that it requires more specialized knowledge and that running the analysis may be a relatively time-consuming process and require a high amount of data processing power. Choosing appropriate prior distributions may also be challenging. In our model, we applied priors based on the most recent publications reporting data on most of our markers (online Supplemental Table 4) (11, 12). For APTT, the $\mu CV_{P(i)}$ was clearly lower than the chosen prior. In our analysis, the priors were given low weight, with a SD of 10%. Thus, these priors had little influence on the results, and rerunning the analysis with a lower prior for APTT, based on (11), did not change the results for APTT (data not shown).

For most measurands in our study, the $CV_{P(i)}$ was not related to the concentration of the homeostatic set point (online Supplemental Fig. 1). This is an important finding with clinical implications, indicating that a common CV_I estimate is relevant for participants with different homeostatic set points, which is particularly important when calculating personalized reference intervals (3). Individual $\mu[CV_{P(i)}]$ of fibrinogen and protein S free appeared related to the homoeostatic set point and to each other (online Supplemental Figs. 1 and 2). The concentrations of these markers may be influenced by acute phase (16, 17). A previous EuBIVAS publication found transient increases in C-reactive protein in some participants (15), possibly caused by subclinical acute phase episodes. It is thus possible that this may have influenced also the results of fibrinogen and protein S free in these participants. However, a similar pattern is not seen for factor VIII, also an acute phase protein, suggesting that our observation may be coincidental.

The samples in our study were stored at −80 °C for 2 years before analysis, but data indicate that most coagulations markers are stable for up to at least 24 months at −74 °C (18). Furthermore, all the samples from the same participant were thawed simultaneously and analyzed in one run. Only 1 manufacturer's reagent was used, but it is unlikely that this will affect BV estimates, as long as the same measurand is targeted (19).

PREVIOUSLY PUBLISHED BV ESTIMATES FOR COAGULATION MARKERS

Previous studies on coagulation markers have described different approaches regarding the statistical analysis of

the data. Some have followed the protocol of Fraser (5), excluding outlying data points and assessing variance homogeneity (11, 20–22), whereas others have not (12). These studies also differed with regard to study design, with de Maat (12) and Kristoffersen (11) applying monthly samplings, and Falay (20) and our study applying weekly samplings. The CV_I estimates reported in these studies are quite similar, except for APTT, where de Maat (12) reported a higher CV_I estimate than other studies. This discrepancy could be explained by the lack of outlier exclusion in the de Maat study. Our CV_I estimates for protein C, protein S free, and FVIII were also lower than those reported from the other studies. This may be related to our study design or differences in data analysis. We also found small, but significantly lower CV_I estimates in men than women $\leq$50 years for APTT, protein C, and protein S free. Differences between sexes have previously not been reported. Previous studies, including 20–40 participants, may not have been sufficiently powered to assess sex differences. This is of particular relevance for coagulation markers where hormonal factors play a central role, and where the higher CV_I estimates in fertile women may be related to the variation of hormones during the menstrual cycle (23).

APPLICATION OF BV DATA FOR COAGULATION MARKERS

The coagulation markers assessed in our study have different clinical applications. These include diagnosis for inherited thrombophilia, hemophilia, and venous thromboembolism, for evaluation of liver function, and as part of diagnostic algorithms such as for, e.g., disseminated intravascular coagulation (DIC). The clinical context in which these markers are used must be taken into account when considering the applicability of the BV data. For most of the markers, data were homogenous for the overall population. Thus, setting APS based on these estimates, for the purpose of general diagnosis and monitoring, appears appropriate. Our finding that the $CV_{P(i)}$ was not dependent on the homeostatic set point (online Supplemental Fig. 1), also indicates that these APS may be relevant for patients with different measurand concentrations, but further studies are required to confirm this. For heterogeneous data, different percentiles of the $dCV_{P(i)}$ may be used to set APS. D-dimer displayed high heterogeneity, but is usually applied as a rule-out test in patients with low pretest probability for venous thromboembolism (24), and, thus, a clinical outcome model is recommended for setting APS (25).

All the markers included in the study showed marked individuality (Table 1). In this setting, the RCV provides a measure, with a given probability, for assessing whether a difference in consecutive results can be expected given the CV_I and CV_A for a participant in steady state conditions. For RCV calculation, laboratories' own long-term CV_A estimates must be included. For markers with differences in CV_I between subgroups such as for APTT and protein C, different RCVs should be applied for women and men. For protein S free, different CV_I estimates were found between men, and women $\leq$50 years and >50 years. Data for the women $\leq$50 years were homogeneous, but not for men or women >50 years (Table 1). When data are heterogeneous, instead of using a mean CV_I estimate, different percentiles of the predicted distribution of the $CV_{P(i)}$ may be used for RCV calculation, depending on the clinical purpose. Thus, in a clinical situation where it is important to identify a change, e.g., for a rule-in approach for further diagnostic work-up, the $dCV_{P(i)}\_20$ can be used to deliver a RCV, which increases the sensitivity. For a rule-out approach, the $dCV_{P(i)}\_80$ will deliver a RCV aimed at increasing specificity. Regarding D-dimer, data were so heterogeneous that applying an average CV_I estimate in RCVs to be used as part of, e.g., DIC assessment should be avoided. Furthermore, the BV estimates for D-dimer are based on low concentrations found in healthy individuals, often close to the limit of detection, and may thus not be appropriate for monitoring of patients.

Conclusions

In this study, we deliver updated BV estimates for commonly requested coagulation markers by the use of a Bayesian method, based on samples from the large-scale, highly powered EuBIVAS. Our Bayesian model, which is robust to extreme observations and includes prior information, is of particular value when assessing relatively heterogeneous data. Our results indicate that for most coagulation markers, a common CV_I estimate is applicable, whereas for others, such as APTT, protein C, and protein S free, sex-specific RCVs should be applied. Furthermore, our study illustrates that D-dimer data are so heterogeneous that the use of a mean CV_I estimate is not recommended for, e.g., RCV calculations. The use of a Bayesian model to deliver BV estimates with individual $CV_{P(i)}$ allows for improved interpretation and application of the data.

Supplemental Material

Supplemental material is available at *Clinical Chemistry* online.

Nonstandard Abbreviations: BV, biological variation; EuBIVAS, European Biological Variation Study; $CV_{P(i)}$, within-participant BV estimate; APTT, activated partial thromboplastin time; PT,

prothrombin time; AT, antithrombin; FVIII, factor VIII; APS, analytical performance specifications; CV_I, within-subject biological variation; CV_G, between-subject biological variation; RCV, reference change value; CV_A, analytical variation; $\mu[CV_{P(i)}]$, mean $CV_{P(i)}$; $\sigma[CV_{P(i)}]$, SD of $CV_{P(i)}$; $d[CV_{P(i)}]$, distribution of $CV_{P(i)}$; $dCV_{P(i)}\_20$, 20th percentile of the $d[CV_{P(i)}]$; $dCV_{P(i)}\_50$, 50th percentile of the $d[CV_{P(i)}]$; $dCV_{P(i)}\_80$, 80th percentile of the $d[CV_{P(i)}]$; DIC, disseminated intravascular coagulation; ANOVA, analysis of variance

Author Contributions: *All authors confirmed they have contributed to the intellectual content of this paper and have met the following 4 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; (c) final approval of the published article; and (d) agreement to be accountable for all aspects of the article thus ensuring that questions related to the accuracy or integrity of any part of the article are appropriately investigated and resolved.*

A.K. Aarsand, statistical analysis; A. Coşkun, provision of study material or patients; P. Fernandez-Calle, provision of study material or patients; J. Díaz-Garzón, provision of study material or patients; E. Guerra, provision of study material or patients; T. Røraas, statistical analysis; A. Carobene, statistical analysis, provision of study material or patients.

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: Werfen provided materials and a dedicated instrument for the coagulation analyses. T. Røraas (grant number 912232) and A.H. Kristoffersen (grant number 911956), Western Norway Regional Health Authority.
Expert Testimony: None declared.
Patents: A.K. Aarsand, PCT/TR2020/051103 (application filed); A. Coşkun, PCT/TR2020/051103 (application filed).

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

Acknowledgments: We thank Werfen for providing materials and an instrument for analysis, the study participants and the EuBIVAS partners for their contributions: Gerhard Barla, William Bartlett, Giulia Cajano, Mario Plebani, Una Ørvim Sølvik, Marit Sverresdotter Sylte, Mustafa Serteser, Francesca Tosato, and Ibrahim Unsal, and the Western Norway Regional Health Authority for supporting Thomas Røraas and Ann Helen Kristoffersen with postdoctoral fellowships.

References

1. Fraser CG. Reference change values: the way forward in monitoring. Ann Clin Biochem 2009;46:264–5.
2. Fraser CG, Sandberg S. Biological variation. In: N Rifai, AR Horvath, CT Wittwer, editors. Tietz textbook of clinical chemistry and molecular biology. 6th Ed. St. Louis (MO): Elsevier, 2017. p. 157–70.
3. Coskun A, Sandberg S, Unsal I, Cavusoglu C, Serteser M, Kilercik M, Aarsand AK. Personalized reference intervals in laboratory medicine: a new model based on within-subject biological variation. Clin Chem 2020;67: 256–64.
4. Fraser CG. Biological variation: from principles to practice. Washington (WA): AACC Press; 2001:18–22.
5. Fraser GG, Harris EK. Generation and application of data on biological variation in clinical chemistry. Crit Rev Clin Lab Sci 1989;27:409–37.
6. Roraas T, Stove 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.
7. Aarsand AK, Roraas T, Fernandez-Calle P, Ricos C, Diaz-Garzon J, Jonker N, et al. The Biological Variation Data Critical Appraisal Checklist: a standard for evaluating studies on biological variation. Clin Chem 2018;64:501–14.
8. Roraas T, Sandberg S, Aarsand AK, Stove B. A Bayesian approach to biological variation analysis. Clin Chem 2019;65:995–1005.
9. 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.
10. Gelman A, Carlin JB, Stern HS, Dunson DB, Vehtari A, Rubin DB. Bayesian data analysis, 3rd Ed. Boca Raton (FL): Chapman & Hall/CRC Texts in Statistical Science; 2013:435–45.
11. Kristoffersen AH, Petersen PH, Sandberg S. A model for calculating the within-subject biological variation and likelihood ratios for analytes with a time-dependent change in concentrations; exemplified with the use of D-dimer in suspected venous thromboembolism in healthy pregnant women. Ann Clin Biochem 2012;49: 561–9.
12. de Maat MP, van Schie M, Kluft C, Leebeek FW, Meijer P. Biological variation of hemostasis variables in thrombosis and bleeding: consequences for performance specifications. Clin Chem 2016;62:1639–46.
13. Harris EK, Brown SS. Temporal changes in the concentrations of serum constituents in healthy men. Distributions of within-person variances and their relevance to the interpretation of differences between successive measurements. Ann Clin Biochem 1979;16: 169–76.
14. Carpenter B, Gelman A, Hoffman MD, Lee D, Goodrich B, Betancourt M, et al. STAN: a probabilistic programming language. J Stat Softw 2017;76:1–32.
15. Carobene A, Aarsand AK, Guerra E, Bartlett WA, Coskun A, Diaz-Garzon J, et al. European Biological Variation Study (EuBIVAS): within- and between-subject biological variation data for 15 frequently measured proteins. Clin Chem 2019;65:1031–41.
16. Kaba NK, Francis CW, Hall WJ, Falsey AR, Smith BH. Protein S declines during winter respiratory infections. J Thromb Haemost 2003;1:729–34.
17. Khalil RH, Al-Humadi N. Types of acute phase reactants and their importance in vaccination. Biomed Rep 2020; 12:143–52.
18. Woodhams B, Girardot O, Blanco MJ, Colesse G, Gourmelin Y. Stability of coagulation proteins in frozen plasma. Blood Coagul Fibrinolysis 2001;12: 229–36.
19. Carobene A, Marino I, Coskun A, Serteser M, Unsal I, Guerra E, et al. The EuBIVAS project: within- and between-subject biological variation data for serum creatinine using enzymatic and alkaline picrate methods and implications for monitoring. Clin Chem 2017;63: 1527–36.
20. Falay M, Senes M, Korkmaz S, Turhan T, Okay M, Ozturk BA, et al. Biological variation estimates of prothrombin time, activated partial thromboplastin time, and fibrinogen in 28 healthy individuals. Int J Lab Hematol 2018; 40:721–5.
21. Kristoffersen AH, Petersen PH, Bjorge L, Roraas T, Sandberg S. Within-subject biological variation of activated partial thromboplastin time, prothrombin time, fibrinogen, factor VIII and von Willebrand factor in pregnant women. Clin Chem Lab Med 2018;56: 1297–308.
22. Kristoffersen AH, Petersen PH, Roraas T, Sandberg S. Estimates of within-subject biological variation of protein C, antithrombin, protein S free, protein S activity, and activated protein C resistance in pregnant women. Clin Chem 2017;63:898–907.
23. Chaireti R, Gustafsson KM, Bystrom B, Bremme K, Lindahl TL. Endogenous thrombin potential is higher during the luteal phase than during the follicular phase of a normal menstrual cycle. Hum Reprod 2013;28: 1846–52.
24. Tripodi A. D-dimer testing in laboratory practice. Clin Chem 2011;57:1256–62.
25. 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.