

# Clinical Application of Confidence Interval for Monitoring Changes in Tumor Markers to Determine the Responsiveness to Cancer Treatment

Jooyoung Cho

Yonsei University Wonju College of Medicine <https://orcid.org/0000-0002-9628-2334>

Dong Min Seo

Yonsei University Wonju College of Medicine

Young Uh (✉ [u931018@yonsei.ac.kr](mailto:u931018@yonsei.ac.kr))

Wonju Severance Christian Hospital, Yonsei University Wonju College of Medicine <https://orcid.org/0000-0002-2879-7870>

---

## Research article

**Keywords:** Tumor marker, Biological variation, Confidence interval, Intra-individual variation, Delta check, Cancer treatment

**DOI:** <https://doi.org/10.21203/rs.3.rs-66846/v1>

**License:** © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

---

## Abstract

**Background:** Tumor markers are used to monitor disease progression and determine the responsiveness to cancer treatment. However, there are no standardized criteria for monitoring serial tumor marker measurements. Herein, we have developed our own monitoring system for interpreting changes in tumor markers using overlapping 95% confidence intervals (CIs) to determine whether the changes are significant.

**Methods:** Two-year data, including 117,289 results for 11 tumor markers in our laboratory, were analyzed. The distributions of absolutely delta% and cut-off values for certain percentiles were calculated. CI ranges for each tumor marker were set based on biological variation, and data were analyzed for each patient assessed at health check-ups and clinics, individually and overall.

**Results:** Most tumor markers had low indices of individuality, with between inter-individual variability. The 95th percentile cut-offs for each tumor marker were much higher in the health check-up group than in the clinic group. In decreasing order, the percentages of results with no overlap in 95% CIs were thyroglobulin antigen, 14.9%; protein induced by vitamin K absence-II (PIVKA), 11.9%; prostate-specific antigen, 9.8%; and cancer antigen 72-4, 8.7%. After correction using the reference interval, the percentages decreased to less than 5%, except for PIVKA (10.9%).

**Conclusions:** We suggest that our own monitoring system can serve as a criterion for delta check and auto-verification of tumor markers. Further studies are required to validate and demonstrate this concept in real clinical situations using actual clinical data reflecting disease progression in cancer patients and responsiveness to cancer treatment.

## Background

In recent decades, there have been many improvements in laboratory medicine. However, it is difficult to determine the significance of changes between two consecutive laboratory results. Obviously, many laboratory results have variability in many aspects, including pre-analytical, analytical, and post-analytical factors [1–6]. Thus, clinicians, as well as laboratory physicians, should be aware of the potential risk of false interpretation.

Together with these confounding factors, biological variation is an important factor to consider when interpreting the changes in serial laboratory measurements. The term “biological variation” represents the physiological and metabolic properties according to the diverse status of healthy or non-healthy individuals and reflects fluctuations in serial measurements [7–9].

Comparing the changes in serial laboratory results is important for assessing the clinical status of a patient and responsiveness to therapy as well as predicting the point when additional interventions are necessary [6–11]. Conventional reference intervals (RIs) may also provide information about these points, but they are often of limited value [2, 11–13]. For analytes with a large RI, two consecutive results may be within the RI even if they differ significantly. In addition, when one of two consecutive results is outside the RI and another is inside, the clinical decision can be mistaken [11]. For these reasons, monitoring changes in serial laboratory results can be more helpful than conventional RIs in actual clinical situations [10]. Moreover, the decision limit should be set based on the individual properties of an analyte, factors associated with pre- and post-analysis, medical need, and clinical decision point [6].

For tumor markers, monitoring changes also plays an important role in patient management. The increasing interest in tumor markers as a less-invasive diagnostic tool for determining malignancy has encouraged their clinical use along with other diagnostic tools [2, 13]. Ideal tumor markers should reflect early detection, a differential diagnosis, response to therapy, prognosis, and progression or recurrence of malignancy and provide accurate and reliable results [9]. However, the changes in tumor marker values can show wide fluctuations over time and can be difficult to access. The analytical uncertainty and unreliability of tumor marker values determined by different methods are also obstacles for access [1, 2]. Moreover, there are no uniform or standardized criteria for determining the clinical significance of a difference between two consecutive marker results [3, 14]. Hence, it is difficult to set decision limits for interpreting the clinical significance of changes in tumor markers.

In clinical chemistry, biological variation is an important concept for explaining the variability of serial laboratory results. Biological variation includes analytical variation and within- and between-subject variation, thus, reflecting the variation of each

analyte [10, 12]. Moreover, numerous studies have examined the biological variations in tumor markers [1–3, 9, 13, 14]. One widely used evaluation tool based on biological variation is the reference change value (RCV), which has been used to interpret changes in serial laboratory results [8, 10, 11, 15] and as a criterion for delta check and auto-verification [4, 16–18]. However, RCV is a one-sided comparison method and still has limitations for analytes with a large intra-individual variation [11, 18].

When comparing two independent values, the confidence interval (CI) assigns a statistical significance for each value [19, 20]. The CI indicates a range of values that is likely to contain the parameter of interest with a specified probability, most often 95%. In two-sided comparison, the CIs for two certain values may overlap if they are not significantly different. In our previous study, we introduced the concept of CI into the interpretation of our clinical chemistry results [21].

To be best of our knowledge, there is no decision limit yet for conducting a delta check of tumor markers. Furthermore, there is no study regarding the association of overlapping CIs and laboratory results. In this study, we aimed to (i) apply overlapping CIs to interpret changes in two consecutive tumor marker values, (ii) estimate the clinical significance in each tumor marker, and (iii) consider the possibility of practical clinical application in routine clinical laboratories.

## Method

### Study design and data collection

This study was conducted at Wonju Severance Christian Hospital, a tertiary university-affiliated hospital located in Wonju, Republic of Korea. Two-year data of tumor marker test results between 2017 and 2018 were collected retrospectively from our laboratory information system (LIS). The tumor marker test items included in our study were as follows: alpha-fetoprotein (AFP), beta-human chorionic gonadotropin (HCG), cancer antigen 15 – 3 (CA15-3), cancer antigen 19 – 9 (CA19-9), cancer antigen 72 – 4 (CA72-4), cancer antigen 125 (CA125), carcinoembryonic antigen (CEA), protein induced by vitamin K absence-II (PIVKA), prostate-specific antigen (PSA), squamous cell carcinoma antigen (SCC), and thyroglobulin antigen (TG).

Tumor markers were measured using the following automated analyzers: the ADVIA Centaur® XP immunoassay system (Siemens Healthineers, Erlangen, Germany) for CA19-9, CA125, CEA, and PSA; the Cobas® e 601 module (Roche Diagnostics, Basel, Switzerland) for CA15-3, CA72-4, HCG, SCC, and TG; and the uTASwako i30 (Wako Diagnostics, Osaka, Japan) for AFP and PIVKA. This study was approved by the institutional review board (IRB) of Wonju Severance Christian Hospital (IRB no. CR319041) and the requirement for informed consent was waived.

### Data analysis

To calculate the changes in two consecutive tumor marker test results, pairs of test results were collected from each patient. The statistical methods are described in detail in our previous publication [21]. In brief, the following statistical processes were used. (i) The absolute values of the percent difference between two consecutive test results were calculated using the formula, Absolute delta% =  $|\text{Current result} - \text{Previous result}| / \text{Previous result} \times 100$  (%). (ii) The distributions of absolute delta% were assessed to determine the median, interquartile range, and some specified percentile values (i.e., 90th, 95th, and 97.5th). (iii) Overlapping CIs were used to monitor changes in serial tumor marker values, and CI ranges based on biological variation were calculated using the formula,

$$\text{Range of CI} = Z_{\text{test}} \times \sqrt{CV_A^2 + CV_I^2},$$

where  $Z_{\text{test}}$  indicates the Z score at which the degree of overlapping CIs at a given probability can reject a null hypothesis,  $CV_A$  is the analytical coefficient of variation obtained from the internal quality control program of our laboratory, and  $CV_I$  is the within-subject variation from the Westgard database of biological variation with the desired specifications of biological variation [22]. We also calculated the index of individuality, which is the ratio of  $CV_I$  to between-subject variation ( $CV_G$ ). Serial changes are known to be more important for an analyte with a low index (< 0.6), but conventional RI is diagnostically valuable for an analyte with a high index (> 1.4) [23, 24].

As mentioned in our previous study, the chance of non-overlapping 95% CIs is not at a type I error probability of  $\alpha = 0.05$ , which instead corresponds to that of 83.4% CIs [21], but at  $\alpha = 0.0056$ . In other words, the chance of non-overlapping 83.4% CIs corresponds to that exceeding  $RCV_{95\%}$ . In this study, we used the  $Z_{\text{test}}$  values of 1.39, 1.65, and 1.98 at CIs of 83.4%, 90%, and 95%, respectively.

All statistical analyses were performed using SPSS version 25.0 (IBM Corp., Armonk, NY, USA) and the add-in Analyse-it version 5.11 (Analyse-it Software Ltd., Leeds, UK) in Microsoft Excel 2016 (Microsoft Corp, Redmond, WA, USA).

## Development of a monitoring system

We developed a monitoring system using overlapping CIs for interpreting serial tumor marker values. The detailed processes are described in our previous study[21], and Fig. 1 is an example of our new system. If the 95% CIs of the previous and current test values do not overlap, then the two results are significantly different from each other. However, even in this case, if both values are within the RI, they have low clinical significance. Therefore, we conducted further analysis after correction using the RI. In other words, we proceeded with further analysis only if at least one previous or current test result deviated from the RI, and clinical significance was determined according to whether their CIs do not overlap.

Furthermore, through an administration menu of our system, we input our own  $CV_A$  values obtained annually from our one-year internal quality control data.  $CV_I$  is a fixed value for each tumor marker, but it can also be changed if the Westgard database is changed.

## Results

In total, 117,289 tumor marker results were analyzed over the study period, and 65,112 (55.5%) results were paired. Table 1 shows the test numbers for each tumor marker. We further analyzed the results by dividing them into health check-up and clinic (including both in- and out-patient) groups. AFP, CA125, CEA, and PSA, which are routinely tested tumor markers in our hospital, all showed higher percentages of paired results in the clinic group than in the health check-up group ( $p < 0.05$  for all).

Table 1  
Test numbers of tumor markers between 2017 and 2018.

Analyte		Overall	Paired	%
AFP	(Overall)	28,148	15,524	55.2%
	(Health check-up)	11,971	4,632	38.7%
	(Clinics)	16,177	10,892	67.3%
CA15-3		2,595	2,030	78.2%
CA19-9		17,048	8,236	48.3%
CA72-4		2,092	1,548	74.0%
CA125	(Overall)	8,048	3,554	44.2%
	(Health check-up)	3,490	1,022	29.3%
	(Clinics)	4,558	2,532	55.6%
CEA	(Overall)	28,924	14,978	51.8%
	(Health check-up)	9,560	4,036	42.2%
	(Clinics)	19,364	10,942	56.5%
HCG		2,180	1,550	71.1%
PIVKA		3,263	1,794	55.0%
PSA	(Overall)	15,303	7,726	50.5%
	(Health check-up)	5,902	2,136	36.2%
	(Clinics)	9,401	5,590	59.5%
SCC		1,502	1,086	72.3%
TG		8,186	7,086	86.6%

AFP, alpha-fetoprotein; CA15-3, cancer antigen 15-3; CA19-9, cancer antigen 19-9; CA72-4, cancer antigen 72-4; CA125, cancer antigen 125; CEA, carcinoembryonic antigen; HCG, beta-human chorionic gonadotropin; PIVKA, protein induced by vitamin K absence-II; SCC, squamous cell carcinoma antigen; TG, thyroglobulin antigen.

Most tumor markers had low ( $< 0.6$ ) indices of individuality, except for SCC (1.10). For HCG,  $CV_I$  was derived from the allowable limits of performance for biochemistry specified by the Royal College of Pathologists of Australasia [25]. For CA72-4 and PIVKA,  $CV_I$  was determined in our laboratory because there were no references of allowable limits for these items. We could not calculate indices of individuality for CA72-4, HCG, and PIVKA, because they had no available  $CV_G$  values. Including these items, the CI ranges according to various confidence levels are listed in Table 2.

Table 2  
Coefficients of variation and CI ranges for each tumor marker (%).

Analyte	CV <sub>A</sub>	CV <sub>I</sub>	CV <sub>G</sub>	Index of individuality	CI <sub>83.4%</sub>	CI <sub>90%</sub>	CI <sub>95%</sub>
AFP	7.28	12.2	45.6	0.27	29.58	35.12	41.72
CA15-3	3.97	6.1	62.9	0.10	10.12	12.01	14.27
CA19-9	6.61	16.0	130.5	0.12	24.06	28.56	33.93
CA72-4*	2.61	20.0†	-	-	28.04	33.28	39.53
CA125	5.35	24.7	54.6	0.45	35.13	41.70	49.53
CEA	6.90	12.7	55.6	0.23	20.09	23.85	28.33
HCG†	2.91	15.0*	-	-	21.24	25.21	29.95
PIVKA*	4.94	20.0†	-	-	28.64	33.99	40.38
PSA	5.36	18.1	72.4	0.25	26.24	31.15	37.00
SCC	4.59	39.4	35.7	1.10	55.14	65.45	77.75
TG	4.59	14.0	39.0	0.36	20.48	24.31	28.88
CV <sub>A</sub> is the analytical coefficient of variation obtained from the internal quality control program of our clinical chemistry laboratory; CV <sub>I</sub> and CV <sub>G</sub> are the within- and between-subject variation from the Westgard database with the desired specifications of biological variation, respectively. Index of individuality is equal to CV <sub>I</sub> divided by CV <sub>G</sub> .							
*For CA72-4 and PIVKA, CV <sub>I</sub> was determined by our laboratory because there were no references of allowable limits for these items.							
†For HCG, CV <sub>I</sub> was derived from the allowable limits of performance for biochemistry from the Royal College of Pathologists of Australasia.							
CI, confidence interval. See Table 1 for other abbreviations.							

Table 3 shows the statistical characteristics for the absolute value of each delta% in each tumor marker. All tumor markers showed non-normality (non-parametric), which was confirmed by the Kolmogorov–Smirnov test ( $p < 0.05$  for all). Of all tumor markers analyzed, CA15-3 (48.2%) and CEA (65.4%) showed the lowest 95th percentile cut-off value. However, CA19-9 (135.8%), SCC (185.3%), CA72-4 (188.7%), PSA (413.0%), TG (422.0%), and PIVKA (589.6%) showed high 95th percentile cut-off values of more than 100%. The 95th percentile cut-offs for all tumor markers were much higher in the health check-up group than in the clinic group.

Table 3  
Statistical characteristics for the absolute value of each delta% in each tumor marker (%).

Analyte		Median	(IQR)	Percentile value		
				90th	95th	97.5th
AFP	(Overall)	10.0	(0.0-32.5)	65.0	93.4	125.9
	(Health check-up)	10.4	(0.0-34.8)	65.0	90.0	114.9
	(Clinics)	9.7	(0.0-31.5)	64.8	95.0	133.3
CA15-3		6.9	(2.8-14.6)	28.0	48.2	145.2
CA19-9		15.7	(6.5-37.5)	76.9	135.8	597.1
CA72-4		28.0	(6.1-58.0)	111.4	188.7	294.6
CA125	(Overall)	22.2	(9.9-40.9)	70.8	93.2	161.7
	(Health check-up)	22.9	(11.4-36.5)	52.0	67.4	85.2
	(Clinics)	21.5	(9.4-43.9)	77.5	110.7	189.1
CEA	(Overall)	0.0	(0.0-12.3)	36.7	65.4	91.8
	(Health check-up)	0.0	(0.0-0.0)	11.7	21.7	28.5
	(Clinics)	0.0	(0.0-17.7)	46.0	75.4	117.0
HCG		54.5	(14.8-86.5)	98.8	99.7	99.9
PIVKA		26.3	(11.8-77.9)	204.4	589.6	> 1000
PSA	(Overall)	29.8	(11.1-70.4)	138.1	413.0	> 1000
	(Health check-up)	11.9	(6.0-21.9)	36.3	54.7	78.4
	(Clinics)	45.5	(18.8-88.1)	227.8	723.2	> 1000
SCC		20.2	(7.6-51.7)	125.0	185.3	274.7
TG		7.8	(0.0-50.9)	134.9	422.0	> 1000
IQR, interquartile range. See Table 1 for other abbreviations.						

Many tumor markers showed right-skewed patterns in the histogram, but for some tumor markers (CA72-4, HCG, PIVKA, PSA, SCC, and TG), a higher proportion (absolute delta > 100%) was noted. The absolute delta% distributions for each tumor marker are illustrated in Supplementary Fig. 1.

Table 4 shows the comparison of the percentages of test results with non-overlapping CIs. The percentages of non-overlapping 95% CIs were 14.9%, 11.9%, 9.8%, 8.7%, 7.6%, 5.7%, 2.8%, 1.9%, 1.9%, 0.5%, and 0.0% for TG, PIVKA, PSA, CA72-4, CA15-3, CA19-9, CEA, CA125, AFP, HCG, and SCC, respectively. After correction using the RI, the percentages decreased to less than 5%, except for PIVKA (10.9%). For the 95th percentile cut-offs, the percentages of non-overlapping 95% CIs were much higher in the health check-up group than in the clinic group for all analytes ( $p < 0.05$ ).

Table 4

Comparison of the percentages of test results with non-overlapping CIs before and after correction using the reference range according to CI level (%; Mean and 95% CI for each).

Analyte	CI range	83.4% CI		90% CI				95% CI					
		Correction	Before	After	Before	After	Before	After	Before	After			
AFP	(Overall)	6.2	(5.7–6.8)	1.6	(1.3–1.9)	3.6	(3.2–4.1)	1.3	(1.1–1.7)	1.9	(1.6–2.2)	1.0	(0.8–1.2)
	(Health check-up)	5.8	(4.9–6.8)	0.4	(0.2–0.7)	3.0	(2.3–3.8)	0.4	(0.2–0.7)	1.3	(0.8–1.7)	0.3	(0.1–0.5)
	(Clinics)	6.4	(5.7–7.0)	1.2	(1.0–1.5)	3.9	(3.4–4.4)	1.1	(0.9–1.3)	2.1	(1.8–2.6)	1.2	(0.9–1.5)
CA15-3		13.7	(11.6–16.0)	3.7	(2.7–4.9)	10.7	(9.1–12.6)	3.6	(2.6–4.8)	7.6	(6.0–9.3)	3.1	(2.1–4.1)
CA19-9		6.6	(5.9–7.4)	3.0	(2.5–3.5)	6.2	(5.5–7.1)	2.9	(2.4–3.4)	5.7	(4.9–6.4)	2.8	(2.3–3.3)
CA72-4		13.7	(11.2–16.1)	5.7	(4.1–7.4)	11.1	(9.0–13.4)	4.9	(3.5–6.6)	8.7	(6.8–10.9)	4.5	(3.1–6.1)
CA125	(Overall)	4.2	(3.3–5.2)	2.3	(1.6–3.0)	3.1	(2.3–4.0)	2.0	(1.4–2.6)	1.9	(1.3–2.6)	1.2	(0.7–1.8)
	(Health check-up)	1.6	(0.6–2.7)	0.4	(0.0–1.0)	1.2	(0.4–2.2)	0.4	(0.0–1.0)	0.6	(0.0–1.4)	0.1	(0.0–0.1)
	(Clinics)	5.2	(4.1–6.5)	3.1	(2.2–4.0)	3.9	(2.8–4.9)	2.6	(1.7–3.5)	2.4	(1.7–3.3)	1.6	(0.9–2.4)
CEA	(Overall)	4.5	(4.0–4.9)	3.3	(2.9–3.7)	3.8	(3.3–4.2)	3.0	(2.7–3.4)	2.8	(2.4–3.2)	2.5	(2.1–2.8)
	(Health check-up)	0.4	(0.1–0.7)	0.1	(0.0–0.2)	0.2	(0.0–0.5)	0.1	(0.0–0.1)	0.1	(0.0–0.2)	0.0	(0.0–0.1)
	(Clinics)	6.0	(5.4–6.6)	4.5	(4.0–5.1)	5.1	(4.5–5.7)	4.1	(3.7–4.7)	3.7	(3.3–4.3)	3.4	(2.9–3.9)
HCG		0.5	(0.1–1.0)	0.5	(0.1–1.0)	0.5	(0.1–1.0)	0.5	(0.1–1.0)	0.5	(0.1–1.0)	0.5	(0.1–1.0)
PIVKA		16.1	(13.8–18.7)	12.9	(10.7–15.3)	13.9	(11.8–16.4)	12.0	(9.9–14.2)	11.9	(9.8–14.0)	10.9	(8.9–13.0)
PSA	(Overall)	13.5	(12.5–14.6)	5.4	(4.7–6.2)	11.5	(10.5–12.5)	4.9	(4.1–5.6)	9.8	(8.9–10.7)	4.3	(3.6–5.0)
	(Health check-up)	2.2	(1.3–3.0)	0.3	(0.0–0.6)	1.9	(1.1–2.7)	0.3	(0.0–0.6)	1.3	(0.7–2.0)	0.3	(0.0–0.6)
	(Clinics)	17.9	(16.4–19.4)	7.4	(6.4–8.4)	15.2	(13.8–16.5)	6.6	(5.7–7.5)	13.1	(11.8–14.3)	5.8	(4.9–6.7)
SCC		3.3	(1.8–5.0)	1.8	(0.7–3.1)	0.9	(0.2–1.7)	0.7	(0.2–1.5)	0.0	(0.0–0.0)	0.0	(0.0–0.0)
TG		20.5	(19.2–21.8)	3.6	(3.0–4.2)	17.5	(16.3–18.8)	3.1	(2.6–3.7)	14.9	(13.7–16.0)	2.6	(2.1–3.1)

CI, confidence interval. See Table 1 for other abbreviations.

## Discussion



A tumor marker is a substance produced by a tumor or elevated by the presence of a tumor. The tumor marker concentration may be influenced by treatment or surgery, and these markers allow for differentiating patients with malignancy from healthy individuals [26]. Tumor markers have been increasingly used in clinical applications. Changes in serial tumor marker values should reflect the clinical status and response to cancer treatment, as well as provide relevant information for treatment decisions [2]. However, previous studies of tumor markers have only focused on the detection, differentiation, and prognosis of cancer and responsiveness to therapy. There is little to no consensus regarding the interpretation of tumor marker results. Additionally, few studies have been conducted on the biological variation of tumor markers [1–3, 9, 13, 14]. The major obstacle in interpreting serial tumor marker values is the development of tools that can overcome existing limitations by considering random errors, analytical factors, and biologic variations [1, 2, 13, 14].

In our previous study, we introduced our new monitoring system using overlapping 95% CIs in serial clinical chemistry test results. Thereafter, in this study, we attempted to utilize our system to monitor serial tumor marker values, including  $CV_A$  and  $CV_I$ . Furthermore, we have applied overlapping 95% CIs to monitor changes in two consecutive tumor marker results in our previous publication concerning clinical chemistry items [21]. The coefficients of variation,  $CV_A$ ,  $CV_I$ , and  $CV_G$ , for some tumor markers have been added to the Westgard database [22, 25]. Here, we have applied overlapping CIs and CI ranges based on the biological variation for each tumor marker. Biological variation is important when considering tumor markers, especially for interpreting two consecutive results to assess cancer progression or responses to therapy [1, 3, 14].

In our study, the indices of individuality were low ( $< 0.6$ ) in most tumor markers, except for SCC, indicating that many markers fluctuate in the same individual. Therefore, when interpreting two consecutive tumor marker results, a monitoring tool is more valuable than RI or a specified cut-off value [23, 24]. Considering each  $CV_I$ , the application of overlapping CIs as a two-sided comparison of two consecutive tumor marker results can be clinically meaningful for monitoring changes.

We have integrated this concept into our LIS system for monitoring our routine tumor marker tests. After application in our laboratory, we collected the data from our LIS server and analyzed the statistical characteristics. We compared the percentages of test results with non-overlapping CIs before and after correction using the RI. As mentioned in the introduction, in the case of tumor markers with a large RI range, if both values are within the RI, it may not be clinically meaningful even if they show a significant difference [11]. This is essential for tumor markers because their cut-off values are important for diagnosing the presence of malignancy. However, since the significance of the RI may vary for each tumor marker, it is necessary to apply it according to the characteristics of each tumor marker or the clinical situation of each laboratory, such as in the case of clinical chemistry [21].

Our study has some limitations. First, we collected only tumor marker results from patients to gain an exemption for informed consent from the IRB. Therefore, we did not consider clinical information, such as whether a patient had a tumor or received a treatment. Instead, we estimate that most patients in the health check-up group are cancer free. Further prospective study is needed to compare the relationship between our system and the real clinical course of cancer patients and healthy individuals.

Second, we did not consider the time variable in our new system. Although the time difference is considered in delta checks for some routine clinical chemistry parameters, thus far, there are no tumor markers with a time-dependent role. Furthermore, the application of delta checks for tumor markers is not yet available for most clinical laboratories.

## Conclusions

In conclusion, we have applied overlapping CIs to interpret changes in two consecutive tumor marker results. We suggest that this approach can overcome the limitations of existing RCVs and replace them, especially for analytes with a large  $CV_I$  value. Further studies are required to validate and demonstrate this approach in real clinical situations using actual clinical data reflecting the cancer status and treatment responsiveness of the patient.

## Abbreviations

RI: Reference interval; RCV: Reference change value; CI: Confidence interval; LIS: Laboratory information system; AFP: Alpha-fetoprotein; HCG: Beta-human chorionic gonadotropin; CA15-3: Cancer antigen 15-3; CA19-9: Cancer antigen 19-9; CA72-4: Cancer antigen 72-4; CA125: Cancer antigen 125; CEA: Carcinoembryonic antigen; PIVKA: Protein induced by vitamin K absence-II; PSA: Prostate-specific antigen; SCC: Squamous cell carcinoma antigen; TG: Thyroglobulin antigen; IRB: Institutional review board; CV<sub>A</sub>: Analytical coefficient of variation obtained from the internal quality control program; CV<sub>I</sub>: Within-subject variation; CV<sub>G</sub>: Between-subject variation;

## Declarations

## Author's contributions

Jooyoung Cho and Young Uh designed and directed the study. Dong Min Seo developed a monitoring system and applied to our laboratory information system. Dong Min Seo collected the data and Jooyoung Cho analysed the data. Young Uh supervised the findings of this work and helped to the final version of the manuscript. All authors discussed the results and contributed to the writing of the manuscript.

## Funding

None applicable.

## Availability of data and materials

The dataset used and analyzed during the current study is available from the corresponding author on reasonable request.

## Ethics approval and consent to participate

This study was approved by the institutional review board (IRB) of Wonju Severance Christian Hospital (IRB no. CR319041) and the requirement for informed consent was waived.

## Competing interests

The authors declare no potential conflicts of interest

## References

1. Tuxen MK, Soletormos G, Petersen PH, Schioler 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(1):12–22.
2. 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(3):212–8.
3. Soletormos G, Duffy MJ, Hayes DF, Sturgeon CM, Barak V, Bossuyt PM, Diamandis EP, Gion M, Hyltoft-Petersen P, Lamerz RM, et al. Design of tumor biomarker-monitoring trials: a proposal by the European Group on Tumor Markers. *Clin Chem.* 2013;59(1):52–9.
4. Fernandez D, Avinash S, Malathi M, Shivashankara A, Kumar A, Fernandez P. Establishing the reference change values (RCVs) and validating the delta check auto-verification in a clinical biochemistry laboratory. *Muller J Med Sci Res.* 2017;8(1):42–6.
5. Pavlou MP, Diamandis EP, Blasutig IM. The long journey of cancer biomarkers from the bench to the clinic. *Clin Chem.* 2013;59(1):147–57.
6. Fraser CG, Petersen PH. Desirable standards for laboratory tests if they are to fulfill medical needs. *Clin Chem.* 1993;39(7):1447–53. discussion 1453 – 1445.

7. Widjaja A, Morris RJ, Levy JC, Frayn KN, Manley SE, Turner RC. Within- and between-subject variation in commonly measured anthropometric and biochemical variables. *Clin Chem*. 1999;45(4):561–6.
8. Nunes LA, Brenzikofer R, de Macedo DV. Reference change values of blood analytes from physically active subjects. *Eur J Appl Physiol*. 2010;110(1):191–8.
9. 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(10):2077–83.
10. Fraser CG. Making better use of differences in serial laboratory results. *Ann Clin Biochem*. 2012;49(Pt 1):1–3.
11. Fraser CG. Reference change values: the way forward in monitoring. *Ann Clin Biochem*. 2009;46(Pt 3):264–5.
12. Fraser CG. Data on biological variation: essential prerequisites for introducing new procedures? *Clin Chem*. 1994;40(9):1671–3.
13. Qi Z, Zhang L, Chen Y, Ma X, Gao X, Du J, Zhang F, Cheng X, Cui W. Biological variations of seven tumor markers. *Clin Chim Acta*. 2015;450:233–6.
14. Trape J, Perez de Olaguer J, Buxo J, Lopez 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(1):219–22.
15. Harris EK, Yasaka T. On the calculation of a "reference change" for comparing two consecutive measurements. *Clin Chem*. 1983;29(1):25–30.
16. Fernandez-Grande E, Valera-Rodriguez C, Saenz-Mateos L, Sastre-Gomez A, Garcia-Chico P, Palomino-Munoz TJ. Impact of reference change value (RCV) based autoverification on turnaround time and physician satisfaction. *Biochem Med (Zagreb)*. 2017;27(2):342–9.
17. Lee J, Kim SY, Kwon HJ, Lee HK, Kim Y, Kim Y. Usefulness of biological variation in the establishment of delta check limits. *Clin Chim Acta*. 2016;463:18–21.
18. Ko DH, Park HI, Hyun J, Kim HS, Park MJ, Shin DH. Utility of Reference Change Values for Delta Check Limits. *Am J Clin Pathol*. 2017;148(4):323–9.
19. Austin PC, Hux JE. A brief note on overlapping confidence intervals. *J Vasc Surg*. 2002;36(1):194–5.
20. du Prel JB, Hommel G, Rohrig B, Blettner M. Confidence interval or p-value?: part 4 of a series on evaluation of scientific publications. *Dtsch Arztebl Int*. 2009;106(19):335–9.
21. Cho J, Seo DM, Uh Y. Clinical Application of Overlapping Confidence Intervals for Monitoring Changes in Serial Clinical Chemistry Test Results. *Ann Lab Med*. 2020;40(3):201–8.
22. Westgard QC. Desirable Biological Variation Database specifications. Available at: <https://www.westgard.com/biodatabase1.htm>, Accessed 13 Aug 2019.
23. Petersen PH, Fraser CG, Sandberg S, Goldschmidt H. The index of individuality is often a misinterpreted quantity characteristic. *Clin Chem Lab Med*. 1999;37(6):655–61.
24. Jhang JS, Lifshita MS: Poastanalysis: Medical decision making. In: Henry's clinical diagnosis and management by laboratory methods. 23rd edition edn. Edited by McPherson RA, Pincus MR. St. Louis, Missouri: Elsevier Inc.; 2017: 73–83.
25. Westgard QC. RCPA Allowable Limits of Performance for Biochemistry. Available at: <https://www.westgard.com/rcpa-biochemistry.htm>, Accessed 13 Aug 2019.
26. Jain S, Pincus MR, Bluth MH, McPherson RA, Bowne WB, Lee P: Diagnosis and management of cancer using serologic and other body fluid markers. In: Henry's clinical diagnosis and management by laboratory methods. 23rd edition edn. Edited by McPherson RA, Pincus MR. St. Louis, Missouri: Elsevier Inc.; 2017: 1432–1449.

## Figures

(A)

Test item		Value	Unit	HL/D/P/I	Reference interval	95% CI	95% CI_O
CEA	:	< 2	ng/mL		(<5)	(<2.7)	
CA 19-9	:	52.8	U/mL	▲	(.~<37.0)	(34.9~70.7)	

(B)

Test item		Value	Unit	HL/D/P/I	Reference interval	95% CI	95% CI_O
CEA	:	3.35	ng/mL		(<5)	(2.4~4.3)	-
CA 19-9	:	35.4	U/mL		(.~<37.0)	(23.4~47.4)	-

(C)

Test item		Value	Unit	HL/D/P/I	Reference interval	95% CI	95% CI_O
CEA	:	8.69	ng/mL	▲	(<5)	(6.2~11.2)	▲
CA 19-9	:	92.0	U/mL	▲	(.~<37.0)	(60.8~123.2)	▲

**Figure 1**

An example of our monitoring system for changes in two consecutive tumor marker results using overlapping 95% confidence intervals. (A) Initial result, (B) and (C) follow-up result of the patient. "HL/D/P/I" indicates the reference range (high, low), delta, panic flag, and serum index, respectively, showing a bias of more than  $\pm 10\%$  for hemolysis, icterus, or lipemia. "95% CI" indicates the 95% confidence interval for the current result, and "95% CI\_O" indicates whether the CIs of current and previous results overlap with each other. An arrow at "HL/D/P/I" indicates whether the test value has increased or decreased relative to the reference interval; a hyphen at "95% CI\_O" indicates that the CIs of the current result have not changed significantly relative to the CIs of the previous result (B), and an arrow at "95% CI\_O" indicates whether the CIs of the current result have increased or decreased significantly relative to the CIs of the previous result (C). See Table 1 for abbreviations.

## Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [SupplFig1Tumormarker.jpg](#)