Computer Methods and Programs in Biomedicine (2005) **77**, 165—173

Computer Methods and Programs in Biomedicine

www.intl.elsevierhealth.com/journals/cmpb

Power calculation for the likelihood ratio-test when comparing two dependent intraclass correlation coefficients

B. Giraudeau[a,b,*], R. Porcher[b,c], J.Y. Mary[b,c]

[a] *INSERM CIC 202, Faculté de Médecine, 10 Bd Tonnellé, BP 3223, 37032 Tours Cedex 1, France*
[b] *Biostatistique et Epidémiologie Clinique, INSERM ERM 0321, Université Paris 7, Paris, France*
[c] *Département de Biostatistique et Informatique Médicale, Assistance Publique des Hôpitaux de Paris, Hôpital Saint-Louis, Paris, France*

Received 3 February 2004; received in revised form 29 September 2004; accepted 1 October 2004

KEYWORDS
Reproducibility;
Intraclass correlation coefficient;
Likelihood ratio-test;
Power

Summary Comparing the reproducibility level of two devices with continuous outcome on a unique sample of subjects (each subject being assessed several times with both devices) comes down to compare two dependent intraclass correlation coefficients (ICCs). When planning such a reproducibility study, one has to specify both the number of subjects to be included and the number of replicates per subject associated to each device. We propose SAS and S-plus macros, which allow power calculations by implementing a simulation study where dependent ICCs are compared by means of a likelihood ratio-test.
© 2004 Elsevier Ireland Ltd. All rights reserved.

1. Introduction

Reproducibility (also called reliability) is a metrological property, which refers to the consistency between several measurements realised either by the same reader (intra-reader reproducibility) or by different readers (inter-reader reproducibility). For continuous outcomes, the intraclass correlation coefficient (ICC) is the coefficient generally used for its assessment [1]. This coefficient admits a dual definition and can thus be defined either as the proportion of the total variance due to the between-subject variance or as the correlation that exists between two distinct measures realised on the same subject [2]. When developing a new device, its reproducibility has to be studied. If the device is to be compared to another device already in use, one is thus led to compare two ICCs. In such a situation, two elements can motivate the recruitment of a unique sample of subjects who are then assessed several times with each device. First, paired comparisons are of higher power than comparisons between independent samples. Second, the ICC is known to be a variance-dependent index [3] and the recruitment of a unique sample thus prevents

* Corresponding author. Tel.: +33 2 47 36 61 42; fax: +33 2 47 36 62 00.
E-mail address: giraudeau@med.univ-tours.fr (B. Giraudeau).

0169-2607/$ — see front matter © 2004 Elsevier Ireland Ltd. All rights reserved.
doi:10.1016/j.cmpb.2004.10.001

from a discrepancy of inter-subject variance between two independent samples. Comparing the reproducibility of the two devices therefore comes down to the comparison of two dependent ICCs, which can be handled by means of a likelihood ratio-test. However, as acknowledged by Donner et al. [4] there seems to be no explicit expression for the maximum likelihood estimate of a common ICC (under the null hypothesis H_0). This prevents from deriving explicitly the likelihood ratio statistic and therefore to calculate an analytic expression of the power of the test. We therefore propose SAS and S-plus macros, which use numerical optimisation and thus allow power calculations by means of simulation studies. These macros are described in Section 3, while the underlying statistical model is presented in Section 2. An example illustrates the use of these macros in Section 4.

2. Computational methods and theory

2.1. Statistical model

Let $X_i = \begin{pmatrix} X_i^{(1)} \\ X_i^{(2)} \end{pmatrix}$ where $(i = 1, \ldots, n)$, the p-vector of measures associated to subject i with $X_i^{(k)} = (X_{i1}^{(k)}, X_{i2}^{(k)}, \ldots, X_{ip_k}^{(k)})' = (X_{ij_k}^{(k)})'$ where $(j_k = 1, \ldots, p_k)$, being the p_k-vector of measures realised with device k (k = 1, 2). The total number of measures associated to each of the n subjects, noted p, is fixed and equals $p_1 + p_2$.

We assumed that the following model holds:

$$X_i \sim N_p(M, \Sigma), \qquad i = 1, \ldots, n \tag{1}$$

where

- N_p refers to the p-variate normal distribution,
- $M = \begin{pmatrix} \mu^{(1)} 1_{p_1} \\ \mu^{(2)} 1_{p_2} \end{pmatrix}$ with $\mu^{(k)}$, the overall mean in sample k and 1_{p_k} the p_k-vector containing only 1's, and
- $\Sigma = \begin{pmatrix} \Sigma_{p_1}^{(1)} & \Sigma_{p_1 p_2} \\ \Sigma_{p_2 p_1} & \Sigma_{p_2}^{(2)} \end{pmatrix}$ with
  - o $\Sigma_{p_k}^{(k)} = \sigma_{X^{(k)}}^2 \left\{ (1 - \rho^{(k)}) I_{p_k} + \rho^{(k)} J_{p_k} \right\}$, $\sigma_{X^{(k)}}^2$ referring to the variance of $X_{ij_k}^{(k)}$, $\rho^{(k)}$ to the common correlation between $X_{il}^{(k)}$ and $X_{im}^{(k)}$ $(l \neq m)$, whatever i ($\rho^{(k)}$ thus refers to the ICC associated to device k), I_{p_k} being the $p_k \times p_k$ identity matrix and J_{p_k} the $p_k \times p_k$ matrix containing only 1's;
  - o $\Sigma_{p_1 p_2} = \delta J_{p_1 p_2}$ and $\Sigma_{p_2 p_1} = \delta J_{p_2 p_1}$, δ referring to $cov(X_{ij_1}^{(1)}, X_{ij_2}^{(2)})$, whatever i, j_1 and j_2, and $J_{p_1 p_2}$ and $J_{p_2 p_1}$ being, respectively, the $p_1 \times p_2$ and $p_2 \times p_1$ matrices containing only 1's.

The model thus defined assumed that for any subject i, the p_k measures realised using device k share a common correlation, which we note $\rho^{(k)}$. It moreover assumes that two measures realised with distinct devices on the same subject (say $X_{ij_1}^{(1)}$ and $X_{ij_2}^{(2)}$) share also a common correlation, which is defined as $\eta = \frac{\delta}{\sigma_{X^{(1)}} \sigma_{X^{(2)}}}$.

2.2. Likelihood ratio-test

We are interested in testing whether the ICCs associated to instrument 1 and 2 are equal and the test hypotheses can therefore be stated as follows:

$$H_0 : \rho^{(1)} = \rho^{(2)} = \rho \qquad H_1 : \rho^{(1)} \neq \rho^{(2)} \tag{2}$$

To assess the likelihood ratio statistic, maximum likelihood estimates have to be derived both under H_0 and H_1. Under H_1, these estimates have been derived by Elston [5] (Appendix A). On the contrary, under H_0, an explicit expression of the maximum likelihood estimate of the common ICC is unlikely, as acknowledged by Donner et al. [4]. Numerical optimisation has, therefore, to be performed and in Appendix A we derive the log-likelihood function to be maximised and the partial derivatives.

2.3. Method for power calculation

We propose to assess power from a simulation study using the likelihood ratio-test. Parameters $\rho^{(1)}$ (under H_0 and H_1) and $\rho^{(2)}$ (under H_1) have to be specified, thus, determining the test hypotheses. The interclass correlation η has also to be specified. As for global means ($\mu^{(1)}$ and $\mu^{(2)}$) and variances ($\sigma_{X^{(1)}}^2$ and $\sigma_{X^{(2)}}^2$), no hypotheses have to be made regarding these parameters since they are of no influence on correlation estimates. Then, for a fixed sample size n, and for fixed number of replicates p_1 and p_2, we propose to calculate power to test the null hypothesis of a common intraclass correlation (i.e. $H_0 : \rho^{(1)} = \rho^{(2)} = \rho$) from a simulation study. For that, data-sets are generated according to the simulation algorithm specified in Appendix B. For each of them, we use the Newton—Raphson optimisation technique (PROC NLP in SAS/OR and the nlminb function in S-plus) to assess maximum likelihood estimates under H_0 and then perform a likelihood ratio-test. Power is then calculated as the proportion of data-sets for which a significant

Download English Version:

https://daneshyari.com/en/article/10345482

Download Persian Version:

https://daneshyari.com/article/10345482

Daneshyari.com