

Federica Braga*, Simona Ferraro, Francesca Ieva, Anna Paganoni and Mauro Panteghini

A new robust statistical model for interpretation of differences in serial test results from an individual

DOI 10.1515/cclm-2014-0893

Received September 9, 2014; accepted December 10, 2014; previously published online January 24, 2015

Abstract

Background: Population-based reference intervals have very limited value for the interpretation of laboratory results when analytes display high biological individuality. In these cases, the longitudinal evaluation of individual results using the reference change value (RCV) is the recommended approach. However, the traditional model for RCV calculation requires a Gaussian frequency distribution of data and risks to overestimate the parameter if a correlation between within-subject serial measurements is present.

Methods: We propose and validate an alternative non-parametric statistical model for interpretation of differences in serial results from an individual, overcoming data distribution and correlation issues.

Results: After describing the traditional and newly proposed statistical models, we compared them with each other using a simulation on three specific analytes displaying different concentration distributions in biological setting. We demonstrated that when analyte concentrations followed a Gaussian frequency distribution, as in the case of glycated hemoglobin, both methods can be used equally. On the contrary, if analyte concentrations present a bimodal (e.g., chromogranin A) or skewed (e.g., C-reactive protein) distribution, the information obtained by two statistical methods is different.

Conclusions: The proposed statistical approach may be more appropriate in assessing difference between serial measurements when individual data are not normally distributed.

Keywords: biological variability; laboratory tests; reference change value.

Introduction

The most common approach used to interpret individual laboratory results is the comparison with population-based reference intervals (RI) [1]. However, this method suffers from important limitations, potentially misleading the test interpretation by physicians and laboratory professionals. In this regard, the index of individuality (II) yields objective information about the utility of conventional RI. It derives from biological variation (BV) of the analyte and is expressed as the ratio between average within-subject total variance (SD_s^2) and between-subject biological variance (SD_G^2) [2]. When the ratio between these two sources of variance is <0.6 , RI are of very limited value in the detection of unusual results for a particular individual. Conversely, when II is >1.4 , the individual result can be usefully compared with the corresponding RI [3]. Unfortunately, it has been demonstrated that for most analytes assayed in laboratory medicine, the within-subject biological variation (CV_i) is much smaller than between-subject biological variation (CV_G) [1]. In this situation, individuals may have highly unusual results, but this information can be ignored by clinicians as it still lies within the RI [4]. On the other hand, in some individuals results can change from inside the RI to outside (and vice versa) without any clinical significance: laboratories would conventionally flag the results outside the reference limits, potentially initiating some unnecessary follow-up activity [5].

When analytes display high biological individuality (i.e., low II), the recommended way to interpret their results is to use the reference change value (RCV), derived from BV data (in addition to the analytical variation

*Corresponding author: Federica Braga, UOC Patologia Clinica, Azienda Ospedaliera 'Luigi Sacco', Via GB Grassi 74, 20157 Milano, Italy, Phone: +39 0239042743, Fax: +39 0250319835, E-mail: federica.braga@unimi.it; and Department of Biomedical and Clinical Sciences 'Luigi Sacco', University of Milan, Milan, Italy

Simona Ferraro and Mauro Panteghini: Department of Biomedical and Clinical Sciences 'Luigi Sacco', University of Milan, Milan, Italy

Francesca Ieva: Department of Mathematics, University of Milan, Milan, Italy

Anna Paganoni: Department of Mathematics, Politecnico di Milano, Milan, Italy

of measurements), particularly from CV_1 [6]. The RCV concept was first introduced by Harris and Yasaka to identify the change needed between two serial results from the same individual to be statistically significantly different [7]. It has been increasingly popular in laboratory medicine and currently there is a general agreement in considering RCV as the best tool for monitoring, but also for making diagnosis when the RI approach is inadequate [8]. RCV has also been applied to set objective criteria for use in Δ -checking quality control techniques and in auto-verification and auto-validation strategies [9].

The estimation of RCV has sparked wide debate and discussion regarding the statistical approach that should be applied. Indeed, the Fraser's traditional approach assumes a normal frequency distribution of data and derives the RCV from the estimation of intra-individual variability as CV_1 [3]. By definition, the CV concept refers to parametric statistics resulting from the ratio between SD and mean of data distribution. However, in biology a Gaussian frequency distribution of data occurs for a limited number of quantities. As Franzini already pointed out 15 years ago [10], if the CV exceeds 33.3% then either the variable can assume negative values (but this is not the case for laboratory results) or the distribution is not normal. In the latter case, mean and SD are no longer appropriate parameters to describe the distribution. When a not normal data distribution is present, a log-normal approach is often recommended [11], but this approach does not always solve the distribution problems and/or work appropriately. We recently discussed the difficulty of deriving RCV when the biologic behavior of markers does not meet assumptions needed to apply statistical parametric models recommended to investigate the different sources of variability [12]. In the example of C-reactive protein (CRP), the symmetric distribution of individual data is seldom achieved, even after logarithmic transformation of data, and this could prevent the estimate of variability coefficients [13].

Although generally neglected, an additional relevant issue is the correlation between within-subject serial measurements, likely causing an overestimation of intra-individual variability [14]. To avoid biased estimations, this correlation necessitates a statistical analysis that appropriately accounts for the dependency among measurements within the same experimental unit, resulting in a more accurate and powerful statistical evaluation [15]. Furthermore, the adequacy of sample size is an issue that should be tackled both considering the number of individuals to recruit and the number of serial measurements per subject to perform. A high sample size should assure more precise estimate of intra-individual variability, but a

low number of serial measurements might imply biased estimates [16]. The application of Bootstrap methods might reasonably aid in increasing precision and accuracy of estimates.

In this study, we propose and preliminarily validate a statistical model alternative to the already employed RCV approaches to assess the significance of difference between serial measurements in an individual. We describe the traditional and new statistical models and compare them with each other in a simulation using experimental data from three different analytes [glycated hemoglobin (HbA_{1c}), serum chromogranin A (CgA) and serum CRP].

Materials and methods

Traditional parametric statistical model

Differences in serial test results from an individual may be due to clinical improvement or deterioration, but also to three inherent sources of variation, namely, pre-analytical (CV_p), analytical (CV_A) and CV_1 . In principle, the difference between two serial results from an individual becomes clinically significant if it is greater than the sum of these three sources of variance. Particularly, the minimal significant change between two consecutive determinations is calculated as: $RCV=2^{1/2} \times Z \times (CV_p^2 + CV_A^2 + CV_1^2)^{1/2}$, where $2^{1/2}$ is due to the variation of each of the two measurements with the same CV_1 and Z is the number of SD appropriate to the desired probability of the Gaussian distribution (e.g., 1.96 for $p < 0.05$ and 2.58 for $p < 0.01$) [5]. When preparation of the individual for sample drawing and sample collection, handling and storage prior to analysis are optimized, as they should be by good training of staff and adherence to standard operating procedures, CV_p becomes minimal and the above formula reduces to: $RCV=2^{1/2} \times Z \times (CV_A^2 + CV_1^2)^{1/2}$ [17]. It seems relevant to recognize that the above reported Z-score values are two-sided and should only be used when both a rise and fall are being considered together as a change [18]. If the real clinical requirement is just the evaluation of an increase or decline of test results, then one-sided Z-scores must be used: these are 1.65 for $p < 0.05$ and 2.33 for $p < 0.01$ [4].

New non-parametric statistical model

Let $y_i(t)$ be the measure of an analyte in subject $i=1, \dots, I$ at time $t_j, j=1, \dots, J$. For every subject we compute $J-1$ differences of analyte measurements along time, i.e., $d_{ij}=y_i(t_{j+1})-y_i(t_j)$. The usual sample variance estimator is an unbiased estimator of the measurement differences variance, we obtain an estimate for each intra-individual variance,

$$S_i^2 = \frac{1}{J-2} \sum_{j=1}^{J-1} (d_{ij} - \bar{d})^2, \quad i=1, \dots, I.$$

A rule based on the interquartile range is performed for outlier identification [19]. After discarding the outliers, generally due to subjects with too high intra-measurements variability, we compute $\delta_{1-\alpha}^2$, i.e., the empirical quantile of order $(1-\alpha)$ of this distribution [19]. In

the following, we set $\alpha=0.05$ and the threshold value we consider is $\delta_{0.95} = \sqrt{\delta_{0.95}^2}$.

Let us observe that if: 1) the variance of the first n differences is less or equal to $\delta_{1-\alpha}^2$; and 2) the difference between the following measure of the analyte and the current one is close to the mean difference at most $\delta_{0.95}$ then we can prove that the variability of the measured differences is again less or equal to $\delta_{1-\alpha}^2$.

Assume that $|d_{n+1} - \bar{d}| < \delta_{(1-\alpha)}$, then $\text{Var}(d_1, \dots, d_{n+1}) < \delta_{1-\alpha}^2$. In fact:

$$\begin{aligned} \text{Var}(d_1, \dots, d_{n+1}) &= \frac{1}{n} \sum_{j=1}^{n+1} (d_j - \bar{d})^2 = \frac{1}{n} \left[\sum_{j=1}^n (d_j - \bar{d})^2 + (d_{n+1} - \bar{d})^2 \right] \\ &< \frac{1}{n} [(n-1)\delta_{(1-\alpha)}^2 + \delta_{(1-\alpha)}^2] \end{aligned}$$

Let us observe that the threshold value is an empirical quantile and then it is data and analyte dependent.

Model application

We selected three analytes with different behavior in the distribution of individual data obtained during BV studies and for each of them we applied the two statistical models to derive the first result being significantly lower/upper (P2) when compared with baseline value (P1). In particular, experimental data for HbA_{1c} and CgA were derived from previous BV studies published by our group [20, 21], whereas CRP data were from an unpublished study employing the same samples collected for the CgA study. Detailed information on the methodology and employed protocols is available in the quoted references. CRP was measured on the Cobas 6000 platform using high sensitive latex immunoturbidimetric assay (Roche Diagnostics), with a limit of detection (LoD) of 0.15 mg/L.

We have simulated the practical use of the two statistical models at three different concentrations for each analyte (around the main diagnostic cut-off and at other concentrations relevant for clinical decision making). For convenience, the traditional parametric model and the proposed non-parametric one were called Method 1 and Method 2, respectively. In order to derive P2 we set up the following calculations:

Method 1: $P2 = P1 \pm P1 * RCV/100$,

Method 2: $P2 = P1 + (Pm \pm 2 \delta_{0.95})$, where Pm is the mean of differences between all samples in all subjects.

In addition, to effectively compare results obtained by the two Methods in a well defined and controlled ideal scenario, we performed a further simulation study by considering a hypothetical analyte measured in 30 subjects at time $t_j = j$, for $j=1, 2, \dots, 8$. We generated a random sequence of 30 numbers with a normal distribution ($\mu=50$; $\sigma=1$). Each number represents the mean of normal distribution of each of the eight serial measurements performed in each subject (matrix 30×8). We considered that the variance $[1+0.5*(j-1)]$ could increase from t_1 to t_8 . The simulation was performed $M=1000$ times. In this scenario, σ obtained with Method 1 should be comparable with that obtained by Method 2 (even with increasing variances).

Results

Figure 1 shows the distribution of HbA_{1c}, CgA and CRP values of all samples in recruited individuals. Table 1

shows the results of P2 by the two statistical models for the three analytes studied at different selected concentrations.

Example 1: HbA_{1c}

In 2011, we published a reevaluation of BV components of HbA_{1c} [20]. By recovering data from this study, we evaluated the frequency distribution of HbA_{1c} values characterizing the subject population (Figure 1A). In Figure 2A we have reported for each subject ($n=18$) the concentration of HbA_{1c} of each serial sample ($n=5$). No observations were removed as statistical outliers. In particular, we characterized three groups of subjects with mean HbA_{1c} concentrations around 34, 36 and 38 mmol/mol, but no difference in intra-individual variability at visual inspection. After checking the normality of frequency distribution by Shapiro-Wilk test, we assumed HbA_{1c} as example of an analyte normally distributed ($p=0.76$) among the sample population and compared the behaviour of Method 1 and Method 2 under this condition.

The RCV to apply Method 1 was that derived in the original study, i.e., 9.5% [20]. To apply Method 2, we firstly derived $\delta_{0.95}$ and Pm, obtaining 1.76 and 0.11 mmol/mol, respectively. To better clarify the variability of the differences between serial samples, the density distribution and box plot of differences variance are shown (Figure 2B and C).

The P1 concentrations considered for the simulation were 37, 50 and 70 mmol/mol. At all of those concentrations, the P2 results obtained by Method 1 and Method 2 overlapped (Table 1). As a result, if the analyte concentrations in individuals follow a Gaussian frequency distribution, as in the case of HbA_{1c}, both evaluated methods can be used equally for deriving P2.

Example 2: CgA

More recently, we derived BV data for CgA [21]. In this case, even when one within-subject variance was removed as outlier (Cochran's statistical test), the frequency distribution of obtained CgA values was not normal and further difficult to transform in Gaussian being bimodal (Figure 1B). The profile of each subject ($n=21$), the density distribution and box plot of variability of the differences between serial samples are shown in Figure 3. The RCV to be used in Method 1 was 46% [21]. For Method 2, we obtained a $\delta_{0.95}$ of 11.34 $\mu\text{g/L}$ and a Pm of 1.2 $\mu\text{g/L}$, respectively. Interestingly, while P2 values obtained by two statistical methods resulted quite similar at P1 of 50 $\mu\text{g/L}$

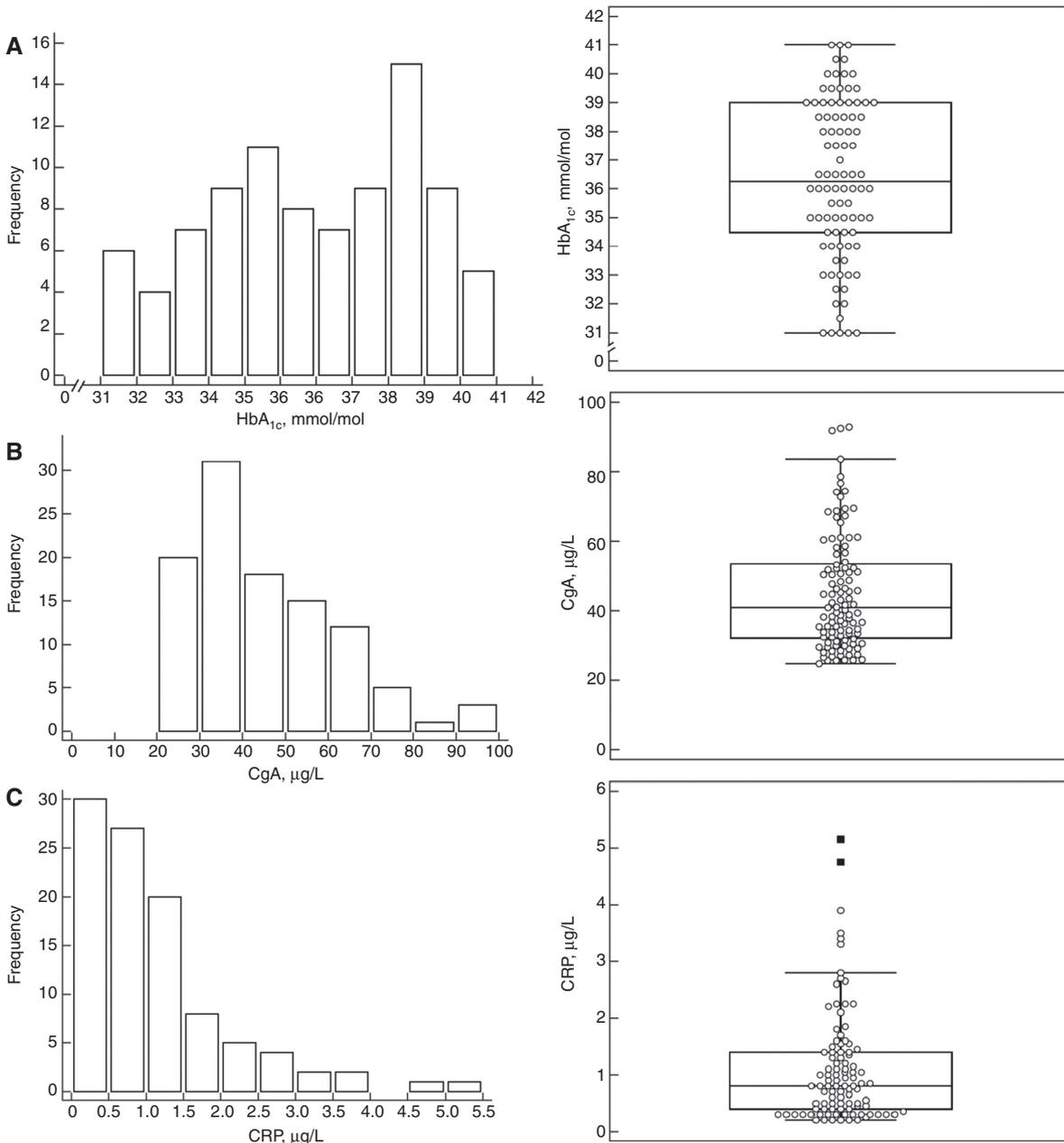


Figure 1 Distributions of glycated hemoglobin (HbA_{1c}), chromogranin A (CgA) and C-reactive protein (CRP) values of all samples from recruited individuals.

For each analyte, a frequency distribution (left) and a box-and-whisker plot (right) were shown. In the box-and-whisker plot, the central box represents the values from the lower to upper quartile (25th–75th percentile). The middle line represents the median. The horizontal line extends from the minimum to the maximum value, excluding ‘outside’ and ‘far out’ values, which are displayed as separate points.

(corresponding approx. to the average concentration of a reference population [22]), for P1 concentrations around the upper reference limit (90 $\mu\text{g/L}$) or frankly abnormal (200 $\mu\text{g/L}$) the P2 estimate significantly differed, indicating a higher sensitivity of the Method 2 for detecting significant changes in CgA values (Table 1).

Example 3: CRP

It is well known that the biologic behavior of CRP does not meet assumptions needed to apply statistical parametric models [12, 13]. Furthermore, for CRP the symmetric distribution of individual data is seldom achieved even after

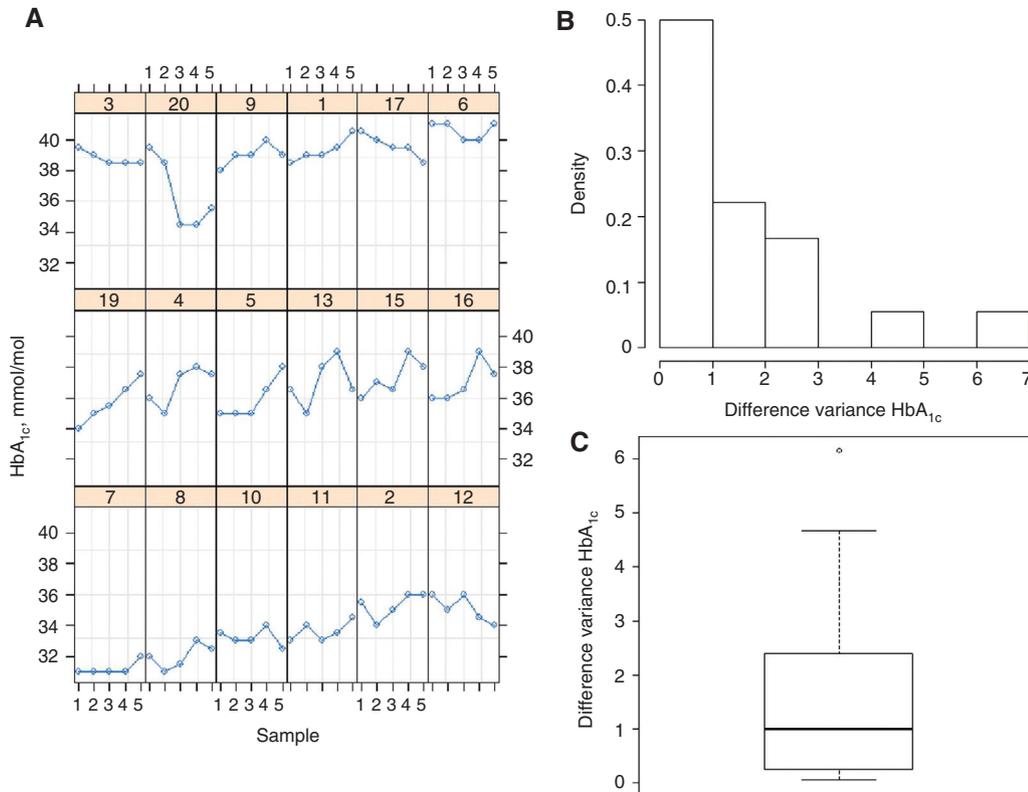


Figure 2 Biological profile of each individual (A), density distribution (B) and box plot (C) of variability of the differences among serial blood samples for glycated hemoglobin (HbA_{1c}). Numerical data taken from Braga et al. [20].

their logarithmic transformation and this may imply an incorrect estimate of variability coefficients due to the persistent non-normal data distribution [13]. Our study involved collection of 110 serum specimens (five from each of 22 apparently healthy volunteers), each assayed

Table 1 First result (P2) being significantly lower/upper when compared with baseline result (P1) obtained by the traditional parametric statistical model (Method 1) and the new non-parametric statistical model (Method 2) for glycated hemoglobin (HbA_{1c}), chromogranin A (CgA) and C-reactive protein (CRP) at three different concentrations.

Analyte	P1	P2	
		Method 1	Method 2
HbA _{1c} , mmol/mol	37	34/41	34/41
	50	45/55	47/54
	70	63/77	67/74
CgA, µg/L	50	27/73	29/74
	90	49/131	69/114
	200	108/292	179/224
CRP, mg/L	3.0	-2.2/8.2	0.5/5.7
	10.0	-7.4/27.4	7.5/12.7
	20.0	-14.8/54.8	17.5/22.7

in duplicate. One case was eliminated as three CRP values out of five were <LoD. Furthermore, one within-subject variance was detected by Cochran’s test as statistical outlier and corresponding data removed. After removing the outlier subject, the Cochran’s test was repeated and no further within-subject variance was found as outlier (Cochran’s test value, 0.265; p>0.01).

Figure 1C shows the distributions of CRP values of all remaining specimens. The profile of each subject (n=20), the density distribution and box plot of variability of the differences between serial samples are displayed in Figure 4. Visual inspection clearly shows huge differences in intra-individual variability of CRP. The Shapiro-Wilk test failed to accept the hypothesis of normality for the distribution of 25% of intra-individual CRP values even after their logarithmic transformation. Consequently, we assumed CRP as example of an analyte not normally distributed among the sample population and compared the behaviour of Method 1 and Method 2 under this condition.

The RCV obtained by Method 1 was 174%. By applying the Method 2, we obtained a $\delta_{0.95}$ of 1.30 mg/L and a Pm of 0.06 mg/L, respectively. The three CRP concentrations considered for simulation were 3, 10 and 20 mg/L

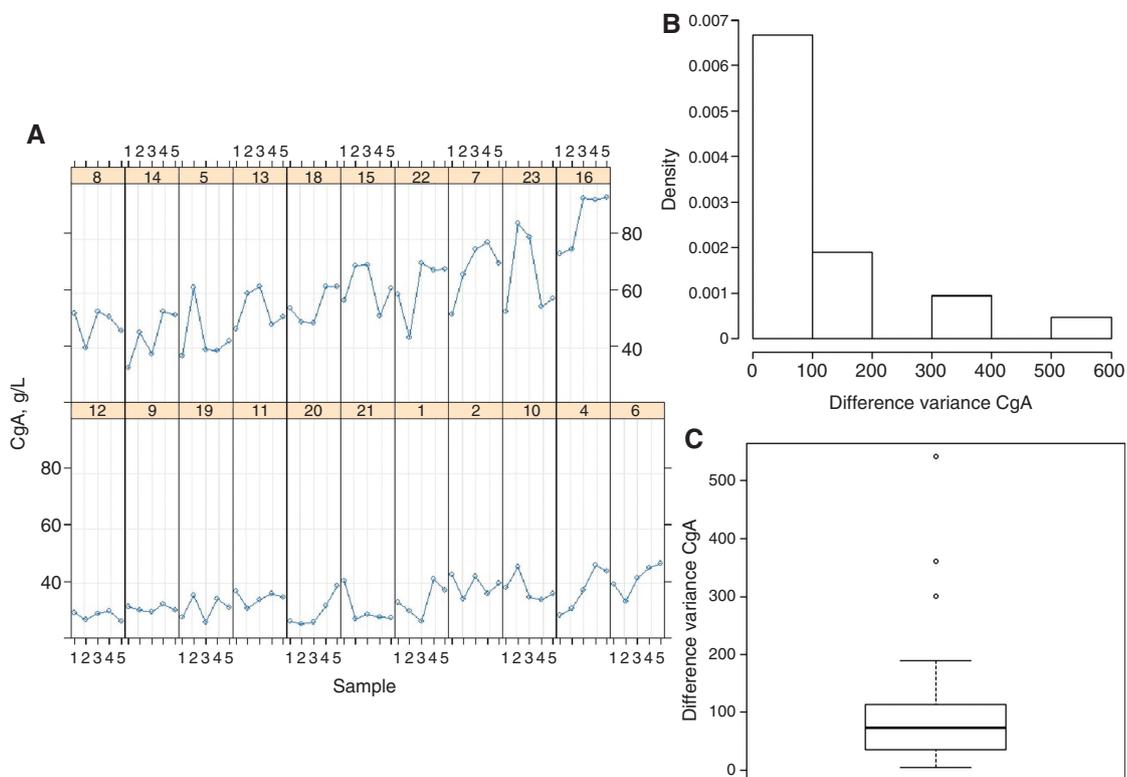


Figure 3 Biological profile of each individual (A), density distribution (B) and box plot (C) of variability of the differences among serial serum samples for chromogranin A (CgA). Numerical data taken from Braga et al. [21].

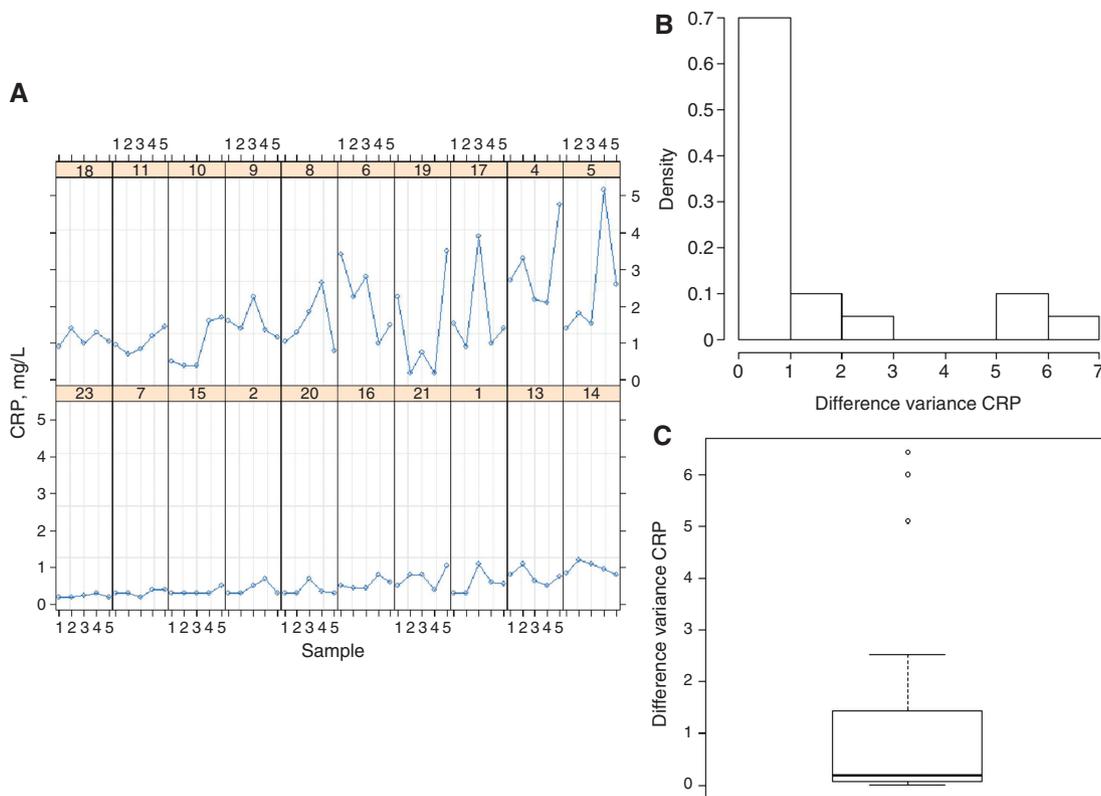


Figure 4 Biological profile of each individual (A), density distribution (B) and box plot (C) of variability of the differences among serial serum samples for C-reactive protein (CRP).

(Table 1). As expected, P2 values derived from Method 1 were unreliable and likely to be clinically impractical.

Data simulation study

According to the data simulation study, we obtained a σ of 0.49 and 0.66 for Method 1 and Method 2, respectively. Since the estimated σ are not statistically different in a context of normal distribution and increasing variance, we are rather confident on the statistical robustness of Method 2 with respect to the Method 1.

Discussion

For most analytes CV_1 is much smaller than CV_G , hindering the RI use for the result interpretation. In these cases, RCV is considered the best tool for interpreting laboratory tests. Arguments supporting the use of RCV are: 1) clinicians need to objectively differentiate pathophysiological changes in test results from analytical/BV; 2) RCV information is available for ~300 analytes; and 3) RCV is also valuable for clinical validation of laboratory results [23]. The traditional approach to derive RCV requires, however, a number of assumptions including: 1) the random variation follows a Gaussian distribution; 2) pre-analytical variation is negligible; 3) analytical variation is constant and independent of the analyte concentration; and 4) BV is constant and independent of other biological characteristics, e.g., age, fertility status, etc. [24]. These assumptions are not always met, especially that about the normal distribution of data. Indeed, many analytes of current interest in laboratory medicine appear to have a variation in their concentrations over time in individuals that is not normally distributed. The simple log-transformation, although it may be sometimes appropriate to solve problems related to a non-Gaussian distribution of values, may create confusion in interpreting the obtained RCV as, for individual patient care, the use of transformed results is impractical [13].

Fokkema et al. [11], demonstrating that the distributions of brain natriuretic peptide values in healthy individuals are log-normal rather than normal, adopted a rather different approach to calculate RCV. Although this has been used by other investigators to estimate RCV in the setting of not normally distributed data [25], it is not always applicable for a number of drawbacks. First, the log-transformation tends to contain the variability, with the risk of underestimating it. Furthermore, not always this approach solves the distribution problems and/

or works. For instance, for CRP the symmetric distribution of individual data is seldom achieved, even after logarithmic transformation. More recently, deGoma et al. have employed an alternative model for calculating RCV in the case of CRP [26]. Although their methodological approach is interesting in trying to overcome limitation of parametric protocols, some important pitfalls about, among others, sample analysis, assessment of analytical variability, study duration and sampling frequency have been highlighted [12]. With regard to the statistical analysis of data, these authors applied a linear mixed effects model for longitudinal data, previously adopted by Glynn et al. in the study of intra-class correlation coefficient (ICC) of CRP [27]. Once again, the log-transformation of CRP results, which is instrumental to employ the ICC estimate, may not assure the normalization of the distribution. Moreover, when the number of observations for each subject is very low and when the variability within individuals looks very different, suggesting the existence of two or more subpopulations, the usual hypotheses of linear and non-linear mixed effects models are questionable and in general not satisfied.

In this study, we proposed a new non-parametric statistical model for interpretation of differences in serial test results from an individual. In addition to the distribution problem, our approach also overcomes the correlation between within-subject serial measurements, which may cause an overestimation of intra-individual variability, allowing for the first time to establish reliable interpretative criteria for assessment of results of biologically complex analytes, such as CRP, and potentially contributing to set aside previously raised perplexities about their clinical utility [28]. In particular, we validated this new statistical model by comparing it with the traditional parametric one in three settings using analytes displaying different data distributions. When analyte concentrations in individuals clearly followed a Gaussian distribution, as in the case of HbA_{1c} , the two statistical approaches were interchangeable. On the contrary, if the analyte presented a bimodal (as CgA) or skewed (as CRP) distribution, the interpretation obtained by the two models was different. In these cases, our approach based on non-parametric statistics seems to be more suitable in assessing if the last laboratory result is significantly different from the former.

The weakness of the approach is that the CV_A is not separately considered in the calculations. Thus, the presented data can be generalized only to those laboratories working with similar analytical imprecision. Furthermore, a validation of the new proposed model in specific clinical contexts is required before its final application.

Acknowledgments: The skilful technical assistance of Fatima Calderaro (Luigi Sacco University Hospital, Milano, Italy) is gratefully acknowledged.

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

Financial support: None declared.

Employment or leadership: None declared.

Honorarium: None declared.

Competing interests: The funding organization(s) played no role in the study design; in the collection, analysis, and interpretation of data; in the writing of the report; or in the decision to submit the report for publication.

References

- Cerioti F, Hinzmann R, Panteghini M. Reference intervals: the way forward. *Ann Clin Biochem* 2009;46:8–17.
- Harris EK. Effects of intra- and inter-individual variation on the appropriate use of normal range. *Clin Chem* 1974;20:1535–42.
- Fraser CG, Harris EK. Generation and application of data on biological variation in clinical chemistry. *Crit Rev Clin Lab Sci* 1989;27:409–37.
- Fraser CG. Making better use of differences in serial laboratory results. *Ann Clin Biochem* 2012;49:1–3.
- Fraser CG. Reference change values. *Clin Chem Lab Med* 2012;50:807–12.
- Fraser CG. Biological variation: from principles to practice. Washington, DC: AACC Press, 2001.
- Harris EK, Yasaka T. On the calculation of a “reference change” for comparing two consecutive measurements. *Clin Chem* 1983;29:25–30.
- Ferraro S, Panteghini M. Is serum human epididymis protein 4 ready for prime time? *Ann Clin Biochem* 2014;51:128–36.
- Fraser CG, Stevenson HP, Kennedy IM. Biological variation data are necessary prerequisites for objective autoverification of clinical laboratory data. *Accred Qual Assur* 2002;7:455–60.
- Franzini C. Need for correct estimates of biological variation: the example of C-reactive protein. *Clin Chem Lab Med* 1998;36:131–2.
- Fokkema MR, Herrmann Z, Muskiet FA, Moecks J. Reference change values for brain natriuretic peptides revisited. *Clin Chem* 2006;52:1602–3.
- Braga F, Ferraro S, Lanzoni M, Szoke D, Panteghini M. Estimate of intraindividual variability of C-reactive protein: a challenging issue. *Clin Chim Acta* 2013;419:85–6.
- Braga F, Panteghini M. Biologic variability of C-reactive protein: is the available information reliable? *Clin Chim Acta* 2012;413:1179–83.
- Sullivan LM. Repeated measures. *Circulation* 2008;117:1238–43.
- Peters JL, Mengersen KL. Meta-analysis of repeated measures study designs. *J Eval Clin Pract* 2008;14:941–50.
- Røraas T, Petersen H, Sandberg S. Confidence intervals and power calculations for within-person biological variation: effect of analytical imprecision, number of replicates, number of samples and number of individual. *Clin Chem* 2012;58:1306–13.
- Fraser CG. Reference change values: the way forward in monitoring. *Ann Clin Biochem* 2009;46:264–5.
- Garner AE, Lewington AJ, Barth JH. Detection of patients with acute kidney injury by the clinical laboratory using rises in serum creatinine: comparison of proposed definitions and a laboratory delta check. *Ann Clin Biochem* 2012;49:59–62.
- Bland JM. An introduction to medical statistics, 3rd ed. Oxford: Oxford University, 2000.
- Braga F, Dolci A, Montagnana M, Pagani F, Paleari R, Guidi GC, et al. Reevaluation of biological variation of glycated hemoglobin (HbA1c) using an accurately designed protocol and an assay traceable to the IFCC reference system. *Clin Chim Acta* 2011;412:1412–6.
- 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.
- Ferraro S, Mozzi R, Michelazzo C, Basco D, Panteghini M. Reference intervals for the new Brahms chromogranin A (CGA) assay. *Biochim Clin* 2013;37:S169.
- Ricós C, Cava F, García-Lario JV, Hernández A, Iglesias N, Jiménez CV, et al. The reference change value: a proposal to interpret laboratory reports in serial testing based on biological variation. *Scand J Clin Lab Invest* 2004;64:175–84.
- Cooper G, DeJonge N, Ehrmeyer S, Yundt-Pacheco J, Jansen R, Ricós C, et al. Collective opinion paper on findings of the 2010 convocation of experts on laboratory quality. *Clin Chem Lab Med* 2011;49:793–802.
- Wu AH, Lu QA, Todd J, Moecks J, Wians F. Short- and long-term biological variation in cardiac troponin I measured with a high-sensitivity assay: implications for clinical practice. *Clin Chem* 2009;55:52–8.
- deGoma EM, French B, Dunbar RL, Allison MA, Mohler III ER, Budoff MJ. Intraindividual variability of C-reactive protein: the multi-ethnic study of atherosclerosis. *Atherosclerosis* 2012;224:274–9.
- Glynn RD, MacFadyen JG, Ridker PM. Tracking of high-sensitivity C-reactive protein after an initially elevated concentration: the JUPITER Study. *Clin Chem* 2009;55:305–12.
- Campbell B, Badrick T, Flatman R, Kanowsli D. Limited clinical utility of high-sensitivity plasma C-reactive protein assays. *Ann Clin Biochem* 2002;39:85–8.