Opinion Paper


Rationale for using data on biological variation

DOI 10.1515/cclm-2014-1142
Received November 21, 2014; accepted March 31, 2015

Abstract: The aims of this study are: 1) to use the data included in the biological variation (BV) database to address the usability of BV estimates; and 2) to use different examples from the authors’ laboratories to illustrate the use and the usefulness of BV data in laboratory medicine. The BV database is an essential tool for laboratory management. Examples of application of data derived from BV are given in this paper, such as analytical performance specifications that have been included in various quality control software designed to optimize operative rules; also they have been incorporated as acceptability limits in external quality assurance reports. BV data from pathological status are of utmost interest for monitoring patients and differences between the intra-individual coefficients of variation (CVi) estimated from healthy and patients are shown. However, for a number of analytes there are limited data available and for many there are no data, consequently new studies should be encouraged at an international level. In addition, developing international criteria to evaluate publications dealing with the estimation of BV components would be of the utmost interest. We are ready and willing to collaborate with such worthy initiatives. The first EFLM strategic conference on analytical performance specifications is an excellent opportunity for debating these ideas.

Keywords: biological variation; patient safety; quality assurance; quality specification.

Introduction

Laboratory medicine is the science focused on the knowledge of health status assessment by testing the body fluids composition. The components of human samples are variable “per se” and one of the sources of variation...
is biological. This biological variation (BV) represents the fluctuation of the concentration of constituents around their homeostatic set point. It is called “within-subject biological variation”. Differences between individual set-points are named “between-subject biological variation”.

Both components of BV are usually expressed in terms of coefficients of variation: between-subject (CVG) and within-subject (CVI), respectively [1, 2].

The Analytical Quality Commission of the Spanish Society of Clinical Biochemistry and Molecular Pathology (SEQC) created a BV database in 1998 by compiling available papers in scientific journals. This database is updated every 2 years with new published articles [3, 4].

The 2014 edition includes the data of 358 analytes: CVI and CVG estimates, number of subjects studied, age, sex, number of samples per subject, time length of the study, mean value obtained, analytical imprecision (CVG), analytical procedure and mathematical model used to estimate the components of BV, type of population (healthy volunteers or patients), name of first author and bibliographic reference of the study.

The purposes of this work are: 1) to use the data included in the BV database to address the transferability of BV estimates; and 2) to use different examples from the authors’ laboratories to illustrate the use and the usefulness of BV data in laboratory medicine.

Materials and methods

The material used in this work is:

- The last edition of our BV database (December 2013).
- The information contained in the database concerns CVI and CVG values, estimated for a number of analytes in healthy subjects [4].
- A second database focused on non-healthy situations, also updated at December 2013.
- A report of the SEQC external quality assurance (EQA) program.
- Results for serum creatinine obtained in a study performed with Dutch Foundation for Quality Assessment in Medical Laboratories (SKML), the Dutch EQA organizer, which uses a set of commutable controls with values assigned by certified reference methods.

The method consists of description of:

- How articles are searched for when the BV database is updated.
- The selection criteria used prior to inclusion, explained in a recent paper [5], that includes:
  - Paper has to be specifically designed to estimate the components of BV.
  - Mathematical model for estimating the components of variance (CVG, CVG, CVG) has to be based on ANOVA test.
  - CVG has to be no higher than 0.5*CVI.
  - Examples of applications of BV-derived data in laboratories of our setting.

Results

Data included in the databases

The BV components have been estimated from healthy individuals or patients suffering diverse pathologies and, therefore are separated in two different BV databases. These databases have been updated every 2 years since 1999, by searching at Pubmed website using biological variation as the keyword.

The number of papers rejected has evolved over time, at the beginning the main reasons for rejection were not using the mathematical ANOVA test and showing a CVG excessively high compared with CVI. In recent years, because an increasing number of papers dealing with non-healthy situations were published, the main reason for rejection was a non-specific design of the study to determine the components of BV.

As was already seen by Fraser [6], CVI from young people, adults and the elderly are quite similar. Children constitute a specific group; in a recent study it has been seen that CVI values of the pediatric population seems to be different compared to those from adults, being higher for ceruloplasmin (11.3% vs. 5.8%) and glucose (11.4% vs. 5.6%) and lower for C-reactive protein (CRP) (19% vs. 42%) and γ-glutamyltransferase (GGT) (2.7% vs. 13%) [7].

Table 1 shows an extract of information contained in the database for serum glucose. The highest CVI does not correspond to the larger number of subjects studied (n=1105), neither to the longer study (365 days) nor to the larger number of samples taken per subject (Ss=12). Moreover, the wide variation in CVI shown in this table cannot be related to the year of publication of the study, thus results seem to be independent of the evolution of technology for glucose testing, but might be due to different methods (e.g., outlier exclusion, homogeneity etc.) used to treat the data and calculate the CVI.

It should be emphasized that studies obtaining samples within a day, show the lowest CVI values. For example, serum cholesterol the lowest CVI=1.5%
corresponds to a paper that takes samples every 8 h within a day, whereas the median CV$_i$ is 6.1%. Furthermore, serum cardiac troponin I shows the lowest CV$_I$ = 3.4% for a frequency of sampling of 4 h, whereas the median CV$_i$ is 12.9%. For this reason values from 1-day studies have been excluded from the final estimations of CV$_I$ and CV$_G$ in the database from healthy subjects.

Our second database deals with patients diagnosed with different pathologies. The first publication in 2007 compiled 66 analytes and included 34 disease states [13], whereas the 2014 update contains 97 analytes and 41 diseases, covering many types of pathologies. For the majority of analytes, CV$_I$ values from patients seem to be similar to those from healthy individuals, as was already seen by Fraser [14]. An example extracted from our database is CV$_i$ for albumin in patients with diabetes mellitus type I, chronic renal impairment and chronic hepatopathy are 2.8%, 2.9% and 3.3%, respectively, whereas in healthy people CV$_i$ is 3.1%.

However, in some pathologies, the CV$_i$ of the analytes seems to be higher than the CV$_i$ of healthy subjects, as shown in Table 2. This should be taken into account when applying the reference change value (RCV) to consecutive results of an analyte, either for automatic verification of results [26] or for reporting significant changes in a patient’s health status [27].

### Use of databases in laboratory medicine

The use of BV data from healthy people covers many purposes, the most widely accepted being to derive quality specifications for analytical imprecision, bias (B$_A$) and total error [28, 29]. For monitoring purposes analytical imprecision has to be maintained below 0.5*CV$_i$. In this situation the contribution of laboratory error to the total variation is calculated as 12% [30]. For diagnosis, case finding and screening purposes, B$_A$ has to be maintained below 0.25*(CV$_i$+CV$_G$) in order to share population-based

---

**Table 1: BV database. Extract of serum glucose.**

<table>
<thead>
<tr>
<th>CV$_i$</th>
<th>CV$_o$</th>
<th>CV$_A$</th>
<th>N</th>
<th>T$_{days}$</th>
<th>S$_s$</th>
<th>Mean</th>
<th>Year [references]</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.2</td>
<td>11</td>
<td>2.4</td>
<td>40</td>
<td>28</td>
<td>3</td>
<td>5.5</td>
<td>1994 [8]</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>4.5</td>
<td>5.8</td>
<td>1.4</td>
<td>15</td>
<td>70</td>
<td>10</td>
<td>5.5</td>
<td>1988 [9]</td>
</tr>
<tr>
<td>4.7</td>
<td>5.4</td>
<td>2.4</td>
<td>27</td>
<td>140</td>
<td>10</td>
<td>5.2</td>
<td>1989 [6]</td>
</tr>
<tr>
<td>5.0</td>
<td>7.7</td>
<td>3.4</td>
<td>20</td>
<td>365</td>
<td>12</td>
<td>5.2</td>
<td>1989 [10]</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>6.5</td>
<td>8.7</td>
<td>2.2</td>
<td>1105</td>
<td>60</td>
<td>9</td>
<td>4.8</td>
<td>1978 [11]</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>13.1</td>
<td>3.2</td>
<td>3.0</td>
<td>10</td>
<td>5</td>
<td>5</td>
<td>4.8</td>
<td>1993 [12]</td>
</tr>
</tbody>
</table>

Each row show the values reported in one article. CV$_i$, within-subject coefficient of variation; CV$_o$, between-subject coefficient of variation; CV$_A$, analytical coefficient of variation. All of them are expressed in percentages; N, number of subjects studied; S$_s$, number of samples obtained per subject; T, length of the study, expressed in days. Dot-lines: rows existing in the database but not shown here. The total number of rows in the 2014 database for serum glucose is 20.

**Table 2: BV from patients. Differences in CV$_i$ between patients and healthy.**

<table>
<thead>
<tr>
<th>Pathology</th>
<th>Analyte</th>
<th>CV$_i$ healthy</th>
<th>CV$_i$ patients [references]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Breast carcinoma</td>
<td>Alkaline phosphatase</td>
<td>6.4</td>
<td>17.3 [15]</td>
</tr>
<tr>
<td></td>
<td>Calcium</td>
<td>1.9</td>
<td>6.5 [15]</td>
</tr>
<tr>
<td></td>
<td>CA 15.3</td>
<td>6.2</td>
<td>17.3 [16]</td>
</tr>
<tr>
<td></td>
<td>CEA</td>
<td>13</td>
<td>26.9 [16]</td>
</tr>
<tr>
<td></td>
<td>Osteocalcin</td>
<td>5.5</td>
<td>43 [15]</td>
</tr>
<tr>
<td>Cirrhosis Hepatocellular carcinoma</td>
<td>α-Fetoprotein</td>
<td>12</td>
<td>40 [17]</td>
</tr>
<tr>
<td>Chronic liver disease</td>
<td>α-Glutamyltransferase</td>
<td>14</td>
<td>4.7 [18]</td>
</tr>
<tr>
<td>Chronic renal disease in children</td>
<td>Creatinine</td>
<td>4.3</td>
<td>8.9 [19]</td>
</tr>
<tr>
<td>Diabetes mellitus type I</td>
<td>Glucose</td>
<td>5.7</td>
<td>30 [9]</td>
</tr>
<tr>
<td></td>
<td>HbA$_1c$</td>
<td>1.9</td>
<td>8.8 [20]</td>
</tr>
<tr>
<td></td>
<td>Lipoprotein a</td>
<td>8.5</td>
<td>26 [21]</td>
</tr>
<tr>
<td></td>
<td>Microalbumin</td>
<td>36</td>
<td>61 [22]</td>
</tr>
<tr>
<td>Hepatitis B</td>
<td>α2-Macroglobulin</td>
<td>3.4</td>
<td>7.6 [23]</td>
</tr>
<tr>
<td>Lung carcinoma</td>
<td>CA 19.9</td>
<td>16</td>
<td>24.5 [24]</td>
</tr>
<tr>
<td></td>
<td>CEA</td>
<td>13</td>
<td>23.6 [24]</td>
</tr>
<tr>
<td>Ovarian carcinoma</td>
<td>CA 19.9</td>
<td>16</td>
<td>24.5 [24]</td>
</tr>
<tr>
<td>Paget’s disease</td>
<td>Alkaline phosphatase</td>
<td>6.4</td>
<td>12.4 [25]</td>
</tr>
</tbody>
</table>

CV$_i$ values for healthy subjects are those appearing in the 2014 database [4]. CV$_i$ values for patients have been obtained from different papers given in the references.
reference intervals [31]. These are the most common specifications which are known as “desirable” specifications. Other levels of quality namely “minimum” (more permissive) and “optimum” (more restrictive) have also been suggested [32].

The 2014 updated database shows these three levels of quality specifications for 358 analytes [4].

BV data from healthy subjects are also important to design internal quality control procedures for the analytical phase, which should include four steps:
- To define total error allowable (TEA);
- To measure analytical imprecision (CVA) and BA;
- To calculate $\sigma$ metrics: $\sigma = (\text{TEA} - \text{BA})/\text{CVA}$;
- To search for an appropriate control rule.

Selection of TEA is a key point, because it will result in the use of a permissive or restrictive rule for accepting analytical runs. When applying the Stockholm [28] and Milan [33] hierarchical criteria for defining quality specifications, if a “top criterion” is applied, $\sigma$ could be small and this could lead to a restrictive operative rule being used. On the contrary, if criterion at the bottom of the hierarchy is used, $\sigma$ will usually be larger and a more relaxed operative rule would be appropriate. For example, creatinine measured in one of our laboratories using Jaffe method has $\text{CVA} = 2.2\%$ and $\text{BA} = 4.3\%$. If the quality specification for total error is based on BV, TEA = 8.9%, $\sigma = 2.1$ and a multi-rule with several controls per run should be used. If the specification is based on the state of the art, TEA = 20%, $\sigma = 7.1$ and a simple rule with two controls per run would be used. TEA = 20% is the minimum level of quality defined by consensus of four Spanish scientific EQA organizers, based on the results obtained by the participants in the four programs [34]; this value is similar to other organizations, such as Clinical Laboratory Improvement Amendments (CLIA) [35] and RiliBÄK [36]. It seems to be clear that defining analytical quality specifications could have an important impact on patient safety.

Another application of BV data from healthy people is to interpret the EQA reports. Figure 1 illustrates an example for serum creatinine in a report of the SEQC-EQA program. When a laboratory has a percentage deviation versus the peer-group mean higher than the BV-derived acceptance limit, this could initiate an alert for a corrective action.

**Figure 1:** SEQC-EQA monthly report for serum creatinine for a laboratory. Left side figure: Frequency histogram of results for a control sample. Left side table: total and accepted number of results from all laboratories, method-group and peer-group (same method and instrument) of laboratories, as well as mean and standard deviation for these three groups. Right side figure: last 12 results of the individual laboratory compared with the standard deviation index related to the peer-group mean (horizontal dot-lines) and the limits derived from BV for total error (gray shadows). Right side table: laboratory result in SI and conventional units, its deviation related to the peer-group mean, expressed in standard deviation index and in percentage. In bold: desirable deviation for total error derived from BV for creatinine. SEQC-EQA uses stabilized (non-commutable) control sera, targeted by the peer-group mean.
Although some colleagues believe that reaching the BV-derived specifications may be extremely difficult, Figure 2 illustrates that 80%–100% of the majority of analyte results included in the SEQC-EQA serum biochemistry scheme of 2013 achieved the goal; even when considering the most “difficult” analyte (sodium), 55% of the results were also within the BV limits.

From another perspective, BV data are also useful to assess whether there is any current available method providing laboratory information that could potentially compromise patient safety. Figure 3 shows serum creatinine results obtained by 23 Spanish laboratories participating in a SKML pilot study using commutable controls with values assigned by reference methods. The samples are targeted with reference methods, undertaken in either the Joint Committee for Traceability in Laboratory Medicine (JCTLM) listed reference laboratories or in International Federation of Clinical Chemistry and Laboratory Medicine (IFCC) network laboratories [37]. All laboratories using the alkaline picrate kinetic (Alk picr k) (method gave unacceptably high results at the clinical decision level, while various laboratories using the compensated method produced dispersed results. Only the enzymatic method gave all results within the limits derived from BV for total error [38]. It has to be emphasized that this can be firmly stated because this EQA uses commutable samples that are targeted with reference methods.

BV can also be applied to see if a change in serial results can be explained by within-subject BV and analytical variation only. In this case, the RCV has to be used for a change between two [2, 39, 40] or more [41] consecutive results. The formula for a change between two results is: \[ RCV = z \times 2^{1/2} (CVA^2 + CVI^2)^{1/2} \] where \( z \) is the probability that a change between the two results can be explained by biological and analytical variation only. The probability for detecting changes increases when considering two analytes together. An example can be seen in Figure 4 where RCV of various renal post-transplanted patients for creatinine and urate combined reveals significant

**Figure 2:** Percentage of results satisfying BV limits for TE<sub>c</sub>.
SEQC-EQA report. This figure corresponds to the 2013 serum basic biochemistry program, which had 888 participating laboratories. X-axis: analytes included in the serum biochemistry program with quality specifications for total error allowable (TE<sub>c</sub>) derived from BV are published [4]. Alb, albumin; ALP, alkaline phosphatase; ALT, alanin-aminotransferase; Amy, α-amylase; AST, aspartate-aminotransferase; bil, bilirubin; Cal, calcium; Cho, cholesterol; Cl, chloride; Cre, creatinine; Fe, iron; GGT, γ-glutamyltransferase; Glu, glucose; HDLc, HDL-cholesterol; K, potassium; LDH, lactate dehydrogenase; Li, lithium; Mg, magnesium; Na, sodium; Osm, osmolality; P, phosphate; Pro, total protein; Tri, triglycerides; Ura, urate; Ure, urea; Y-axis: percentage of results reaching the quality specifications for TE<sub>c</sub>.

**Figure 3:** Serum creatinine.
Percentage deviation to reference-method value. X-axis: serum creatinine concentration, expressed in mmol/L, of 11 control samples. Y-axis: percentage deviation of results with respect to the reference-method value. Methods used by participating laboratories: Alk picr k, alkaline picrate kinetic (Jaffé); Alk picr co, alkaline picrate kinetic compensated; Enz, enzymatic method.
changes in patients suffering complications, whereas no changes are seen for patients with no complications [42].

**Discussion**

Quality specifications based on BV are widely used and have been included in various quality control software designed to optimize operative rules; also they have been incorporated in various EQA reports. Additionally, the use of RCV allows us to detect changes in patients’ health status. Regarding all these purposes, the BV databases are an essential tool for healthcare.

The main advantages of our BV databases are:

- Using agreed criteria for accepting papers.
- Updating the database every 2 years, since the first compilation of 1999.

The weaknesses are:

- There is no data for many analytes and limited data for a great number of them (27 out of the 358 have more than 10 publications, 129 analytes have between 2 and 9 publications and 202 only 1 paper).
- Lack of confidence intervals for the derived parameters, in order to better interpret the wide dispersion of CVs observed for certain analytes.
- Derived quality specifications are too restrictive for some analytes as compared with current technological capability (s-sodium, albumin, chloride and blood HbA1c).

Another criticism recently published is the need for more stringent criteria for selection of papers and for defining an international standard for performing and reporting studies on BV [43, 44].

**Conclusions**

There are some ways to improve the BV database, than can be summarized as:

- Knowing the difficulty that studying BV represent for laboratories, international guidelines should help or promote other initiatives including analysis based on data mining processes.

- Developing an international criterion to select the more reliable publications dealing with the estimation of BV components would be of utmost interest.

- We are ready and willing to collaborate with such worthy initiatives. After the first European Federation of Laboratory Medicine (EFLM) strategic conference [33], EFLM has generated a task and finish group (TFG) in which SECQ participate aiming to improve the evaluation of the papers on BV and to generate a new database.

**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


41. Lund F, Petersen PH, Fraser CG, Soléïtromos G. Calculation of limits for significant bidirectional changes in two or more serial biomarker results based on a computer simulation model. Ann Clin Biochem 2014 Sep 23 pii: 0004563214555163.

