POWER CALCULATION FOR THE LIKELIHOOD RATIO-TEST
WHEN COMPARING TWO DEPENDENT INTRACLASS
CORRELATION COEFFICIENTS

Author’s name(s): Giraudeau B\textsuperscript{1,2} *, Porcher R\textsuperscript{2,3} and Mary JY\textsuperscript{2,3}
Affiliation(s): \textsuperscript{1}INSERM CIC 202, Faculté de Médecine, 2bis Bd Tonnellé, 37032 Tours, France; \textsuperscript{2}Biostatistique et Epidémiologie Clinique, INSERM ERM 0321, Université Paris 7, Paris, France; \textsuperscript{3}Département de Biostatistique et Informatique Médicale, Assistance Publique des Hôpitaux de Paris, Hôpital Saint-Louis, Paris, France

Phone: 33 2 47 36 61 42 ; Fax: 33 2 47 36 62 00
E-mail: giraudeau@med.univ-tours.fr

Corresponding author: Bruno Giraudeau

Keywords: Reproducibility, Intraclass correlation coefficient, Power.
Topic area of the submission: Agreement

Abstract

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 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.

References:


