

# On the Index of Repeatability: Estimation and Sample Size Requirements

Maha Al-Eid<sup>1</sup>, Mohamed M. Shoukri<sup>2\*</sup>

<sup>1</sup>Department of Biostatistics, Epidemiology and Scientific Computing, King Faisal Specialist Hospital and Research Center, Riyadh, KSA

<sup>2</sup>Department of Epidemiology and Biostatistics, Schulich School of Medicine and Dentistry, London, Ontario, Canada

Email: \*Mahaeid@kfshrc.edu.sa, mmshoukr@uwo.ca

**How to cite this paper:** Al-Eid, M. and Shoukri, M.M. (2019) On the Index of Repeatability: Estimation and Sample Size Requirements. *Open Journal of Statistics*, 9, 530-541.

<https://doi.org/10.4236/ojs.2019.94035>

**Received:** July 30, 2019

**Accepted:** August 17, 2019

**Published:** August 20, 2019

Copyright © 2019 by author(s) and Scientific Research Publishing Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

## Abstract

**Background:** Repeatability is a statement on the magnitude of measurement error. When biomarkers are used for disease diagnoses, they should be measured accurately. **Objectives:** We derive an index of repeatability based on the ratio of two variance components. Estimation of the index is derived from the one-way Analysis of Variance table based on the one-way random effects model. We estimate the large sample variance of the estimator and assess its adequacy using bootstrap methods. An important requirement for valid estimation of repeatability is the availability of multiple observations on each subject taken by the same rater and under the same conditions. **Methods:** We use the delta method to derive the large sample variance of the estimate of repeatability index. The question related to the number of required repeats per subjects is answered by two methods. In first methods we estimate the number of repeats that minimizes the variance of the estimated repeatability index, and the second determine the number of repeats needed under cost-constraints. **Results and Novel Contribution:** The situation when the measurements do not follow Gaussian distribution will be dealt with. It is shown that the required sample size is quite sensitive to the relative cost. We illustrate the methodologies on the Serum Alanine-aminotransferase (ALT) available from hospital registry data for samples of males and females. Repeatability is higher among females in comparison to males.

## Keywords

Measurements Errors, Functions of Variance Components, Delta Method, Lagrange Multiplier, Bootstrap Technique

## 1. Introduction

Repeatability and reproducibility are ways of measuring precision, particularly

in the fields of biochemistry, radiology, and medical diagnoses. In general, scientists perform the same experiment several times in order to confirm their findings. These findings may show variations. In the context of an experiment, repeatability measures the variation in measurements taken by a single instrument or person under the same conditions, while reproducibility measures whether an entire study or an experiment can be reproduced. There has been confusion in the literature about the way that repeatability and reproducibility are quantified. Both concepts were often reported as either standard deviations or coefficient of variations.

The main focus of this paper is on the concept of repeatability, which was first introduced by Bland and Altman [1]. For repeatability to be established, the following conditions must be in place: the measurements should be taken in the same location; the same measurement procedure; the same observer; the same measuring instrument, used under the same conditions; and repetition over a short period of time.

What's known as "the repeatability" is in fact a measurement of precision, which denotes the absolute difference between a pair of repeated test results. We note that when we have more than two readings per subject the idea of pairing produces several repeatability coefficients and the concept becomes unclear.

Repeatability is also known as test-retest reliability indicating the closeness of the agreement between the results of successive measurements of the same measurand carried out under the same conditions of measurement. A less-than-perfect test-retest reliability causes test-retest variability. Such variability can be caused by, for example, intra-individual variability and intra-observer variability. A measurement may be said to be *repeatable* when this variation is smaller than a pre-determined acceptance criterion. A complete account on the reliability literature can be found in Shoukri [2] [3].

One of the most important applications of the concept of repeatability is in the construction of the normal range or reference range in clinical medicine, which relies on the availability of a large sample of healthy individuals. Research has shown that the distribution of these measurements is affected by two main sources of variations: the between subjects and the within subject-components of variations.

This paper has three-fold objectives: Firstly; we define a proposed index of repeatability, as the ratio of the within-subjects' variations to the between subject variation. The within subject variation is expected to be quite small relative to the between subjects-variations. To formalize the presentation, we assume that a single measurement  $y_i$  from subject  $i = 1, 2, \dots, k$  is written as:

$$y_i = \mu + s_i + e_i \quad (1)$$

Hence  $s_i$  represents the sources of between-subjects biological variation, and  $e_i$  represents sources of within subject variations, while  $\mu$  denotes the population mean. Note that the assumption of additivity of components is made to simplify the presentation. However, a multiplicative model may be made addi-

tive under the logarithmic transformation. Following Harries and De Mets [4] it is further assumed that  $s_i \sim N(0, \sigma_s^2)$ , and  $e_i \sim N(0, \sigma_e^2)$  and that  $s_i \perp e_i$  for all  $i$ . We define the “Repeatability Index Parameter” (RIP) as  $\theta = \sigma_e^2 / \sigma_s^2$ .

The salient point is that  $\theta$  cannot be estimated unless we have at least two repeated measurements on any subject in the study.

In Section 2 we specify the model generating the observations and discuss a general method of estimating RIP from a sample of  $k$  subjects when there is an opportunity to have  $n$  repeated samples per subject. In Section 3, we provide two alternatives for the sampling strategies. The first, we assume that the investigator has decided to acquire on total number of measurements  $N = kn$ , and the question becomes; what is the best split between  $(k, n)$ , that maximizes the accuracy of estimating RIP?

One of the biggest obstacles in clinical studies is the cost constraints. Therefore, the second strategy is to find the optimal split of  $N = kn$ , so that IRP is estimated with maximum precision under cost restrictions (constrained optimization). The third objective of the study is to address the issue of estimating RIP when the assumption of the Gaussian distribution of observation is not tenable.

## 2. Model Specifications and Parameter Estimation

We assume that for subject  $i$ ,  $n$  replicates of the same variable of interest  $y_{ij}$  are taken by the same instrument at the same time, so that

$$y_{ij} = \mu + s_i + e_{ij} \quad (2)$$

$$i = 1, 2, \dots, k$$

$k$  is the number of subjects

$$j = 1, 2, \dots, n$$

$n$  is the number of replications per subject

We further assume that the components of the model described by (2) are such that:

$s_i \sim N(0, \sigma_s^2)$ , are independently distributed random variables measuring the subjects effect, are independently distributed of the within subjects variation denoted by  $e_{ij} \sim N(0, \sigma_e^2)$ .

Under the additivity assumption of the model components, we have:

$$\text{var}(y_{ij}) = \sigma_s^2 + \sigma_e^2 = \sigma_s^2 (1 + \sigma_e^2 / \sigma_s^2) = \sigma_s^2 (1 + \theta)$$

The parameter  $\theta = \sigma_e^2 / \sigma_s^2$  is the target parameter of interest, named “Repeatability Index Parameter” (RIP). The components of variation of the model set-up can be estimated using the well-known one-way Analysis of Variance (ANOVA) with random effects (Table 1).

S.O.V = Source of variation, DF = Degrees of freedom associated with the corresponding sum of squares, S.O.S = Corrected sums of squares, MS = Mean square error = S.O.S/DF, EMS = Expected mean square. The sample statistics needed for the ANOVA computations based on the available observations are given as:

**Table 1.** The ANOVA set-up.

S.O.V	DF	S.O.S	M.S	EMS
Between subjects	$k - 1$	$SSB$	$MSB$	$\sigma_e^2 + n\sigma_s^2$
Within subjects	$N - k$	$SSW$	$MSW$	$\sigma_e^2$
Total	$N - 1$			

$$\bar{y}_{..} = \frac{1}{nk} \sum_i \sum_j y_{ij}$$

$$SSW = \sum_i \sum_j (y_{ij} - \bar{y}_{i.})^2$$

$$\bar{y}_{i.} = \frac{1}{n} \sum_{j=1}^n y_{ij}$$

$$SSB = \sum_{i=1}^k n(\bar{y}_{i.} - \bar{y}_{..})^2$$

The moment estimator of the parameter  $\theta$ , and hence the maximum likelihood estimator (under balanced design) is given by:

$$\hat{\theta} = n \left[ \frac{MSW}{MSB - MSW} \right] \quad (3)$$

The parameter estimator  $\hat{\theta}$  given in Equation (3) is a nonlinear function of the sample statistics, and therefore an exact expression for its variance is not available. We use the delta method (Kendall and Stuart, 1989) [5], to obtain the first approximation of the variance of  $\hat{\theta}$  given by:

$$\begin{aligned} \text{var}(\hat{\theta}) = & \text{var}(MSB) \left( \frac{\partial \hat{\theta}}{\partial MSB} \right)^2 + 2 \text{cov}(MSB, MSW) \left( \frac{\partial \hat{\theta}}{\partial MSB} \right) \left( \frac{\partial \hat{\theta}}{\partial MSW} \right) \\ & + \text{var}(MSW) \left( \frac{\partial \hat{\theta}}{\partial MSW} \right)^2 \end{aligned} \quad (4)$$

Substituting the required quantities in (4) and simplifying we get the first order approximation of the variance of  $\hat{\theta}$  as:

$$v = \text{var}(\hat{\theta}) = \frac{2\theta^2 (n + \theta)^2}{N(n-1)(1+\theta)^8} \quad (5)$$

An  $(1-\alpha)100\%$  approximate Wald's confidence interval on  $\theta$  may be constructed as:

$$\hat{\theta} \pm z\sqrt{v} \quad (6)$$

where  $z$  in Equation (6) is the  $(1-\alpha/2)100\%$  cut-off point from the standard normal table.

### 3. How Many Repeats Do We Need?

Our first approach to estimate the optimal number of replications is to assume

that  $N = kn$  is fixed a priori, and one needs to determine the number of replicates that minimizes the variance  $v$ , as given in (5).

Minimizing  $v$  with respect to the number of replications per subjects and solving for  $n$  we get:

$$\frac{\partial v}{\partial n} = 0 \Rightarrow n = 2 + \theta \tag{7}$$

$$\frac{\partial^2 v}{\partial n^2} = \frac{1}{2(1+\theta)^2} > 0$$

This means that  $\text{var}(\hat{\theta})$  is minimized (*i.e.* maximizing precision) when at least 2 repeats are attained from each subject as shown in Equation (7). When  $\theta = 0$  (no within subject-variation) then  $n = 2$  precisely.

We may also estimate the number of repeats for fixed width confidence interval as follows:

Suppose that we have decided on the number of subjects  $k$ . The question now is how many repeats per-subject are needed to estimate  $\theta$  with 95% confidence such that the width of the confidence interval has a maximum given length  $w$ .

Since the length of the Wald's confidence interval is given as:

$$w = 2(1.96)\sqrt{\text{var}(\hat{\theta})}$$

$$w^2 = \frac{4(1.96)^2 \cdot 2\theta^2 (n + \theta)^2}{(1 + \theta)^8 kn(n - 1)}$$

Let  $A = \frac{kw^2 (1 + \theta)^8}{8(1.96)^2 \theta^2} > 1$

Solving for  $n$  we have:

$$n = \frac{(A + 2\theta) + [A(4\theta^2 + 4\theta + A)]^{1/2}}{2(A - 1)} \tag{8}$$

This closed form expression is quite simple, and the computation of  $n$  from Equation (8), is straight forward. Substituting  $\theta = 0.25$ ,  $k = 100$ , and  $w = 1$  in (8), then  $n = 156$ .

#### 4. Estimating the Number of Repeats under Cost Constraints

It is an extremely expensive, and in some circumstances, it is a difficult task to obtain repeated samples from each subject. Some of these difficulties are related to cost and time (which may be translated into cost). Clearly too small a sample may lead to a study that produces many false negatives, too large a sample may result in many false positives and additional cost. Thus, a critical decision in constructing accurate estimate of normal range is to balance the cost of recruiting healthy normal with the need to obtain accurate estimate of RIP. In this section we shall address the issue of obtaining the combination  $(n, k)$  that minimizes the variance of  $\hat{\theta}$  subject to cost constraints. The sampling cost depends

primarily on the size of the sample, and includes the data collection costs, subjects recruiting costs, management and technicians' costs. On the other hand, overhead costs remain fixed regardless of the sample size. The total cost is assumed this additive formula

$$T = t_0 + kt_1 + nkt_2 \quad (9)$$

In Equation (9)  $t_0$  is the fixed cost,  $t_1$  is the cost of recruiting a healthy subject, and  $t_2$  is the cost of taking a single measurement. Denoting the variance of  $\hat{\theta}$  by  $V$ , the main objective is to determine the number of repeated measurements that minimize the variance of  $\hat{\theta}$  subject to cost constraints  $T$ . In terms of language of optimization, we construct the objective function

$$Q = V + \lambda(T - t_0 - kt_1 - nkt_2) \quad (10)$$

The parameter  $\lambda$  in Equation (10) is the Lagrange-multiplier. The necessary conditions for minimization of  $Q$  are:

$$\frac{\partial Q}{\partial n} = 0, \frac{\partial Q}{\partial k} = 0, \text{ and } \frac{\partial Q}{\partial \lambda} = 0.$$

Differentiating  $Q$  with respect to  $n$ ,  $k$ , and  $\lambda$  and equating to zero we get:

$$n^3 - n^2(2 + \theta) - nR(1 + 2\theta) + \theta R = 0 \quad (11)$$

Note that from Equation (9) we have:

$$k = \frac{T - t_0}{t_1 + nt_2}, \text{ where } R = t_1/t_2$$

The cubic Equation (11) has an explicit solution given by:

$$n_{opt} = \frac{2 + \theta}{3} + \alpha^{1/3} \left( \frac{1 + \theta}{3} \right) + \beta/3 \quad (12)$$

where

$$\beta = \frac{1}{(1 + \theta)\alpha^{1/3}} \left\{ (2 + \theta)^2 + 3R(1 + 2\theta) \right\}$$

and

$$\begin{aligned} \alpha = \frac{3}{1 + \theta} \left[ 3R \left\{ (R + 1)^2 \frac{1}{(1 + \theta)^4} - (6R^2 + 4R - 2) \frac{1}{(1 + \theta)^3} \right. \right. \\ \left. \left. + 12R(R + 1) \frac{1}{(1 + \theta)^2} - (8R^2 + 10R + 2) \frac{1}{1 + \theta} - R - 1 \right\} \right]^{1/2} \\ + 9R \left( \frac{1}{(1 + \theta)^3} - \frac{1}{(1 + \theta)^2} + \frac{1}{1 + \theta} \right) + \left( \frac{2 + \theta}{1 + \theta} \right)^3 \end{aligned}$$

Equation (12) is the optimum number of replicates per subject that is needed to minimize the variance of the estimated RIP when the total cost of the investigation is held fixed.

Note that, when  $t_1 = 0$  and  $t_2 = 1$  (i.e.  $R = 0$ ), then  $n_{opt} = 2 + \theta$ , as given in Equation (7).

This means that a special cost structure is implied by the optimal allocation procedure discussed in the previous section. Note also, when  $\theta = 0$ ,  $n_{opt} = 1 + (1 + R)^{1/2} \geq 2$ , implying that the ratio  $R = t_1/t_2$  is an important factor in determining the optimal allocation  $(n, k)$ .

Examples:

$$R = 0.1, \theta = 0.1, \text{ then } n = 2$$

$$R = 0.1, \theta = 0.5, \text{ then } n = 3$$

$$R = 0.2, \theta = 3, \text{ then } n = 6$$

$$R = 0.5, \theta = 4, \text{ then } n = 7$$

Remarks:

We set as a bench mark to the value of the estimator of RIP a maximum of 1%. That is if the within subject variation relative to the between-subjects variation is above 1%, then repeatability is low, and visa-versa.

Note also, that the estimator of  $\theta$  is a non-linear function of the sample data, and hence is potentially biased estimator. Moreover, the derived variance is just a first order approximation of the actual variance. Finally, if the measurements are not normally distributed, then construction of confidence interval on the population parameter using the normal quantiles will not be acceptable unless the sample size is quite large. One way to assess the properties of the proposed estimator is to use the nonparametric-bootstrap sampling techniques. We shall address this issue in the data analysis section.

## 5. Effect of Non-Normality of Components of Variations on the Estimated Variance of RIP

Not all biological markers that are measured on continuous scale have Gaussian distributions. In this section we drop the assumptions of normality regarding the distributions of  $b_i$  and  $e_{ij}$ , and evaluate the effect of non-normality on the estimation of the RIP. The immediate consequences of dropping the assumption of non-normality of the measurements are:

- 1) The one-way ANOVA mean squares  $MSB$  and  $MSW$  will not have chi-square distributions.
- 2) The mean squares  $MSB$  and  $MSW$  are no longer independent, and hence the ratio of the mean squares will not have the usual F-distribution.

Relaxing the assumption of normality both the measures of Kurtosis of  $b_i$  and  $e_{ij}$  are needed in the calculation of the asymptotic variance of  $\hat{\theta}$  [6].

Let  $\delta_e$  and  $\delta_b$  denote respectively the coefficients of kurtosis of  $e_{ij}$  and  $b_i$ . These quantities are defined as:

$$\delta_b = \left\{ E(b_i^4) / \sigma_b^4 \right\} - 3$$

$$\delta_e = \left\{ E(e_{ij}^4) / \sigma_e^4 \right\} - 3$$

Using results for the balanced one-way ANOVA [6] we have:

$$\text{var}(MSW) = c_1 \sigma_e^4, \quad \text{var}(MSB) = c_2 \sigma_e^4, \quad \text{and} \quad \text{cov}(MSW, MSB) = c_{12} \sigma_e^4, \quad \text{where}$$

$$c_1 = \{k(n-1)\}^{-1} \left[ 2 + \delta_e \frac{n-1}{n} \right]$$

$$c_2 = 2(k)^{-1} \left[ (\theta)^{-2} n^2 (1 + \delta_b) + 2n(\theta)^{-1} + \left( 1 + \frac{\delta_e}{n} \right) \right]$$

$$c_{12} = (kn)^{-1} \delta_e$$

Using the delta method [5], and substituting in (4) we get, the first order approximation, variance of  $\hat{\theta}$  is:

Simplifying we get:

$$\text{var}(\hat{\theta}) = \frac{1}{n^2} \left\{ [\theta(n+\theta)]^2 c_1 - 2\theta^3 (n+\theta) c_{12} + c_2 \theta^4 \right\} \quad (13)$$

Comments:

The first question that needs to be answered is: which component of variation has the largest effect on the variance of the RIP estimate, and hence on the number of repeats. We answer this question in a heuristic manner. We note from Equation (13) that  $\delta_e$  is divided by the factor  $\{kn\}$  in  $c_1$ ,  $c_2$ , and  $c_{12}$ . The implication is that, as the number of subjects increase, the kurtosis of the error term has negligible effect on the variance of the estimated RIP.

We may also demonstrate the effect of non-normality using tools of probability and power calculations. This can be illustrated through testing of statistical hypotheses on the RIP. Suppose that we need to determine the number of subjects to detect the departure from the null hypothesis  $H_0: \theta = \theta_0$  in the direction of the one-sided alternative  $H_1: \theta = \theta_1 < \theta_0$ , with type-one error rate  $\alpha$  and power  $1 - \beta$ . For fixed  $n$ , we can show that:

$$k = \frac{\left[ z_\alpha \sqrt{v[\theta_0]} \right] + z_\beta \sqrt{v[\theta_1]}^2}{(\theta_0 - \theta_1)^2} \quad (14)$$

If we set the Type I error rate at 5% and power at 80%, for given values of  $\theta_0, \theta_1, n, \delta_b$  and  $\delta_e$ , the estimated values of  $k$  can be easily calculated.

Specifically, for an effect size  $(\theta_0 - \theta_1) = (0.05 - 0.02)$ ,  $\delta_b = 0$ , and  $\delta_e = 6$ , and  $n = 5$ , then from Equation (14) we need to recruit 6 subjects, while for the same range of values of the RIP we need to recruit 21 subjects if  $\delta_b = 6$ , and  $\delta_e = 0$ . The worst situation is when the two components of variation are far from being normal. For example, for the same values under the null and alternative hypotheses, with  $\delta_b = 3$ ,  $\delta_e = 3$ , and  $n = 30$ , then  $k = 67$ . However, when  $\delta_b = \delta_e = 0$ , we need to recruit  $k = 18$  only. These computations illustrate the impact of the departure from normality of the distribution of between and within subject-variations on the sample size requirements.

## 6. Data Analysis and Bootstrap

In this section we apply the methodology presented in this paper on Serum Alanine-aminotransferase (ALT). The ALT is a critical parameter for both the as-

assessment and follow-up of patients with liver disease. Therefore, establishing the repeatability and the precision of ALT measurements as a diagnostic marker are of paramount importance. Regardless of gender or body mass index (BMI) [7], the normal range was most often estimated from a population that included patients with subclinical liver disease, including non-alcoholic fatty liver disease (NAFLD), which is now documented as the greatest prevalent cause of chronic liver disease worldwide [8]. Recent studies have recommended establishing normal ranges for ALT separately in males and females [9].

Furthermore, lately published HBV guidelines suggested that treatment decisions should be based on these new ALT levels [10], with the exception of one recently published Korean study, no earlier reports have established normal liver histology when evaluating reference ALT ranges [11].

From a large tertiary hospital-based registry, the available data were grouped into female group with 20 subjects and another male group of 30 subjects. In both groups, each subject's ALT was evaluated three times according to the rules set in [1]. The data are summarized in **Table 2**, for females and in **Table 3** for males.

#### Bootstrap results

We used R to bootstrap the data. We set the number of bootstrap samples at 1000 for both data sets.

Bootstrap Statistics for females' data:

original bias std. error

0.001 0.00117 0.0004

As can be seen from **Figure 1**, both the histogram and the Q-Q plot show that the large sample distribution of the estimator is skewed to the right. Therefore, one should be careful when constructing Wal's confidence limits of the population RIP

Bootstrap Statistics for males' data:

original bias std. error

0.002 - 0.0001 0.0006

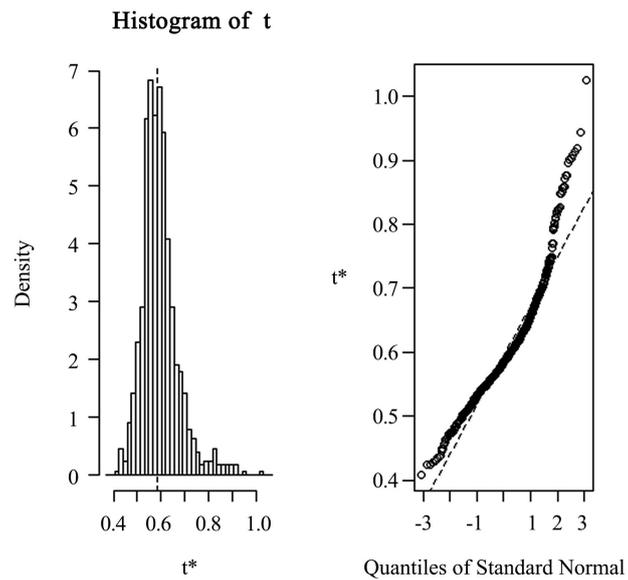
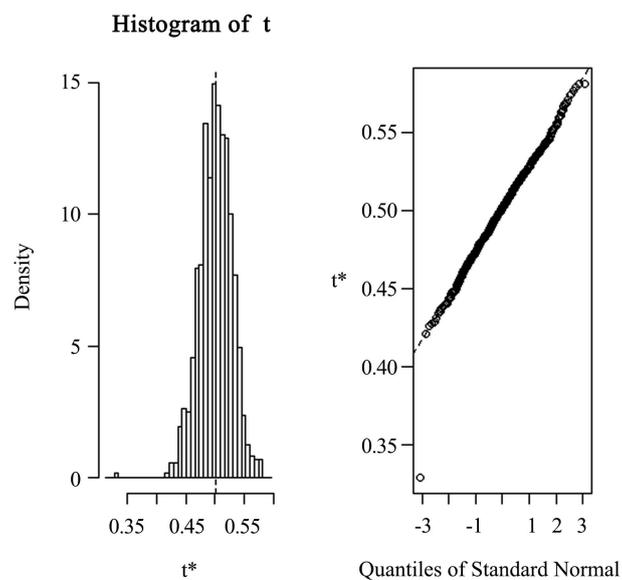
In contrast to females' data, the histogram of the sample statistics as shown in **Figure 2** is skewed to the left, but the Q-Q plot exhibit closer to normality. This may be due to the fact that the males' data is larger than the females' data.

**Table 2.** Descriptive statistics of the female ALT data.

ALT Mean	Standard Deviation	N	Minimum	Maximum
ALT1 75.05	39.16	20	1	141
ALT2 76.65	39.34	20	2	143
ALT3 75.84	39.30	20	1	142

**Table 3.** Descriptive statistics of the male ALT data.

ALT Mean	Standard Deviation	N	Minimum	Maximum
ALT1 106.97	81.91	30	6	317
ALT2 108.87	81.96	30	8	319
ALT3 108.00	82.06	30	7	318

**Figure 1.** Histogram and the Q-Q plots of the 1000 bootstrap samples of the estimated RIP (females ALT data).**Figure 2.** Males' data histogram and the Q-Q plots of the 1000 bootstrap samples of the estimated RIP.

## 7. Comments and Summary

As can be seen from the histograms and the Q-Q plots, the distribution of the estimated  $RIP = t1^*$  is far from being normally distributed. But we expect that the distributional properties may be closer to normality when the number of subjects is much larger than the number of have here. When one attempts to establish the population-based reference range of health populations, the number of subjects is typically in the hundreds, and the issue of normality may be irrelevant.

Further investigations for the case of categorical measurements and when the number of replications per subject is not fixed, are needed.

## Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

## References

- [1] Bland, M. and Altman, D. (2010) Statistical-Methods for Assessing Agreement between 2 Methods of Clinical Measurement. *International Journal of Nursing Studies*, **47**, 931-936. <https://doi.org/10.1016/j.ijnurstu.2009.10.001>
- [2] Shoukri, M. (2010) Measures of Interobserver Agreement and Reliability. 2nd Edition, CRC Press, Taylor and Frances, Boca Raton.
- [3] Shoukri, M.M. (2015) Measures of Agreement. Invited Contribution to Wiley Statistical References. Stat05301. <https://doi.org/10.1002/9781118445112.stat05301.pub2>
- [4] Harris, E.K. and DeMets, D. (1972) Estimation of Normal Ranges and Cumulative Proportions by Transforming Observed Distributions to Gaussian Form. *Clinical Chemistry*, **18**, 605-612.
- [5] Stuart, A. and Ord, J.K. (1987) Kendall's Advanced Theory of Statistics, Volume 1. 5th Edition, Charles Griffin, London.
- [6] Shoukri, M.M., Tracy, D.S. and Mian, I.U.H. (1990) The Effect of Kurtosis in Estimation of the Parameters of the One-Way Random Effects Model from Familial Data. *Computational Statistics and Data Analysis*, **10**, 339-345. [https://doi.org/10.1016/0167-9473\(90\)90016-B](https://doi.org/10.1016/0167-9473(90)90016-B)
- [7] Kaplan, M.M. (2002) Alanine Aminotransferase Levels: What's Normal? *Annals of Internal Medicine*, **137**, 49. <https://doi.org/10.7326/0003-4819-137-1-200207020-00012>
- [8] Lazo, M. and Clark, J.M. (2008) The Epidemiology of Nonalcoholic Fatty Liver Disease: A Global Perspective. *Seminars in Liver Disease*, **28**, 339-350. <https://doi.org/10.1055/s-0028-1091978>
- [9] Prati, D., Taioli, E., Zanella, A., Della Torre, E., Butelli, S., Del Vecchio, E., Conte, D., *et al.* (2002) Updated Definitions of Healthy Ranges for Serum Alanine Aminotransferase Levels. *Annals of Internal Medicine*, **137**, 1-10. <https://doi.org/10.7326/0003-4819-137-1-200207020-00006>
- [10] Sanai, F.M., Helmy, A., Dale, C., Al-Ashgar, H., Abdo, A.A., Katada, K., Hashem, A., *et al.* (2011) Updated Thresholds for Alanine Aminotransferase Do Not Exclude Significant Histological Disease in Chronic Hepatitis C. *Liver International*, **31**, 1039-1046. <https://doi.org/10.1111/j.1478-3231.2011.02551.x>

- [11] Keeffe, E.B., Dieterich, D.T., Han, S.H.B., Jacobson, I.M., Martin, P., Schiff, E.R. and Tobias, H. (2008) A Treatment Algorithm for the Management of Chronic Hepatitis B Virus Infection in the United States: 2008 Update. *Clinical Gastroenterology and Hepatology*, **6**, 1315-1341. <https://doi.org/10.1016/j.cgh.2008.08.021>