



**University of Dundee**

## **Critical review and meta-analysis of biological variation estimates for tumor markers**

Marques-Garcia, Fernando; Boned, Beatriz; González-Lao, Elisabet; Braga, Federica; Carobene, Anna; Coskun, Abdurrahman

*Published in:*  
Clinical Chemistry and Laboratory Medicine (CCLM)

*DOI:*  
[10.1515/cclm-2021-0725](https://doi.org/10.1515/cclm-2021-0725)

*Publication date:*  
2022

*Document Version*  
Publisher's PDF, also known as Version of record

[Link to publication in Discovery Research Portal](#)

*Citation for published version (APA):*  
Marques-Garcia, F., Boned, B., González-Lao, E., Braga, F., Carobene, A., Coskun, A., Díaz-Garzón, J., Fernández-Calle, P., Perich, M. C., Simon, M., Jonker, N., Aslan, B., Bartlett, W. A., Sandberg, S., Aarsand, A. K., & Task Group for the Biological Variation Database (2022). Critical review and meta-analysis of biological variation estimates for tumor markers. *Clinical Chemistry and Laboratory Medicine (CCLM)*, 60(4), 494-504. <https://doi.org/10.1515/cclm-2021-0725>

### **General rights**

Copyright and moral rights for the publications made accessible in Discovery Research Portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

### **Take down policy**

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Fernando Marques-Garcia\*, Beatriz Boned, Elisabet González-Lao, Federica Braga, Anna Carobene, Abdurrahman Coskun, Jorge Díaz-Garzón, Pilar Fernández-Calle, Maria Carmen Perich, Margarida Simon, Niels Jonker, Berna Aslan, William Alexander Bartlett, Sverre Sandberg and Aasne K. Aarsand, on behalf of the European Federation of Clinical Chemistry and Laboratory Medicine Working Group on Biological Variation and Task Group for the Biological Variation Database

# Critical review and meta-analysis of biological variation estimates for tumor markers

<https://doi.org/10.1515/cclm-2021-0725>

Received June 24, 2021; accepted February 1, 2022;

published online February 10, 2022

## Abstract

**Objectives:** Biological variation data (BV) can be used for different applications, but this depends on the availability of robust and relevant BV data. In this study, we aimed to summarize and appraise BV studies for tumor markers, to examine the influence of study population characteristics and concentrations on BV estimates and to discuss the applicability of BV data for tumor markers in clinical practice.

**Methods:** Studies reporting BV data for tumor markers related to gastrointestinal, prostate, breast, ovarian, haematological, lung, and dermatological cancers were

identified by a systematic literature search. Relevant studies were evaluated by the Biological Variation Data Critical Appraisal Checklist (BIVAC) and meta-analyses were performed for BIVAC compliant studies to deliver global estimates of within-subject ( $CV_I$ ) and between-subject ( $CV_G$ ) BV with 95% CI.

**Results:** The systematic review identified 49 studies delivering results for 22 tumor markers; four papers received BIVAC grade A, 3 B, 27 C and 15 D. Out of these, 29  $CV_I$  and 29  $CV_G$  estimates met the criteria to be included in the meta-analysis. Robust data are lacking to conclude on the relationship between BV and different disease states and tumor marker concentrations.

**Conclusions:** This review identifies a lack of high-quality BV studies for many tumor markers and a need for delivery of BIVAC compliant studies, including in different disease

---

\*Corresponding author: **Fernando Marques-Garcia**, PhD, Biochemistry Department, Metropolitan North Clinical Laboratory (LCMN), Germans Trias i Pujol University Hospital, Badalona, Barcelona, Spain; and Spanish Society of Laboratory Medicine (SEQC<sup>ML</sup>), Analytical Quality Commission, Barcelona, Spain, Phone: +34934651200, E-mail: f.marg@hotmail.es

**Beatriz Boned**, Spanish Society of Laboratory Medicine (SEQC<sup>ML</sup>), Analytical Quality Commission, Barcelona, Spain; and Royo Villanova Hospital, Zaragoza, Spain

**Elisabet González-Lao**, Spanish Society of Laboratory Medicine (SEQC<sup>ML</sup>), Analytical Quality Commission, Barcelona, Spain; and Quality Healthcare Consulting, Grupo ACMS, Barcelona, Spain

**Federica Braga**, Research Centre for Metrological Traceability in Laboratory Medicine (CIRME), University of Milan, Milan, Italy. <https://orcid.org/0000-0003-3562-7180>

**Anna Carobene**, Servizio Medicina di Laboratorio, Ospedale San Raffaele, Milan, Italy

**Abdurrahman Coskun**, School of Medicine, Acibadem Mehmet Ali Aydınlar University, Atasehir, Istanbul, Turkey

**Jorge Díaz-Garzón and Pilar Fernández-Calle**, Spanish Society of Laboratory Medicine (SEQC<sup>ML</sup>), Analytical Quality Commission, Barcelona, Spain; and Department of Laboratory Medicine, La Paz University Hospital, Madrid, Spain. <https://orcid.org/0000-0002-3171-1505> (J. Díaz-Garzón)

**Maria Carmen Perich**, Spanish Society of Laboratory Medicine (SEQC<sup>ML</sup>), Analytical Quality Commission, Barcelona, Spain

**Margarida Simon**, Spanish Society of Laboratory Medicine (SEQC<sup>ML</sup>), Analytical Quality Commission, Barcelona, Spain; and Consortium of Laboratory Intercomarcal Alt Penedès and Garraf l'Anoia, Vilafranca del Penedès, Spain

**Niels Jonker**, Certe-Wilhelmina Ziekenhuis Assen, Assen, The Netherlands

**Berna Aslan**, Institute for Quality Management in Healthcare (IQMH), Centre for Proficiency Testing, Toronto, Ontario, Canada

**William Alexander Bartlett**, School of Medicine, University of Dundee, Dundee, Scotland, UK

**Sverre Sandberg**, Department of Medical Biochemistry and Pharmacology, Norwegian Porphyria Centre, Haukeland University Hospital, Bergen, Norway; Norwegian Organization for Quality Improvement of Laboratory Examinations (NOKLUS), Haralds plass Deaconess Hospital, Bergen, Norway; and Department of Global Health and Primary Care, Faculty of Medicine, University of Bergen, Bergen, Norway

**Aasne K. Aarsand**, Department of Medical Biochemistry and Pharmacology, Norwegian Porphyria Centre, Haukeland University Hospital, Bergen, Norway; and Norwegian Organization for Quality Improvement of Laboratory Examinations (NOKLUS), Haralds plass Deaconess Hospital, Bergen, Norway

states and tumor marker concentrations. As of yet, the state-of-the-art may still be the most appropriate model to establish analytical performance specifications for the majority of tumor markers.

**Keywords:** biological variation; critical review; meta-analysis; tumor marker.

## Introduction

Current clinical practice in oncology focuses on early diagnosis, proper prognostication, and screening in asymptomatic groups to reduce the burden of cancer [1]. Measurements of tumor markers in serum and other biological fluids may be instrumental for this purpose; being used to screen selected populations to enable early diagnosis, provide tumor classification and staging, aid in prognosis and to monitor progression or recurrence [2, 3]. Sequential tumor marker measurements are primarily used to deliver early information about response to therapy and early detection of recurrence [3]. Some are more appropriate for follow-up and others for the early detection of recurrence disease.

When monitoring patients over time, random biological variation (BV) will influence the results. There are two components of random BV: within-subject BV ( $CV_I$ ); describing the variation around the homeostatic set point of an individual and between-subject BV ( $CV_G$ ); describing the variation between the set points of different individuals [4].

BV data have many different applications in laboratory medicine, including to set analytical quality specifications for relevant measurands [4], to calculate the Index of Individuality (II) which provides information about the utility of using population based reference intervals for test interpretation [5] and to aid in interpretation of variations in serial results. For this latter purpose, the reference change value (RCV), calculated on estimates of  $CV_I$  and analytical variation ( $CV_A$ ) can be used [6]. The RCV enables the assessment of whether the difference between two consecutive measurements can be explained by analytical and biological variation, with a given probability. However, BV data are often derived from studies on healthy volunteers. In healthy individuals, most tumor markers are present in low concentrations, thus potentially delivering limitations for users intending to apply BV data clinically in patients [7].

Many published BV studies, including for tumor markers, were performed several decades ago. Improvement in analytical methods for tumor markers has led to

the delivery of more accurate results, and it is thus likely that historically reported estimates may be less relevant for today's practice. Thus, for some tumor markers, a limited number of studies of acceptable quality may be available and data for relevant subpopulation and sampling intervals may be scarce. The Biological Variation Data Critical Appraisal Checklist (BIVAC) provides an objective appraisal of the quality of BV publications by verifying whether all essential elements that may impact the reliability and of the derived BV estimates are present [8, 9].

This study aimed to summarize and evaluate, by the use of the BIVAC, BV studies for tumor markers related to gastrointestinal, prostate, breast, ovarian, haematological, lung, and dermatological cancers and to examine the effect of study populations' characteristics and measurand concentrations on the BV estimates. We also provide updated BV estimates delivered by meta-analysis of BIVAC-compliant studies with similar study design and discuss potential applications of BV data for tumor markers.

## Materials and methods

Studies included in this review were identified in a historical BV database [10], and by a systematic literature search in PubMed published until October 2020, using an approach previously described by Aarsand et al. [9] for the following tumor markers:  $\alpha$ -fetoprotein (AFP), cancer antigen-19.9 (CA19.9), cancer antigen-125 (CA125), Cancer antigen-15.3 (CA15.3), cancer antigen-54.9 (CA54.9), carcinoembryonic antigen (CEA), Cyfra 21.1 antigen, human epididymis protein 4 (HE-4), mucinous carcinoma associated antigen (MCA), neuron-specific enolase (NSE), pepsinogen I and II, pro-gastrin releasing peptide (ProGRP), tissue polypeptide antigen (TPA), tissue polypeptide specific antigen (TPS), total prostatic specific antigen (tPSA), conjugated prostatic specific antigen (cPSA), free prostatic specific antigen (fPSA),  $\beta$ -2-microglobulin, cancer antigen 72.4 (CA-72.4), chromogranin A and S100 $\beta$ .

The publications in this review are identified by the article number assigned in the EFLM BV Database [11] and summarized with their BIVAC grade in Supplementary Table 1. An overview of the included BV studies is provided in Table 1. When one study included estimates for several different subpopulations, such as based on sex, age, or health status, a subscript (a, b, c or d) was added to the article number. Appraisal of publications by the BIVAC quality items (QI) (Supplementary Table 2) was independently performed by two reviewers [9]. In this process, each study was assessed and graded against 14 QIs, with alternative scores A, B, C, or D awarded to verify whether the essential elements required to deliver a valid study of BV were present, as documented previously [12–14]. The lowest score given to any of the QI determined the overall BIVAC grade for the study. When a study included estimates for different study subgroups, such as health status, age and gender, appraisal was performed independently for each subset. If there was disagreement for any of the QIs, a third assessor performed a full new review [9].

**Table 1:** Number of papers identified by the systematic review, with number of subgroups and BIVAC grade.

Measurands	Number of papers	Number of subgroups	BIVAC grades (number of papers)			
			A	B	C	D
AFP	4	8	1	0	2	1
β-2-microglobulin	4	4	1	0	1	2
CA125	9	11	1	1	2	5
CA15.3	9	14	1	0	2	6
CA19.9	5	9	1	0	4	0
CA54.9	1	1	0	0	0	1
CA72.4	2	2	0	1	1	0
CEA	15	23	1	0	8	6
Chromogranin A	2	4	0	1	1	0
Cyfra 21.1 antigen	3	7	1	0	1	1
HE-4	2	6	1	1	0	0
MCA antigen	2	2	0	0	0	2
NSE	6	8	1	2	3	0
Pepsinogen I	1	1	0	0	1	0
Pepsinogen II	1	1	0	0	1	0
ProGRP	1	1	0	0	1	0
cPSA	3	5	1	0	2	0
fPSA	9	14	1	0	6	1
tPSA	16	38	1	0	13	2
S100β	3	3	1	0	2	0
TPA	5	5	0	0	4	1
TPS	1	1	0	0	1	0

AFP, α-fetoprotein; CEA, carcinoembryonic antigen; CA, cancer antigen; HE-4, human epididymis protein 4; MCA, mucinous carcinoma associated antigen; NSE, neuron-specific enolase; ProGRP, pro-gastrin releasing peptide; cPSA, conjugated prostatic specific antigen; fPSA, free prostatic specific antigen; tPSA, total prostatic specific antigen; TPA, tissue polypeptide antigen; TPS, tissue polypeptide specific antigen.

Ninety five percent confidence intervals (CIs) for the individual BV estimates were calculated as described by Burdick [15], if the required information was provided, including the mean number of subjects and samples and  $CV_A$  estimate. Meta-analyses for each tumor marker were performed on BIVAC compliant studies (BIVAC grades A, B or C) which had been performed in healthy adults (18–75 years) with sampling intervals from biweekly to monthly including at least three samples per subject [9]. The method applied was the weighted median approach, by this, global  $CV_I$  and  $CV_G$  estimates are calculated by using the BIVAC quality grade (grade A weighted by 4, B by 2 and C by 1) and the inverse of the CI as weights, as previously described [9]. Since the 95% CI is used as a weight, only studies where CI was reported or could be calculated were included [9]. Finally, a percentile bootstrap with the weighted median was used for calculating the CIs [16].

APS, II and RCV, based on meta-analysis derived BV data, were calculated for all tumor markers following the currently recommended formulas [17–19]:

$$CV_{APS} = 0.5 \times CV_I$$

$$Bias_{APS} = 0.25 \times \sqrt{CV_I^2 + CV_G^2}$$

$$II = \frac{CV_I}{CV_G}$$

RCV were estimated as

$$SD_{A, \log}^2 = \text{Log}_e(CV_A^2 + 1)$$

$$SD_{I, \log}^2 = \text{Log}_e(CV_I^2 + 1)$$

$$SD^* = \sqrt{SD_{A, \log}^2 + SD_{I, \log}^2}$$

$$RCV\% = 100\% \times e^{((\pm Z_{\alpha} \times \sqrt{2} \times SD^*)^{-1})}$$

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

## Results

The systematic review identified 49 studies (Supplementary Table 1) delivering results for 22 different tumor markers (Table 1). The highest number of studies was available for tPSA and CEA, 16 and 15, papers respectively. For five measurands (CA54.9, pepsinogen I and II, proGRP and TPS), only a single study was identified. For the other tumor markers, four or less papers were identified (Table 1). Most BV studies had been performed in healthy volunteers (23 papers) and 14 in unhealthy subjects.

Appraisal with the BIVAC identified four papers with grade A, 3 as B, and 27 papers as C (Table 1). Fifteen papers were graded as D, which included all studies reporting BV estimates for MCA (2 papers) and CA54.9 (1 paper).

The main reason for a BIVAC grade B was the failure to examine steady state (QI7). QIs eliciting grade C were mainly related to the statistical study items, i.e. variance homogeneity (QI10), testing of outliers (QI8) and whether the study provided the number of results that were used to estimate the BV components (QI13). A BIVAC grade D was assigned due to an insufficient description of the study population (number of participants, sex, age) (QI2) or for having applied obsolete analytical methods (QI4) [20].

A total of 137  $CV_I$  and 79  $CV_G$  estimates were obtained from papers with BIVAC grade A, B or C. Of these, 29  $CV_I$  and 29  $CV_G$  estimates met the criteria to be included in the meta-analysis (Table 2 and Supplementary Table 3). The main reasons for exclusion of studies from of meta-analyses were: study population included unhealthy individuals; and less than three samples collected number of samples per subject less than three or sampling intervals more frequent than es not collected between 2/week and/or less frequent than 1/month (“not fulfilling sampling interval/number of samples criteria”).

**Table 2:** CV<sub>I</sub> and CV<sub>G</sub> estimates with 95% CI derived from meta-analysis of BIVAC compliant studies with similar study design.

Measurands	N	n	Nm	nm	Meta-analysis derived BV estimates	
					CV <sub>I</sub> (CI 95%)	CV <sub>G</sub> (CI 95%)
AFP	4	8	2	2	4.6 (4.1–26.7)	54.8 (43.6–57.7)
β-2-microglobulin	4	3	2	2	4.1 (4–5.9)	11.5 (11.2–15.5)
CA125	9	11	3	3	6.6 (6.3–23.3)	42.5 (10.6–70.6)
CA15.3	9	14	1	1	4.4 <sup>a</sup>	36.8 <sup>a</sup>
CA19.9	5	9	3	3	4.3 (4–27.2)	57.6 (56–102.2)
CA72.4	2	2	1	1	50.3 <sup>a</sup>	103.4 <sup>a</sup>
CEA	15	23	3	3	6.8 (6.3–30.9)	59.2 (37.1–59.8)
Cyfra 21.1 antigen	3	7	1	1	19.7 <sup>a</sup>	29.5 <sup>a</sup>
Chromogranin A	2	4	2	2	14.8 (12.8–16.3)	30.7 (26.3–33.5)
HE-4	2	6	2	2	7.1 (6.7–9.7)	17.9 (16.4–18.5)
NSE	6	8	4	4	10.9 (10.1–21.5)	16.6 (11.5–28.8)
cPSA	3	5	1	1	8.8 <sup>a</sup>	57.7 <sup>a</sup>
fPSA	9	14	1	1	7.1 <sup>a</sup>	46.2 <sup>a</sup>
tPSA	16	38	1	1	6.8 <sup>a</sup>	42 <sup>a</sup>
S100β	3	3	1	1	10.2 <sup>a</sup>	32.7 <sup>a</sup>
TPA	5	5	1	1	27.1 <sup>a</sup>	62.4 <sup>a</sup>

N, number of papers; n, number of subgroups; Nm, papers include in meta-analysis; nm, subgroups include in meta-analysis; AFP, α-fetoprotein; CEA, carcinoembryonic antigen; CA, cancer antigen; HE-4, human epididymis protein 4; NSE, neuron-specific enolase; cPSA, conjugated prostatic specific antigen; fPSA, free prostatic specific antigen; tPSA, total prostatic specific antigen; TPA, tissue polypeptide antigen. Tumor markers without papers included in the meta-analysis were not included in this Table. <sup>a</sup>Markers with only one paper fulfilling criteria to be included in meta-analysis, thus no CI is reported.

The appraised papers mainly included tumor markers analysed in serum, with only three studies having used plasma for analysis (paper 201: CA125, CEA, Cyfra21.1 antigen; paper 208: chromogranin A, and paper 353: tPSA, fPSA).

An overview of CV<sub>I</sub> and CV<sub>G</sub> estimates for studies with BIVAC grades A, B, and C are provided in Figures 1 and 2 for CEA, Figures 3 and 4 for tPSA, and Supplementary Figures 1–27 for the rest of the tumor markers, except for CA54.9, pepsinogen I and II, proGRP and TPS, where only a single study was available. Results for II and RCV are summarized in Table 3.

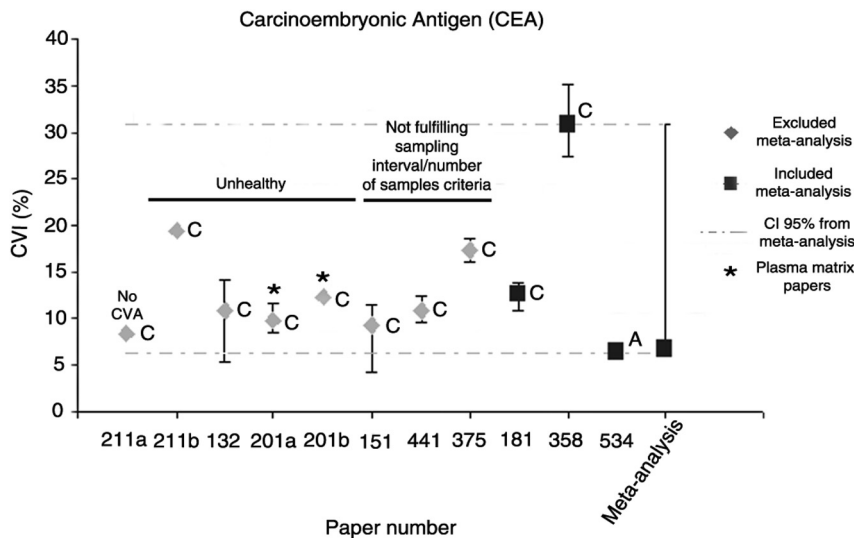
## Discussion

Applications of BV data for laboratory and clinical applications demand high-quality and well characterised BV estimates derived from relevant study populations [12]. For many measurands, including, as shown in our study, tumor markers, there is a lack of reliable data, making it challenging to obtain robust global estimates of CV<sub>I</sub> and CV<sub>G</sub>. This systematic review provides an overview of published BV data for tumor markers and addresses both the complexity and quality issues that could limit the use of tumor marker BV data in clinical practice.

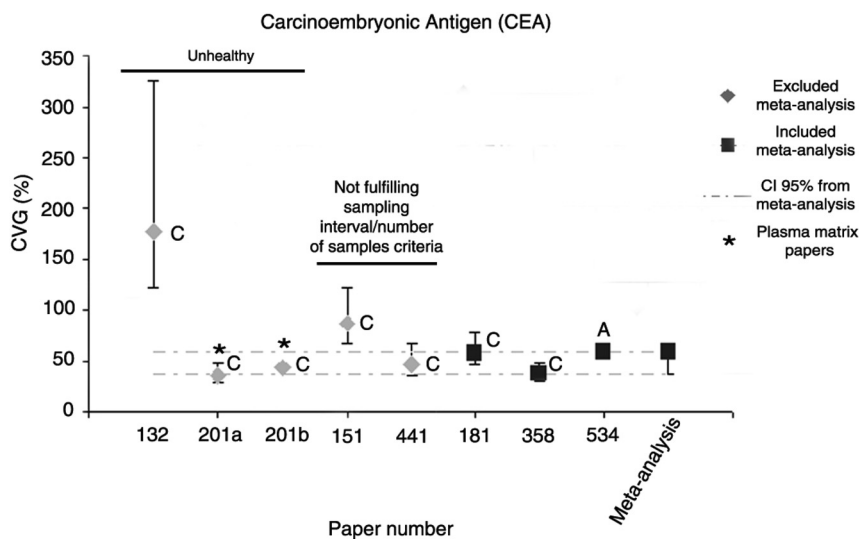
### Quality and study characteristics of BV studies for tumor markers

Generally, the quality of reported BV estimates is often compromised by variations in study design and the applied methodology. For tumor markers, there is an additional issue in that several historical studies have used older analytical methods, which may have delivered less accurate estimates. For some markers such as tPSA and CEA, our review identified a relatively high number of studies, but in total only four studies received the highest quality grade. These four studies, all derived from the European Biological Variation Study (EuBIVAS) [21, 22], comprised BV data for NSE, S100β [23], PSA [24], β-2-microglobulin [25], and more recently for AFP, CEA, CA125, CA15.3, CA19.9, HE4, and Cyfra21.1 antigen [26]. There were four papers graded B, providing data for chromogranin A, CA125, HE4, NSE, and CA72.4. All of the remaining papers were graded C (27 papers) or D (15 papers) (Table 1). Generally, more high-quality studies are warranted for tumor markers, in particular for CA54.9, pepsinogen I and II, proGRP and TPS where data were derived from only single studies and for CA54.9 and MCA, and where only grade D studies are available.

We also evaluated the effect of study population characteristics and concentrations of the different tumor markers on the BV estimates. There are, however, insufficient data to draw any robust conclusions on the influence of age and sex on CV<sub>I</sub> estimates. Some data are available for markers related to gastrointestinal tumors. For CEA and CA19.9, no significant differences between CV<sub>I</sub> estimates derived from studies performed only in females [26–28] and mixed-gender population [29–32] were observed, except for in paper 358 (mixed) [33] (Figure 1). For AFP, Coskun et al.



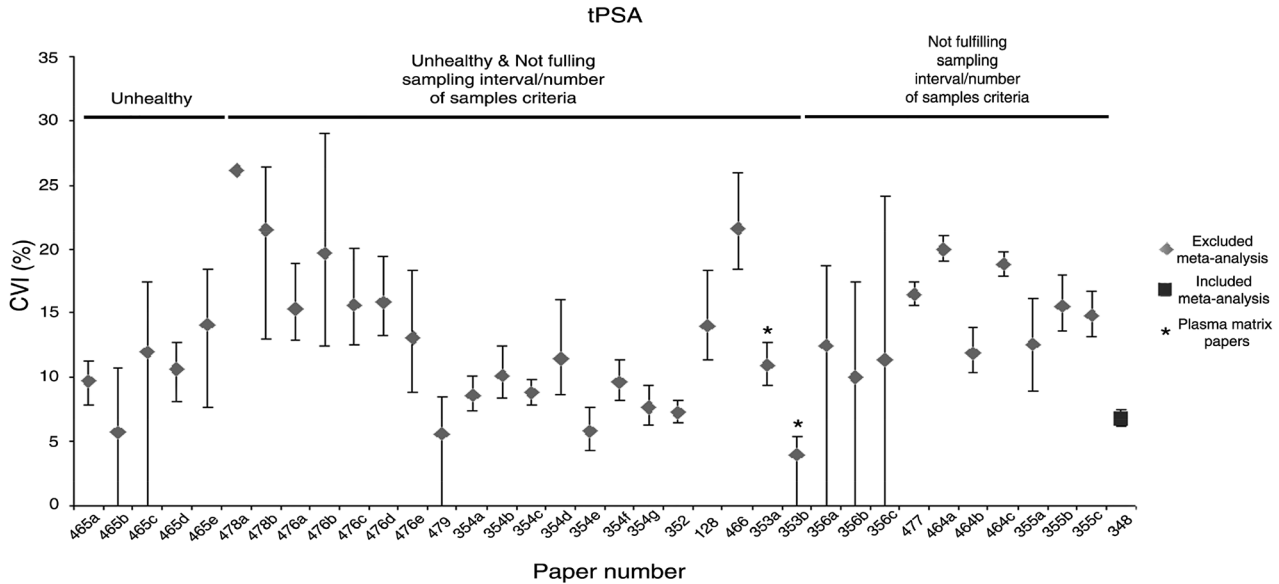
**Figure 1:** Estimates of within-subject biological variation ( $CV_1 \pm CI\ 95\%$ ) for carcinoembryonic antigen (CEA). Symbols indicate which studies have been included and excluded from the meta-analysis, grouped based on the cause of exclusion. The X-axis numbers represent the reference identification number in the EFLM biological variation database (as listed in Supplementary Table 1) and letters different subgroups. CI for individual papers are shown as error bars around the  $CV_1$  data point. Capital letters indicate the BIVAC grade given to the study from which the estimate has been derived from. In text causes of meta-analysis exclusion, No  $CV_A$ : the study did not report any analytical CV estimate. Not fulfilling sampling interval/number of samples criteria: less than three samples collected per subject or sampling intervals more frequent than 2/week or less frequent than 1/month.



**Figure 2:** Estimates of between-subject biological variation ( $CV_G \pm CI\ 95\%$ ) for carcinoembryonic antigen (CEA). Symbols indicate which studies have been included and excluded from the meta-analysis, grouped based on the cause of exclusion. The X-axis numbers represent the reference identification number in the EFLM biological variation database (as listed in Supplementary Table 1) and letters different subgroups. CI for individual papers are shown as error bars around the  $CV_G$  data point. Capital letters indicate the BIVAC grade given to the study from which the estimate has been derived from. In text causes of meta-analysis exclusion, not fulfilling sampling interval/number of samples criteria: less than three samples collected per subject or sampling intervals more frequent than 2/week or less frequent than 1/month.

have lately observed significant differences in  $CV_1$  estimates between males and females [26] and for CA19.9, significantly lower  $CV_1$  in postmenopausal women than

younger females; however, they did not identify a significant difference between  $CV_1$  estimates for CEA in these two females groups, as Tuxen et al. [7] reported previously.



**Figure 3:** Estimates of within-subject biological variation ( $CV_1 \pm CI$  95%) for total PSA (tPSA).

Symbols indicate which studies have been included and excluded from the meta-analysis, grouped based on the cause of exclusion. All the included studies were graded as BIVAC C, except 248 (grade A). The X-axis numbers represent the reference identification number in the EFLM biological variation database (as listed in Supplementary Table 1) and letters different subgroups. CI for individual papers are shown as error bars around the  $CV_1$  data point. In text causes of meta-analysis exclusion, not fulfilling sampling interval/number of samples criteria: less than three samples collected per subject or sampling intervals more frequent than 2/week or less frequent than 1/month.

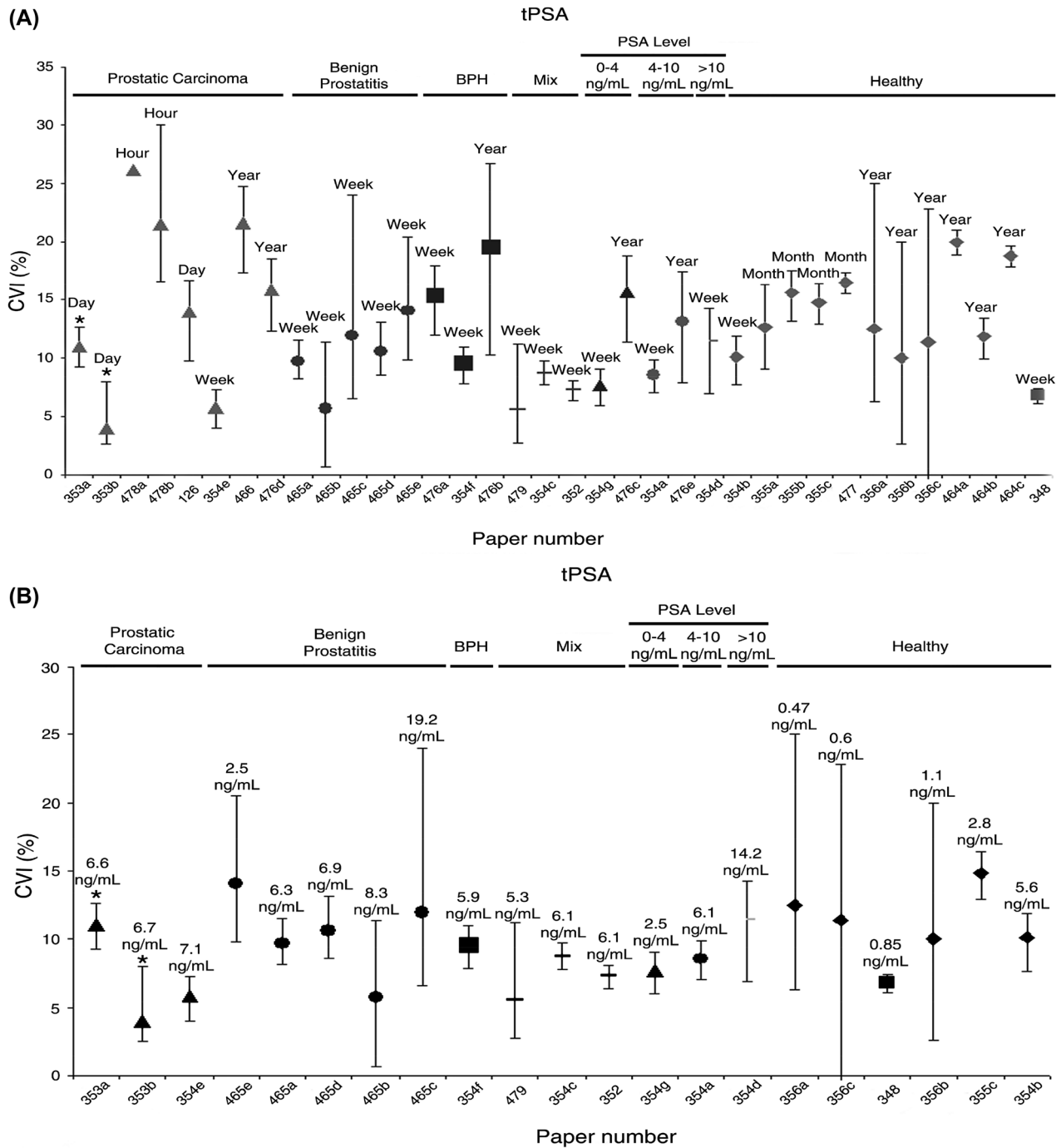
Braga et al. [34] did not identify differences in  $CV_1$  estimates between genders for NSE or chromogranin A. Regarding changes related to life phases, Braga et al. in another study [35] and the EuBIVAS [26] observed lower  $CV_1$  estimates for HE4 in postmenopausal healthy women compared to fertile women. Lower  $CV_1$  estimates in postmenopausal women compared with younger females were also observed for Cyfra21.1 antigen, CA125, and CA15.3 in the EuBIVAS. These findings should be taken into account when applying BV data.

For the studies included in our review, different sampling intervals were applied, ranging from hourly to yearly. Differences in sampling intervals could influence  $CV_1$  estimates [36]. A higher number of studies were available for tPSA with  $CV_1$  estimates that appear to increase slightly depending on the duration of the study, from a week (short-term) to a year (long-term), in different disease states as well as in a healthy population (Figure 4A). Understanding the characteristics of this variation could have a clinical impact influencing the definition of follow-up periods between medical examinations in these patients.

Another question arising is whether there is a relationship between  $CV_1$  estimates and tumor marker concentration. There were some differences in  $CV_1$  estimates relative to CEA concentration and health status, but no clear conclusion can be drawn (Supplementary Figure 13).

In healthy individuals with CEA concentrations ranging from 1.01 to 1.88  $\mu\text{g/L}$ , different results have been reported;  $CV_1$  estimates around 10% [26, 28, 30], Erden et al. [33] reported a  $CV_1$  of 30.9% and the EuBIVAS the lowest  $CV_1$ , at 6.3% (6.0–6.7%) [26]. Tuxen et al. [28] reported a  $CV_1$  estimate of 13.9% in another group of healthy individuals, where the mean CEA concentration was 3.8  $\mu\text{g/L}$ . Dittadi et al. [31] reported in breast cancer patients with CEA concentration ranging from 1.0 to 2.6  $\mu\text{g/L}$ , a  $CV_1$  of 8.4% and Plebani et al. a  $CV_1$  of 10.9% in patients with colon adenocarcinoma, with a median CEA concentration of 7.4  $\mu\text{g/L}$  [30]. These same authors did not find statistically significant differences when evaluating BV of CA19.9 and TPA in healthy and in colorectal cancer patients [29].

Estimates for tPSA were also grouped according to the type of study population, and for each group, estimates were ranked from the lowest to the highest concentration (Figure 4B). With this data, there was a huge variation in the  $CV_1$  estimates, but no apparent differences were observed in  $CV_1$  estimates with different concentrations of tPSA, both by type of pathology or between healthy and unhealthy individuals. A previous study has indicated that  $CV_1$  for a number of analytes, which are the key determinant in some organ-related pathologies, is larger than the  $CV_1$  estimated in healthy subjects [37].





**Table 3:** Index of individuality (II) and reference change value (RCV) calculated from meta-analysis derived biological variation estimated.

Measurands	II	<sup>a</sup> RCV, %
AFP	0.1	-12 to 13.6
β-2-microglobulin	0.43	-10.8 to 12.1
CA125	0.18	-16.4 to 19.6
CA15.3	0.13	-10.8 to 12.1
CA19.9	0.13	-15.5 to 18.3
CA72.4	0.49	-66.9 to 202.0
CEA	0.14	-17.4 to 21.1
Cyfra 21.1 antigen	0.67	-36.6 to 57.8
Chromogranin A	0.49	-29.3 to 41.4
HE-4	0.4	-15.3 to 18.1
NSE	0.7	-23.7 to 31.1
cPSA	0.17	-20.6 to 26.0
fPSA	0.16	-15.4 to 18.3
tPSA	0.18	-16.1 to 19.1
S100β	0.32	-21.5 to 27.4
TPA	0.46	-47.9 to 91.9

II, index of individuality; <sup>a</sup>RCV asymmetrical, reference change value calculated at  $Z=1.65$  for the probability level of significant change set at 95%.  $CV_A$  calculated from the mean  $CV_A$  of the papers included in the meta-analysis for each tumor marker.

robustly assess the relationship between  $CV_I$ , tumor marker concentrations and disease states.

## APS for tumor markers

APS are required to assure and ensure that the performance characteristics of laboratory methods adopted are fit for clinical application. Three models have been proposed for setting APS [38]. Generally, the use of a BV model is recommended in the absence of clinical outcome studies [39]. According to Ceriotti et al., the BV model is better applied to measurands under homeostatic control or measurands with a stable measurable concentration. The BV estimates of tumor markers in healthy people depend upon the physiological fluctuation in concentration due to variability in production in the noncancerous cells and to variability in release and elimination from the circulation [27, 30].

In principle, the BV model for APS is not the most appropriate for tumor markers; generally low measurand values are observed in non-diseased subjects, and these markers are not under a strict homeostatic control. Nevertheless, in the absence of relevant clinical outcome data, BV data may still have some utility in setting APS. However, the huge variability in  $CV_I$  estimates for some tumor markers (Table 2) indicates that, based on presently

available BV data, the uncertainty in these APSs will be large [39]. Generally, it is important that users recognise that the APS we present are suggestions rather than definitive specifications and must be applied with care. In Supplementary Table 4, desirable CV and bias derived on BV meta-analysis data are presented. Desirable  $CV_A$  ( $<0.5 \cdot CV_I$ ) for AFP, CA125, CA19.9, CEA, NSE, tPSA and S100β are stricter than the general goal suggested by the National Academy of Clinical Biochemistry (NACB) for analytical imprecision in relation to monitoring ( $<10\%$ ) [40]. Analytical requirements for bias for tumor markers have not yet been published. We did not calculate APS for bias for AFP, CA125, CEA, CA19.9 and tPSA, as the  $CV_G$  estimates were higher than 33%, indicating the skewed distribution of the homeostatic points of the individuals [41, 42]. For some tumor markers such as CA15.3, CA72.4, Cyfra21.1, tPSA, S100β and TPA, only one study fulfilled the criteria set for being included in the meta-analysis; and the APS are based on estimates from that single dataset. Thus, considering the lack of reliable BV data for many measurands, state-of-the-art analytical performance may yet be the best candidate to establish APS for many tumor markers.

## Assessment of serial changes in tumor markers

Recommendations for interpretation of changes in tumor markers concentrations, by different scientific groups such as the European Group on Tumor Marker (EGTM), the National Academy of Clinical Biochemistry (NACB) and the American Society of Clinical Oncology (ASCO) are based on studies that use cut-off values or empirically established increases and none of the recommendations include BV data [43–45]. BV data derived from healthy individuals with normal tumor marker concentrations are likely to have limited utility in many patients with cancer, but for patients in remission who are monitored for biochemical relapse, the RCV concept may add useful information [46, 47]. The II is useful for deciding whether it is suitable to use population based reference intervals for evaluating individual patient test results when monitoring, or if RCV is the better approach [17]. When II is below 0.6, the usual range of results for an individual subject spans only a small part of the population reference interval. Hence, a low index suggests that reference intervals have little value in detecting unexpected results for an individual subject [17, 47]. The II of the tumor markers included in our study varied, ranging from 0.1 for AFP to 0.7 for NSE. For most, the II is lower than 0.6, indicating that RCV is suitable to be

applied as a monitoring tool. For Cyfra 21.1 antigen and NSE, on the other hand, the  $CV_I$  estimate is very close to the  $CV_G$  value, indicating a low individuality.

We calculated RCVs based on the one-sided asymmetrical approach since in a setting with a patient in remission with normal concentrations, we are looking for an increase in the tumor marker concentration [19, 47]. An upward increase could be interpreted as a potential increased release of a tumor marker. In our study, the RCVs were calculated using estimates derived from the meta-analysis and  $CV_A$  estimates calculated from the mean  $CV_A$  of the included papers. In practice,  $CV_A$  estimates from the laboratory's own internal long-term quality control should be used to derive RCVs.

The meta-analysis based RCV derived from our study, based on data from healthy subjects, is for tPSA 19.1% (Table 3) [24] similar to the RCV value based on data from patients with PCa (19.4%) [48]. For metastatic prostate cancer, ASCO has established as a criterion of progression an increase of the tPSA of 25% plus criterion of nadir + 2  $\mu\text{g/L}$  [45].

Regarding CEA, although serial measurements are widely used in surveillance, there is a lack of agreement on the magnitude of concentration change that constitutes a clinically significant increase [3]. The NACB and EGTM recommendations describe an increase in CEA as significant, if the elevation is at least 30% over that of the previous value [1, 44]. Plebani et al. [30] suggested a RCV for detecting disease progression in patients with ovarian and colorectal cancer of 32.7%. Our lower RCV of 21.1% is established on data from healthy individuals, whose serum CEA concentration is within the normal range, and may not be representative for patients with high concentrations. Soletormos et al. [49] evaluated sequential concentrations of CA15.3, CEA, and TPA during monitoring of patients with breast cancer, observing that the variability of concentrations during steady state and the rate of increments during progression differed between the secretory markers CA15.3 and CEA and the cell turnover marker TPA. The authors suggested that when monitoring patients, both the variation in steady state and the rate of increase in progression must be taken into account for each marker [49]. An alternative approach for monitoring individual patients, is to establish personalized reference interval, which can be derived based on previous test results from steady state and estimates of analytical and biological variation [50]. This model for deriving personalized reference interval is a recently introduced concept, and further work is required to show whether the use of personalized reference intervals can

lead to earlier detection of recurrence of cancer and improved outcomes.

## Limitations

Currently, published studies on BV in tumor markers have important limitations. Though BV data are available for a large number of tumor markers (22 different markers), much of these data are derived from a limited number of studies. Based on the BIVAC criteria [9], only papers published by the EuBIVAS group (4 papers) [23–26] received an A grade, reporting BV data for 13 tumor markers. For eight of the remaining 16 tumour markers for which there were eligible data for meta-analysis, only one study fulfilled the inclusion criteria. This small number of eligible studies prevents us from delivering global BV estimates with sufficient robustness for most tumor markers.

Generally, BV estimates obtained for tumor markers have limitations for use in calculating APS, and in this article, BV-based APSs are described only as a suggestion.

Furthermore, our study reports meta-analysis based BV data derived from healthy individuals. There are a lack of robust data from relevant patient groups and our results have limited applicability for patients with cancer. The usefulness of BV estimates for tumor markers for the interpretation of changes in serial determinations based on RCV has recently been questioned, with van Rossum et al. [51] describing the limitations of using RCV and what factors need to be considered to overcome the limitations of RCV.

## Conclusions

In this study, we have summarized published BV data for a wide range of tumor markers. Though a substantial number of BV studies for tumor markers has been published, it is necessary to increase the number of high-quality studies performed in relevant populations, derived by the use of recent analytical technologies, for many measurands. There are also a lack of data to draw any robust conclusions on the influence of population characteristics and tumor marker concentrations on BV. Due to questions regarding the appropriateness of the BV model to set APS for these measurands, as well as the limitations in the quality of available BV data, state-of-the-art may still be the most appropriate model to establish APS for the majority of tumor markers. Application of BV data for delivery of personalized reference intervals and for RCVs, based on data from relevant study populations, may theoretically be

useful in monitoring of patients for biochemical relapses, but as of yet, further studies are required to demonstrate their clinical applicability.

**Research funding:** None declared.

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

**Competing interests:** Authors state no conflict of interest.

**Informed consent:** Not applicable.

**Ethical approval:** Not applicable.

## References

1. Sturgeon CM, Diamantis EP, Hoffman BR, Chan DW, Chang SO, et al. NACB: practice guidelines and recommendations for use of tumor markers in the clinic: quality requirements [Section 2]. Washington, DC: National Academy of Clinical Biochemistry; 2008.
2. Sokoll LJ, Chan DW. Clinical chemistry: tumor markers. In: Abeloff MD, Armitage JO, Niederhuber JE, Kastan MB, McKenna WG, editors. In Abeloff: clinical oncology, 3rd ed. Pennsylvania: Elsevier Churchill Livingstone; 2004.
3. Duffy ML. Tumor markers in clinical practice: a review focusing on common solid cancers. *Med Princ Pract* 2013;22:4–11.
4. Fraser CG. Biological variation: from principles to practice. Washington, DC: AACC Press; 2001:1–28 pp.
5. Harris EK. Effects of intra- and inter-individual variation on the appropriate use of normal range. *Clin Chem* 1974;20:1535–42.
6. Fraser CG. Reference change values: the way forward in monitoring. *Ann Clin Biochem* 2009;46:264–5.
7. 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.
8. Bartlett WA, Braga F, Carobene A, Coskun A, Prusa R, Fernández-Calle P, et al. Biological Variation Working Group, European Federation of Clinical Chemistry and Laboratory Medicine (EFLM). A check-list for critical appraisal studies of biological variation. *Clin Chem Lab Med* 2015;53:879–85.
9. Aarsand AK, Røraas T, Fernandez-Calle P, Ricós C, Díaz-Garzón 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.
10. 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. <https://www.westgard.com/biodatabase1.htm> [Accessed Mar 2021].
11. Aarsand AK, Fernandez-Calle P, Webster C, Coskun A, Gonzales-Lao E, Diaz-Garzon J, et al. The EFLM biological variation database. <https://biologicalvariation.eu/> [Accessed Mar 2021].
12. Gonzalez-Lao E, Corte Z, Simon M, Ricos C, Coskun A, Braga F, et al. Systematic review of the biological variation data for diabetes related analytes. *Clin Chim Acta* 2018;488:61–7.
13. Díaz-Garzón J, Fernández-Calle P, Minchinela J, Aarsand AK, Bartlett WA, Aslan B, et al. Biological variation data for cardiovascular risk assessment biomarkers. A systematic review using the critical appraisal checklist (BIVAC). *Clin Chim Acta* 2019;495:467–75.
14. Coskun A, Braga F, Carobene A, Tejedor Ganduxe X, Aarsand AK, Fernández-Calle P, et al. Systematic review and meta-analysis of within-subject and between-subject biological variation estimates of 20 haematological parameters. *Clin Chem Lab Med* 2019;58:25–32.
15. Burdick RK, Graybill FA. Confidence intervals on variance components. In: *Statistics: textbooks and mono-graphs*. New York (NY): Marcel Dekker; 1992, vol 127:78–115 pp.
16. Tu D, Shao J. *The Jackknife and bootstrap*, 1st ed. New York: Springer Series in Statistics; 1995.
17. Fraser CG, Harris EK. Generation and application of data on biological variation in clinical chemistry. *Crit Rev Clin Lab Sci* 1989;27:409–37.
18. Fokkema MR, Herrmann Z, Muskiet FA, Moecks J. Reference change values for brain natriuretic peptides revisited. *Clin Chem* 2006;52:1602–3.
19. Aarsand AK, Røraas T, Bartlett WA, Coşkun A, Carobene A, Fernandez-Calle P, et al. Harmonization initiatives in the generation, reporting and application of biological variation data. *Clin Chem Lab Med* 2018;56:1629–36.
20. Aarsand AK, Røraas T, Sandberg S. Biological variation-reliable data is essential. *Clin Chem Lab Med* 2015;53:153–4.
21. 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.
22. Carobene A, Aarsand AK, Bartlett WA, Coskun A, Díaz-Garzón J, Fernandez-Calle P, et al. The European Biological Variation Study (EuBIVAS): a summary report. *Clin Chem Lab Med* 2022; 60:505–17.
23. 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.
24. 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.
25. 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.
26. Coşkun A, Aarsand AK, Sandberg S, Guerra E, Locatelli M, Díaz-Garzón J, et al. Within- and between-subject biological variation data for tumor markers based on the European Biological Variation Study. *Clin Chem Lab Med* 2022;60: 543–52.
27. Soletormos G, Schiler S, 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.

28. Tuxen MK, Sölétormos G, Petersen PH, Schiøler V, Dombrowsky 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.
29. 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.
30. Plebani M, Giacomini A, Beghi L, De Paoli M, Roveroni G, Galleotti F, et al. Serum tumor markers in monitoring patients: interpretation of results using analytical and biological variation. *Anticancer Res* 1996;16:2249–52.
31. Dittadi R, Peloso L, Gion M. Within-subject biological variation in disease: the case of tumour markers. *Ann Clin Biochem* 2008;45:226–8.
32. 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.
33. 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.
34. Braga F, Ferraro S, Mozzi R, Dolci A, Panteghini M. Biological variation of neuroendocrine tumor markers chromogranin A and neuron-specific enolase. *Clin Biochem* 2013;46:148–51.
35. 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.
36. Sölétormos G, Semjonow A, Sibley PEC, Lamerz R, Petersen PH, Albrecht W, et al. Biological variation of prostate specific antigen: a survey of published estimates and consequences for clinical practice. *Clin Chem* 2005;51:1342–51.
37. Ricós C, Iglesias N, García-Lario JV, Simón M, Cava F, Hernandez A, et al. Within-subject biological variation in disease: collated data and clinical consequences. *Ann Clin Biochem* 2007;44:343–52.
38. 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.
39. 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.
40. Sturgeon CM, Hoffman BR, Chan DW, Chang SL, Hammond E, Hayes DF, et al. National Academy of Clinical Biochemistry Laboratory Medicine practice guidelines for use of tumor markers in clinical practice: quality requirements. *Clin Chem* 2008;54:e1–10.
41. Carobene A. Reliability of biological variation data available in an online database: need for improvement. *Clin Chem Lab Med* 2015;53:871–7.
42. Braga F, Panteghini M. Generation of data on within-subject biological variation in laboratory medicine: an update. *Crit Rev Clin Lab Sci* 2016;53:313–25.
43. Duffy MJ, van Dalen A, Haglund C, Hansson L, Klapdor R, Lamerz R, et al. Clinical utility of biochemical markers in colorectal cancer: European Group on Tumor Markers (EGTM) guidelines. *Eur J Cancer* 2003;39:718–27.
44. Van Poznak C, Somerfield MR, Bast RC, Cristofanilli M, Goetz MP, Gonzalez-Angulo AM, et al. Use of biomarkers to guide decisions on systemic therapy for women with metastatic breast cancer. American Society of Clinical Oncology Clinical practice guideline. *J Clin Oncol* 2015;33:2695–704.
45. Locker GY, Hamilton S, Harris J, Jessup JM, Kemeny N, Macdonald JS, et al. ASCO. ASCO 2006 update of recommendations for the use of tumor markers in gastrointestinal cancer. *J Clin Oncol* 2006;24:5313–27.
46. 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.
47. Fraser CG. Inherent biological variation and reference values. *Clin Chem Lab Med* 2004;42:758–64.
48. Nixon RG, Wener MH, Smith KM, Parson RE, Strobel SA, Brawer MK. Biological variation of prostate specific antigen levels in serum: an evaluation of day-to-day physiological fluctuations in a well-defined cohort of 24 patients. *J Urol* 1997;157:2183–90.
49. Sölétormos G, Petersen PH, Dombernowsky P. Assessment of CA 15.3, CEA and TPA concentrations during monitoring of breast cancer. *Clin Chem Lab Med* 2000;38:453–63.
50. 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.
51. Van Rossum HH, Meng QH, Ramanathan LV, Holdenrieder S. A word of caution on using tumor biomarker reference change values to guide medical decisions and the need for alternatives. *Clin Chem Lab Med* 2022;60:553–5.

**Supplementary Material:** The online version of this article offers supplementary material (<https://doi.org/10.1515/cclm-2021-0725>).