Abdurrahman Coşkun*, Aasne K. Aarsand, Sverre Sandberg, Elena Guerra, Massimo Locatelli, Jorge Díaz-Garzón, Pilar Fernandez-Calle, Ferruccio Ceriotti, Niels Jonker, William A. Bartlett and Anna Carobene, on behalf of the European Federation of Clinical Chemistry and Laboratory Medicine Working Group on Biological Variation

Within- and between-subject biological variation data for tumor markers based on the European Biological Variation Study

Abstract

Objectives: Reliable biological variation (BV) data are required for the clinical use of tumor markers in the diagnosis and monitoring of treatment effects in cancer. The European Biological Variation Study (EuBIVAS) was established by the EFLM Biological Variation Working Group to deliver BV data for clinically important measurands. In this study, EuBIVAS-based BV estimates are provided for cancer antigen (CA) 125, CA 15-3, CA 19-9, carcinoembryonic antigen, cytokeratin-19 fragment, alpha-fetoprotein and human epididymis protein 4.

Methods: Subjects from five European countries were enrolled in the study, and weekly samples were collected from 91 healthy individuals (53 females and 38 males; 21–69 years old) for 10 consecutive weeks. All samples were analyzed in duplicate within a single run. After excluding outliers and homogeneity analysis, the BVs of tumor markers were determined by CV-ANOVA on trend-corrected data, when relevant (Røraas method).

Results: Marked individuality was found for all tumor markers. CYFRA 21-1 was the measurand with the highest index of individuality (II) at 0.67, whereas CA 19-9 had the lowest II at 0.07. The CV_Is of HE4, CYFRA 21-1, CA 19-9, CA 125 and CA 15-3 of pre- and postmenopausal females were significantly different from each other.

Conclusions: This study provides updated BV estimates for several tumor markers, and the findings indicate that marked individuality is characteristic. The use of reference change values should be considered when monitoring treatment of patients by means of tumor markers.

Keywords: biological variation; cancer; tumor markers.

*Corresponding author: Abdurrahman Coskun, MD, EFLM Working Group on Biological Variation, Milan, Italy; EFLM Task Group for the Biological Variation Database, Milan, Italy; and Department of Medical Biochemistry Atasehir, School of Medicine, Acibadem Mehmet Ali Aydınlar University, Istanbul, Turkey, Phone: +90 532 744 66 83, Fax: +90 216 576 51 20, E-mail: Coskun2002@gmail.com
Aasne K. Aarsand, EFLM Working Group on Biological Variation, Milan, Italy; EFLM Task Group for the Biological Variation Database, Milan, Italy; Department of Medical Biochemistry and Pharmacology, Norwegian Porphyria Centre (NAPOS), Haukeland University Hospital, Bergen, Norway; and Norwegian Organization for Quality Improvement of Laboratory Examinations (NOKLUS), Haraldsplass Deaconess Hospital, Bergen, Norway
Sverre Sandberg, EFLM Working Group on Biological Variation, Milan, Italy; EFLM Task Group for the Biological Variation Database, Milan, Italy; Department of Medical Biochemistry and Pharmacology, Norwegian Porphyria Centre (NAPOS), Haukeland University Hospital, Bergen, Norway; Norwegian Organization for Quality Improvement of Laboratory Examinations (NOKLUS), Haraldsplass Deaconess Hospital, Bergen, Norway; and Department of Global Health and Primary Care, Faculty of Medicine, University of Bergen, Bergen, Norway
Elena Guerra and Massimo Locatelli, Servizio di Medicina di Laboratorio, Ospedale San Raffaele, Milan, Italy
Jorge Díaz-Garzón and Pilar Fernandez-Calle, EFLM Working Group on Biological Variation, Milan, Italy; EFLM Task Group for the Biological Variation Database, Milan, Italy; Department of Laboratory Medicine, La Paz University Hospital, Milan, Italy; and Analytical Quality Commission, Spanish Society of Laboratory Medicine (SEQCML), Barcelona, Spain
Ferruccio Ceriotti, Central Laboratory, Fondazione IRCCS Ca' Granda, Ospedale Maggiore Policlinico, Milan, Italy. https://orcid.org/0000-0002-0958-5354
Niels Jonker, EFLM Working Group on Biological Variation, Milan, Italy; EFLM Task Group for the Biological Variation Database, Milan, Italy; and Certe, Wilhelmina Ziekenhuis Assen, Assen, The Netherlands
William A. Bartlett, EFLM Working Group on Biological Variation, Milan, Italy; EFLM Task Group for the Biological Variation Database, Milan, Italy; and Blood Sciences, Ninewells Hospital & Medical School, Dundee, UK
Anna Carobene, EFLM Working Group on Biological Variation, Milan, Italy; EFLM Task Group for the Biological Variation Database, Milan, Italy; and Servizio Medicina di Laboratorio, Ospedale San Raffaele, Milan, Italy

Introduction

According to World Health Organization (WHO), cancer is the second leading cause of death in the world and globally responsible for one in six deaths [1]. Despite high mortality rates, early diagnosis of cancer may be lifesaving. Noninvasive laboratory tests such as tumor markers (TM) may play crucial roles in the management of cancer including screening, detection, differential diagnosis, staging, planning of treatments, monitoring and detection of recurrences [2]. However, when using TMs in the management of cancer, understanding the sources of variations such as pre-analytical, analytical and biological variations (BV), are essential. BV components constitute the within-subject (CV_I) and between-subject (CV_G) BV, where CV_I describes the fluctuation of a measurand around its homeostatic set point in a steady-state condition and CV_G the variations between the homeostatic set points among different individuals [3]. BV data has numerous uses in laboratory medicine including 1) to calculate the reference change value (RCV) which can be used to evaluate the significance of changes between serial measurements when monitoring patients over time, 2) to calculate the index of individuality (II) to evaluate the utility of population based reference intervals (popRI), 3) to calculate personalized reference intervals (prRI) which allow comparison of patients' test results with their own prRI [4], and 4) to set the analytical performance specifications (APS) of measurement procedures [3]. However, all these applications require that BV data are reliable and relevant to the population in which the BV applications will be used.

Over time, concern has been raised about the quality of BV data presented in the literature [5, 6]. For TMs, most of studies on BV are from the 1980s and 1990s, and results were delivered by analytical methods now considered obsolete and which may have targeted another measurand than methods in use today. In order to deliver updated BV data, the European Federation of Clinical Chemistry and Laboratory Medicine (EFLM) Biological Variation Working Group set up the European Biological Variation Study (EuBIVAS); a highly powered multicenter study which collected samples from five different countries following a stringent protocol [7]. The EuBIVAS has provided updated BV estimates for a high number of measurands [8–12]. In this study, we report EuBIVAS-based BV data for the following frequently requested TMs; cancer antigen (CA) 125, CA 15-3, CA 19-9, carcinoembryonic antigen (CEA), cytokeratin-19 fragment (CYFRA 21-1), alpha fetoprotein (AFP) and human epididymis protein (HE4), in order to improve on the use of these markers in the diagnosis and management of cancer.

Materials and methods

The demographic characteristics, health status, exclusion and inclusion criteria of subjects enrolled in EuBIVAS and a detailed study protocol have been previously reported in detail [7]. Demographic characteristics of subjects are provided in Supplemental Table 1.

Sample collection and handling

In total, 91 healthy volunteers (53 females, 38 males; age interval, 21–69 years) from six centers in five different countries (Italy, Norway, The Netherlands, Spain, Turkey) were enrolled in the EuBIVAS (Supplemental Table 1). The study protocol was approved by the Institutional Ethical Review Board of San Raffaele Hospital and Ethical Board/Regional Ethics Committee of each participating laboratory in agreement with the World Medical Association Declaration of Helsinki. All volunteers provided written informed consent prior to the study procedure. Fasting blood samples were drawn weekly, between 08.00 and 10.00 am, for 10 consecutive weeks (April–June 2015). Serum samples were stored at −80 °C until all samples were collected and then sent, frozen in dry ice, to the coordinating center, −San Raffaele Hospital in Milan, Italy and stored in frozen at −80 °C until analysis.

Analytical methods

All measurements were performed on the Roche Cobas e801 (Roche Diagnostics) at San Raffaele Hospital. Samples from the same person were measured in duplicate in a single run. The measurement system, reagents, calibrators and control materials used are detailed in Supplemental Table 2.

Data analysis

The Røraas method [13] was used to derive estimates of CV_I and CV_G as described in the following. Firstly, measurement results lower than limit of quantitation (LOQ) and thereafter, outliers as identified based on Dixon-Q test, were excluded for each TM. To verify that all participants were in steady-state, a linear regression was performed on the mean group value over the whole study period for each measurand. If the 95% confidence interval (CI) of the slope of the regression line included 0, the participants were considered to be in steady state, and if not, data were adjusted by applying the inverse of the regression formula to all measurement results at each week.

CV_I was estimated by using CV-ANOVA [13]. The homogeneity of analytical (between replicates) and within-subject variabilities were verified on CV-transformed data using Cochran and Bartlett tests, respectively. The normality of the residuals was verified using the Kolmogorov Smirnov test. The CV_G was estimated on natural-logarithmic transformed data using ANOVA after excluding outliers (Supplemental Table 3) and verifying the normality of data for all measurands.

We arbitrarily accepted 50 years of age as the cut-off value for premenopausal and postmenopausal women and performed separate analysis for women below and over of 50 years. We also

estimated the CV_I and CV_G of CEA and HE4 for smokers and non-smokers, separately [14, 15].

The 95% CIs of all BV estimates were calculated as described by Burdick and Graybill [16]. The lack of overlap between the 95% CIs was used to identify significant differences between BV estimates.

Kruskall-Wallis and Mann-Whitney U tests were used to evaluate the difference between the median concentrations of TMs of different countries. The correlation between the concentrations of TMs and ages of subjects were analyzed by Pearson (parametric variables) and Spearman r (non-parametric variables) tests.

APS (based on BV data) for analytical imprecision (CV_{APS}), and bias (B_{APS}), RCV, II and the number of samples required to estimate the homeostatic set points (NHSPs) based on BV data were calculated using the equations given below. For most measurands, BV estimates derived from all subjects were used, but when CV_I estimates between men and women were significantly different, as judged by the lack of overlap of the 95% CIs, APS, RCV and II were calculated for each subgroup separately. APS for bias was not calculated when the CV_G estimate of the analyte was above 33% which indicates the skewed distribution of homeostatic set points of individuals [17]. Additionally, when the mean concentration of a TM in the subgroups of men and women were significantly different, the lowest CV_G estimate was used to calculate the APSs.

$$CV_{APS} = 0.5 \times CV_I \qquad (1)$$
$$Bias_{APS} = 0.25 \times \sqrt{CV_I^2 + CV_G^2} \qquad (2)$$
$$II = \frac{CV_I}{CV_G} \qquad (3)$$

RCV were estimated using the equations given below [13]:

$$SD_{A,log}^2 = Log_e\left(CV_A^2 + 1\right) \qquad (4)$$
$$SD_{I,log}^2 = Log_e\left(CV_I^2 + 1\right) \qquad (5)$$
$$SD^\star = \sqrt{SD_{A,log}^2 + SD_{I,log}^2} \qquad (6)$$
$$RCV\% = 100\% \times e^{\left(\left(\pm Z_a \times \sqrt{2} \times SD^\star\right) - 1\right)} \qquad (7)$$

where $SD_{A,log}$ is the analytical SD calculated from the back-log transformation of CV_A obtained from the duplicate measurement of study samples; the $SD_{I,log}$, is the within-subject SD calculated from the CV_I estimates; and the $SD^\star$ is the combination of the $SD_{A,log}$ and $SD_{I,log}$. For a significant unidirectional change, z value was accepted as 1.65 for the probability level at 95%.

NHSP were estimated using the following formula:

$$NHSP = \left(z \times \frac{\sqrt{CV_I^2 + CV_A^2}}{D}\right)^2 \qquad (8)$$

where D is the allowed percentage deviation from the true homeostatic set point (HSP). NHSPs were calculated for 5, 10 and 15% deviations from the true homeostatic set points, using CV_A estimates from the duplicate analysis of study samples.

Results

The number of measurements results lower than LOQ and results identified as outliers by Dixon-Q, Bartlett and Cochran tests are given in Supplemental Table 3. Negative significant slopes were found for CA 125 and Cyfra 21-1 and

the trends were corrected by adding $0.1019x(S-1)$ and $0.006x(S-1)$ respectively to each measurements results (S: week number).

Mean concentrations, BV estimates and NHSP (within 5, 10 and 15% of the actual value) for all participants and subgroups are given in Table 1. Significant differences in mean concentrations between males and females were observed for AFP and CEA. Except for AFP; the mean concentrations of TMs were significantly different between postmenopausal and premenopausal women (Table 1).

Using data from the overall study population, the NHSPs (within 10% of the actual value) was 1 for AFP, CA 19-9, CA 15-3, indicating that one sample is sufficient to estimate the HSP within ±10% deviations for these TMs. For Cyfra 21-1, on the other hand, 16 samples were necessary to estimate the NHSP with ±10% deviation (Table 1).

The measurement results for HE4 are presented in Figure 1, for CYFRA21-1 in Figure 2 and for the other measurands in Supplemental Figures 1–5. In female subjects, there were significant correlations between the age of the subjects and their median concentrations of AFP (r=0.296; p=0.041), HE4 (r=0.621; p=0.001) (Figure 3) and CEA (r=0.345; p=0.010). TM results for participants classified by center and country are presented in Supplemental Figures 6–12 and Supplemental Table 4. The median of AFP results from subjects of laboratory 1 were significantly lower than those from subjects of laboratories 2, 5 and 6 and subjects of laboratory 3 from the subjects of laboratory 4 (Supplemental Table 4).

The measurement results of CEA and HE4 for smoker and non-smoker subjects, sorted according to increasing age, are presented in Supplemental Figures 13, 14. A significant effect of smoking was not observed on the BV of CEA and HE4. However, there was a significant difference between the mean concentrations of smokers and non-smokers for HE4 (Table 1).

Significantly different CV_I estimates were observed between females above and below 50 years for HE4 and CA 125 and between males and females for AFP and therefore separate APS, RCV and II were calculated (Table 2). The CV_G of AFP, CA 125, CEA, CA 19-9 and CA 15-3, were higher than 33% and therefore the APS for bias were not calculated (Table 2).

The CV_Is of HE4, CYFRA 21-1, CA 19-9, CA 125 and CA 15-3 of pre- and postmenopausal females were significantly different from each other. However, this was not the case for CV_G (Table 1).

Marked individuality was observed for all TMs. Except CYFRA 21-1, the II of all TMs were lower than 0.6 (Table 2). For CYFRA 21-1 highest II was found, at 0.67, whereas CA 19-9 had the lowest II, at 0.07.

Table 1: Within- and between-subject BV estimates for tumor markers with 95% CIs for all subjects, males (M), all females (F), females below (F<50 years) and above (F>50 years) years and for selected measurands, smokers (S) and non-smokers (NS) and the numbers of samples required to estimate the homeostatic set points (NHSP) within 5, 10 and 15% deviation.

Measurands		Number of individuals	Total number of results	Mean concentration, 95% CI	CV_A%, 95% CI	CV_I %, 95% CI	CV_G, 95% CI[c]	NHSP, 5%	NHSP, 10%	NHSP, 15%
AFP, µg/L	All	81	1,513	**3.1 (3.0–3.2)**	**3.0 (2.9–3.2)**	4.1 (3.9–4.4)	57.7 (48.6–69.8)	4	1	1
	M	34	633	2.5[a] (2.4–2.6)		3.4[a] **(3.0 – 3.8)**	**41.9 (33.1–56.5)**	4	1	1
	F	46	866	3.1 (3.0–3.2)		**4.6 (4.1–5.0)**	57.5 (45.8–74.4)	5	2	1
	F<50 years	37	707	3.0 (2.8–3.2)		5.3 (4.8–5.8)	60.6 (47.2–82.0)	6	2	1
	F>50 years	9	178	3.3 (3.0–3.5)		4.0 (3.4–4.9)	50.1 (32.6–112.1)	4	1	1
HE4, pmol/L	F	52	971	50.4 (49.7–51.1)	**1.1 (1.0–1.2)**	**6.7 (6.3–7.2)**	**18.5 (15.3–22.9)**	8	2	1
	F_S	11	213	47.9 (47.0–48.8)		7.9 (6.9–9.2)	9.9 (6.6–17.7)	10	3	2
	F_{NS}	41	761	51.2 (50.3–52.0)		6.6 (6.1–7.2)	20.1 (16.3–25.8)	7	2	2
	F<50 years	42	795	47.8 (47.2–48.5)		8.1 (7.6–8.8)	15.9 (13.0–20.4)	11	3	2
	F>50 years	10	191	59.5[b] (58.6–60.4)		4.5[b] (3.9–5.2)	8.9 (6.0–16.4)	4	1	1
CA 125, kIU/L	F	48	895	**11.4 (11.1–11.8)**	**2.3 (2.2–2.5)**	**8.6 (8.0–9.3)**	**40.9 (33.7–52.6)**	13	4	2
	F<50 years	39	717	12.2 (11.8–12.6)		9.4 (8.6–10.2)	38.4 (30.8–50.8)	15	4	2
	F>50 years	10	197	8.8[b] (8.2–9.3)		6.1[b] (5.3–7.3)	44.8 (30.1–92.1)	7	2	1
CEA, µg/L	All	89	1,685	**1.9 (1.8–1.9)**	**2.0 (1.9–2.1)**	**6.4 (6.0–6.7)**	59.8 (51.8–73.4)	7	2	1
	All_S	19	370	1.9 (1.8–2.0)		6.8 (6.1–7.7)	50.1 (36.7–79.1)	8	2	1
	All_{NS}	70	1,318	1.8 (1.8–1.9)		6.4 (6.0–6.8)	62.9 (53.5–79.8)	7	2	1
	M	38	717	2.0[a] (1.9–2.0)		6.7 (6.2–7.3)	**45.5 (36.4–60.7)**	8	2	1
	F	51	980	1.8 (1.7–1.9)		6.6 (6.2–7.1)	67.8 (55.7–90.3)	8	2	1
	F<50 years	41	782	1.7 (1.6–1.8)		6.7 (6.2–7.2)	67.0 (54.0–93.0)	8	2	1
	F>50 years	9	180	2.0[b] (1.8–2.1)		7.2 (6.2–8.6)	41.2 (27.2–88.5)	9	3	1
Cyfra 21-1, µg/L	All	89	1,690	**1.5 (1.5–1.6)**	**2.5 (2.4–2.6)**	**19.7 (18.7–20.7)**	**29.5 (25.3–35.0)**	61	16	7
	M	37	693	1.53 (1.5–1.6)		18.7 (17.4–20.3)	29.9 (23.6–39.3)	55	14	7
	F	52	996	1.6 (1.5–1.6)		21.4 (20.1–22.9)	28.2 (22.8–35.2)	72	18	10
	F<50 years	42	810	1.5 (1.4–1.5)		23.0 (21.4–24.8)	27.1 (21.6–35.2)	83	21	10
	F>50 years	10	196	1.7[b] (1.6–1.8)		17.4[b] (15.0–20.2)	26.4 (17.4–50.5)	48	12	6
CA 19-9, kIU/L	All	74	1,349	**10.0 (9.7–10.3)**	**2.3 (2.2–2.5)**	**4.0 (3.7–4.2)**	**56.0 (46.9–68.1)**	4	1	1
	M	33	619	9.7 (9.2–10.1)		4.1 (3.8–4.6)	54.4 (42.5–75.4)	4	1	1
	F	40	705	10.4 (9.9–10.8)		4.1 (3.8–4.5)	58.1 (45.8–76.8)	4	1	1
	F<50 years	32	571	9.4 (9.0–9.9)		5.1 (4.7–5.6)	59.1 (45.9–82.3)	5	2	1
	F>50 years	10	192	12.5[b] (11.6–13.4)		3.3[b] (2.8–4.0)	52.0 (34.9–111.5)	3	1	1
CA 15-3, kIU/L	All	81	1,554	**16.3 (15.9–16.6)**	**2.2 (2.1–2.3)**	**4.4 (4.1–4.6)**	**36.8 (31.6–44.1)**	4	1	1
	M	29	561	16.7 (16.3–17.2)		4.3 (3.9–4.8)	29.4 (23.6–41.3)	4	1	1
	F	52	997	16.0 (15.6–16.4)		4.5 (4.19–4.9)	40.4 (33.3–50.8)	4	1	1
	F<50 years	42	807	15.3 (14.8–15.8)		5.0 (4.6–5.4)	39.6 (32.0–51.4)	5	2	1
	F>50 years	10	200	18.6[b] (17.5–19.8)		3.5[b] (2.9–4.2)	40.6 (27.3–81.3)	3	1	1

Results in bold indicate the estimates used for calculations of APSs, RCVs and II (see Table 2). [a]Significantly different from females (All). [b]Significantly different from females <50 years. [c]Normality assumption was not fulfilled for all individuals by using Kolmogorov-Smirnov test (p>0.05). S, smoker; NS, non-smoker.
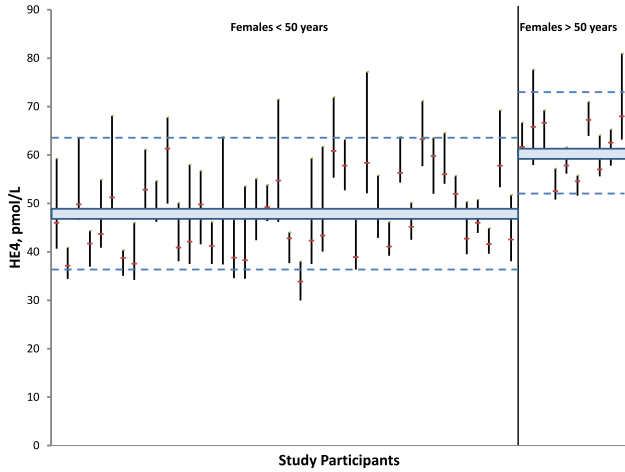
Figure 1: The median (minimum-maximum) concentrations of HE4 for each individual ordered by increasing age.
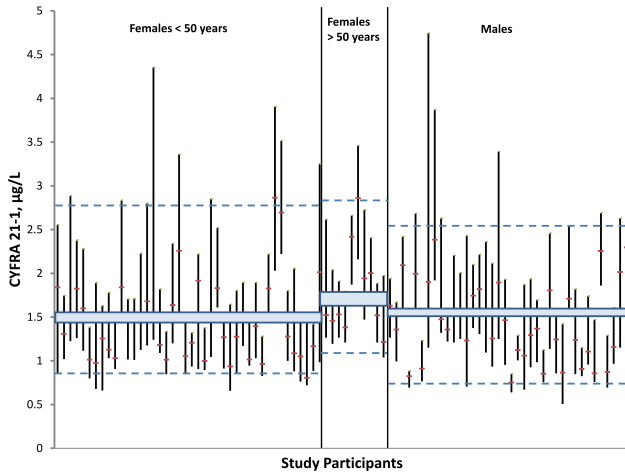Gray bar indicates the mean ± CI (95%); dashed lines indicate 5th and 95th percentiles.

Figure 2: The median (minimum-maximum) concentrations of CYFRA21-1 for each individual ordered by increasing age.
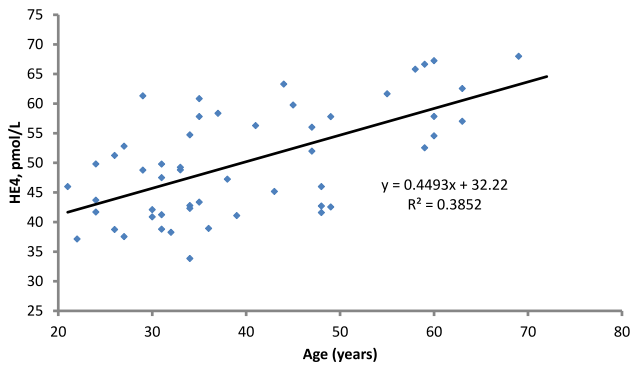Gray bar indicates the mean ± CI (95%); dashed lines indicate 5th and 95th percentiles.

Figure 3: Correlation between the age and median levels of HE4 for female subjects (r=0.621; p=0.001).

Table 2: Analytical performance specifications for imprecision (CV$_{APS}$) and for bias (Bias$_{APS}$), asymmetrical reference change values (RCV) for decrease and increase and index of individuality (II) of tumor markers derived using the data given in Table 1.

Measurands		CV$_{APS}$, %	Bias$_{APS}$, %	RCV, % decrease; increase	II
AFP[a]	M	1.69		−10.0; 11.1	0.08
	F	2.28		−11.9; 13.6	0.08
HE4	F<50 years	4.65	4.46	−17.9; 21.0	0.51
	F>50 years	2.24	2.50	10.2; 11.3	0.50
CA 125[a]	F<50 years	4.68		−20.0; 25.1	0.24
	F>50 years	3.06		−14.1; 16.4	0.14
CEA[a]	All	3.17		−14.3; 16.7	0.11
Cyfra 21	All	9.83	8.87	−36.7; 57.9	0.67
CA 19-9[a]	All	1.99		−10.2; 11.3	0.07
CA 15-3[a]	All	2.19		−10.8; 12.0	0.12

[a]APS for bias was not calculated when the CV$_G$ estimate of the analyte was above 33% which indicates the skewed distribution of homeostatic set points of individuals.

Discussion

Despite great improvements in the understanding of tumor biology, metabolism and treatment in the last four decades, high quality studies on BV of TMs are scarce [18]. Although TMs are included in diagnosis, screening, case finding, evaluating of prognosis, staging of tumors and planning the treatments, they are mainly used for monitoring of patients, including in detecting recurrences. Our results support the use of TMs for monitoring, considering the marked individuality of the included TMs. The II provides a criterion for assessing the utility of popRIs. If the II of a test is lower than 1.4 and particularly if it is lower than 0.6, the popRI is not recommended for monitoring of patients and RCV should be used instead of reference limits as thresholds for actions [3]. In our study, all II were lower than 1.4; the highest was 0.67 for CYFRA 21-1 and the lowest was 0.07 for CA 19-9 (Table 2).

The CV$_I$s of TMs reported in this study and the corresponding RCVs based on these estimates, were lower than those reported in previously published studies. These differences may be caused by differences in data handling, such as lack of outlier and homogeneity analysis, as reported by a systematic review of BV studies for TM [19]. Furthermore, some historical studies applied outdated analytical methods, which may target a different measurand [19]. There are also differences in study design. Some studies were performed in unhealthy individuals [20, 21],

included only a small number of samples per subjects [22], and performed sampling at longer time intervals [23], all of which may have influenced the reported CV_Is of TMs.

The sex- and age-related significant differences in CV_I estimates that were found in our study indicate that sex- and age-specific RCVs should be applied when monitoring patients with TMs. This is relevant particularly for postmenopausal females. Although an increase in serum concentrations of TMs is considered as a possible indication of presence or recurrence of tumors, decreasing concentrations should also be considered, particularly in patients under treatment. Therefore, both the decreasing and increasing RCV, i.e. a two-sided RCV should be calculated and applied for TMs. The CV_A estimates used as basis for RCV calculations in our study, were based on the duplicate analysis of the study samples. When using RCVs in clinical practice, these must be calculated using long-term CV_A estimates based on relevant concentrations from the laboratory in question, for these to be representative.

Knowledge on the NHSP is important, in particular in a setting where using TM to monitor cancer patient treatment and progression, and where assessment of the patient's status is usually performed on the basis of a measurement result of a single sample. Data of this study shows that for AFP, CA 19-9 and CA 15-3, the result of only single measurement of single sample (with the given APS) is adequate to predict the HSP within 10%. However, this is not the case for the other TMs which require replicate sampling (Table 1), and particularly for Cyfra 21-1, a high number of samples is required.

The EuBIVAS is a highly powered large-scale study, thus for many tumor markers, our data represents a clear improvement compared to previously published data. In the following, all included TMs and the BV results are discussed in detail.

Alpha-fetoprotein

Increased AFP concentration has been associated with poor prognosis of hepatocellular carcinoma and also recurrence after treatment [24–26]. The marked individuality of AFP supports that RCV should be used in monitoring of patients rather than cut-offs such as popRI. Previously published BV estimates and associated RCV of AFP are based on studies in healthy [22], testicular cancer [21] and subjects with hepatic disease [20]. Erden et al. measured AFP in adult healthy subjects and reported a CV_I estimate with 95% CI of 26.7% (23.8–30.3) [22], i.e. significantly and substantially higher than our estimate of 4.1% (3.9–4.4). The lack of outlier and

homogeneity analysis and the small number of samples (4 samples per subjects) in the Eden study [22] might explain this discrepancy. Trape et al. also have measured the BV of AFP in patients with testicular cancer [21] and hepatic disease [20] and reported CV_I estimates of 12.5 and 38%, and RCV of 33.3 and 88.9%, respectively. In our study, we found significantly lower CV_Is in males (3.4% [3.0–3.8]) than females (4.6% [4.1–5.0]) (Table 1), which indicates that different RCV should be applied, however, the difference between is quite small (RCV_increase males; 11.1%, females; 13.6%) (Table 2).

Human epididymis protein 4

HE4 is widely used as a biomarker for ovarian cancer and elevated concentrations of HE4 have also been observed in lung, endometrial, breast, gastrointestinal and renal cancers [27].

Braga et al. [28] analyzed HE4 in serum in 14 premenopausal and 14 postmenopausal healthy women sampled monthly for 4 consecutive months and reported CV_I estimates for HE4 of 12.1 and 6.5%, respectively, and a common CV_G estimates for both groups of 16.4%. We found a significantly lower CV_I estimate in postmenopausal (4.5% [3.9–5.2]) than premenopausal women (8.1% [7.6–8.8]), but similar CV_G estimates as Braga (Table 1). The IIs, reported by both Braga's (0.59) and this study (0.35) indicated marked individuality for HE4, which supports the use of this marker in monitoring of patients.

Braga et al. [28] found no difference in median concentration for HE4 in postmenopausal females as compared to younger females, whereas Urban et al. [15] reported increasing concentrations of HE4 with age. In our study, we found lower CV_I, but higher concentrations of HE4 in females above 50 years, as compared to those below (Table 1). Additionally, there was a significant, positive correlation between the median concentration of HE4 and age of all subjects (Figure 3). In consecutive measurement of HE4, we found that RCV_increasing is 21% for young females and 11% in postmenopausal women. The lower CV_Is in postmenopausal women reported both by Braga et al. [28] and this study indicate that smaller changes in HE4 concentrations may be considered significant in postmenopausal than in fertile women.

Higher HE4 concentrations have previously been reported in female smokers [15] and therefore, we analyzed BV components separately for smoking (n=11) and non-smoking (n=42) females, but no significant differences were observed (Table 1).

Cancer antigen 125

CA 125 is considered the most robust serum biomarker for ovarian cancer, but elevated concentrations of CA 125 are also observed in benign gynecological diseases and even in heart failure and liver cirrhosis [29, 30].

Braga et al. [28] have reported a CV_I estimate of 9.09% in healthy subjects, whereas Tuxen et al. [31] reported a much higher results, of 23.3%. The study of Tuxen et al. was conducted in 1999, used radioimmunoassay to measure CA 125 and did not perform analysis of homogeneity of variances. Trape et al. [23] measured CA 125 in two groups; surgically treated colon adenocarcinoma (control group, Astler–Coller classification stages A or B1) and non-small cell lung cancer patients (subjects with complete remission treated with chemotherapy and total tumor resection). They reported CV_I estimates of 21.1 and 22.5% for these two groups, respectively, and a common RCV of 53.1%. In our study, we found CV_I of 8.6% (8.0–9.3) and $RCV_{increase}$ of 23%, which are compatible with the data reported by Braga et al. [28]. Additionally, we found that the CV_I for CA 125 was significantly lower in postmenopausal females (6.1% [5.3–7.3]) than in younger subjects (9.4% [8.6–10.2]) (Table 1). Braga et al. [28] reported similar results (9.11 and 9.07%) for the two groups, but Tuxen et al. [32] found CV_I estimates of 28.9 and 12.6% for young and postmenopausal women, respectively. Similar to HE4, smaller changes in serum CA 125 concentrations might be of significance in postmenopausal than fertile women.

Carcinoembryonic antigen

CEA is widely used in the management of gastrointestinal tumors. However, it is not specific to gastrointestinal system tumors and is also used in combination with other TMs such as CA 125 in the management of ovarian cancers [33]. Different results have been reported for the CV_I of CEA in healthy and cancer patients. In studies on healthy subjects, Erden et al. [22], Sölétormos et al. [34] and Dittadi et al. [35] have reported CV_I estimates of 30.9, 9.3 and 8.4%, respectively. In study of Eden et al. the lack of outlier and homogeneity analysis and the small number of samples (four samples per subjects) might explain the large discrepancy. In cancer patients (in remission), such as breast cancer [35], surgically treated colon adenocarcinoma and non-small cell lung cancer [23], CV_I estimates have been reported at 19.3, 9.9 and 11.9%, respectively. We found lower CV_I estimates for CEA in all subjects and subgroups than all previous publications (Tables 1 and 2), but similar CV_I and CV_G estimates in fertile and postmenopausal women, in line with

that reported by Tuxen et al. [36]. The marked individuality, as also reported by others [34, 37], supports the use of RCV instead of popRI for monitoring of CEA concentrations by serial measurements.

Smoking is one of the main reasons of lung cancer and also increases the CEA concentrations in healthy subjects [38]. In our study, we found higher concentrations of CEA in smoking subjects (Table 1), however, the difference was not significant. Furthermore, estimates of CV_I and CV_G were similar in smoking and non-smoking subjects.

Cytokeratin-19 fragment

Serum CYFRA 21-1 concentrations is used in the management of non-small cell lung cancers, particularly for squamous cell tumors [39]. BV for CYFRA 21-1 has previously not been studied in healthy subjects. Based on surgically treated colon adenocarcinoma patients Trape et al. [23] reported the CV_I and CV_G of CYFRA 21-1 as 22.5 and 38.1%, respectively, and similar estimates in lung cancer patients (in remission). We found a CV_I estimate (19.7% [18.79–20.7]), in line with these estimates, in our healthy study population, but a lower CV_G estimate (29.5% [25.3–35.0]). Similar to some of the other markers included in our study, the CV_I of CYFRA 21-1 was lower in postmenopausal women than in younger subjects (Table 1).

The CV_I estimate of CYFRA 21-1 was higher than those of the other TMs included in our study, thus resulting in higher RCV and a higher II. Both CEA and CYFRA 21-1 are used in the monitoring of lung cancer. Data of this study shows that CEA is associated with a higher individuality (II=0.11) and smaller RCV ($RCV_{increase}$=16.7%) than CYFRA 21-1. This indicates that smaller increases in CEA, than in CYFRA 21-1, might be of significance when monitoring patients. Additionally, unlike the other TMs, a marked asymmetry was evident in the RCV of CYFRA 21-1, where the $RCV_{increase}$ was 57.9% and the $RCV_{decrease}$ was −36.7%.

Cancer antigen 19-9

Serum concentration of CA 19-9 is mainly used in the management of pancreatic cancer, in addition to gastric and colorectal cancers. In the literature, the reported data of CV_I for CA 19-9 is controversial. Vestergaard et al. [40] recruited subjects with specific genotypes and reported the CV_I and CV_G as 15.8 and 102.2%, respectively. Erden et al. [22] recruited 38 subjects with four samples per subjects and reported CV_I and CV_G of 27.2 and 64.2%, respectively. Within the published data, for healthy subjects the lowest

BV estimates were reported by Qi et al. [41]; CV_I and CV_G estimates of 6.8 and 49.6%, respectively. In our study, we found a CV_I estimate of 4.0 [3.7–4.2]), and the lowest results for II and RCV of CA 19-9 of all the TMs included in our study.

Pancreatic cancer is a highly aggressive cancer. In reflection of the low RCV found in our study, smaller increases in CA 19-9 concentrations should be considered as outside the expected, than for other TMs.

In this study, we observed higher mean concentration, but lower CV_I for CA 19-9 in postmenopausal women than in younger females (Table 1), which has previously not been reported.

Cancer antigen 15-3

CA 15-3 is used in the management of breast cancer [42], but is not specific, and elevated concentrations of CA 15-3 have been reported also in lung, ovary, colon, kidney and pancreatic cancers. Based on studies in healthy women, CV_I estimates for CA 15-3 have been reported by Söletormos et al. [43] and Dittadi et al. [35] as 6.2 and 6.0%, respectively. In addition, Dittadi et al. [35] and Hölzel et al. [44] reported CV_I of 17.3 and 11.2%, respectively, in brast cancer patients.

We found similar but lower CV_I for CA 15-3 at 4.4% (4.1–4.6) in our healthy population. A significantly lower CV_I for CA 15-3 was found in postmenopausal women compared with younger females, an observation that has previously not been published. Also CA 15-3 displays marked individuality (II=0.12) (Table 2), and the RCV should be used instead of popRI when monitoring patients.

Limitations of the study

The data of this study were obtained from healthy subjects and the derived BV of TMs represents the variation under physiological conditions, and may therefore differ from the BV of cancer patients. Additionally, subjects whose data were lower than LoQ were excluded and for some measurands, such as AFP, a relatively high number of results were excluded to fulfill statistical criteria (Supplemental Table 3). The exclusion of these data may lead to lower BV estimates and make our BV estimates less representative for the general population. This must be taken into account when using our data for RCV and other applications. Furthermore, our BV estimates were based on weekly samples collected over 10 consecutive weeks. These sampling intervals are

likely different from the monitoring period applied in most cancer forms. Consequently, RCVs based on BV estimates derived from the healthy participants may need to be evaluated for fitness for the intended use in the monitoring of cancer patients.

Although the mean concentrations and CV_Is of most of TMs were found to be different in pre- and postmenopausal females, the number of postmenopausal females was only 10, and these estimates are therefore associated with large uncertainty. New studies with larger numbers of postmenopausal subjects are needed to verify our findings.

Conclusions

The data presented in this study delivers updated BV data for TMs based on the highly powered EuBIVAS study, which includes samples from five different countries and in which strict pre-analytical protocols and updated statistical techniques have been applied. Except for CYFRA 21-1, all the included markers displayed marked individuality indicating that conventional popRIs are not suitable for monitoring. Furthermore, the CV_Is of HE4, CYFRA 21-1, CA 19-9, CA 125 and CA 15-3 were significantly different between pre- and postmenopausal females. It is recommended that RCVs are used, and additionally, for female patients, different RCVs for pre- and postmenopausal conditions should be considered.

References

1. World Health Organization (WHO). Cancer. Available from: https://www.who.int/health-topics/cancer#tab=tab_1 [Accessed March 2021].

2. Duffy MJ, Sturgeon CM, Soletormos G, Barak V, Molina R, Hayes DF, et al. Validation of new cancer biomarkers: a position statement from the European group on tumor markers. Clin Chem 2015;61: 809–20.

3. Fraser CG. Biological variation: from principles to practice. Washington, DC: AACC Press; 2001.

4. Coşkun A, Sandberg S, Unsal I, Cavusoglu C, Serteser M, Kilercik M, et al. Personalized reference intervals in laboratory medicine: a new model based on within-subject biological variation. Clin Chem 2021;67:374–84.

5. Aarsand AK, Røraas T, Sandberg S. Biological variation – reliable data is essential. Clin Chem Lab Med 2015;53:153–4.

6. Carobene A. Reliability of biological variation data available in an online database: need for improvement. Clin Chem Lab Med 2015; 53:871–7.

7. 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.

8. Aarsand AK, Díaz-Garzón J, Fernandez-Calle P, Guerra E, Locatelli M, Bartlett WA, et al. The EuBIVAS: within- and between-subject biological variation data for electrolytes, lipids, urea, uric acid, total protein, total bilirubin, direct bilirubin, and glucose. Clin Chem 2018;64:1380–93.

9. Carobene A, Guerra E, Locatelli M, Ceriotti F, Sandberg S, Fernandez-Calle P, et al. Providing correct estimates of biological variation—not an easy task. The example of S100-β protein and neuron-specific enolase. Clin Chem 2018;64:1537–9.

10. Carobene A, Guerra E, Locatelli M, Cucchiara V, Briganti A, Aarsand AK, et al. Biological variation estimates for prostate specific antigen from the European Biological Variation Study; consequences for diagnosis and monitoring of prostate cancer. Clin Chim Acta 2018;486:185–91.

11. Carobene A, Marino I, Coşkun 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.

12. Carobene A, Røraas T, Sølvik UØ, Sylte MS, Sandberg S, Guerra E, et al. Biological variation estimates obtained from 91 healthy study participants for 9 enzymes in serum. Clin Chem 2017;63: 1141–50.

13. 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.

14. Alexander JC, Silverman NA, Chretien PB. Effect of age and cigarette smoking on carcinoembryonic antigen levels. JAMA 1976;235:1975–9.

15. Urban N, Thorpe J, Karlan BY, Mcintosh MW, Palomares MR, Daly MB, et al. Interpretation of single and serial measures of HE4 and CA125 in asymptomatic women at high risk for ovarian cancer. Cancer Epidemiol Biomark Prev 2012;21:2087–94.

16. Burdick RK, Graybill F. Confidence intervals on variance components, 1st ed. New York: Marcel Dekker, Inc; 1992.

17. Carobene A, Aarsand AK, Guerra E, Bartlett WA, Coşkun A, Díaz-Garzón 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.

18. Aarsand AK, Fernandez-Calle P, Webster C, Coskun A, Gonzales-Lao E, Diaz-Garzon J, et al. The EFLM biological variation database. Available from: https://biologicalvariation.eu/ [Accessed March 2021].

19. Marqués-García F, Boned B, González-Lao E, Braga F, Carobene A, Coskun A, et al. Critical review and meta-analysis of biological variation estimates for tumor markers. Clin Chem Lab Med 2021, [submitted].

20. Trapé J, Botargues JM, Porta F, Ricós C, Badal JM, Salinas R, et al. Reference change value for α-fetoprotein and its application in early detection of hepatocellular carcinoma in patients with hepatic disease. Clin Chem 2003;49:1209–11.

21. Trapé J, Franquesa J, Sala M, Domenech M, Montesinos J, Catot S, et al. Determination of biological variation of α-fetoprotein and choriogonadotropin (β chain) in disease-free patients with testicular cancer. Clin Chem Lab Med 2010;48:1799–801.

22. Erden G, Barazi AO, Tezcan G, Yildirimkaya MM. Biological variation and reference change values of CA 19-9, CEA, AFP in serum of healthy individuals. Scand J Clin Lab Invest 2008;68: 212–8.

23. Trapé J, Pérez de Olaguer J, Buxó J, López L. Biological variation of tumor markers and its application in the detection of disease progression in patients with non-small cell lung cancer. Clin Chem 2005;51:219–22.

24. Ikai I, Arii S, Kojiro M, Ichida T, Makuuchi M, Matsuyama Y, et al. Reevaluation of prognostic factors for survival after liver resection in patients with hepatocellular carcinoma in a Japanese nationwide survey. Cancer 2004;101:796–802.

25. Bruix J, Cheng AL, Meinhardt G, Nakajima K, De Sanctis Y, Llovet J. Prognostic factors and predictors of sorafenib benefit in patients with hepatocellular carcinoma: analysis of two phase III studies. J Hepatol 2017;67:999–1008.

26. Galle PR, Foerster F, Kudo M, Chan SL, Llovet JM, Qin S, et al. Biology and significance of alpha-fetoprotein in hepatocellular carcinoma. Liver Int 2019;39:2214–29.

27. Galgano MT, Hampton GM, Frierson HF. Comprehensive analysis of HE4 expression in normal and malignant human tissues. Mod Pathol 2006;19:847–53.

28. Braga F, Ferraro S, Mozzi R, Panteghini M. The importance of individual biology in the clinical use of serum biomarkers for ovarian cancer. Clin Chem Lab Med 2014;52:1625–31.

29. Nolen BM, Lokshin AE. Protein biomarkers of ovarian cancer: the forest and the trees. Future Oncol 2012;8:55–71.

30. Ferraro S, Panteghini M. Making new biomarkers a reality: the case of serum human epididymis protein 4. Clin Chem Lab Med 2019;57:1284–94.

31. Tuxen MK, Sölétormos G, Petersen PH, Schiøler V, Dombernowsky P. Assessment of biological variation and analytical imprecision of CA 125, CEA, and TPA in relation to monitoring of ovarian cancer. Gynecol Oncol 1999;74: 12–22.

32. Tuxen MK, Sölétormos G, Petersen PH, Dombernowsky P. Interpretation of sequential measurements of cancer antigen 125 (CA 125), carcinoembryonic antigen (CEA), and tissue polypeptide antigen (TPA) based on analytical imprecision and biological variation in the monitoring of ovarian cancer. Clin Chem Lab Med 2001;39:531–8.

33. Sagi-Dain L, Lavie O, Auslander R, Sagi S. CEA in evaluation of adnexal mass: retrospective cohort analysis and review of the literature. Int J Biol Markers 2015;30:394–400.

34. Soletormos G, Schioler V, Nielsen D, Skovsgaard T, Dombernowsky P. Interpretation of results for tumor markers on the basis of analytical imprecision and biological variation. Clin Chem 1993;39:2077–83.

35. Dittadi R, Peloso L, Gion M. Within-subject biological variation in disease: the case of tumour markers. Ann Clin Biochem 2008;45: 226–7.

36. Tuxen MK, Sölétormos G, Petersen PH, Dombernowsky P. Interpretation of sequential measurements of cancer antigen 125 (CA 125), carcinoembryonic antigen (CEA), and tissue polypeptide antigen (TPA) based on analytical imprecision and biological variation in the monitoring of ovarian cancer. Clin Chem Lab Med 2001;39:531–8.

37. Tuxen MK, Sölétormos G, Petersen PH, Dombernowsky P. Interpretation of sequential measurements of cancer antigen 125 (CA 125), carcinoembryonic antigen (CEA), and tissue polypeptide antigen (TPA) based on analytical imprecision and biological

variation in the monitoring of ovarian cancer. Clin Chem Lab Med 2001;39:531–8.

38. Bjerner J, Høgetveit A, Wold Akselberg K, Vangsnes K, Paus E, Bjøro T, et al. Reference intervals for carcinoembryonic antigen (CEA), CA125, MUC1, Alfa-foeto-protein (AFP), neuron-specific enolase (NSE) and CA19.9 from the NORIP study. Scand J Clin Lab Invest 2008;68:703–13.

39. Molina R, Marrades RM, Augé JM, Escudero JM, Viñolas N, Reguart N, et al. Assessment of a combined panel of six serum tumor markers for lung cancer. Am J Respir Crit Care Med 2016;193:427–37.

40. Vestergaard EM, Hein HO, Meyer H, Grunnet N, Jørgensen J, Wolf H, et al. Reference values and biological variation for tumor marker CA 19-9 in serum for different Lewis and secretor genotypes and evaluation of secretor and Lewis genotyping in a Caucasian population. Clin Chem 1999;45:54–61.

41. Qi Z, Zhang L, Chen Y, Ma X, Gao X, Du J, et al. Biological variations of seven tumor markers. Clin Chim Acta 2015;450:233–6.

42. Lin DC, Genzen JR. Concordance analysis of paired cancer antigen (CA) 15-3 and 27.29 testing. Breast Canc Res Treat 2018;167: 269–76.

43. Soletormos G, Schioler V, Nielsen D, Skovsgaard T, Dombernowsky P. Interpretation of results for tumor markers on the basis of analytical imprecision and biological variation. Clin Chem 1993;39:2077–83.

44. Hölzel WGE, Beer R, Deschner W, Griesmacher A, Müller MM. Individual reference ranges of CA 15-3, MCA and CEA in recurrence of breast cancer. Scand J Clin Lab Invest 1995;55:93–101.