A note on the calculation of reference change values for two consecutive normally distributed laboratory results

Citation for published version (APA):

DOI:
10.1016/j.chemolab.2017.10.008

Document status and date:
Published: 15/12/2017

Document Version:
Publisher's PDF, also known as Version of Record (includes final page, issue and volume numbers)

Please check the document version of this publication:
• A submitted manuscript is the author's version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
• The final author version and the galley proof are versions of the publication after peer review.
• The final published version features the final layout of the paper including the volume, issue and page numbers.

Link to publication

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

• Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
• You may not further distribute the material or use it for any profit-making activity or commercial gain
• You may freely distribute the URL identifying the publication in the public portal.

Take down policy
If you believe that this document breaches copyright please contact us:
openaccess@tue.nl
providing details. We will immediately remove access to the work pending the investigation of your claim.
A note on the calculation of reference change values for two consecutive normally distributed laboratory results

M. Regis a,*, Th.A. Postma b, E.R. van den Heuvel a

a Eindhoven University of Technology, Department of Mathematics and Computer Science, P.O. Box 513, 5600 MB Eindhoven, The Netherlands
b Gino Software, Stationsweg 3D, 9726 AC Groningen, The Netherlands

ARTICLE INFO
Keywords:
Monitoring
Change-point detection
Heterogeneity
Heteroscedasticity
Critical difference
Reference change values

ABSTRACT
Population reference limits are inadequate for personalized analyses of medical laboratory results. Reference change values have been recommended as a valid alternative in assessing individual changes across sequential measurements. In this paper, we investigate the accuracy (type I error) and power (complement of type II error) of reference change values under three different statistical modeling scenarios and show that oversimplified hypotheses lead to misinterpretation of laboratory results. The power is strongly affected by the statistical modeling assumptions: it is shown that positive shifts in the individual average health condition are difficult to detect, while it is much easier to identify negative shifts.

1. Introduction

Medical laboratory results are traditionally compared with normal reference limits, i.e. ranges of values that are expected for healthy persons. They are typically defined by the lower and upper 5% quantiles of a reference group, i.e. subjects for which no morbidity is assessed [7]. These population-based reference ranges are mere cut-off values and can lead to false positives and false negatives. The classification of a normal measurement does not guarantee that the value is normal for the specific patient and alternatively an abnormal measurement does not necessarily imply disease alert, in particular when the value is close to the critical threshold. The reason is that measurements in individuals are affected by true condition’s shifts, but also by some other inherent causes, such as pre-analytical, analytical, between- and within-subject biological variations [5]. Population-based reference ranges do not separate these sources of variation.

The considerations above have led to the development of alternative methods for the interpretation of medical laboratory results, in particular recently with an increased interest in personalized medicine. Modern methods aim at understanding changes in individuals, rather than comparing results with respect to population-based references. Reference change values (RCV) or critical differences are a popular method for the assessment of laboratory results, introduced by Ref. [6]. Several manuscripts can be found in this field, for example [3,9–11]. The approach defines criteria for normal variation in two sequential measurements, which mathematically are often defined as

$$\text{RCV} = \pm z_{\alpha} \sqrt{\sigma_A^2 + \sigma_I^2};$$

(1)

with $\sigma_A^2$ the analytical variability, $\sigma_I^2$ the intra-individual variability, and $z_{\alpha}$ a quantile from the standard normal distribution. Typical values for $z_{0.025}$ are 1.645 for $\alpha = 0.05$, 1.960 for $\alpha = 0.025$ or even 2.58 for $\alpha = 0.005$. A good review on the subject is given by Ref. [4].

The calculation of RCVs for many different laboratory outcomes is quite simple since all laboratories know approximately the analytical and intra-individual variabilities for medical laboratory results from samples from the healthy population. A possible disadvantage of the RCVs in (1) is that the analytical variation is assumed to be independent of the health status, an assumption which might be incorrect, as raised by Refs. [8] and [1]. To overcome this [8], proposed a modification of the traditional RCV to an RCV that changes with the level of the quantity it tries to measure, assuming a constant coefficient of variation. However, the work of [8] lacks mathematical rigor, does not specify the statistical assumptions, and does not discuss the consequences of the simplified hypotheses on the detection of possible shifts in the health status between two sequential measurements.

In this paper, we give a rigorous formulation of the reference change values under various statistical hypotheses, and demonstrate that the result of [8] is a special case of our general framework. We also show that...
the power to detect true health changes is counterintuitive when the analytical variability depends on the health state. In Section 2 we introduce the modeling framework and provide a general definition of the reference change values. We consider three different scenarios, corresponding to diverse dependencies of the variances. In Section 3 we discuss how the model can be generalized to allow for a real non-physiological change in the health status and study the power of reference change values in detecting real changes. We end the paper with a discussion.

2. Reference change values

Let \( Y_i \) be a measured medical laboratory quantity for subject \( i = 1, 2, \ldots, n \) at two consecutive moments in time \( j = 1, 2 \). We will assume, possibly after a suitable mathematical transformation (e.g. logarithm), that these laboratory results can be described by an additive structure such as

\[
Y_i = \mu_i + V_i + E_i, \tag{2}
\]

where \( \mu_i \) is the true mean of subject \( i \), \( V_i \) is the random intra-subject variability and \( E_i \) is the random analytical variability. The subject-specific mean \( \mu_i \) should be considered random for population reference, normally distributed with zero mean and variance \( \tau_i^2 \), conditionally on \( V_i \). Furthermore, the two measurements are assumed to be distant enough in time, such that \( (V_i, E_i) \) can be assumed independent of \( (V_1, E_1) \).

To include variability that depends on the true health condition, we will assume that the variances \( \sigma_i^2 \) and \( \tau_i^2 \) may depend on the true health values \( \mu_i \) and \( X_i \),

\[
\sigma_i^2 = \sigma^2(\mu_i), \quad \tau_i^2 = \tau^2(\mu_i, X_i) \tag{4}
\]

Note that the intra-subject variability may depend on the mean health state only, but the analytical variability may also be affected by the levels of \( V_i \). This general formulation implies that the difference \( Y_2 - Y_1 \) is no longer normally distributed if \( V \) depends on \( V \). The cumulative distribution function of the difference \( Y_2 - Y_1 \) is

\[
P(Y_2 - Y_1 \leq y) = E \left[ \Phi \left( \frac{y - (Y_2 - Y_1)}{\sqrt{\sigma^2(\mu_i, V_i) + \tau^2(\mu_i, V_i)}} \right) \right], \tag{5}
\]

where \( \Phi \) is the cumulative distribution of a standard normal variable, and the expectation is with respect to \( V_i \) and \( V_2 \) (see Appendix A.1 for the derivation). The mean of the difference is zero and the variance is given by

\[
\text{Var}(Y_2 - Y_1) = 2(\sigma^2(\mu_i) + E[\tau^2(\mu_i, V_i)]), \tag{6}
\]

with \( \sigma \sim \mathcal{N}(0, \sigma^2(\mu_i)) \). Based on this variance and the definition in (1), a first approximation for the reference change values becomes

\[
L_a = -z_a \sqrt{2(\sigma^2(\mu_i) + E[\tau^2(\mu_i, V_i)])}, \quad U_a = z_a \sqrt{2(\sigma^2(\mu_i) + E[\tau^2(\mu_i, V_i)])} \tag{7}
\]

where \( z_a \) denotes again the upper \( a \)-quantile of the standard normal distribution. A more precise RCV would be determined by two reference change values \( L_a \) and \( U_a \) such that

\[
P(Y_2 - Y_1 \leq L_a) = 1 - P(Y_2 - Y_1 \leq U_a) = \alpha, \tag{8}
\]

but the boundaries \( L_a \) and \( U_a \) may still depend on the unknown parameter \( \mu_i \) through \( \sigma(\mu_i) \) and \( \tau(\mu_i, V_i) \). A possible alternative is to use the first observation \( Y_1 \) in the lower and upper bounds, i.e. \( L_a(Y_1) \) and \( U_a(Y_1) \), such that

\[
P(Y_2 - Y_1 \leq L_a(Y_1)) = 1 - P(Y_2 - Y_1 \leq U_a(Y_1)) = \alpha. \tag{9}
\]

The inclusion of \( Y_1 \) in the RCV may help eliminate the parameter \( \mu_i \) and in a way it is used as an estimator of \( \mu_i \). The following cases provide the boundaries under certain assumptions on \( \sigma(\mu_i) \) and \( \tau(\mu_i, V_i) \).

2.1. Case I. \( \sigma(\mu_i) = \sigma_0 \) and \( \tau(\mu_i, V_i) = t_0 \)

This is the traditional setting for computing reference change values. Under the stated normal distribution and independence assumptions, the RCVs are directly determined by (1) and (6),

\[
L_a = -z_a \sqrt{2(\sigma_0^2 + t_0^2)} \quad \text{and} \quad U_a = z_a \sqrt{2(\sigma_0^2 + t_0^2)}. \tag{10}
\]

This case may seem trivial, and in a way it is, but it can include the cases of distributions other than normal. If for instance the original data \( Y \) follows a log-normal distribution, it is possible to compute reference change values for the measurements in the logarithmic scale, i.e. \( Y_2 - Y_1 = \log(Y_2) - \log(Y_1) \). These limits can be transformed back to the limits for the ratio \( Y_2/Y_1 \) in the original scale. In fact, when introducing model (2), we have considered the additive normal structure possibly after some transformations of the original data, and that can include more general distributions.

2.2. Case II. \( \sigma(\mu_i) = c_0\mu_i \) and \( \tau(\mu_i, V_i) = c_0\mu_i \)

Let \( c_i > 0 \) and \( c_0 > 0 \) denote the intra-subject and measurement coefficients of variation respectively, expressed as fractions. Since the variance of the measurement error \( \tau_i^2 \) is independent of the random term \( V_i \), the difference \( Y_2 - Y_1 \) is normally distributed with variance \( \text{Var}(Y_2 - Y_1) = 2c_0^2\mu_i^2 \), where \( c_i = \sqrt{c_i^2 + c_0^2} \) denotes the total coefficient of variation. Thus the reference change values, without using \( Y_1 \) in the computation, would be given by \( \pm z_a \sqrt{2c_0^2\mu_i^2} \). While the total coefficient of variation of laboratory results is usually known from experimental studies, the individual mean \( \mu_i \) is unknown. Instead, we propose the limits \( L_a(Y_1) = \sqrt{2c_0^2\mu_i^2} \) and \( U_a(Y_1) = \sqrt{2c_0^2\mu_i^2} \), where \( L_a^0 > 0 \) and \( U_a^0 > 0 \) are constants chosen such that the equalities (9) hold. In this case \( L_a^0 \) and \( U_a^0 \) can be determined in closed-form expression

\[
L_a^0 = -\frac{\sqrt{2}}{2} \sqrt{2 - 2c_i^{-2} - 2c_0^{-2}}, \quad \text{and} \quad U_a^0 = \frac{\sqrt{2}}{2} \sqrt{2 - 2c_i^{-2} - 2c_0^{-2}}, \tag{11}
\]

with restrictions \( c_i \leq \sqrt{2c_0^{-1}} \) for the lower bound and \( c_i < \sqrt{2c_0^{-1}} \) for the upper bound. Note that the lower bound \( L_a^0 \) for \( c_i = \sqrt{2c_0^{-1}} \) is defined by its continuity extension, since the limit exists and is finite,

\[
\lim_{c_i \to \sqrt{2c_0^{-1}}} L_a^0 = -\frac{\sqrt{2}}{2}, \tag{12}
\]

while the upper bound diverges to infinity when \( c_i \) approaches \( \sqrt{2c_0^{-1}} \). See Appendix A.2 for the derivation of the bounds (11) and Appendix A.3 for
the derivation of their limits.

The reference change values found under the assumptions of this section satisfy a property of reversibility. It means that, under the assumption of \( Y_2 \geq Y_1 \), given a first measurement \( Y_1 \) we can determine the upper bound \( Y_1 (1 + \sqrt{2c_i U_{i}^2}) \) for \( Y_2 \) but alternatively, given \( Y_2 \), it is possible to determine the lower bound \( Y_2 (1 + \sqrt{2c_i U_{i}^2}) \) for \( Y_1 \). If the observations \( Y_1 \) and \( Y_2 \) are such that they would equal these bounds, we would obtain the reversibility criterion

\[
(1 + \sqrt{2c_i U_{i}^2}) (1 + \sqrt{2c_i U_{i}^2}) = 1.
\]

(13)

Substituting (11) in (13) shows that this reversibility criterion is indeed satisfied.

The bounds \( L_i(Y_1) \) and \( U_i(Y_1) \) with \( L_i^0 \) and \( U_i^0 \) given in (11) were also determined by Ref. [8] using an intuitive approach. The author likewise mentions the restriction on the value \( c_i \) for which it is not possible to determine the upper bound, and these limitations are in line with the formulas provided here. For example, with the commonly used significance level \( \alpha = 0.05 \), the total coefficient of variation should satisfy \( c_i < 0.8597 \) for the lower bound, and \( c_i < 0.6079 \) for the upper bound.

For how they are defined, the boundaries can be rewritten as bounds on the ratio \( Y_2/Y_1 \). They are given by \( 1 + \sqrt{2c_i U_{i}^2} \) and \( 1 + \sqrt{2c_i U_{i}^2} \) respectively, with \( L_i^0 \) and \( U_i^0 \) given in (11). In certain cases, this might be easier to understand than a range on the difference.

2.3. Case III. \( \sigma(\mu_i) = c_i \mu_i \) and \( \tau(\mu_i, \sigma_i) = c_m \sigma_i \)

In practice, the intra-subject variability makes the true value change over time. Thus, if the variance of the measurement error depends on the value it tries to measure, then it will depend on the true value \( X_0 \), rather than on the average value \( \mu_i \), since \( \mu_i \) is not known at the moment that \( X_0 \) is being measured. Case III models this more realistic setting by assuming that \( \sigma(\mu_i) = c_i \mu_i \) and \( \tau(\mu_i, \sigma_i) = c_m \sigma_i \). Under these hypotheses, the variance of the difference \( Y_2 - Y_1 \) is given by

\[
\text{Var}(Y_2 - Y_1) = 2\mu_i^2 (c_i^2 + c_m^2 + c_s^2) = 2\mu_i^2 (c_i^2 + c_s^2).
\]

(14)

and it is larger than the variance \( 2\mu_i^2 c_s^2 \) that was obtained in Section 2.2 and by Ref. [8]. However, the product \( c_s^2 \) may be small enough to still use the same limits from Case II. On the other hand, the distribution of the difference is no longer normally distributed and the RCV cannot be derived from (1). Thus we consider the reference change values in the form \( L_i(Y_0) = \sqrt{2c_i U_{i}^2} Y_1, U_i(Y_0) = \sqrt{2c_i U_{i}^2} Y_1 \) as in Section 2.2, and compute the probabilities (9). The cumulative distribution function of the difference \( Y_2 - \hat{D}_i Y_1 \) is given by

\[
P(Y_2 - \hat{D}_i Y_1 \leq 0) = \mathbb{E} \left[ \Phi \left( \frac{\left( (1 + c_i Z_j) + \hat{D}_i (1 + c_i Z_j) \right)}{c_m \sqrt{(1 + c_i Z) \frac{1}{2} + \hat{D}_i(1 + c_i Z_j)^2}} \right) \right],
\]

(15)

where \( Z_1 \) and \( Z_2 \) are independent and standard normally distributed random variables, and \( \hat{D}_i = 1 + \sqrt{2c_i U_{i}^2} \) or \( \hat{D}_i = 1 + \sqrt{2c_i U_{i}^2} \) for the lower and upper bounds respectively. The full derivation is available in Appendix A.4. Note that the distribution (15) is independent of the mean \( \mu_i \), but does depend on the parameters \( c_i \) and \( c_m \), which would be approximately known in practice.

Solving the equalities in (9) with (15) does not lead to a closed-form expression of \( L_i^0 \) and \( U_i^0 \), but the bounds can be obtained numerically, with the values computed in (11) as starting point for iterative procedures.

In Table 1 we show the lower and upper bounds for Case II and Case III, when \( \alpha = 0.05 \). It is interesting to note that including measurement-dependent analytical variation results in limits that are no longer symmetric. Furthermore, under the assumptions of Case III the property of reversibility introduced in Section 2.2 is no longer exactly satisfied. The reason is that now there is a stochastic component affecting the variance of the measurement error, and this influences the measurement differently at the two points in time. The values found numerically for Case III are monotone functions of both \( c_i \) and \( c_m \), as in Case I and Case II. More precisely, the values of both \( L_i^0 \) and \( U_i^0 \) increase with the intra-subject and analytical coefficients of variation. More interesting is the difference in the upper bounds for Case II and III. When intra-individual and analytical variation increase, the absolute difference in the upper bounds for Case II and III becomes substantial. This suggests that the criteria proposed by Ref. [8] are too conservative when the analytical variation depends on the true value it tries to measure, as confirmed by the type I error \( a \) when the bounds from Case II are applied to Case III. However, this becomes clinically relevant when \( c_i \) and \( c_m \) are large.

2.4. Example

To illustrate the calculation of reference change values we use the example of cardiac troponin T (cTnT) in Ref. [2] and take \( \alpha = 0.05 \). Here the authors provide hourly and weekly coefficients of variation from two different assays. We consider the weekly total and biological variation

<table>
<thead>
<tr>
<th>( c_i )</th>
<th>( c_m )</th>
<th>( L_i^0 )</th>
<th>( U_i^0 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( 0.1 )</td>
<td>( 0.2 )</td>
<td>( 0.3 )</td>
<td>( 0.4 )</td>
</tr>
<tr>
<td>( 0.1 )</td>
<td>( 0.2 )</td>
<td>( 0.3 )</td>
<td>( 0.4 )</td>
</tr>
<tr>
<td>( -1.1 )</td>
<td>( -1.2 )</td>
<td>( -1.3 )</td>
<td>( -1.4 )</td>
</tr>
<tr>
<td>( -1.1 )</td>
<td>( -1.2 )</td>
<td>( -1.3 )</td>
<td>( -1.4 )</td>
</tr>
<tr>
<td>( -1.1 )</td>
<td>( -1.2 )</td>
<td>( -1.3 )</td>
<td>( -1.4 )</td>
</tr>
<tr>
<td>( -1.1 )</td>
<td>( -1.2 )</td>
<td>( -1.3 )</td>
<td>( -1.4 )</td>
</tr>
</tbody>
</table>

\( c_i \) and \( c_m \) are the coefficients of variation for the measurement and analytical variations, respectively. For each combination of \( c_i \) and \( c_m \), we calculate the lower and upper bounds for Case II and Case III. The results are presented in Table 1.

<table>
<thead>
<tr>
<th>( c_i )</th>
<th>( c_m )</th>
<th>( L_i^0 )</th>
<th>( U_i^0 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( 0.1 )</td>
<td>( 0.2 )</td>
<td>( 0.3 )</td>
<td>( 0.4 )</td>
</tr>
<tr>
<td>( 0.1 )</td>
<td>( 0.2 )</td>
<td>( 0.3 )</td>
<td>( 0.4 )</td>
</tr>
<tr>
<td>( -1.1 )</td>
<td>( -1.2 )</td>
<td>( -1.3 )</td>
<td>( -1.4 )</td>
</tr>
<tr>
<td>( -1.1 )</td>
<td>( -1.2 )</td>
<td>( -1.3 )</td>
<td>( -1.4 )</td>
</tr>
<tr>
<td>( -1.1 )</td>
<td>( -1.2 )</td>
<td>( -1.3 )</td>
<td>( -1.4 )</td>
</tr>
<tr>
<td>( -1.1 )</td>
<td>( -1.2 )</td>
<td>( -1.3 )</td>
<td>( -1.4 )</td>
</tr>
</tbody>
</table>
from the Elecsys 2010 assay, given by a median total coefficient of variation $c_t = 32\%$ and a median intra-subject coefficient of variation $c_s = 30\%$ respectively, from which we derive the measurement coefficient of variation, $c_m = \sqrt{c_t^2 - c_s^2} = 11.14\%$ (see Table 1 in Ref. [2]). Since there are no absolute values for the intra-individual and analytical variances given in Ref. [2], the RCV in (10) for Case I is formulated as a relative difference with respect to the first measurement, i.e. the RCV for $(Y_2 - Y_1)/Y_1$, presented in percentages, is given by $\text{RCV}(\%) = 100\%c_m/\sqrt{2}c_t$. The lower and upper limits for Case I are given by $-74.44\%$ and $+74.44\%$, so that an acceptable range for the ratio $100\%Y_2/Y_1$ is given by $[25.6\%, 174.4\%]$. Note that these bounds differ from the ones provided in Ref. [2] due to a sample size correction. Using (11), the lower and upper limits for Case II are given by $L^0_{0.05} = -1.2649$ and $U^0_{0.05} = +2.9587$, respectively. This leads to an acceptable range of $[-42.8\%, 233.9\%]$ for the ratio $100\%Y_2/Y_1$ by applying $100\%(1 + \sqrt{2}c_t^0_{0.05})$ and $100\%(1 + \sqrt{2}c_t^0_{0.05})$. This demonstrates the asymmetry in reference change values when analytical variation is proportional to the measurand, emphasized by Ref. [8]. These limits are much more realistic than the ones obtained under Case I, normally calculated by researchers. However, when the measurement error is truly proportional to the measurand, both the lower and upper limits, obtained by solving equation (9) with (15) through numerical integration, are equal to $L^0_{0.05} = -1.2360$ and $U^0_{0.05} = +2.8280$ respectively. These bounds lead to an acceptable range of $[44.1\%, 228.0\%]$ for the ratio $100\%Y_2/Y_1$ in Case III. This confirms what is observed in the previous section, that the reference change values under the assumptions of [8] are more conservative than those obtained under the more realistic assumption of measurement error dependent on the measurand. The difference in this case is not dramatic though, due to a relatively low analytical coefficient of variation ($c_t = 11.14\%$). Larger differences may exist when the analytical variation becomes larger (see Table 1).

3. Power of reference change values

So far we have assumed the measurements to be stable across the true mean value $\mu_i$ except for random deviations around it, and these random deviations have been incorporated in the reference change values. In this section, the assumption of stability is relaxed and the power of the different bounds in detecting a change is determined. We focus on one-directional detection, i.e. we study the power of the upper bound for detecting positive changes, and respectively the power of the lower bound for detecting negative shifts in the three cases introduced in the previous sections.

Let $\Delta$ denote a shift between the first and second measurements due to a non-physiological variation in the patient. We will assume that the values can still be modeled by an additive structure,

$$Y_i = \mu_i + V_i + E_i,$$

with $\mu_i = \mu_0 + \Delta [Y]_i$ and $1_{A}(z)$ an indicator function equal to one when $z \in A$ and zero otherwise. We choose the shift to be proportional to the intra-subject deviation,

$$\Delta = k\sigma(\mu_i), \quad k \in \mathbb{R},$$

with $k > 0$ for positive shifts and $k < 0$ for negative ones. Note that the mean value at the second time point is given by $\mu_0 + \Delta$, and that this also affects the variances at the second time point in Case II and III. As a consequence of this, we need to restrict $k$ in Case II and III to

$$k > -1/c_s,$$

to ensure positive mean values and acceptable standard deviations also at the second time point.

3.1. Case I. $\sigma(\mu_i) = \sigma_0$ and $\tau(\mu_i, Y_i) = \tau_0$

Under the assumption of constant variances, the powers in detecting positive and negative shifts of magnitude $\Delta = k\sigma_0$ can be obtained from

$$P^-(L_u) = \Phi\left(\frac{L_u - k\sigma_0}{\sqrt{2(\sigma_0^2 + \tau_0^2)}}\right)$$

and

$$P^+(U_u) = 1 - \Phi\left(\frac{U_u - k\sigma_0}{\sqrt{2(\sigma_0^2 + \tau_0^2)}}\right),$$

with $L_u$ and $U_u$ provided by (10) (see Appendix B.1 for more details). Both powers converge to one when the absolute value of $k$ diverges to infinity as in this case there are no restrictions on the parameter, and they are symmetric as shown in Fig. 1.

Furthermore, the power increases with the intra-subject variability $\sigma_0$ reaching an asymptotic value, and decreases with higher variances of the measurement error $\tau_0^2$ as expected.

3.2. Case II. $\sigma(\mu_i) = \epsilon_0 \mu_i$ and $\tau(\mu_i, Y_i) = \epsilon_0 \mu_i$

The results shown for Case I are of course trivial, but they function as a benchmark for a comparison with the other two cases. When both variances are assumed to be proportional to the true mean value and the
change is assumed to be proportional to the intra-subject variability \( \Delta = kc_m \), the power can be obtained from

\[
P^*(L_n) = \Phi \left( \frac{-k c_i + \sqrt{2c_i L_n^0}}{c_i \sqrt{(1 + k c_i)^2 + (1 + \sqrt{2c_i L_n^0})^2}} \right),
\]

\[
P^*(U_n) = 1 - \Phi \left( \frac{-k c_i + \sqrt{2c_i U_n^0}}{c_i \sqrt{(1 + k c_i)^2 + (1 + \sqrt{2c_i U_n^0})^2}} \right),
\]

with \( L_n^0 \) and \( U_n^0 \) given in (11). For the full derivation, see Appendix B.2. Note that the shift ends up also in the variance, due to the assumption \( \sigma(q) = c_m q \), and this demonstrates the need for (18). On the other hand, the power is still independent of the mean health status \( \mu_i \). The values are visualized in Fig. 2, as we can see, the power is no longer symmetric for positive and negative shifts, and in particular it is more difficult to detect changes upwards, but the monotonicity with respect to \( c_m \) is retained. While in Case I the power stabilizes at its maximum value when the intra-subject variability increases, this is completely different for Case II. For a negative shift we see different three profiles, depending on the value of \( k \), which are monotonic functions of \( c_i \) (see Fig. 2a); for a positive shift the power seems to increase with \( c_i \) up to a maximum and then seems to decline (see Fig. 2b), which can be explained by the formulas in (20).

Case II seems to detect more quickly a downward shift than in Case I. This is not surprising, since a downward change would also reduce the variation of the measurement error and would thus make it easier to detect changes. However, it is at the expense of upwards shifts, since these increase the measurement error and are thus hard to detect.

### 3.3. Case III. \( \sigma(\mu_i) = c_i \mu_i \) and \( \tau(\mu_i, v_0) = c_i x_0 \)

Under the more realistic assumption of dependence on the true value of the error variance, the power in detecting a change \( \Delta = kc_m \), can be computed with

\[
P^*(L_n) = E \left[ \Phi \left( \frac{-L_n + (1 + c_i Z_1) + D_m(1 + c_i Z_1)}{c_i \sqrt{(1 + k c_i)^2 + (1 + \sqrt{2c_i L_n^0})^2}} \right) \right],
\]

\[
P^*(U_n) = 1 - E \left[ \Phi \left( \frac{-U_n + (1 + c_i Z_2) + D_m(1 + c_i Z_2)}{c_i \sqrt{(1 + k c_i)^2 + (1 + \sqrt{2c_i U_n^0})^2}} \right) \right],
\]

(21)

with \( Z_1 \) and \( Z_2 \) independent and standard normally distributed random variables, and \( D_m = 1 + \sqrt{2c_i L_n^0} \) and \( D_m = 1 + \sqrt{2c_i U_n^0} \) for the lower and upper limits, respectively.

Fig. 2. Case II. Power of the reference change values as function of \( c_i \) for different values of \( k \), \( c_m = 0.1 \). For \( k = -3 \), the power of \( L_n \) is computed for \( c_i < 0.33 \) according to constraint (18).

Fig. 3. Case III. Power of the reference change values as function of \( c_i \), \( c_m = 0.1 \). For \( k = -3 \), the power of \( L_n \) is computed for \( c_i < 0.33 \) according to constraint (18).
upper bounds respectively, as derived in Appendix B.3. Considerations similar to Case II can be made: first of all, the power is still independent of the mean value \( \mu_i \). Furthermore, although the limits computed under the settings of Case III seem to detect shifts with higher probability, the plots in Figs. 2 and 3 do not show a significant difference in the power in these two cases. This is in fact due to the choice of the coefficient of variation of the measurement error. For larger values of \( c_m \), the difference between the powers increases (see the figures in Appendix C), since the discrepancy between bounds is larger for larger values of \( c_m \), as shown in Table 1.

4. Discussion

In this paper we showed that oversimplified assumptions on the distribution of laboratory results lead to misinterpretation of sequential measurements. While the shortcomings of population-based reference values are fairly well accepted, the effect of assuming constant variances when computing reference change values is a matter which has been poorly addressed [8,4]. [8] raised the topic of computing critical differences when the intra-subject and measurement standard deviations vary across measurements in time, and correctly identified that traditional bounds are unacceptable when the measurement error is proportional to the measurand, unless variances are low. However, he did not study the more realistic setting that the measurement error would depend on the true value it is trying to measure, instead of the true average value. We investigated this more general case and studied the consequences on detecting changes. The extended analysis showed that the reference change values of [8] are too conservative in particular for less precise measurement methods and this affects negatively the power of the upper bound when the measurement error is larger. Reference change values in the two non-traditional scenarios have difficulties in revealing an upward shift, while downward shifts are more easily distinguished, even more easily than under traditional settings. This is explained by the fact that positive shifts substantially increase measurement variation, while negative ones reduce it. The cases studied here can be extended assuming more general variance structures, relaxing normality assumptions or the independence between the intra-subject term and the measurement error. However, this simple case already shows the necessity of further investigation in this field. The understanding of medical laboratory measurements is a real issue which needs to be addressed and requires more research in the world of statistics. The accuracy of reference change values might be increased by taking into account more than two sequential measurements, for example using control charts, but we leave this for further investigations. For the purpose of this paper, we limited the analysis to two measurements and showed that the current assessment methods, with measurement error proportional to the measured quantity, can easily detect negative shifts in the individual’s health condition but fail to capture big positive ones.

Acknowledgments

This research was performed within the framework of the strategic joint research program on Data Science between TU/e and Philips Electronics Nederland B.V. We would like to thank Dr. Marleen Schippers for the valuable discussions.

Appendix A. Derivation of the reference change values

Appendix A.1 Bounds (5)

The cumulative distribution function of the difference \( Y_2 - Y_1 \) can be computed by conditioning on the intra-individual values \( V_i \),

\[
P(Y_2 - Y_1 \leq y) = \mathbb{E}[P(Y_2 - Y_1 \leq y|V_i, V_2)]
= \mathbb{E}[P(E_2 - E_1 \leq y - (V_2 - V_1)|V_i, V_2,C_2)].
\] (A.1)

The difference \( E_2 - E_1 \) is normally distributed conditionally on \( (V_i, V_2) \) with mean and variance

\[
\mathbb{E}(E_2 - E_1|V_i, V_2) = \tau^2(\mu, V_1) + \tau^2(\mu, V_2),
\]

and thus (A.1) becomes

\[
P(Y_2 - Y_1 \leq y) = \mathbb{E}\left[ \Phi\left( \frac{y - (V_2 - V_1)}{\sqrt{\tau^2(\mu, V_1) + \tau^2(\mu, V_2)}} \right) \right]
\]

\[
= \int_{-\infty}^{\infty} \Phi\left( \frac{y - (v_2 - v_1)}{\sqrt{\tau^2(\mu, v_1) + \tau^2(\mu, v_2)}} \right) \phi\left( \frac{v_1}{\sigma(\mu)} \right) \phi\left( \frac{v_2}{\sigma(\mu)} \right) \frac{1}{\sigma(\mu)} dv_1 dv_2,
\] (A.2)

which is in fact (5).

Appendix A.2 Case II. Bounds (11)

With the choice of \( D_a(Y_1) = \sqrt{2}c_\alpha D_a^0Y_1 \) made in Section 2.2, and with the constant \( D_a^0 \) being either \( L_a^0 \) or \( U_a^0 \), the probability in (9) becomes

\[
P(Y_2 - Y_1 \leq D_a(Y_1)) = P\left( Y_2 - \left( 1 + \sqrt{2}c_\alpha D_a^0 \right) Y_1 \leq 0 \right).
\] (A.3)

Since the random variable is still normally distributed with mean and variance

\[
\mathbb{E}\left( Y_2 - \left( 1 + \sqrt{2}c_\alpha D_a^0 \right) Y_1 \right) = -\sqrt{2}c_\alpha D_a^0 \mu,
\]

\[
\text{Var}\left( Y_2 - \left( 1 + \sqrt{2}c_\alpha D_a^0 \right) Y_1 \right) = 2\mu^2c_\alpha^2\left( 1 + \sqrt{2}c_\alpha D_a^0 + (c_\alpha D_a^0)^2 \right).
\]
the probability in (A.3) is simply

\[ P(Y_2 - Y_1 \leq D_a(Y_1)) = \Phi \left( \frac{\sqrt{2}c_1 \mu_i D_{2}^0}{\sqrt{2\mu_i^2 c_1^2 + (c_1 D_{2}^0)^2}} \right). \]

Thus, equating this probability to \( \alpha \) and \( 1 - \alpha \) for the lower and upper bounds respectively, leads to the two equations

\[ L_0^0 = \frac{c_1}{\sqrt{2}} \left( 1 + \sqrt{2c_1^2 L_{22}^0} \right), \]

\[ U_0^0 = \frac{c_1}{\sqrt{2}} \left( 1 + \sqrt{2c_1^2 U_{22}^0} \right), \]

from which the closed-form expressions (11) are obtained using standard algebraic calculations.

**Appendix A.3 Case II. Limits of the bounds**

When \( c_i \) approximates the critical value \( z_\alpha^{-1} \), the lower limit converges to a finite value,  

\[ \lim_{c_i \to z_\alpha^{-1}} \frac{\sqrt{2}c_1}{2} \frac{\sqrt{2 - z_\alpha^{-1} c_i}}{1 - z_\alpha^{-1}} = -\frac{\sqrt{2}}{2}. \]

\[ \lim_{c_i \to z_\alpha^{-1}} \frac{\sqrt{2}c_1}{2} \frac{\sqrt{2 - z_\alpha^{-1} c_i}}{1 - z_\alpha^{-1}} = +\infty. \]

**Appendix A.4 Case III. Cumulative density function (15)**

Similarly to A.2, with the same choice of \( D_a(Y_1) = \sqrt{2c_1 D_{2}^a Y_1}, D_{2}^a \) being either \( L_0^0 \) or \( U_0^0 \), the probabilities in (9) become again as in (A.3). However, the random variable \( Y_2 - (1 + \sqrt{2c_1 D_{2}^a Y_1}) \) is no longer normally distributed, and the cumulative density function is obtained by conditioning on the true values \( X_1 = \mu_i + V_1 \) and \( X_2 = \mu_i + V_2 \). Taking \( D_a = \sqrt{2c_1 D_{2}^a} \), the probability in (9) becomes

\[ \mathbb{E} \left[ P(E_2 - \tilde{D}_a E_1 \leq -X_2 + \tilde{D}_a X_1 | X_1, X_2) \right]. \]

Now, \( E_2 - \tilde{D}_a E_1 \) is normally distributed conditionally on \( X_1 \) and \( X_2 \) with mean and variance

\[ \mathbb{E} \left[ (E_2 - \tilde{D}_a E_1 | X_1, X_2) = 0 \right. \]

\[ \text{Var} \left( (E_2 - \tilde{D}_a E_1 | X_1, X_2) = c_i^2 \left( \frac{X_2^2}{c_i^2} + \frac{\tilde{D}_a^2 X_1^2}{c_i^2} \right) \right. \]

Thus (A.5) becomes

\[ \int \Phi \left( \frac{-x_2 + \tilde{D}_a x_1}{c_i \sqrt{\frac{x_2^2}{c_i^2} + \tilde{D}_a^2 x_1^2}} \frac{1}{\sigma(\mu_i)} \phi \left( \frac{x_1 - \mu_i}{\sigma(\mu_i)} \right) \phi \left( \frac{x_2 - \mu_i}{\sigma(\mu_i)} \right) dx_1 dx_2. \]

Rewriting (A.7) using \( \sigma(\mu_i) = c_i \mu_i \) leads to

\[ \int \Phi \left( \frac{-c_i (x_2 + \tilde{D}_a (1 + c_i z_2))}{c_i \sqrt{1 + c_i z_2^2 + \tilde{D}_a^2 (1 + c_i z_2)^2}} \right) \phi(z_1) \phi(z_2) dz_1 dz_2, \]

which is in fact (15).
Appendix B. Derivation of the of the power of RCVs

Appendix B.1 Case I. Powers (19)

Assuming (16), the difference \( Y_2 - Y_1 \) is normally distributed with mean and variance

\[
\mathbb{E}[Y_2 - Y_1] = \Delta, \\
\text{Var}(Y_2 - Y_1) = 2(\sigma_i^2 + \sigma_d^2).
\]

Thus the powers \( P^-(L_a) = \mathbb{P}(Y_2 - Y_1 \leq L_a) \) and \( P^+(U_a) = \mathbb{P}(Y_2 - Y_1 \geq U_a) \) in detecting a shift \( \Delta = k\sigma_0 \) can be computed using the probability

\[
\mathbb{P}(Y_2 - Y_1 \leq D_a) = \Phi\left( \frac{D_a - \Delta}{\sqrt{2(\sigma_i^2 + \sigma_d^2)}} \right), \tag{B.1}
\]

with \( D_a \) denoting either \( L_a \) or \( U_a \) in (10).

Appendix B.2 Case II. Powers (20)

In Case II the difference \( Y_2 - Y_1 \) is still normally distributed, with mean \( \Delta \). However, in this case the variance is affected by the change in the true mean health status, leading to

\[
\text{Var}(Y_1) = c_i^2 \sigma_i^2 \quad \text{and} \quad \text{Var}(Y_2) = c_i^2 (\mu_i + \Delta)^2.
\]

The reference change value is dependent on the first measurement, thus the power in detecting a change \( \Delta = kc_i \mu_i \) is computed with

\[
\mathbb{P}(Y_2 - Y_1 \leq D_a(Y_i)) = \mathbb{P}(Y_2 - \tilde{D}_a Y_1 \leq 0), \tag{B.2}
\]

with \( D_a(Y_i) = \sqrt{2c_i D_i^0} Y_i \) and \( \tilde{D}_a = 1 + \sqrt{2c_i D_i^0} \). The random variable \( Y_2 - \tilde{D}_a Y_1 \) is normally distributed with mean and variance

\[
\begin{align*}
\mathbb{E}[Y_2 - \tilde{D}_a Y_1] &= \Delta - \mu_i \sqrt{2c_i D_i^0}, \\
\text{Var}(Y_2 - \tilde{D}_a Y_1) &= c_i^2 \left( (\mu_i + \Delta)^2 + \tilde{D}_a^2 \sigma_i^2 \right).
\end{align*}
\]

Thus (B.2) becomes

\[
\mathbb{P}(Y_2 - Y_1 \leq D_a(Y_i)) = \Phi\left( \frac{-k \mu_i + \sqrt{2c_i D_i^0}}{c_i \sqrt{(1 + k \sigma_0)^2 + (1 + \sqrt{2c_i D_i^0})^2}} \right), \tag{B.3}
\]

and (20) is now obtained by using (B.3) for \( P^-(L_a) = \mathbb{P}(Y_2 - Y_1 \leq L_a(Y_i)) \) and \( P^+(U_a) = \mathbb{P}(Y_2 - Y_1 \geq U_a(Y_i)) \).

Appendix B.3 Case III. Powers (21)

The power in detecting a change \( \Delta = kc_i \mu_i \) is again computed with the probability (B.2). However, the random variable \( Y_2 - \tilde{D}_a Y_1 \) is no longer normally distributed and we need to condition on the true values \( X_1 \) and \( X_2 \). The probability in (B.2) is then

\[
\mathbb{P}(Y_2 - \tilde{D}_a Y_1 \leq 0) = \mathbb{E}\left[ \mathbb{P}(E_2 - \tilde{D}_a E_1 \leq -X_2 + \tilde{D}_a X_1 \mid X_1, X_2) \right]. \tag{B.4}
\]

Now the difference \( E_2 - \tilde{D}_a E_1 \) is normally distributed conditionally on \( X_1 \) and \( X_2 \), with mean and variance given in (A.6). Taking into account that \( \mathbb{E}[X_2] = \mu_i + \Delta \), (B.4) becomes

\[
\mathbb{E}\left[ \Phi\left( \frac{-X_2 + \tilde{D}_a X_1}{c_i \sqrt{X_2^2 + \tilde{D}_a^2 X_1^2}} \right) \right] = \mathbb{E}\left[ \Phi\left( \frac{-(\mu_i + \Delta)(1 + c_i Z_2) + \tilde{D}_a(\mu_i + c_i Z_2 \mu_i)}{c_i \sqrt{(\mu_i + \Delta)^2(1 + c_i Z_2)^2 + \tilde{D}_a^2(\mu_i + c_i Z_2 \mu_i)^2}} \right) \right] = \mathbb{E}\left[ \Phi\left( \frac{-(1 + k \sigma_0)(1 + c_i Z_2) + \tilde{D}_a(1 + c_i Z_2)}{c_i \sqrt{(1 + k \sigma_0)^2(1 + c_i Z_2)^2 + \tilde{D}_a^2(1 + c_i Z_2)^2}} \right) \right], \tag{B.5}
\]

where \( Z_1 \) and \( Z_2 \) are two independent standard normally distributed random variables. Applying this in \( P^-(L_a) = \mathbb{P}(Y_2 - Y_1 \leq L_a(Y_i)) \) and \( P^+(U_a) = \mathbb{P}(Y_2 - Y_1 \geq U_a(Y_i)) \), the powers in (21) can be easily obtained.
Appendix C. Additional plots for the power of RCVs

Fig. C.4. Power of the reference change values $L_{\alpha}(Y_i)$ as function of $c_i$ for different values of $k$, $c_m = 0.2$. For $k = -3$, the power of $L_\alpha$ is computed for $c_i < 0.33$ according to constraint (18).

Fig. C.5. Power of the reference change values $U_{\alpha}(Y_i)$ as function of $c_i$ for different values of $k$, $c_m = 0.2$.

Fig. C.6. Power of the reference change values $L_{\alpha}(Y_i)$ as function of $c_i$ for different values of $k$, $c_m = 0.3$. For $k = -3$, the power of $L_\alpha$ is computed for $c_i < 0.33$ according to constraint (18).
Fig. C.7. Power of the reference change values $U_{\alpha}(Y_1)$ as function of $c_0$ for different values of $k, c_0 = 0.3$.

References


