Similarity between the concordance and intraclass correlation cofficients

Hiroshi Ohtsu[1], Joe Matsuoka[2],Tatsuo Kagimura[3] and Tadashi Kusunoki[4]

[1] Department of Clinical Trial Data Management, The University of Tokyo,Japan
[2] Quintiles Transnational Japan K.K.,Japan
[3] Nippon Borhringer Ingelheim Co., Ltd. , Japan
[4] Japanese Society for Pharmacoepimiology

The concordance correlation coefficient (CCC) and the intraclass correlation coefficient (ICC) are used for interrater reliability studies. Although it seems that their definitons are equivalent, the estimates are different.
Letting Y_{ij} be continuous data for subject i ($i = 1, \cdots, n$) assessed by randomly selected rater j ($j = 1, \cdots, m$), usual two-way ANOVA model is $Y_{ij} = \mu + a_i + b_j + e_{ij}$, where $a_i \sim N(0, \sigma_S^2)$, $b_j \sim N(0, \sigma_R^2)$, and $e_{ij} \sim N(0, \sigma_e^2)$, respectively, and the ICC was defined by $\rho^{ICC} = \sigma_S^2/(\sigma_S^2+\sigma_R^2+\sigma_e^2)$ [1]. While, CCC considering fixed effect for raters is defined by $\rho^{CCC} = 2\Sigma_j\Sigma_{k>j}\sigma_{jk}/\{(m-1)\Sigma\sigma_j^2 + m\Sigma(\mu_j - \mu)^2\}$, where σ_j^2 and σ_{jk} are variance and covariance of data from rater j and k ($j \neq k$), and μ_j and μ are the mean for rater j and overall mean. The equality $\rho^{ICC} = \rho^{CCC}$ was deduced using $\theta_R^2 = \frac{1}{m-1}\Sigma(\mu_j - \mu)^2$ in place of σ_R^2 in ρ^{ICC} [2, 3]. Conversely, the equality also holds for randomly selected raters if we use $(m-1)\sigma_R^2 = E\{\Sigma(b_j - b)^2\}$ in place of $\Sigma(\mu_j - \mu)^2$. The unbiased estimator for ρ^{ICC} is $r^{ICC} = (V_S - V_e)/\{V_S + (m-1)V_e + \frac{m}{n}(V_R - V_e)\}$, where V_S, V_R, and V_e are the mean squares for subject, rater, and error, respectively. While, same formula as ρ^{CCC} is usually used for the estimate r^{CCC} by replacing σ_{jk},σ_j^2, μ_j, and μ by sample statistics s_{jk}, s_j^2, $\bar{Y}_{.j}$, and $\bar{\bar{Y}}$; applying $S_{jk} = \Sigma_i(Y_{ij} - \bar{Y}_{.j})(Y_{ik} - \bar{Y}_{.k})$ and $S_{jj} = \Sigma_i(Y_{ij} - \bar{Y}_{.j})^2$, $s_{jk} = \frac{1}{n}S_{jk}$ and $s_j^2 = \frac{1}{n}S_{jj}$ are used. However, alternative devisor $n-1$ may also be used for statistical inference and we write corresponding estimate as r^{CCC*}.
Applying $\Sigma S_{jj} = (n-1)\{V_S + (m-1)V_e\}$ and $2\Sigma_j\Sigma_{k>j}S_{jk} = (n-1)(m-1)(V_S - V_e)$, we have alternative expression for r^{CCC} and r^{CCC*} which is very similar to r^{ICC}; using $(V_S - V_e)$ as common numerator, denominator is $V_S + (m-1)V_e + \frac{m}{n-1}V_R$ for r^{CCC} and $V_S + (m-1)V_e + \frac{m}{n}V_R$ for r^{CCC*}, respectively. Assuming that $\rho^{CCC} = \rho^{ICC}$ and taking expectations, although $E(V_S - V_e) = m\sigma_S^2$ is unbiased, denominators of r^{CCC} and r^{CCC*} include extra term as compared to unbiased $m(\sigma_S^2 + \sigma_R^2 + \sigma_e^2)$; $m\{\sigma_S^2 + \sigma_R^2 + \sigma_e^2 + \frac{1}{n-1}(\sigma_R^2 + \sigma_e^2)\}$ for r^{CCC}, and $m(\sigma_S^2 + \sigma_R^2 + \sigma_e^2 + \frac{1}{n}\sigma_e^2)$ for r^{CCC*}, respectively. Thus r^{CCC} and r^{CCC*} are biased, however, E(r^{CCC}) and E(r^{CCC*}) approach to $\rho^{CCC} = \rho^{ICC}$ as $n \to \infty$. Under a regular condition of $V_R \geq V_e$, although always $r^{ICC} \geq r^{CCC*} \geq r^{CCC}$, a large difference is unlikely if interrater bias V_R is negligible and/or n is sufficiently large. As $n \to \infty$, they are asymptotically equal and approach to a common upper limit $(V_S - V_e)/\{V_S + (m-1)V_e\}$, another version of ICC and CCC representing an ideal goal of reliability studies in which interrater bias is ignorable.
Finally, we will present illustrative examples of reliability studies conducted by our Scaling Keys of Evaluation Techniques for Cerebrovascular patients Heterogeneity (SKETCH) study group to examine relationship between the two-way layout version ICC and the CCCs.

References

[1] Bartko JJ (1966). The intraclass correlation coefficient as a measure of reliability. Psychological reports 19: 3–11.

[2] Barnhart HX, Harber M, and Song J (2002). Overall concordance correlation coefficient for evaluating agreement among multiple observers.Biometrics 58: 1020–1027.

[3] CarrascoJL and Jover L (2003). Estimating the generalized concordance correlation coefficient through variance components. Biometrics 59: 849–858.