Impact of calibration fitting models on the clinical value of chromogranin A

Simona Ferraro[1],*, Giuseppe Marano[2], Laura Ciardi[3], Chiara Vendramin[3], Angelo S. Bongo[1], Giorgio Bellomo[3], Patrizia Boracchi[2] and Elia M. Biganzoli[2]

[1] SCDO, Cardiologia 2, Ospedale Maggiore, Novara, Italy

[2] Istituto Statistica Medica e Biometria, Università degli Studi, Milano, Italy

[3] SCDU Laboratorio Ricerche Chimico Cliniche, Ospedale Maggiore, Novara, Italy

Abstract

Background: The clinical relevance of chromogranin A (CgA) concentrations depends on the analytical performance of the assay. The goal of the present study was to define the clinical involvements in CgA calibration models by evaluating the confidence intervals (CIs) for values from patients who were undergoing monitoring for disease.
Methods: Thirty calibration curves for the CgA assay [immunoradiometric assay (IRMA), (CIS-BIO)] were built using linear regression (LR), and four-parameter logistic models were used to estimate CIs for patient concentrations.
Results: We reported the inadequacy of the LR curve estimation procedure. We showed: 1) no evidence that the straight calibration line could fit the average responses, 2) non-constant and non-uniform variance of the replicated calibration responses. All tests performed in the analysis of variance and CI calculation for the calibration curve should be invalidated. The four-parameter logistic function yielded results for 16 curves only; this result could be due to the low number and inappropriate concentration of calibrators. This suggests that some aspects of the assay design should be reviewed. However, using the variance function estimated in this model, we could assess the CI for calibration curves and patient samples.
Conclusions: We showed that the four-parameter logistic calibration model with estimated variance function should better support clinical interpretation of marker concentration changes in patients serially tested.
Clin Chem Lab Med 2009;47:1297–303.

Keywords: calibration; chromogranin A; four-parameter logistic model; immunoradiometric assay.

*Corresponding author: Dott. Simona Ferraro, Ospedale Maggiore della Carità, SCDO, Cardiologia 2, C.so Mazzini 18, 28100 Novara, Italy
Phone: +39/3471598351, Fax: +39/03213733516,
E-mail: simona.ferraro@maggioreosp.novara.it

Introduction

Increased concentrations of circulating chromogranin A (CgA) are characteristic of neuroendocrine tumors (NETs), although the diagnostic and prognostic usefulness of this marker is widely debated (1, 2). Some studies reported low sensitivity and specificity of CgA concentrations for the diagnosis of NET. Concentrations may be dependent on tumor features, such as: type, secretory activity, neuroendocrine differentiation degree and total burden (1, 2). However, several lines of evidence show that the diagnostic accuracy is affected primarily by the type of assay used to measure CgA (1, 3–5). As a matter of fact, it has been reported that in healthy subjects, there is a discordance rate of ~20% between CgA concentrations measured using the two commercially available methods: immunoradiometric assay (IRMA) and enzyme-linked immunosorbent assay (ELISA) (1). The discordance has been ascribed to the different antibodies that are used to measure CgA-derived peptides (1). In addition, the Italian program for CgA external quality control described the pattern of differences between marker concentrations measured by these assays. There was a five-fold increase in the coefficient of variation (CV) between the IRMA and ELISA methods (6). Besides poor standardization of the assays that are clinically available, it is difficult to define the normal reference values for diagnosis of NET due to the high variability in CgA concentrations measured in healthy population (1). The main utility of CgA measurements appears to lie in the monitoring of therapeutic strategies for NET (1, 7, 8).

Considering CgA as a marker of sympathetic activation, recent studies reported its prognostic value in both chronic heart failure and in the setting of acute coronary syndromes (9, 10). These findings revealed new potential clinical applications with respect to cardiovascular disease.

As previously suggested, there is a need to evaluate and potentially improve assay performance due to the dependence of the diagnostic value of CgA with respect to the analytical method employed. The poor results reported from external quality control for the IRMA CgA assay could be partially linked to inappropriate quantitative setup of the assay calibration, which represents a common problem in many clinical assays.

The CgA IRMA method recommends that the standard curve be produced by linear interpolation of the standard tubes vs. the fixed standard concentrations and fitting a straight line regression model. The estimation of parameter models usually involves the ordinary least squares (OLS) method which implies assumptions about the variability of the responses

from the calibration tubes (11–14). The distribution of the response variable should be defined in order to assess the precision of the measured dose (15). To utilize OLS, response errors are assumed to be independent, with a constant variance and following a normal distribution to allow standard statistical inference procedures (14). These assumptions could be applied to an ideal assay system in which: 1) the response from one calibration tube is not affected by other tubes; 2) no systematic effects are present, such as carrier position; 3) no misleading responses are produced by some tubes; 4) the variability of response for fixed standard levels is independent of concentration (13). Any departure from such assumptions could lead to incorrect results (12).

It is widely known that the relationship between concentration and the response [measured as counts per minute (cpm)] is usually not linear, and different models, such as the ''four-parameter logistic'' model have been shown to be more suitable than simple linear regression (LR) in the assessment of the calibration curve (12–19). This model was proposed by Rodbard for two site IRMA and radioreceptor assays (20).

Evaluation of the relationship between variability in the response and the level of response [response-error relationship (RER)], provides a framework for assessing calibration/prediction intervals, and measures of assay performances (21). The original IRMA non-linear calibration approach was integrated to allow for modeling of variability in the assay response, according to a power of mean (POM) function law (12, 14, 22, 23). A critical step in this procedure is the choice of variance function used for calculation of the confidence interval (CI) of the calibration curve and for patient samples (23).

Since CgA is used for monitoring therapy and progression disease, there is a need to investigate the clinical relevance of changing marker concentrations. This evaluation could be performed only after the assessment of measurement imprecision of patient samples.

The goal of the present study was to assess the potential clinical impact of CgA calibration fitting models by evaluating interval estimates for samples from patients submitted to predict development of disease.

Materials and methods

The CgA IRMA assay [(CGA-RIACT, CIS-BIO; Cis-Bio, France)] is a non-competitive two-site immunometric assay, where capture and tracer antibodies are in excess (11). We selected 30 calibration curves from assays performed consecutively over 15 months. These curves were produced using four different batches of five standards, each containing different concentrations of lyophilized recombinant human CgA supplied by the manufacturer. According to the laboratory practice, the standard tubes were measured in duplicate for each standard concentrations and radioactivity was detected with use of a gamma scintillation counter calibrated for measurement of iodine 125. We excluded the first standard, 0 concentration, from computation of the calibration curve models since its measurement should be made for quality control

reasons in order to assess reagent or equipment contamination or high background. The 0 concentration standard should not be used for data processing unless high background counts are present. Patient serum was obtained from venous blood via venipuncture. After centrifugation (15 min, $1500 \times g$ at 4°C), serum were stored at –40°C until analysis, within 2 weeks.

Study design

Our study design consisted of the following steps:

1. Use of LR curve model to build a calibration curve according to standard practice. We checked the adequacy of linear interpolation of standard tube activity vs. fixed standard concentrations, and verified that the responses were consistent with the assumptions for a OLS estimation.
2. Adoption of four-parameter logistic model as an alternative approach in interpolation of dose-response curve. This is described in literature for radioimmunometric assays (RIAs), but it is seldom employed in laboratory practice.
3. Comparison of the results from two models and estimation of CIs for unknown samples.
4. Sampling, from routine laboratory measurements of CgA from patients submitted to monitor the development of disease in order to evaluate the clinical impact of the CI for unknowns. Intra- and inter-run evaluation was performed.

Statistical models and analysis

First, we described dose-response curves and performed dose interpolation using a LR model through OLS, and then by parameterization of the four-parameter logistic model (23, 24). Model (1) accounts for β_1 that is the asymptote, as the concentration x tends to 0 with $\beta_4 > 0$, and β_2 is the asymptote for x tending to ∞, β_3 is the predicted concentration at the response halfway between the two asymptotes and β_4 is related to the slope (23). In addition, we could model the variability in the response if the results were not constant. In this model, we included the estimation of variance function as POM according to the parameter θ. Previous data indicate that for immunoassays, θ generally falls in the range of 0.6–0.9 (23). This last step allows for the evaluation of the CI for calibration curves, and for means of unknown samples. For unknown samples, we chose to report Wald type CIs. Furthermore, in order to define test performances, we constructed an imprecision profile by plotting the estimated CV of the calibrated concentrations vs. concentrations (23). The CV was estimated by inverting the four-parameter logistic function and expressing the variability of concentration (x) as a bi-variate Taylor series expansion in y and β (23). In addition, we calculated the working range of the assay from the imprecision profile.

From the computation of the four-parameter logistic mode we reported:

a. minimal detection concentration (MDC) defined as the lowest concentration where the curve is increasing (decreasing) resulting in an expected response that is significantly greater (less) than the expected response at 0 concentration (20, 23);
b. reliable detection limit (RDL) defined as an increasing (decreasing) curve where the lowest concentration has a high probability of producing a response that is significantly greater (less) than the response at 0 concentration (23);

c. limit of quantitation (LOQ) defined as the lowest concentration where the CV of the dose is less than a threshold value, which was fixed at 20% (23).

We used a statistical tool for calibration of the plate based bioassay, "The Calib Package" was designed by Haaland et al. for the software R-library (25).

$$Y = f(x, \underline{\beta}) + \mu^{\theta}\sigma\varepsilon = \frac{\beta_1 - \beta_2}{1 + (\exp\{\beta_4[\log(x) - \beta_3]\}} + \beta_2 + \mu^{\theta}\sigma\varepsilon$$

$$\mu = f(x, \underline{\beta}) = E(Y)$$

$$Var(Y) = \sigma^2 f(x, \underline{\beta})^{2\theta} \qquad [1]$$

Results

LR model

Considering all the curves from different assay batches, graphical examination of the calibration points showed most to exhibit overlapping patterns. Since duplicate measurements of each standard tube has very close values, we had no evidence that a straight calibration line could fit the average responses (Figure 1A, C, E). We assumed that the variability of the calibration responses did not represent true random error only, as the pattern of the fitted regression straight line was systematically reproduced in

almost all calibrations from all batches. Moreover, we had evidence that the variance of the replicated calibration responses is not uniform but increases with the standard concentrations of the calibration points. The patterns of residuals showed a peculiar plot for every batch; from this we assessed heteroscedasticity (non-constant variance) and inadequacy of the conventional model estimation procedure. Also, systematic deviation ascertained by residual analysis showed the need for more complex models. We confirmed the non-normal distribution of errors (14). As a consequence, tests, such as the t and F tests, performed in the analysis of variance and CI calculation for the calibration curve, should be invalidated.

The four-parameters logistic model

This model yielded results for only 16 curves from 30 calibration curves (Table 1, Figure 1B, D, F). From these results, we could argue that in order to compute the dose-response curves, we should review some statistical aspects of the assay design which would not be evident and result in an empirical linear approach without variance estimation (19). Particularly, we should increase the number of standards and the midpoint concentration should be greater than the concentration related to the β_3 parameter (19). In addition, replicates should reflect true experimental error. The estimates of the parameters β_1, β_2 (inter-
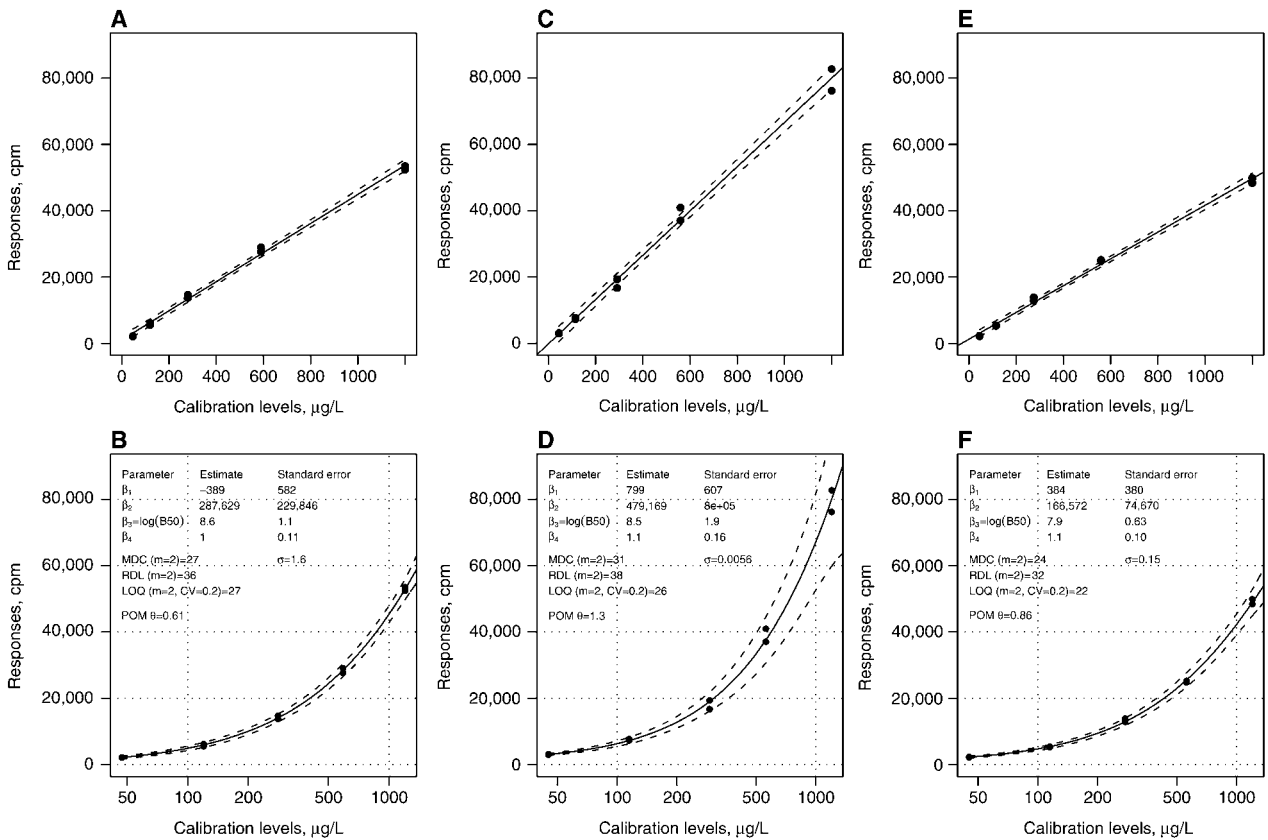

Figure 1 Various regression calibration curves.
(A) Linear regression calibration curve 1, Batch 1; (B) four-parameter regression model calibration curve 1; (C) linear regression calibration curve 18, Batch 3; (D) four-parameter regression model 18 calibration curve; (E) linear regression calibration curve 22, Batch 4; (F) four-parameter regression model calibration curve 22. MDC, minimal detection concentration; cpm, counts per minute; RDL, reliable detection limit; LOQ, limit of quantitation; POM, power of mean.

Table 1 Four-parameter regression model on 16 curves from 3 batches: evaluation of model parameters and related CI.

Curve	Batch[a]	B1 (CI)	B2 (CI) $\times 10^3$	B3 (CI)	B4 (CI)	θ
C1	1	−389.1 (−1530.5, 752.4)	287.6 (−162.9, 738.1)	8.6 (6.4, 10.8)	1.0 (0.8, 1.2)	0.61
C2	1	−114.4 (−1117.6, 888.8)	307.4 (−542.2, 1156.9)	8.8 (5.3, 12.3)	1.0 (0.8, 1.3)	0.72
C3	1	23.3 (−335.2, 381.7)	145.5 (12.5, 278.6)	8.0 (6.8, 9.1)	1.1 (1.0, 1.3)	1.10
C4	1	−227.8 (−681.3, 225.8)	644.6 (−278.0, 1567.1)	9.5 (7.7, 11.2)	1.0 (0.9, 1.1)	0.56
C5	1	727.8 (−295.3, 1751.0)	188.6 (86.3, 290.9)	7.6 (6.8, 8.3)	1.2 (1.0, 1.4)	0.65
C6	1	644.3 (68.3, 1220.3)	118.2 (44.8, 191.5)	7.6 (6.7, 8.5)	1.2 (1.0, 1.4)	0.83
C18	3	799.1 (−391.1, 1989.3)	479.2 (−1077.4, 2035.8)	8.5 (4.8, 12.2)	1.1 (0.8, 1.4)	1.26
C19	3	648.4 (292.0, 1004.8)	130.0 (86.7, 173.3)	7.6 (7.2, 8.1)	1.2 (1.1, 1.3)	0.55
C20	4	521.4 (−119.1, 1161.9)	230.9 (−7.6, 469.5)	8.2 (6.8, 9.5)	1.1 (0.9, 1.3)	0.93
C22	4	383.9 (−361.6, 1129.4)	166.6 (20.2, 312.9)	7.9 (6.7, 9.1)	1.1 (0.9, 1.3)	0.86
C23	4	895.9 (142.9, 1648.9)	70.4 (46.5, 94.4)	6.9 (6.4, 7.5)	1.3 (1.0, 1.5)	0.66
C24	4	−540.7 (−2132.5, 1051.1)	559.6 (−2211.9, 3331.1)	9.6 (3.1, 16.0)	0.9 (0.6, 1.2)	0.48
C27	4	500.3 (56.1, 944.4)	328.4 (142.3, 514.6)	8.4 (7.6, 9.1)	1.1 (1.0, 1.2)	0.73
C28	4	1076.8 (261.1, 1892.4)	137.4 (70.1, 204.6)	7.3 (6.6, 8.0)	1.2 (1.0, 1.5)	0.74
C29	3	1444.3 (51.9, 2836.7)	333.1 (−121.5, 787.7)	7.9 (6.1, 9.6)	1.2 (0.9, 1.4)	1.23
C30	3	1425.3 (593.9, 2256.7)	125.3 (50.1, 200.5)	7.1 (6.3, 7.9)	1.4 (1.1, 1.7)	1.05

[a]Calibrator levels for batch: 47,120,280,590,1200 (Batch 1); 45,115,292,560,1200 (Batch 3); 45,115,275,560,1200 (Batch 4). CI, confidence interval.

cept) and β_3, describing the shape of the estimated curve are reported in Table 1. Comparing these estimates, for all the available intra- and inter-batches curves, we showed stability, particularly for the estimated β_3 and β_4 parameters, independently from a supposed radio-ligand decay. The parameter β_4 (slope), should be related to the rate of change of counts with increasing dose (13). In our curves, estimates of β_4 seemed to not be greatly affected by radioligand quenching.

Setting the intra-batch curves ranking by radioligand quenching, the θ parameter seemed to be affected by radioligand decay: lower radioligand activity seemed to be associated with a higher θ value. The CIs for the calibration curves and MDC, RDL and LOQ are reported in Figure 1B, D, F.

From the imprecision profiles, defined only for 12 curves (with β_1 parameter estimate with a positive value), we could define the intra-run working range of the assay. However, the upper and the lower limits of the calculated range should not be considered because they exceed the upper and the lower concentrations of the calibrators.

Laboratory data

To assess the clinical value of interval estimation for unknowns, we compared the CI from overlapping sample concentrations, obtained experimentally, across the calibration curves fitted with the four-parameter logistic model (Table 2).

We selected samples at the lowest and highest concentrations, as well as around the normal limit value reported for 95% of healthy population using this method. We observed wider variability for the CI from comparable levels, across all curves, independently from the batch. For instance, we showed different CIs for overlapping levels, from the 27th and 28th calibration curves, despite these curves belonged to the same batch and overlapping θ. Taking these results together, we could argue that:

1. the CI for unknown levels is strictly associated with the run, not only to θ but also for all β parameters;

2. owing to the wide inter-run variability of CIs for overlapping patient measures, the CV reported by the assay manufacturer at defined concentrations cannot provide a suitable estimation of the variability of CgA concentrations on calibrated samples from patients. This implies different approaches to assay method development, validation and quality control.

Assessment of changing in marker levels

Previous evaluations on laboratory results should have a great impact on the monitoring of marker concentrations from a single patient during different times and across different runs. Taking into account the variability of the CI for overlapping marker concentrations from unknown samples across all runs, only by its evaluation, we could assess significant changes in marker concentrations. To further illustrate the need for CI for unknowns in order to assess intra-patient changes in marker concentrations, we followed CgA, evaluated in five patients with a diagnosis of myocardial infarction and 15 patients with cancer during follow-up.

Serum CgA from patients with myocardial infarction were measured in the same run (intra-run evaluation) using samples obtained at admission and in the following 6, 24 and 48 h. We showed that only by the interval estimate, could we assess for each patient the relevant changes in serum concentrations of CgA across the 4 time frames (Table 3).

From our experience, the problem of assessing significant increases/decreases in marker concentrations should be more complex when monitoring patient CgA values across different runs (inter-run evaluation). Consequently, we selected 100 CgA measurements from 15 patients being monitored serially for cancer. By assuming that these concentrations could come from the 16 different calibration curves computed using the four-parameter logistic model reported above, we assessed the CI at each level across all calibration curves and compared their inter-run variability. In Table 4, we reported the variability of CI for

Table 2 Laboratory CgA concentrations (median, range) estimated by linear regression model and then computed by four-parameter logistic model (4PL) with related CI (Δ).

Level	Laboratory levels by LR, µg/L	Estimates by 4PL, µg/L	n[a]	Δ CI for estimates by 4PL					
				Level	C2	C3	C4	C5	C6
1	52.2 (50.2, 56.2)	52.6 (50.1, 56.5)	15	1	23.9	8.8	9.3	22.6	18.3
2	107.7 (100.7, 121.4)	107.3 (100.3, 122.9)	15	2	41.4	18.9	14.0	33.8	29.3
3	319.9 (289.2, 363.9)	319.1 (296.7, 358.9)	12	3	87.8	56.2	26.1	NA	66.0
4	506.8 (420.6, 580.7)	512.4 (418.6, 594.7)	9	4	NA[b]	108.3	39.1	NA	101.4
5	753.8 (615.6, 991.2)	747.1 (611.0, 968.4)	5	5	NA	172.4	NA	NA	208.4
					C18	C19	C20	C29	C30
				1	15.4	17.1	13.7	15.1	25.6
				2	37.4	23.1	24.3	30.2	41.9
				3	NA	NA	68.8	100.4	123.0
				4	NA	NA	123.3	199.3	NA
				5	NA	59.1	NA	324.9	NA
					C22	C23	C24	C27	C28
				1	16.7	29.0	31.7	8.6	24.4
				2	27.9	38.2	50.3	13.9	33.9
				3	69.0	88.6	92.3	28.6	71.2
				4	111.1	243.9	121.7	36.3	117.5
				5	NA	NA	NA	NA	NA

[a]Number of concentrations evaluated across the 9 curves; [b]no available concentration for this curve; LR, linear regression; CI, confidence interval.

Table 3 Intra-run CgA concentrations (µg/L) obtained serially from patients with myocardial infarction.

Patient	Time, h	Laboratory levels	Estimates by 4PL	CI on estimates by 4PL	
				Lower	Upper
1	0	78.5	80.1	68.8	91.5
	6	119.8	121.5	104.9	138.1
	24	162.9	163.4	142.4	184.5
	48	106.7	108.5	93.5	123.6
2	0	136.9	138.3	119.8	156.7
	6	329.1	325.4	286.5	364.3
	24	218.3	215.9	189.6	242.2
	48	312.5	308.2	271.5	345.0
3	0	46.1	45.5	37.4	53.6
	6	60.0	60.7	51.8	69.6
	48	92.5	94.5	81.2	107.7
4	0	301.5	296.8	261.5	332.1
	6	308.6	304.2	268.0	340.4
	24	181.2	180.6	157.9	203.4
	48	363.9	361.4	317.9	404.9
5	0	145.4	146.5	127.2	165.8
	6	298.0	293.3	258.3	328.2
	24	181.2	180.9	158.1	203.6[a]
	48	209.1	207.3	181.9	232.7[a]

[a]Overlapping CI (confidence interval); CgA, chromogranin A; 4PL, four-parameter logistic model.

two concentrations obtained from patients with cancer. We confirmed further that the strong variability in CI would affect the intra-patient changes in marker concentrations from the same sample across the different. Such results are dependent on assay design and sampling variability across runs.

Discussion

The serial measurement of circulating CgA concentrations was reported to have value for both prognosis and monitoring response to treatment for several diseases, from cancer to cardiovascular settings, due to its pleiotropic effects as a pro-hormone (9, 10, 26). However, the clinical value of CgA is dependent on the features of the commercially available assays which showed poor measurement agreement due to poor standardization (1, 6). Concerning assay performance, calibration curve fitting is a critical issue that should be carefully evaluated to assess the reliability of CgA measurements. Marker concentrations are affected by measurement error, whose magnitude could exert a significant clinical impact, particularly markers such as CgA that need to be measured serially for monitoring. Thus, we reported the clinical

Table 4 Evaluation of changing CgA concentrations for two sequential measurements used for monitoring patients with cancer.

Patient	Number of total measurements for patient	CgA first level, μg/L	CgA second level, μg/L	Lower–upper CI of the first level, across all runs	Lower–upper CI of the second level, across all runs
1	8	668.0	748.0	44.6–323.1	47.9–374.2
2	5	339.0	277.0	27.2–133.3	24.1–101.9
3	6	602.5	645.0	41.5–283.3	43.6–309.0
4	8	43.6	97.6	7.4–30.8	12.8–46.5
5	6	70.6	91.7	10.1–37.6	12.3–44.7
6	4	140.0	115.8	16.5–58.0	14.5–51.9
7	5	199.0	152.0	20.3–69.3	17.3–60.6
8	9	234.0	246.0	22.0–82.0	22.6–87.4
9	6	335.8	370.0	27.0–131.6	28.8–149.8
10	7	118.0	139.0	14.7–52.5	16.4–57.7
11	5	168.0	104.0	18.5–63.8	13.4–48.5
12	5	291.0	342.0	24.8–108.7	27.3–134.8
13	7	334.4	366.0	26.9–130.8	28.6–147.7
14	8	188.0	144.4	19.7–67.4	16.8–59.0
15	11	89.0	118.8	12.0–43.8	14.8–52.7

CI, confidence interval; CgA, chromogranin A.

impact of interval estimation of CgA concentrations in patients being monitored with serial measurements of the marker. Our data showed that the reliability of estimations for unknown samples, and consequently the clinical value of the laboratory results, was affected by the statistical model used for the calibration curve.

First, we confirmed data from the literature concerning the inadequacy of the straight line regression model, commonly used in laboratory practice to fit calibration curves (11–23, 27). Moreover, in its standard use, this approach did not yield any information about the variance of the dose-response curve. Even in situations of a good straight line fit, adoption of OLS would be suboptimal if the assumption of homogeneity of error variances is not tenable (24). Such a situation, coupled with possible non-normal distribution of errors, prevents the correct application of standard inference procedures for statistical testing and interval estimation (24). Accounting for the non-linear dose-response relationship, suitable statistical methods were required to estimate the variance function parameter θ (27). Furthermore, the use of an assay design with a geometric progression in calibrator concentrations is unsuitable for building a straight line calibration model (28). Consequently, in the assay current application there are no minimal requirements to check performance characteristics, such as accuracy, precision, analytical sensitivity or interval estimation of unknown sample values. Under these conditions, a warning should given in the instructions of the assay methodology because of the possible clinical consequences.

In the second step of our investigation we used a four-parameter logistic model to interpolate immunometric calibration curves, since this model is reported to be the most flexible and reliable model in IRMA, RIA and ELISA calibration framework. However, the attempt failed to provide results from all curves, confirming our suspicion on the suitability of

the IRMA assay design (12–14). The assay design involves a reduced number of standards with a narrow range of concentrations. This results in too few calibrator points for reliable estimation when using the four-parameter logistic model. This could greatly affect calibration performance, preventing optimal estimation of the model parameters and imprecision of unknowns (19), even when using correct model assumptions. We did not consider additional interpolating functions as, for example, polynomials or splines for two major reasons: their lack of biological mechanistic interpretation, unlike the logistic model, and the unavailability of standard software for non-constant variance estimation.

Finally, we provided the assessment of CIs for CgA concentrations from patient samples submitted for serial measurement of the marker. These provided relevant information for the interpretation of changing concentrations during monitoring of disease. From our experience, monitoring the effects of drug therapy on cancer patients, relying on the estimation of serially obtained values, is difficult without CI estimation. We could not provide the magnitude of variability around laboratory data. Previous data reported θ as the critical parameter to define the CIs of the calibration curve, and consequently of unknown sample concentrations (23). We showed that overlapping θ parameters produced very different CIs on overlapping sample levels across the runs. This suggests that this variability is linked directly to all parameters of the calibration curve. Accounting for the wide variability of CI for the overlapping values across different runs, we could not provide the magnitude of error around the CgA measurement by relying on the CV reported by the manufacturer. As a consequence, we showed that for clinical interpretation of variations in marker concentrations in patient samples submitted for serial measurement of CgA, we should provide interval estimation for the samples, owing to the dependence of imprecision from the runs.

In conclusion, the usefulness of this marker could be increased by using a suitable calibration statistical model and improving assay design. Further work is needed for alternative assay development and validation for utilizing reliable interval estimation.

Conflict of interest statement

The authors declare that there are no conflicts of interest.

References

1. Zatelli MC, Torta M, Leon A, Ambrosio MR, Gion M, Tomassetti P, et al. Chromogranin A as a marker of neuroendocrine neoplasia: an Italian multicenter study. Endocr-Relat Cancer 2007;14:473–82.

2. Frolla P, Faggiano A, Mansueto G, Avenia N, Cantelmi MG, Giovenali P, et al. The biological characterization of neuroendocrine tumors: the role of neuroendocrine markers. Endocrinol Invest 2008;31:277–86.

3. Nobels FR, Kwekkeboom DJ, Boiullon R, Lamberts SW. Chromogranin A: its clinical value as marker of neuroendocrine tumors. Eur J Clin Invest 1998;28:431–40.

4. Seregni E, Ferrari L, Bajetta E, Martinetti A, Bombardieri E. Clinical significance of blood chromogranin A in neuroendocrine tumors. Ann Oncol 2001;12:S69–S72.

5. Campana D, Nori F, Piscitelli L, Morselli-Labate AM, Pezzilli R, Corinaldesi, et al. Chromogranin A: is it a useful marker of neuroendocrine tumors? J Clin Oncol 2007;25:1967–73.

6. Verderio P, Dittadi R, Marubini E, Pizzamiglio S, Gion M, De Apollonia L, et al. An Italian program of external quality control for chromogranin A (CgA) assay: performance evaluation of CgA determination. Clin Chem Lab Med 2007;45:1244–50.

7. Zatelli MC, Piccin D, Bondanelli M, Tagliati F, De Carlo E, Culler MD, et al. An in vivo OctreoScan-negative adrenal pheochromocytoma expresses somatostatin receptors and responds to somatostatin analogs treatment in vitro. Horm Metab Res 2003;35:349–54.

8. Zatelli MC, Piccin D, Bottoni A, Ambrosio MR, Margutti A, Padovani R, et al. Evidence for differential effects of selective somatostatin receptor subtype agonists on alpha-subunit and chromogranin a secretion and on cell viability in human non-functioning pituitary adenomas in vitro. J Clin Endocr Metab 2004;89:5181–8.

9. Ceconi C, Ferrari R, Bachetti T, Opasich C, Volterrani M, Colombo B, et al. Chromogranin A in heart failure; a novel neurohumoral factor and a predictor for mortality. Eur Heart J 2002;23:967–74.

10. Omland T, Dickstein K, Syversen U. Association between plasma chromogranin A concentration and long-term mortality after myocardial infarction. Am J Med 2003;114:25–30.

11. Grassi J, Créminon C, Frobert Y, Etienne E, Ezan E, Volland H, et al. Two different approaches for developing immunometric assays of haptens. Clin Chem 1996;42:1532–6.

12. Tiede JJ, Pagano M. The application of robust calibration to radioimmunoassay. Biometrics 1979;35:567–74.

13. Dudley RA, Edwards P, Ekins RP, Finney DJ, McKenzie IG, Raab GM, et al. Guidelines for immunoassay data processing. Clin Chem 1985;31:1264–71.

14. Rodbard D, Lenox RH, Wray HL, Ramseth D. Statistical characterization of the random errors in the radioimmunoassay dose-response variable. Clin Chem 1976;22:350–8.

15. Pilo A, Zucchelli GC, Malvano R, Masini S. Main features of computer algorhythms for RIA data reduction; comparison of some different approaches for the interpolation of the dose-response curve. J Nucl Med Allied S 1982;26:235–48.

16. Pilo A, Zucchelli GC. Automatic treatment of radioimmunoassay data: an experimental validation of the results. Clin Chim Acta 1975;64:1–9.

17. Davis SE, Munson PJ, Jaffe ML, Rodbard D. Radioimmunoassay data processing with a small programmable calculator. J Immunoassay 1980;1:15–25.

18. Hatch KF, Coles E, Busey H, Goldman SC. End-point parameter adjustment on a small desktop programmable calculator for logit-log analysis of radioimmunoassay data. Clin Chem 1976;22:1383–9.

19. Chang PC, Rubin RT, Yu M. Optimal statistical design of radioimmunoassays and competitive protein-binding assays. Endocrinology 1975;96:973–81.

20. Rodbard D. Statistical estimation of the minimal detectable concentration (sensitivity) for radioligand assays. Anal Biochem 1978;90:1–12.

21. Finney DJ. Radioligand assay. Biometrics 1976;32:721–40.

22. Raab GM, Thompson R, McKenzie I. Variance function estimation for immunoassays. Comput Prog Biomed 1980;12:111–20.

23. O'Connel MA, Belanger BA, Haaland PD. Calibration and assay development using the four-parameter logistic model. Chemometr Intell Lab 1993;20:97–114.

24. Draper NR, Smith R. Applied regression analysis, 3rd ed. New York: John Wiley & Sons Inc, 1998.

25. http://cran.r-project.org/web/packages/calib/

26. Tota B, Quintieri AM, Di Felice V, Cerra MC. New biological aspects of chromogranin A-derived peptides: focus on vasostatins. Comp Biochem Physiol A Mol Integr Physiol 2007;147:11–8.

27. Rodbard D, Frazier GR. Statistical analysis of radioimmunoassay data. Method Enzymol 1975;37:3–22.

28. Vineis P, Schulte PA, Vogt RF. Technical variability in laboratory data. In: Schulte PA, Perera FP, editors. Molecular epidemiology: principles and practices. San Diego CA: Academic Press, 1993.