Reliability analysis: Calculate and Compare Intra-class Correlation Coefficients (ICC) in SAS

Li Lu, MS, Nawar Shara, PhD
Department of Biostatistics and Epidemiology, the MedStar Research Institute, Hyattsville, MD

ABSTRACT

Reliability studies are widely used to assess the measurement reproducibility of human observers, laboratory assays or diagnostic tests. For quantitative measures, intra-class correlation coefficient (ICC) is the principal measurement of reliability. In this paper, a SAS macro is provided to calculate the ICC and its confidence limits. The application of the macro is demonstrated in a clinical reliability case study, the comparison of two ICCs is discussed.

INTRODUCTION

Reliability refers to the reproducibility of the measurement when it is randomly repeated for the same study subject. In a clinical trial, the measurement reliability plays a very important role, since it affects the choice of the primary outcome and the choice of eligibility and exclusion criteria. Usually, reliability studies are conducted to assess the measurement reproducibility of human observers, laboratory assays or diagnostic tests. In some of the studies, coefficient of variation has been reported. However, Lachin (2004) [1] has demonstrated that a coefficient of variation does not measure reliability. The best measure of reliability for continuous data is the intra-class correlation coefficient. Currently, there is no SAS procedure or publicly available SAS macro to directly perform the ICC analysis. In this paper, we provide a SAS macro to estimate ICC and its confidence intervals. We demonstrate the use of the macro in a clinical cardiovascular reliability study and further discuss the comparison of two ICCs.

DEFINITION

Let y_{i1} and y_{i2} denote the duplicate measurements of ith subject , i=1 to n.

$$\overline{y_i} = \frac{y_{i1} + y_{i2}}{2}$$, is the mean value of ith subject,

$d_i = (y_{i1} - y_{i2})$, is the ith subject difference,

$S_i^2 = d_i^2 / 2$, is the within subject variance for the ith subject, and

$$\overline{y} = \frac{\sum_{i=1}^{n} \overline{yi}}{n}$$, is the overall mean of all 2n measurements.

$$S_b^2 = \frac{\sum_{i=1}^{n} 2(\overline{y_i} - \overline{y})}{n-1}$$, is the variance between subjects,

$$S_w^2 = \frac{\sum_{i=1}^{n} S_i^2}{n}$$, is the average within subject variance.

Assuming the measurement errors are distributed independently and identically as $N(0, \sigma_\varepsilon^2)$. The intraclass correlation coefficient ρ and its confidence intervals can be estimated as follows:

$$\hat{\rho} \;=\; \frac{S_b^2 - S_w^2}{S_b^2 + S_w^2} \qquad\qquad (1)$$

$$(\hat{\theta}_L, \hat{\theta}_U) = \frac{1}{2}\ln(\frac{1+\hat{\rho}}{1-\hat{\rho}}) \pm \frac{Z_{1-\alpha/2}\sqrt{\hat{V}(\hat{\rho})}}{(1+\hat{\rho})(1-\hat{\rho})} \qquad\qquad (2)$$

$$(\hat{\rho}_L, \hat{\rho}_U) = \frac{e^{2(\hat{\theta}_L, \hat{\theta}_U)} - 1}{e^{2(\hat{\theta}_L, \hat{\theta}_U)} + 1} \qquad\qquad (3)$$

A SAS MACRO TO ESTIMATE ICC AND ITS CONFIDENCE LIMITS
In SAS we can use one-way ANOVA to calculate the between subject variation (S_b^2) and within subject variation (S_w^2), thus calculate the ICC and its confidence limits based on the above equations.
We implement this process in the following SAS macro ICC_SAS:

```
%macro Icc_sas(ds, response, subject);
      ods output OverallANOVA =all;
      proc glm data=&ds;
            class &subject;
            model &response=&subject;
      run;

      data Icc(keep=sb sw n R R_low R_up);
            retain sb sw n;
            set all end=last;
            if source='Model' then sb=ms;
            if source='Error' then do;sw=ms; n=df; end;
            if last then do;
               R=round((sb-sw)/(sb+sw), 0.01);
               vR1=((1-R)**2)/2;
               vR2=(((1+R)**2)/n +((1-R)*(1+3*R)+4*(R**2))/(n-1));
               VR=VR1*VR2;
               L=(0.5*log((1+R)/(1-R)))-(1.96*sqrt(VR))/((1+R)*(1-R));
               U=(0.5*log((1+R)/(1-R)))+(1.96*sqrt(VR))/((1+R)*(1-R));
               R_Low=(exp(2*L)-1)/(exp(2*L)+1);
               R_Up=(exp(2*U)-1)/(exp(2*U)+1);
               output;
             end;
      run;

      proc print data=icc noobs split='*';
            var r r_low r_up;
            label r='ICC*' r_low='Lower bound*' r_up='Upper bound*';
            title 'Reliability test: ICC and its confidence limits';
      run;
%mend;
```

The above macro has three parameters: ds is the input dataset; response is the measurement of interest; subject is the subject id variable. The input dataset should have two observations for each subject.

APPLICATION EXAMPLE
As an application example, we use this Icc_sas macro to analyze a clinical reliability study data. The purpose of the study was to assess the reliability between Cornell lab and Washington Hospital Center

Echo Core lab of measuring Carotid Intima-Media Thickness (IMT) for their patients. The study had 53 IMT images from 27 patients, every image was analyzed once at each lab to measure the CIMT.

```
%Icc_sas(ds=one, response=cimt, subject=subject_id);
```

The above code shows how to call the Icc_sas macro. Among the parameters: *one* is the dataset with 106 CIMT measurement of the 53 images, c*imt* is the measurement we are interested in estimating the reliability, *subject_id* is the id variable of the 53 imt images. The macro generates the following three outputs: the estimated intraclass correlation is 0.93 and the confidence limits are (0.8827, 0.9586).

Additionally, we plot a simple linear regression to check the two lab CIMT measurement reliability in figure1. The regression of IMT_WHC vs. IMT_Conell shows that approximately 91.5% of the variation in IMT_WHC is due to IMT_Conell, which is consistent with our quantified ICC analysis result.
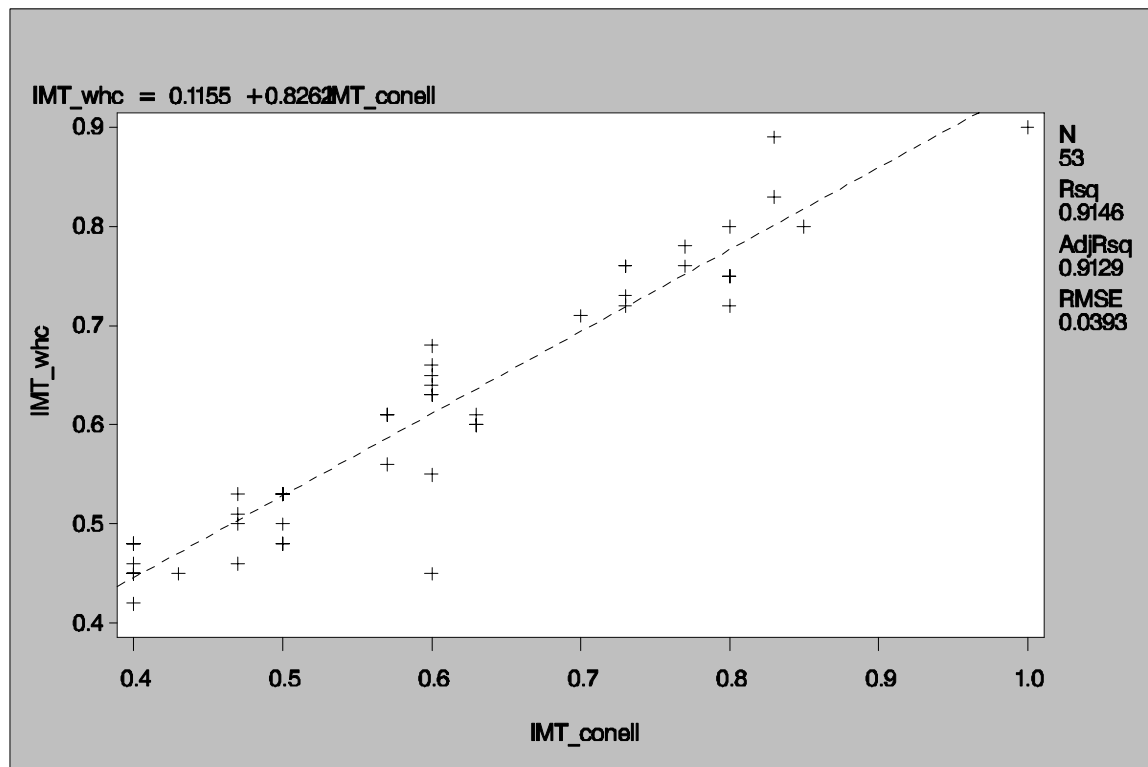

Figure 1. Simple linear regression of CIMT: WHC vs. Cornell.

DISCUSSION
A previous study has found that a within lab ICC of CIMT measurement at Connell lab is 0.98, it is reasonable to assume that the between lab ICC is around 0.9. Based on the given data, we found that the inter-lab correlation is 0.93, which is as good as we have expected.

Practically we often need to compare two ICCs. In the above case, for examples, we might want to compare the inter-lab ICC of 0.93 with previously found within lab ICC of 0.98 and see if they are significantly different. One approach is to use bootstrap method to generate about 1000 sample sets, and call the above SAS macro 1000 times to get 1000 ICCs for both inter-lab and within-lab studies. Based on those 2000 ICCs, we could run a simple T-test to see if the Inter-lab and within-lab ICCs have significant differences. Another approach is to use suitable statistical test. Donner et al (2002) [3] compared the Fisher's Z test, Konishi-Gupta modified Z-test, the likelihood ratio test and Alsawalmeh-Feldt F-test using Monto Carlo

3

simulation studies. Those tests are not easy to perform and their powers need further investigation. The above two approaches both involve intensive computation. Here we use a simple approach of checking confidence limits. The previous report showed that the within lab ICC confidence limits was (0.9237, 0.9989). The overlap between the two confidence limits indicates that there is no significant difference between the two ICCs.

The macro we provided here makes the calculation of ICC based on duplicate measurement as easy as calculating Coefficient of Variation. For m>2 replicates scenarios, the macro can calculate the ICC, the confidence limits estimation need further investigation and improvement.

REFERENCES
[1]. John M Lachin (2004). The role of measurement reliability in clinical trials. Clinical trials 2004, 1: 553-566.
[2]. Smith AB. On the estimation of the intraclass correlation. Ann Human Genet 1956;21:363-73.
[3]. Donner et al. Testing the equality of dependent intraclass correlation coefficients, the Statistician 2002,51: part 3, 367-379.

ACKNOWLEDGMENTS

CONTACT INFORMATION
Your code requests, comments and questions are valued and encouraged. Contact the author at
      Li Lu
      Dept of Epidemiology and Statistics
      MedStar Research Institute
      6495 New Hampshire Avenue
      Hyattsville, MD 20783
      (301) 560-7313
      li.lu@medstar.net