

ABSTRACT

CHEN, CHIA-CHENG. Assessing Agreement with Intraclass Correlation Coefficient and Concordance Correlation Coefficient. (Under the direction of Dr. Huiman X. Barnhart and Dr. Marie Davidian.)

Accurate and precise measurements serve as the basis for studies in bioscience research. Agreement studies are often concerned with assessing whether different observers (e.g. machines, raters, methods, instruments, laboratories, assays, devices, etc.) for measuring responses on the same subject or sample can produce similar results. The intraclass correlation coefficient (ICC) and the concordance correlation coefficient (CCC) are two popular scaled indices (with values between -1 and 1) for assessing agreement (closeness) for continuous measurements, where these two indices may take the systematic shifts into account when assessing reliability between multiple observers. We conducted systematic and in-depth comparisons of these two indices under a general model since ICC depends on specific ANOVA models while CCC does not. Usually, the ICC and CCC are used for data without and with replications based on subject and observer effects only. However, we can not use the methodology if repeated measurements rather than replications are collected. There exist some ICC and CCC type indices for assessing agreement with repeated measurements. However, there is no CCC for random observers and random time points. We consider a new CCC for repeated measures where both observers and time are treated as random effects and also summarize other remaining combinations of random or fixed factors for observers and time. Finally, we compare ICC s and CCC s for data with repeated measurements.

Assessing Agreement with Intraclass Correlation Coefficient and Concordance
Correlation Coefficient

by
Chia-Cheng Chen

A dissertation submitted to the Graduate Faculty of
North Carolina State University
in partial fulfillment of the
requirements for the Degree of
Doctor of Philosophy

Statistics

Raleigh, North Carolina

2009

APPROVED BY:

Dr. Anastasios A. Tsiatis

Dr. Huixia Wang

Dr. Huiman X. Barnhart
Chair of Advisory Committee

Dr. Marie Davidian
Co-Chair of Advisory Committee

DEDICATION

*To
My Parents
and
Wife*

BIOGRAPHY

Chia-Cheng Chen was born in Taipei, Taiwan on May 16, 1977. He earned his Bachelor of Science degree in Applied Mathematics at Chinese Culture University in June, 1999 and Master of Science degree in Statistics at National Tsing Hua University in June, 2001. Two years later, he joined the Ph.D program in Statistics at North Carolina State University and continued his advanced graduate studies towards his Ph.D. degree.

ACKNOWLEDGMENTS

I would like to express my deepest appreciation to my advisors, Dr. Huiman X. Barnhart and Dr. Marie Davidian. Without their guidance, support, and patient throughout my dissertation work, it will be very difficult for me to obtain my Ph.D. degree. Dr. Huiman X. Barnhart especially directed my research work to the area of assessing reliability and agreement. I not only learned a lot from doing research in this field, but also realized how to express my thinking more efficiently and accurately, and how to do an independent research.

I also appreciate my committee members, Dr. Anastasios A. Tsiatis and Dr. Huixia Wang, for their suggestions and teaching during my graduate days at NC State. I appreciate the chance and environment that the Statistics department offers to me for my graduate study at NC State.

Finally, I would like to thank my parents and wife for their endless support and love. There is still a large room for me to make progress in the future, even though I obtain the Ph.D. degree.

TABLE OF CONTENTS

LIST OF TABLES.....	vii
LIST OF FIGURES	x
1 Comparison of ICC and CCC for assessing agreement for data without and with replications	1
1.1 Introduction	1
1.2 Comparisons of ICC and CCC	3
1.2.1 Definitions and estimates of ICC s and CCC	4
1.2.2 Comparisons of $E(ICC)$ and CCC under the general model	7
1.3 Examples	10
1.4 Discussion	19
2 The random-observers concordance correlation coefficient	20
2.1 Introduction	20
2.2 Methodology	22
2.2.1 CCC_R^{inter} for inter-observer agreement	22
2.2.2 CCC_R^{intra} for intra-observer agreement	24
2.2.3 CCC_R^{abs} for absolute agreement	25
2.3 Estimation and Inference	26
2.4 Simulations	28
2.5 Examples	34
2.5.1 Vertebral Body Data	34
2.5.2 Image Data	41
3 Assessing agreement with intraclass correlation coefficient and concordance correlation coefficient with repeated measures	50
3.1 Introduction	50
3.2 Methodology	51
3.2.1 CCC for random observers and fixed times	57
3.2.2 CCC for fixed observers and random times	60
3.2.3 CCC for fixed observers and fixed times	63
3.2.4 ICC for repeated measurements	66
3.3 Estimation and Inference	71
3.3.1 Estimation and Inference for CCC	71
3.3.2 Estimation and Inference for ICC	77
3.4 Data Analysis	80
3.5 Discussion	83
Bibliography	84

Appendices	87
Appendix A	88

LIST OF TABLES

Table 1.1	The ANOVA table for ICC s for blood pressure data by number of replicates	16
Table 1.2	Estimates of the ICC s and CCC s for blood pressure data by number of replicates.....	16
Table 1.3	Estimates of individual parameters for blood pressure data by number of replicates.....	16
Table 1.4	The ANOVA table for ICC s for PEFR data by number of replicates.....	18
Table 1.5	Estimates of the ICC s and CCC s for PEFR data by number of replicates.	18
Table 1.6	Estimates of individual parameters for PEFR data by number of replicates	18
Table 2.1	Three scenarios of the true values for the proposed CCC_R	30
Table 2.2	Results of CCC_R^{abs} based on 1000 Monte Carlo data sets.....	31
Table 2.3	Results of CCC_R^{inter} based on 1000 Monte Carlo data sets.....	32
Table 2.4	Results of CCC_R^{intra} based on 1000 Monte Carlo data sets.....	33
Table 2.5	Description statistics for readings between times by readers and locations in Situation 1 for vertebral body data.....	37
Table 2.6	Description statistics for readings between readers by times and locations in Situation 1 for vertebral body data.....	38

Table 2.7 Description statistics for readings between readers and times by locations in Situation 1 for vertebral body data	39
Table 2.8 Results of the estimates and improved 95% bootstrap- t confidence intervals of CCC_R and the estimates of $\rho_{c,rm}$ for two situations for vertebral body data	40
Table 2.9 Description statistics for readings between times by readers and variables for image data	48
Table 2.10 Description statistics for readings between readers by times and variables for image data	48
Table 2.11 Description statistics for readings between readers and times by variables for image data	49
Table 2.12 Results of the estimates and improved 95% bootstrap- t confidence intervals of CCC_R and the estimates of $\rho_{c,rm}$ for image data	49
Table 3.1 Existing methods (Yes, No) for four cases comparing observer with repeated measures under different approaches and indices	52
Table 3.2 Definitions and results of CCC for inter-observer, intra-observer, and absolute agreement under four cases	54
Table 3.3 Definitions of ICC for inter-observer, intra-observer, and absolute agreement under four cases based on the corresponding ANOVA models ^a	56
Table 3.4 Mean square expectations for three-way ANOVA models based on $Y_{ijk} = \mu + \mu_i^P + \mu_k^T + \mu_j^O + \mu_{ik}^{PT} + \mu_{ij}^{PO} + \mu_{kj}^{TO} + e_{ijk}^{POT}$, where P , O , and T are random ...	78
Table 3.5 Mean square expectations for three-way ANOVA models based on $Y_{ijk} = \mu + \mu_i^P + \mu_k^t + \mu_j^O + \mu_{ik}^{Pt} + \mu_{ij}^{PO} + \mu_{kj}^{tO} + e_{ijk}^{POT}$, where P , O are random and t is fixed	79
Table 3.6 Mean square expectations for three-way ANOVA models based on $Y_{ijk} = \mu + \mu_i^P + \mu_k^T + \mu_j^o + \mu_{ik}^{PT} + \mu_{ij}^{Po} + \mu_{kj}^{To} + e_{ijk}^{PoT}$, where P , T are random and o is fixed	79

Table 3.7	Mean square expectations for three-way ANOVA models based on $Y_{ijk} = \mu + \mu_i^P + \mu_k^t + \mu_j^o + \mu_{ik}^{Pt} + \mu_{ij}^{Po} + \mu_{kj}^{to} + e_{ijk}^{Pot}$, where P is random, and o, t are fixed.	79
Table 3.8	Results of the estimates and improved 95% bootstrap- t confidence intervals of CCC under four cases for image data	81
Table 3.9	Results of the estimates and improved 95% bootstrap- t confidence intervals of ICC under four cases for image data	82

LIST OF FIGURES

Figure 1.1	Plots of ICC and CCC values as function of K with $\rho_{\mu 12} = 0.8$	11
Figure 1.2	Plots of ICC and CCC values as function of K with $\rho_{\mu 12} = 0.6$	12
Figure 1.3	Plots of ICC and CCC values as function of K with $\rho_{\mu 12} = 0.4$	13
Figure 1.4	Plots of ICC and CCC values as function of K with $\rho_{\mu 12} = 0.2$	14
Figure 2.1	Scatter plots between readers for the Anterior at the first time point in Situation 1 for vertebral body data	42
Figure 2.2	Scatter plots between readers for the Anterior at the second time point in Situation 1 for vertebral body data	43
Figure 2.3	Scatter plots between readers for the Posterior at the first time point in Situation 1 for vertebral body data	44
Figure 2.4	Scatter plots between readers for the Posterior at the second time point in Situation 1	45
Figure 2.5	Scatter plots between readers of all T locations in Situation 2 for vertebral body data	46
Figure 2.6	Scatter plots between readers for all L locations in Situation 2 for vertebral body data	47

Chapter 1

Comparison of ICC and CCC for assessing agreement for data without and with replications

1.1 Introduction

Accurate and precise measurement serves as a basis for studies in behavior and medical sciences. Agreement studies are often concerned with assessing whether different machines, raters, or methods give similar results of the measured responses. For simplicity, we use observers to standard for raters, methods, or machines between whom the agreement is assessed. For example, to establish a new method or instrument in clinical studies, one would be interested in investigating whether the new observer can replace an existing gold-standard observer if the new one is less expensive and can reproduce the same outcome, as compared to the standard one. Historically, the Pearson/product-moment correlation coefficient was used as an index to measure the association for continuous outcomes. However, as elaborated by Lin (1989), this correlation coefficient only measures the linear relationship, not agreement, between two quantitative variables. That is, the Pearson's coefficient is inappropriate for assessing the agreement because it fails to detect systematic shifts between observers (Lin, 1989; Müller and Büttner, 1994). For example, a perfect linear trend between two observers does not guarantee that these two observers can produce the same outcome. The concept of good agreement means that the readings between or among the

observers are very close. In this paper, we focus on two popular indices, the intraclass correlation coefficient (*ICC*) and the concordance correlation coefficient (*CCC*), that may take the systematic shifts into account in assessing agreement.

The intraclass correlation coefficient (Fisher, 1925) originated from genetics and was first applied to social science and then to medical science thereafter to assess agreement or reliability between observers. We focus on *ICC*s for assessing agreement between observers where there are only subject and observer effects in the models. Extensions of *ICC* for data with other factors, such as repeated measures, have been proposed by Vangeneugden et al. (2005) and are not considered in this chapter. The original *ICC* was based on the one-way ANOVA model. Extensions of this *ICC* lead to other versions of *ICC*s based on the two-way ANOVA models (Bartko, 1966). Because different versions of *ICC*s can give different results depending on the chosen ANOVA models (Bartko, 1966; Shrout and Fleiss, 1979; Müller and Büttner, 1994; McGraw and Wong, 1996), Müller and Büttner (1994) proposed simple rules to choose a suitable *ICC* with respect to the underlying data setting. However, researchers may compute these *ICC*s without verifying the assumptions and the *ICC* is biased if the ANOVA assumptions are not met. Therefore, there is a need to get a sense of the population parameter that the *ICC* estimator provides under a general setting.

Lin (1989) developed another popular index, concordance correlation coefficient (*CCC*), to assess the agreement without ANOVA assumptions. The *CCC* is well developed for assessing agreement over the past decades, which includes both precision (degree of scatter) and accuracy (degree of systematic location and scale shifts) components. The concordance correlation coefficient (Lin, 1989) is used to evaluate the agreement between paired observers ($J = 2$) for assessing the correlation between two measurements that fall on the 45° line through the origin. It was later extended to more than two observers for data without replications (Lin, 1989; King and Chinchilli, 2001; Lin et al., 2002, Barnhart et al. 2002) and for data with replications (Barnhart et al., 2005; Lin et al., 2007), where there is no reference method in either case and both of these extensions included the original *CCC* for two observers as a special case. Barnhart and Williamson (2001) proposed a generalized estimating equations (GEE) approach to model the *CCC* as a function of covariates. Chinchilli et al. (1996), King et al. (2007a), and King et al. (2007b) extended the *CCC* for paired observers with repeated measurements. Quiroz (2005) extended the *CCC* for

multiple observers with replicated measurements under the two-way ANOVA model without interaction, which is a special case of *CCC* by Barnhart et al. (2005) for data with replication.

If the ANOVA assumptions are met, the *CCC* usually reduces to the *ICC* defined by the ANOVA model (Barnhart et al. 2002). However, it is not clear what the *ICC* estimates are estimating if the ANOVA model does not hold. For data without replications, there were some comparisons of *CCC* and *ICC* when the ANOVA model does not hold (McGraw and Wong, 1996; Nickerson, 1997; Carrasco and Jover, 2003). For example, Nickerson showed that the estimates for the ICCs (ICC_2 and ICC_3 defined here in Section 1.2.1) defined by two-way ANOVA models are similar to the estimate for the *CCC*. Carrasco and Jover (2003) showed that the *CCC* corresponds mathematically to the *ICC* (ICC_2 in Section 1.2.1) defined by two-way ANOVA model without interaction even if the ANOVA assumptions are not met.

In this paper, we provide systematic and in-depth comparisons of three types of *ICC* with *CCC* for both data with and without replications. This includes reproducing the results shown by Nickerson (1997) and Carrasco and Jover (2003). In Section 1.2, we describe the notations used to summarize different versions of *ICCs* as in Barnhart et al. (2007) into three types of *ICCs* used for assessing agreement. The ANOVA models that define the types of *ICCs* and the corresponding moment estimates for data with and without replication are presented. The *CCC* is also presented for data with and without replication under the general model. We then compute the expectations of the *ICC* estimators under the general model and compare them to the *CCC* defined by the general model. A blood pressure example and a peak expiratory flow rate example are used to illustrate the comparisons of the *ICCs* and the *CCC* for data with different number of replications in Section 1.3. Finally, Section 1.4 concludes with some discussions.

1.2 Comparisons of ICC and CCC

The *ICC* is widely used for assessing agreement between observers, which is defined in terms of variance component with respect to the specific ANOVA model. There are various versions of *ICC* (Bartko, 1966, 1974; Shrout and Fleiss, 1979; Müller and Büttner, 1994; Eliasziw, et al. 1994; McGraw and Wong, 1996) shown in the literature under different

ANOVA assumptions for assessing agreement between multiple observers where none of the observers is treated as reference.

Following Barnhart et al. (2007), we use unified notations to summarize these existing versions of *ICCs* in three types under three kinds of ANOVA models for both cases of random and fixed observers. We note that one version of ICC, ICC_{3c} (shown below), is not used for comparison with the CCC because it is a measure of consistency rather than agreement. We present the definitions of the three *ICCs* and the *CCC* for data without replication ($K = 1$) and with replications ($K > 1$) together. However, comparisons of *ICC* estimators with *CCC* is presented separately for data without replication ($K = 1$) and with replications ($K > 1$) in Section 1.2.2. The definitions of three types of *ICC* and *CCC* are presented below. Following the existing literature, the corresponding estimates for *ICCs* are obtained through the method of moments based on the expectation of the mean sums of squares from the ANOVA models.

1.2.1 Definitions and estimates of *ICCs* and *CCC*

Let Y_{ijk} be the k th replicated reading for observer j on the subject i , $i = 1, \dots, N$, $j = 1, \dots, J$, $k = 1, \dots, K$. The first and original *ICC*, ICC_1 , is defined under the one-way ANOVA model given by

$$Y_{ijk} = \mu + \alpha_i + e_{ijk},$$

where μ is the overall effect common to all methods, α_i is the random effect of subject i , which is independent and identically distributed (*i.i.d.*) with normal distribution $N(0, \sigma_\alpha^2)$, and e_{ijk} is the random error, which is *i.i.d.* with normal distribution $N(0, \sigma_e^2)$; α_i and e_{ijk} are mutually independent. The ICC_1 and the corresponding estimator (Bartko, 1966; Shrout and Fleiss, 1979; McGraw and Wong, 1996) are

$$ICC_1 = \frac{\sigma_\alpha^2}{\sigma_\alpha^2 + \sigma_e^2}, \quad \widehat{ICC}_1 = \frac{MS_\alpha - MS_e}{MS_\alpha + (JK - 1)MS_e},$$

where $MS_\alpha = \frac{JK}{N-1} \sum_{i=1}^N (\bar{Y}_{i..} - \bar{Y}_{...})^2$ and $MS_e = \frac{1}{(JK-1)N} \sum_{i=1}^N \sum_{j=1}^J \sum_{k=1}^K (Y_{ijk} - \bar{Y}_{i..})^2$ are the means of the sums of squares from the one-way ANOVA model for between and within subjects, respectively. $\bar{Y}_{i..} = \sum_{j=1}^J \sum_{k=1}^K Y_{ijk} / (JK)$, $\bar{Y}_{...} = \sum_{i=1}^N \sum_{j=1}^J \sum_{k=1}^K Y_{ijk} / (NJK)$.

If $K = 1$,

$$\widehat{ICC}_1 = \frac{MS_\alpha - MS_e}{MS_\alpha + (J - 1)MS_e}.$$

The second and third ICC s are based on two-way ANOVA model without interaction and with interaction, respectively. The two-way ANOVA model without interaction is given by

$$Y_{ijk} = \mu + \alpha_i + \beta_j + e_{ijk},$$

where μ , α_i , and e_{ijk} are as before. If observers are treated as random, then β_j is the random effect of observer j , which is *i.i.d.* with normal distribution $N(0, \sigma_\beta^2)$ and α_i , β_j , and e_{ijk} are mutually independent. If observers are treated as fixed, we use notation $\sigma_\beta^2 = \sum_{j=1}^J \beta_j^2 / (J - 1)$ with constraint $\sum_{j=1}^J \beta_j = 0$. Then ICC_2 and the corresponding estimator (McGraw and Wong, 1996) are

$$ICC_2 = \frac{\sigma_\alpha^2}{\sigma_\alpha^2 + \sigma_\beta^2 + \sigma_e^2}, \quad \widehat{ICC}_2 = \frac{MS_\alpha - MS_e}{MS_\alpha + (JK - 1)MS_e + J(MS_\beta - MS_e)/N},$$

where $MS_\alpha = \frac{JK}{N-1} \sum_{i=1}^N (\bar{Y}_{i..} - \bar{Y}_{...})^2$, $MS_e = \frac{1}{(JK-1)N-J+1} \sum_{i=1}^N \sum_{j=1}^J \sum_{k=1}^K (Y_{ijk} - \bar{Y}_{i..} - \bar{Y}_{.j.} + \bar{Y}_{...})^2$, and $MS_\beta = \frac{NK}{J-1} \sum_{j=1}^J (\bar{Y}_{.j.} - \bar{Y}_{...})^2$. If $K = 1$,

$$\widehat{ICC}_2 = \frac{MS_\alpha - MS_e}{MS_\alpha + (J - 1)MS_e + J(MS_\beta - MS_e)/N}.$$

Considering the interaction between observer and subject, the two-way ANOVA model with interaction is given by

$$Y_{ijk} = \mu + \alpha_i + \beta_j + \gamma_{ij} + e_{ijk},$$

where μ , α_i , β_j , and e_{ijk} are as before; γ_{ij} is the random effect of the observer-subject interaction, which is *i.i.d.* with $N(0, \sigma_\gamma^2)$; α_i , β_j , γ_{ij} , and e_{ijk} are mutually independent if β_j is random. If β_j is fixed, notation $\sigma_\beta^2 = \sum_{j=1}^J \beta_j^2 / (J - 1)$ is used with constraints $\sum_{j=1}^J \beta_j$

and $\sum_{j=1}^J \gamma_{ij} = 0$. The ICC_3 and the corresponding estimator (Bartko, 1966; Shrout and Fleiss, 1979; McGraw and Wong, 1996) are

$$ICC_3 = \frac{\sigma_\alpha^2}{\sigma_\alpha^2 + \sigma_\beta^2 + \sigma_\gamma^2 + \sigma_e^2} \text{ (random } \beta_j), \quad ICC_3 = \frac{\sigma_\alpha^2 - \sigma_\gamma^2 / (J - 1)}{\sigma_\alpha^2 + \sigma_\beta^2 + \sigma_\gamma^2 + \sigma_e^2} \text{ (fixed } \beta_j),$$

$$\widehat{ICC}_3 = \frac{MS_\alpha - MS_\gamma}{MS_\alpha + J(K - 1)MS_e + (J - 1)MS_\gamma + J(MS_\beta - MS_\gamma)/N},$$

where $MS_\alpha = \frac{JK}{N-1} \sum_{i=1}^N (\bar{Y}_{i..} - \bar{Y}_{...})^2$, $MS_e = \frac{1}{NJ(K-1)} \sum_{i=1}^N \sum_{j=1}^J \sum_{k=1}^K (Y_{ijk} - \bar{Y}_{ij.})^2$, $MS_\beta = \frac{NK}{J-1} \sum_{j=1}^J (\bar{Y}_{.j.} - \bar{Y}_{...})^2$, and $MS_\gamma = \frac{K}{(J-1)(N-1)} \sum_{i=1}^N \sum_{j=1}^J (\bar{Y}_{ij.} - \bar{Y}_{i..} - \bar{Y}_{.j.} + \bar{Y}_{...})^2$. If $K = 1$, MS_γ is simply nonestimable and therefore is included in MS_e , thus

$$\widehat{ICC}_3 = \frac{MS_\alpha - MS_e}{MS_\alpha + (J - 1)MS_e + J(MS_\beta - MS_e)/N} = \widehat{ICC}_2.$$

Note that in the case where observers are treated as fixed effect to have a two-way mixed effect model, Bartko (1966) (and later corrected by Bartko (1974) and Shrout and Fleiss (1979)) defined the following version of ICC based on correlation:

$$ICC_{3c} = \frac{\sigma_\alpha^2 - \sigma_\gamma^2 / (J - 1)}{\sigma_\alpha^2 + \sigma_\gamma^2 + \sigma_e^2},$$

McGraw and Wong (1996) called this ICC as the ICC for consistency since the systematic deviation due to method, σ_β^2 , is not included in the denominator. Because this ICC is a measure for consistency and CCC is a measure of agreement, we do not compare ICC_{3c} to CCC in this paper.

We present the CCC below for assessing agreement between J observers for data with K replications ($K = 1$ or $K > 1$) under the general model, $Y_{ijk} = \mu_{ij} + e_{ijk}$ (Barnhart et al., 2005). We use the following minimal assumptions if the K readings measured by the same observer on a subject are true replications. (1) μ_{ij} and e_{ijk} are independent with means of $E(\mu_{ij}) = \mu_j$ and $E(e_{ijk}) = 0$; (2) between-subject and within-subject variances are $Var(\mu_{ij}) = \sigma_{Bj}^2$ and $Var(e_{ijk}) = \sigma_{Wj}^2$, respectively; (3) $Corr(\mu_{ij}, \mu_{ij'}) = \rho_{\mu jj'}$, $Corr(\mu_{ij}, e_{ijk}) = 0$, $Corr(e_{ijk}, e_{ijk'}) = 0$ for all j, j', k, k' . In addition, $\sigma_j^2 = \sigma_{Bj}^2 + \sigma_{Wj}^2$

denotes the total variability of observer j , and $\rho_{jj'} = \text{Corr}(Y_{ijk}, Y_{ij'k'})$ denotes the pairwise correlation between two readings from observers j and j' . The *CCC* is defined as

$$\begin{aligned}\rho_c &= 1 - \frac{\sum_{j=1}^{J-1} \sum_{j'=j+1}^J E(Y_{ijk} - Y_{ij'k'})^2}{\sum_{j=1}^{J-1} \sum_{j'=j+1}^J E_I(Y_{ijk} - Y_{ij'k'})^2} = \frac{2 \sum_{j=1}^{J-1} \sum_{j'=j+1}^J \sigma_j \sigma_{j'} \rho_{jj'}}{(J-1) \sum_{j=1}^J \sigma_j^2 + \sum_{j=1}^{J-1} \sum_{j'=j+1}^J (\mu_j - \mu_{j'})^2} \\ &= \frac{\frac{1}{J-1} \sum_{j=1}^J \sum_{j'=1(j' \neq j)}^J \sigma_{Bj} \sigma_{Bj'} \rho_{\mu jj'}}{\sum_{j=1}^J \sigma_{Bj}^2 + \sum_{j=1}^J \sigma_{Wj}^2 + (\sum_{j=1}^J \mu_j^2 - \frac{1}{J-1} \sum_{j=1}^J \sum_{j'=1(j' \neq j)}^J \mu_j \mu_{j'})},\end{aligned}$$

where E_I is the expectation under independence of $Y_{ijk}, Y_{ij'k'}$. This *CCC* was referred to as a fixed marginal agreement coefficient (FMAC) by Fay (2005). Fay also proposed a random marginal agreement coefficient (RMAC) if the $E_I(Y_{ijk} - Y_{ij'k'})^2$ in the denominator is replaced by $E_{Z_j} E_{Z_{j'}} (Z_j - Z_{j'})^2$, where Z_j and $Z_{j'}$ are independent and identically distributed random variables with a mixture distribution of $0.5Y_j + 0.5Y_{j'}$. This RMAC is closely related to $E(\widehat{ICC}_1)$ as shown in Section 1.2.2.

For the special case when $J=2$ and $K=1$ (without replication), the *CCC* is reduced to the following original *CCC* by Lin (1989).

$$\begin{aligned}\rho_c &= 1 - \frac{E(Y_{i11} - Y_{i21})^2}{E((Y_{i11} - Y_{i21})^2 | \text{corr}(Y_{i11}, Y_{i21}) = 0)} \\ &= \frac{2\sigma_{B1}\sigma_{B2}\rho_{\mu 12}}{\sigma_{B1}^2 + \sigma_{W1}^2 + \sigma_{B2}^2 + \sigma_{W2}^2 + (\mu_1 - \mu_2)^2} \\ &= \frac{2\sigma_1\sigma_2\rho_{12}}{\sigma_1^2 + \sigma_2^2 + (\mu_1 - \mu_2)^2}\end{aligned}$$

with $\rho_{12} = \text{Corr}(Y_{i11}, Y_{i21})$. Note here that the *ICC*s are defined for both random and fixed cases where the observers are treated as, while *CCC* is usually defined for fixed observers.

1.2.2 Comparisons of $E(\widehat{ICC})$ and *CCC* under the general model

We are interested in comparing the expectations under the general model of *ICC* estimators stated in Section 1.2.1 to the *CCC* under the same model $Y_{ijk} = \mu_{ij} + e_{ijk}$. We

approximate $E(\widehat{ICC})$ by taking the expectation of the numerator and the denominator of \widehat{ICC} defined in Section 1.2.1 under this general model. The expectations are presented for data without and with replications. Please note below that even $E(\widehat{ICC}_3) = CCC$ for all K , one may still have $\widehat{ICC}_3 \geq CCC$ as indicated by Nickerson (1997) and Carrasco and Jover (2003). This is because that $(\bar{Y}_{.j} - \bar{Y}_{.j'})^2$ is usually used (Lin, 1989) to estimate $(\mu_j - \mu_{j'})^2$ in CCC . However, $(\bar{Y}_{.j} - \bar{Y}_{.j'})^2$ is a biased estimator of $(\mu_j - \mu_{j'})^2$ as pointed out by Carrasco and Jover (2003). In general, the bias is small for moderate and large sample size. If this bias is corrected, then we have $\widehat{ICC}_3 = CCC$ as indicated by Nickerson (1997) for the case of $K = 1$.

For data without replication, it can be shown that

$$\begin{aligned} E(\widehat{ICC}_1) &\approx \frac{E(MS_\alpha - MS_e)}{E(MS_\alpha + (J-1)MS_e)} \\ &= \frac{(\frac{1}{J-1}) \sum_{j=1}^J \sum_{j'=1(j' \neq j)}^J \sigma_{Bj} \sigma_{Bj'} \rho_{\mu jj'} - \frac{1}{J-1} (\sum_{j=1}^J \mu_j^2 - \frac{1}{J} (\sum_{j=1}^J \mu_j)^2)}{\sum_{j=1}^J \sigma_{Bj}^2 + \sum_{j=1}^J \sigma_{Wj}^2 + \sum_{j=1}^J \mu_j^2 - \frac{1}{J} (\sum_{j=1}^J \mu_j)^2} \\ &\neq CCC. \end{aligned}$$

Similarly,

$$\begin{aligned} E(\widehat{ICC}_2) &\approx \frac{E(MS_\alpha - MS_e)}{E(MS_\alpha + (J-1)MS_e + J(MS_\beta - MS_e)/N)} \\ &= \frac{(\frac{1}{J-1}) \sum_{j=1}^J \sum_{j'=1(j' \neq j)}^J \sigma_{Bj} \sigma_{Bj'} \rho_{\mu jj'}}{\sum_{j=1}^J \sigma_{Bj}^2 + \sum_{j=1}^J \sigma_{Wj}^2 + (\sum_{j=1}^J \mu_j^2 - \frac{1}{J-1} \sum_{j=1}^J \sum_{j'=1(j' \neq j)}^J \mu_j \mu_{j'})} \\ &= CCC \end{aligned}$$

Because $\widehat{ICC}_3 = \widehat{ICC}_2$ when $K = 1$, thus, $E(\widehat{ICC}_3) = E(\widehat{ICC}_2) = CCC$. For the special case with $J = 2$ (two observers), we have

$$\begin{aligned} E(\widehat{ICC}_1) &\approx \frac{E(MS_\alpha - MS_e)}{E(MS_\alpha + MS_e)} \\ &= \frac{2\sigma_1\sigma_2\rho_{12} - \frac{1}{2}(\mu_1 - \mu_2)^2}{\sigma_1^2 + \sigma_2^2 + \frac{1}{2}(\mu_1 - \mu_2)^2} \\ &\neq CCC. \end{aligned}$$

It is interesting to note that this $E(\widehat{ICC_1})$ based on the one-way ANOVA model is equivalent to the RMAC by Fay (2005) when there are two observers and no replication. We think that this is because the scaling term $E_{Z_1}E_{Z_2}(Z_1 - Z_2)^2$ in the denominator of Fay's RMAC equals to two times the total variability used in ICC_1 based on the one way ANOVA model. The fact that Z_1 and Z_2 are *i.i.d.* with a distribution of $0.5Y_1 + 0.5Y_2$ is similar to assuming that there is no observer effect in the one way ANOVA model. In addition, the natural generalization of RMAC to define RMAC for the case of $J > 2$ and $K > 1$ is to replace the term $E_I(Y_{ijk} - Y_{ij'k})^2$ in the definition of CCC by $E_{Z_j}E_{Z_{j'}}(Z_j - Z_{j'})^2$. Because Fay's RMAC is equivalent to ICC_1 for $J = 2$ and $K = 1$, we believe that the generalizations of Fay's RMAC will be the same as ICC_1 with one-way ANOVA model based on the above observation.

For the other two ICCs with $J = 2$, we have,

$$\begin{aligned}
E(\widehat{ICC_2}) &= E(\widehat{ICC_3}) \\
&\approx \frac{E(MS_\alpha - MS_e)}{E(MS_\alpha + MS_e + 2(MS_\beta - MS_e)/N)} \\
&= \frac{2\sigma_1\sigma_2\rho_{12}}{\sigma_1^2 + \sigma_2^2 + (\mu_1 - \mu_2)^2} \\
&= CCC.
\end{aligned}$$

This confirms the same findings shown by Nickerson (1997) and Carrasco and Jover (2003) for data without replications.

For data with replications, the expressions for the expected values of $\widehat{ICC_1}$, $\widehat{ICC_2}$, and $\widehat{ICC_3}$ under the general model $Y_{ijk} = \mu_{ij} + e_{ijk}$ are presented in the Appendix due to the complexity of the formulas. As shown in Appendix A, $E(\widehat{ICC_1})$, which in general does not equal to CCC , depends on the number of observer J and the number of replication K . $E(\widehat{ICC_2})$, which in general does not equal to CCC , depends on the number of subject N , the number of observer J and the number of replication K . However, $E(\widehat{ICC_3})$ is similar to CCC that only depends on the number of observer J . In general, if the data comes from a model that is different from the ANOVA model used to define ICC, the ICC will be biased and will not concur with CCC that is defined without ANOVA model. In addition, if data

came from a model with no interaction, then ICC_2 is equivalent to ICC_3 . However, there may be a loss of efficiency of using ICC_3 in relation to ICC_2 because extra parameters are estimated in ICC_3 .

To better understand the dependency of ICC_1 and ICC_2 on the number of replications, we consider a special case of $J = 2$ where the between-subject variances of two observers are the same ($\sigma_{B1}^2 = \sigma_{B2}^2 = \sigma_B^2$), and the within-subject variances of two observers are the same ($\sigma_{W1}^2 = \sigma_{W2}^2 = \sigma_W^2$), especially when K is large and goes to infinity, then

$$\lim_{K \rightarrow \infty} E(\widehat{ICC_1}) \approx \frac{(1 + \rho_{\mu 12})\sigma_B^2}{2\sigma_B^2 + 2\sigma_W^2 + \frac{1}{2}(\mu_1 - \mu_2)^2} \geq CCC,$$

and

$$\lim_{K \rightarrow \infty} E(\widehat{ICC_2}) \approx \frac{(1 + \rho_{\mu 12})\sigma_B^2}{(2 + \frac{1}{N} - \frac{1}{N}\rho_{\mu 12})\sigma_B^2 + 2\sigma_W^2 + (\mu_1 - \mu_2)^2} \geq CCC.$$

In this special case, $\lim_{K \rightarrow \infty} E(\widehat{ICC_1}) = CCC$ only if $\mu_1 = \mu_2$, and $\lim_{K \rightarrow \infty} E(\widehat{ICC_2}) = CCC$ only if $\rho_{\mu 12} = 1$.

To illustrate the dependency of $E(\widehat{ICC_1})$ and $E(\widehat{ICC_2})$ on the number of replications, we plot $E(\widehat{ICC_1})$, $E(\widehat{ICC_2})$, and CCC verses K in Figures 1.1-1.4. In these figures, we mimic the BP data in the Example Section by setting $N = 85$, $\sigma_{B1}^2 = \sigma_{B2}^2 = 955$, and $\sigma_{W1}^2 = \sigma_{W2}^2 = 63$. In addition, we set $\mu_1 = 128$ and $\mu_2 = 144$. Figures 1.1-1.4 display plots of ICC and CCC values as function of K with $\rho_{\mu 12} = 0.8, 0.6, 0.4$, and 0.2 , respectively. The value of CCC is plotted as a horizontal line because it does not depend on K .

From these figures, we see that $E(\widehat{ICC_1})$ is less than both of $E(\widehat{ICC_2})$ and CCC when $K = 1$ and increases quickly to the limit when $K \rightarrow \infty$, which exceeds both of $E(\widehat{ICC_2})$ and CCC . $E(\widehat{ICC_2})$ equals CCC when $K = 1$ and increases quickly to the limit when $K \rightarrow \infty$, which also exceeds CCC . For this special case, both $E(\widehat{ICC_1})$ and $E(\widehat{ICC_2})$ are increasing functions of K .

1.3 Examples

We first used the systolic blood pressure data from Bland and Altman (1999) to illustrate the estimated values of ICC s and CCC . In this data set, two observers and

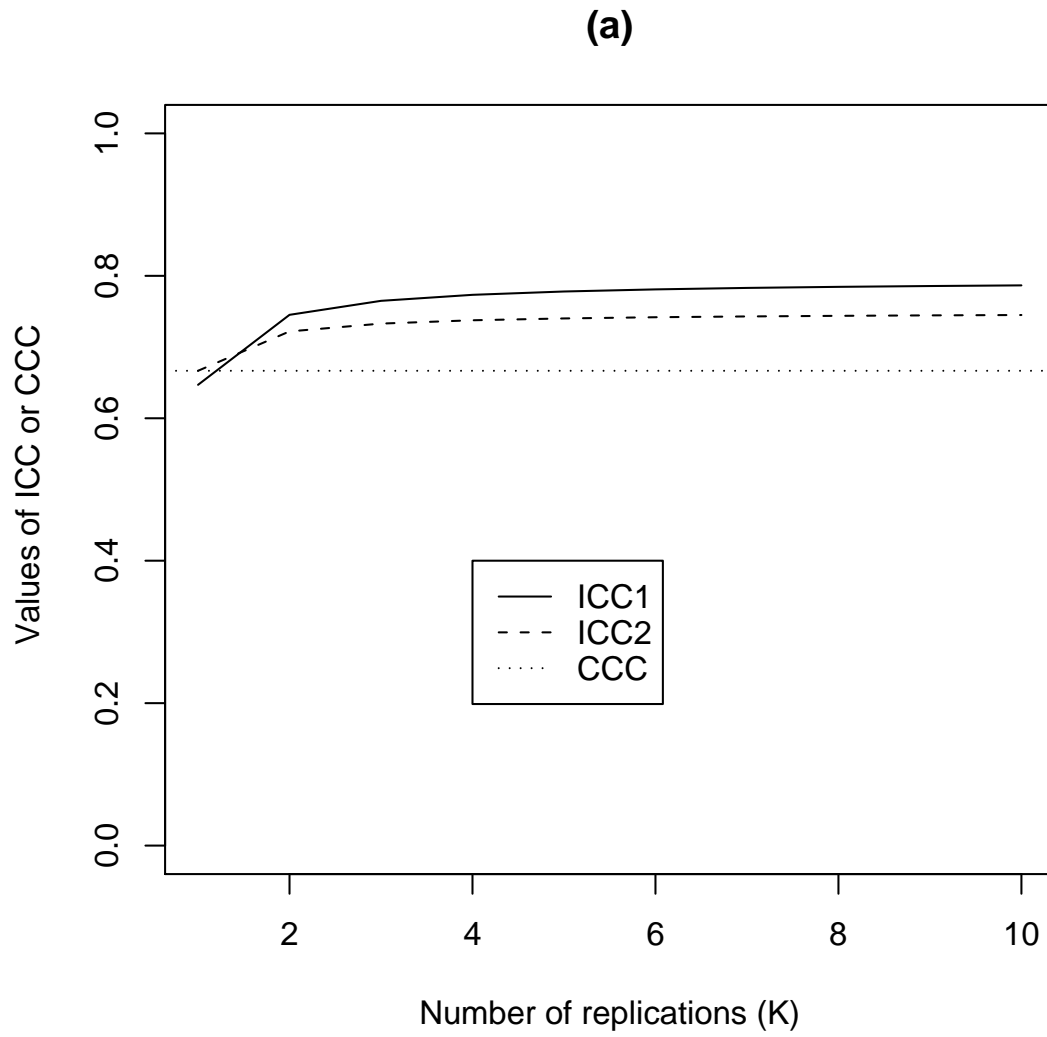


Figure 1.1: Plots of ICC and CCC values as function of K with $\rho_{\mu 12} = 0.8$

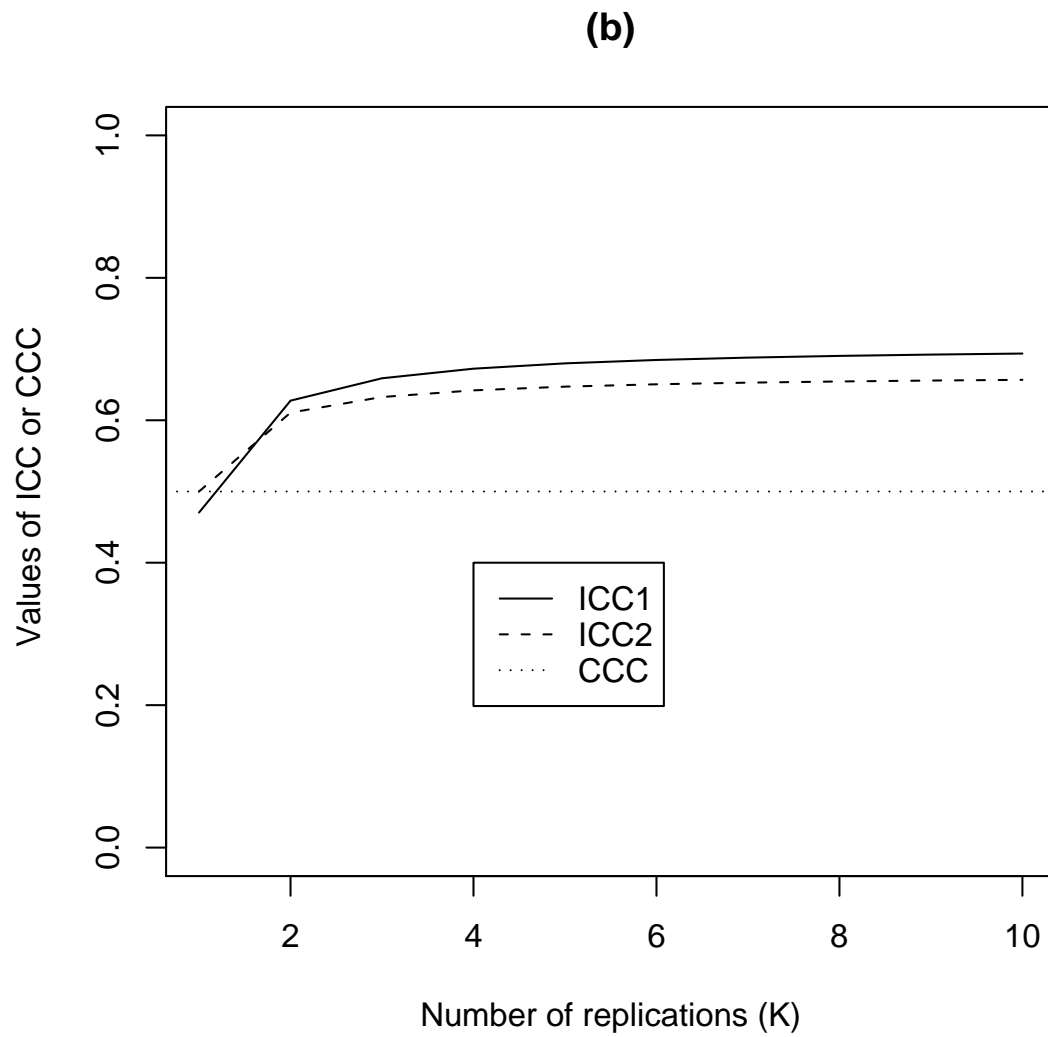


Figure 1.2: Plots of ICC and CCC values as function of K with $\rho_{\mu 12} = 0.6$

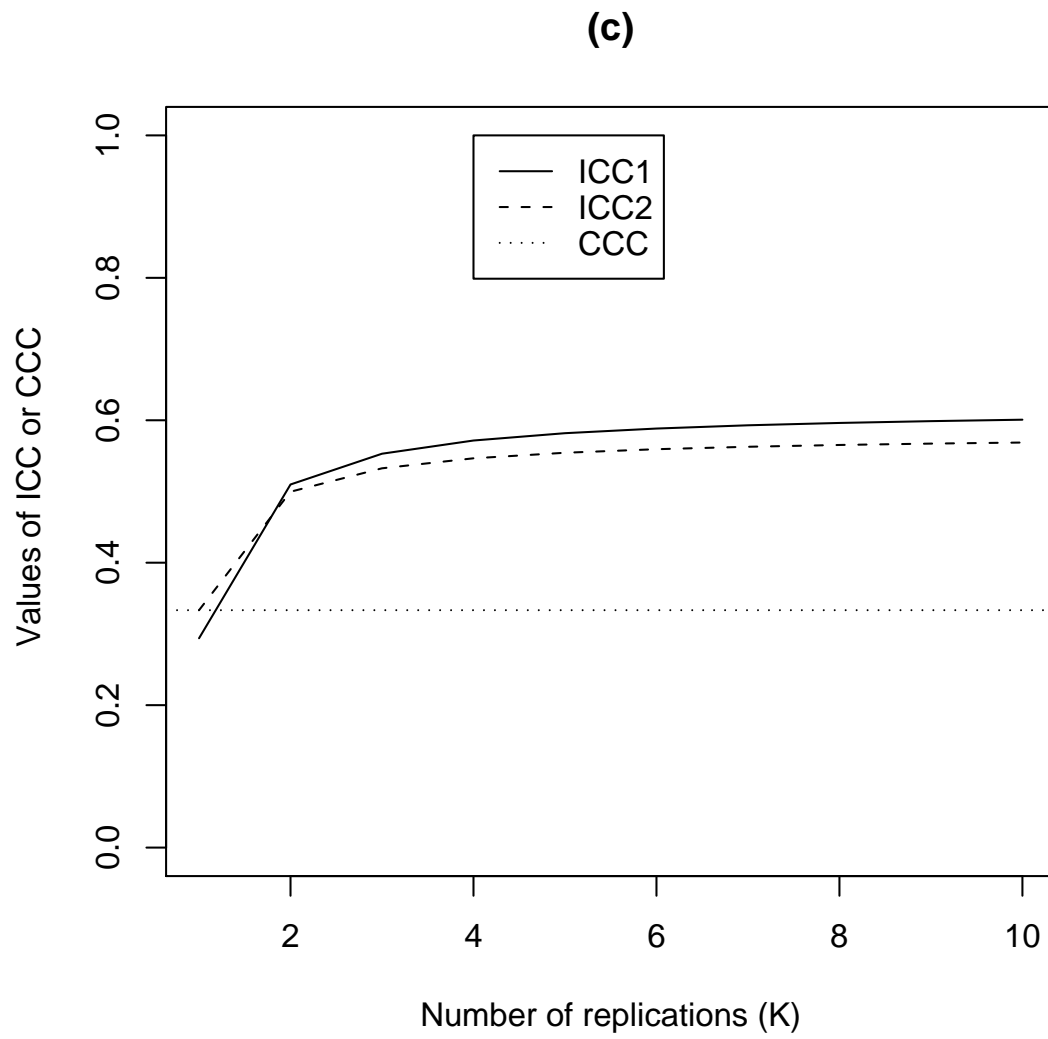


Figure 1.3: Plots of ICC and CCC values as function of K with $\rho_{\mu 12} = 0.4$

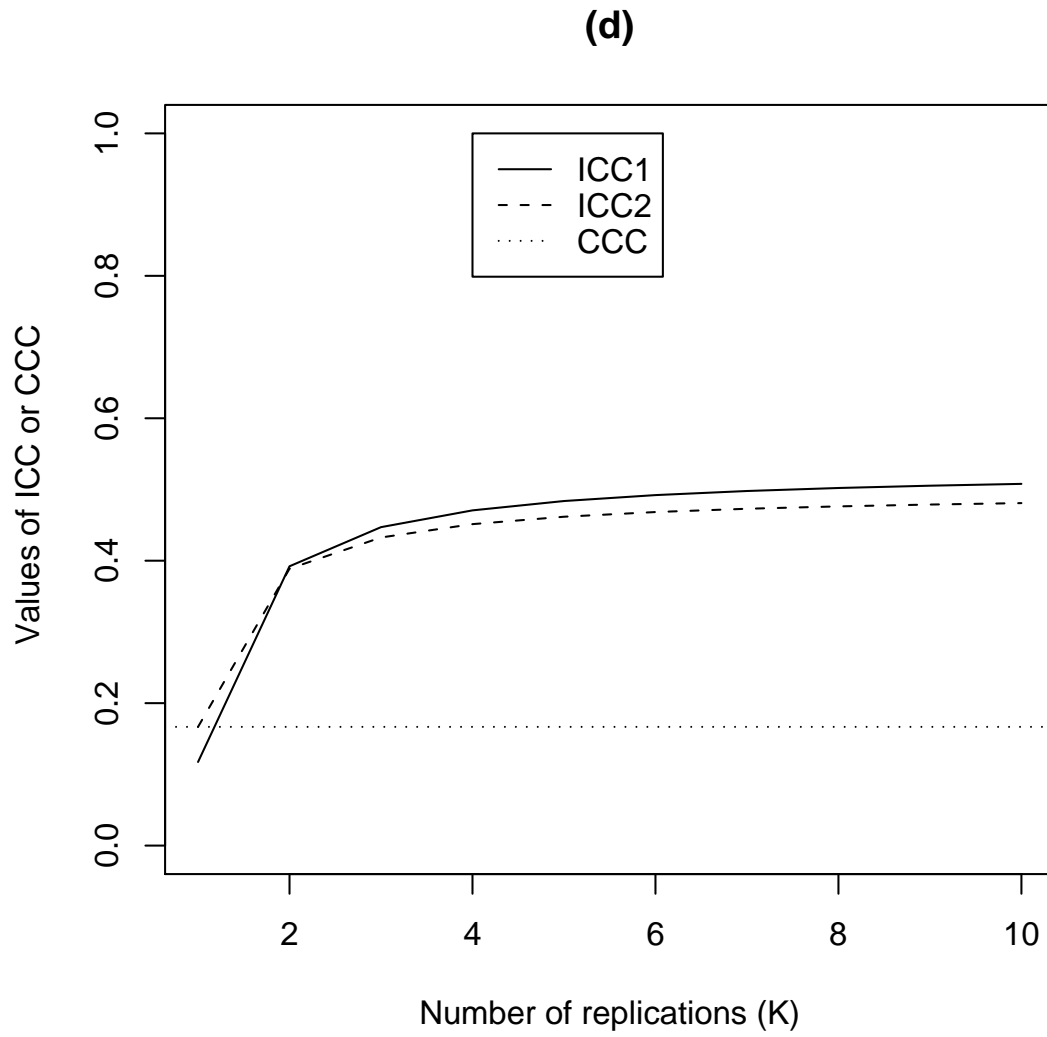


Figure 1.4: Plots of ICC and CCC values as function of K with $\rho_{\mu 12} = 0.2$

an automatic machine made three quick successive observations on 85 subjects' (N=85) systolic blood pressure. For illustration, we compute estimates of three *ICC*s and *CCC* to assess agreement between the first observer and the automatic machine (J=2) for three scenarios:

- (1) K=1 where only the first replication is used.
- (2) K=2 where only the first two replications are used.
- (3) K=3 where all three replications are used.

Table 1.1 shows the means of the sums of squares from the one-way ANOVA model, two-way ANOVA model without interaction, and two-way ANOVA model with interaction for different replications. The point estimations of *ICC*s and *CCC*s are shown in Table 1.2, where \widehat{CCC}^a and \widehat{CCC}^b are the estimates of *CCC* without and with bias corrections for estimating $(\mu_1 - \mu_2)^2$, respectively. Without the bias correction for estimating $(\mu_1 - \mu_2)^2$ in \widehat{CCC} , the \widehat{ICC}^a may be slightly larger than \widehat{CCC} and even theoretically we have $E(\widehat{ICC}) = E(\widehat{CCC})$. We summarize all the estimates of the parameters used in the definition of *CCC* by number of replications in Table 1.3. Please note that when $K = 1$, parameters $\rho_{\mu 12}$, $\sigma_{B_1}^2$, $\sigma_{B_2}^2$, $\sigma_{W_1}^2$, and $\sigma_{W_2}^2$ are nonestimable for data without replications. The \widehat{ICC}_1 is less than \widehat{CCC}^a when $K = 1$ and is greater than \widehat{CCC}^a when $K > 1$. The \widehat{ICC}_2 is close to and greater than \widehat{CCC}^a when $K = 1$ but it is less than \widehat{ICC}_1 when $K > 1$. As expected, $\widehat{ICC}_2 = \widehat{ICC}_3$ that is close to and greater than \widehat{CCC}^a and equivalent to \widehat{CCC}^b when $K = 1$. The \widehat{ICC}_3 is similar to and slightly greater than \widehat{CCC}^a when $K > 1$ even though we have $E(\widehat{ICC}_3) = CCC$. However, the \widehat{ICC}_3 is close to \widehat{CCC}^b because of the bias correction term. These observations confirm the findings in Section 2 and indicate that \widehat{ICC}_3 is similar to \widehat{CCC} while \widehat{ICC}_1 or \widehat{ICC}_2 can be substantially larger than \widehat{CCC} .

Table 1.1: The ANOVA table for ICC s for blood pressure data by number of replicates

Replicates	ANOVA for ICC_1		ANOVA for ICC_2			ANOVA for ICC_3			
	MS_α	MS_e	MS_α	MS_β	MS_e	MS_α	MS_β	MS_γ	MS_e
$K = 1$	1922.24	322.78	1922.24	11283.78	192.30	1922.24	11283.68	192.30	0
$K = 2$	3688.00	249.10	3688.00	21409.42	165.79	3688.00	21409.42	375.69	62.08
$K = 3$	5337.80	227.63	5337.80	31106.45	154.87	5337.80	31106.45	537.74	60.27

Table 1.2: Estimates of the ICC s and CCC s for blood pressure data by number of replicates

	\widehat{ICC}_1	\widehat{ICC}_2	\widehat{ICC}_3	\widehat{CCC}^a	\widehat{CCC}^b
K=1	.712	.728	.728	.727	.728
K=2	.775	.752	.707	.706	.708
K=3	.789	.758	.702	.701	.702

^a Moment estimator of CCC without a bias adjustment^b With adjustment of a bias correction term for estimating $(\mu_1 - \mu_2)^2$

Table 1.3: Estimates of individual parameters for blood pressure data by number of replicates

	$\hat{\rho}_{12}$	$\hat{\rho}_{\mu 12}$	$\hat{\mu}_1$	$\hat{\mu}_2$	$\hat{\sigma}_{B_1}^2$	$\hat{\sigma}_{B_2}^2$	$\hat{\sigma}_{W_1}^2$	$\hat{\sigma}_{W_2}^2$
K=1	.82	-	128.54	144.84	-	-	-	-
K=2	.79	.84	127.92	143.79	957.30	1012.47	35.36	88.79
K=3	.79	.83	127.41	143.03	935.13	983.19	37.41	83.14

As a second example, we used the data of peak expiratory flow rate (PEFR) from Bland and Altman (1986). In this data set, two replications ($K = 2$) were made by each of the two methods ($J = 2$) for each of the 17 subjects ($N = 17$). The two methods are Wright peak flow meter and mini Wright meter. Similar to the first example, we compute estimates of three *ICCs* and *CCC* to assess agreement between the Wright peak flow meter and the mini Wright meter ($J=2$) for two scenarios when (1) $K=1$ where only the first replication is used and (2) $K=2$ where two replications are used.

Table 1.4 shows the means of the sums of squares from the one-way ANOVA model, two-way ANOVA model without interaction, and two-way ANOVA model with interaction for different replications. The point estimations of *ICCs* and *CCC* are shown in Table 1.5, where \widehat{CCC}^a and \widehat{CCC}^b are estimates of *CCC* without and with bias corrections for estimating $(\mu_1 - \mu_2)^2$, respectively. As in the first example, we summarize all the estimates of the parameters used in the definition of *CCC* by number of replications in Table 1.6. When $K = 1$, we see that all *ICCs* and *CCCs* are similar because there is no much systematic shifts between $\hat{\mu}_1$ and $\hat{\mu}_2$. However, when $K = 2$, both \widehat{ICC}_1 and \widehat{ICC}_2 are larger than \widehat{ICC}_3 and \widehat{CCC} due to dependency of ICC_1 and ICC_2 on number of replications. As expected, \widehat{ICC}_3 and \widehat{CCC} are similar.

Table 1.4: The ANOVA table for ICC s for PEFR data by number of replicates

Replicates	ANOVA for ICC_1		ANOVA for ICC_2			ANOVA for ICC_3			
	MS_α	MS_e	MS_α	MS_β	MS_e	MS_α	MS_β	MS_γ	MS_e
$K = 1$	25572.26	709.41	25572.26	38.12	751.37	25572.26	38.12	751.37	0
$K = 2$	51268.85	568.25	51268.85	618.01	567.25	51268.85	618.01	1102.51	315.37

Table 1.5: Estimates of the ICC s and CCC s for PEFR data by number of replicates

	\widehat{ICC}_1	\widehat{ICC}_2	\widehat{ICC}_3	\widehat{CCC}^a	\widehat{CCC}^b
K=1	.946	.946	.946	.943	.946
K=2	.957	.957	.948	.945	.948

^a Moment estimator of CCC without a bias adjustment^b With adjustment of a bias correction term for estimating $(\mu_1 - \mu_2)^2$

Table 1.6: Estimates of individual parameters for PEFR data by number of replicates

	$\hat{\rho}_{12}$	$\hat{\rho}_{\mu 12}$	$\hat{\mu}_1$	$\hat{\mu}_2$	$\hat{\sigma}_{B_1}^2$	$\hat{\sigma}_{B_2}^2$	$\hat{\sigma}_{W_1}^2$	$\hat{\sigma}_{W_2}^2$
K=1	0.94	-	450.35	452.47	-	-	-	-
K=2	0.95	0.97	447.88	453.91	13683	12188	234.29	396.44

1.4 Discussion

There are several versions of intraclass correlation coefficient discussed in the literature and we summarize these versions in three types for data with only factors of subject and observer. Based on the definition of ICC shown in Section 1.2.1, we compare the expected value of the \widehat{ICC} to another popular agreement index CCC under the general model by taking the expectation of the numerator and the denominator of \widehat{ICC} for data without and with replications.

We find that the expectation of \widehat{ICC}_1 and \widehat{ICC}_2 depend on the number of replication K , while the expectation of \widehat{ICC}_3 and CCC do not depend on K . In summary, for data without replication ($K = 1$), $E(\widehat{ICC}_2)$ and $E(\widehat{ICC}_3)$ are essentially the same as CCC , while $E(\widehat{ICC}_1)$ may be greater or less than CCC . For data with replications ($K > 1$), in general, $E(\widehat{ICC}_1)$ and $E(\widehat{ICC}_2)$ are greater than CCC for moderate and large K , and $E(\widehat{ICC}_3) = CCC$ regardless of K . If data comes from a model where there is subject by observer interaction and the ANOVA model does not include such effect, the ICC defined on this model will be biased and will not concur with the CCC defined on a model that accounts for such interaction. Conversely, if the ANOVA model includes all the significant effects the ICC will be the same as the CCC defined on the general model, although influence may be different. In this paper, we only compare the point estimations of ICC s and CCC . Further research is needed to compare the inference of these point estimators to assess the efficiency of estimating ICC s and CCC .

Based on these findings, if ICC is the choice of agreement index, we recommend to use ICC_3 over all other ICC s as it gives similar estimates as the CCC where it does not require ANOVA assumptions. We believe that one does not loss much by recommending ICC_3 other than possibly losing efficiency by estimating extra parameters if simpler ANOVA models hold. Unless one has a large sample size, usually it is difficult to detect interaction effect and thus realistically checking the assumptions is often not possible in practice. If the simpler ANOVA model is correct, the ICC_3 will reduce to the ICC corresponding to the simpler model. However, one may use other ICC s based on the simpler ANOVA model if one believes that the simpler model holds or there is sufficient power to check the assumptions.

Chapter 2

The random-observers concordance correlation coefficient

2.1 Introduction

The intraclass correlation coefficient (ICC) and the concordance correlation coefficient (CCC) are frequently used for assessing agreement between or among observers for data with continuous variables. Usually, the ICC and CCC are used for data without and with replications, and they are based on the subject and observer effects only. However, we can not use the methodology if repeated measurements rather than replications are collected. In practice, a true replication, i.e. nothing changed other than the times of the measurements, is often difficult to obtain since it is not easy to maintain the same condition and environment. Oftentimes, repeated measurements rather than replications are taken over time. The time effect should be considered in the measurement in order to account properly different sources of error.

There exist some ICC and CCC type indices for assessing agreement with repeated measurements. Vangeneugden et al. (2005) proposed a linear mixed model approach for assessing the agreement for data with repeated measures. By applying concepts of generalizability theory, they considered several types of ICC -like agreement indices. In their linear mixed models, both observer and time are treated as random factors. King et al. (2007a; 2007b) proposed a class of agreement indices of CCC for data with repeated measures, where the observer and time are treated as fixed factors. Traditionally, for data with

replications, the *ICC* has been proposed for both random and fixed observers, while the *CCC* is developed for fixed observer where subjects are treated as random effects. However, there is no *CCC* for random observers and random time points, while there is a need to develop *ICC* and *CCC* type agreement indices with different combinations of random or fixed factors for observers and times.

We investigate *CCC* agreement index for data with random factors for observers and times since researchers may be interested in assessing many observers who take measurements at different time points. For example, when measuring blood pressure, we may want to know whether observers (e.g. nurses) can make reliable measurements at any time, and we are interested in whether these observers can be used interchangeably at any time. Unfortunately, it is difficult or even expensive to use all observers to make measurements at any time. Therefore, observers can be randomly selected from the nurse population to make measurements at randomly selected time points. Since blood pressure fluctuates overtime, their readings should be considered as repeated rather than replicated measurements. Consider another example in image study where different observers read the recorded images at different time throughout the day. It is not clear whether the observers can make similar measurements at different times. Thus, there is a need to assess the closeness between measurements on the same image by randomly selected observers at randomly selected time. In longitudinal studies, oftentimes different observers are need to take measurements at different follow-up times and we want to make sure these measurements are accurate. Consider all observers used in the study as the observer population and time points as the time population and we may consider to establish good agreement between observers from the observer population at different time from the time population.

We propose a new index CCC_R for random observers and random time points for data with repeated measurements. In Section 2.2, CCC_R is defined for assessing agreement in three types. The estimation and statistical inference of CCC_R are presented in Section 2.3. In Section 2.4, a simulation study is conducted to evaluate the performance of the CCC_R s. The vertebral body data and image data are used to illustrate the proposed methodology in Section 2.5.

2.2 Methodology

Consider that there are N randomly selected subjects where measurements are taken by J randomly selected observers at K randomly selected time points. For each subject, the random observer and the random time are drawn from observer and time populations, respectively. Let Y_{ijk} denote the observed data of subject i measured by observer j at time k , where $i = 1, \dots, N$, $j = 1, \dots, J$, and $k = 1, \dots, K$. Let $E_i(Y_{ijk}|j, k) = \mu_{jk}$ and $Var_i(Y_{ijk}|j, k) = \sigma_{jk}^2$ be the mean and variance of Y_{ijk} conditional on subject i by observer j at time k , where the parameters μ_{jk} and σ_{jk}^2 are random variables because observers and time are treated as random factors. Notations E_i and Var_i are the expectation and variance respect to subjects for Y_{ijk} conditional on j and k , respectively.

Vangeneugden et al. (2005) derived several *ICC*-type agreement indices to assess test-retest, interrater, and absolute reliability coefficients based on linear mixed models. In this section, we proposed random observers and random times for *CCC* to assess the same three types of agreement, inter-observer, intra-observer and absolute agreement, without assumptions of linear mixed models. The inter-observer agreement is assessing the agreement of an observed reading by one rater to an observed reading by a different observer on the same subject at the same time (*i.e.* Y_{ijk} vs $Y_{ij'k}$). The intra-observer agreement is assessing the agreement of an observed reading on one time to an observed reading on another time by the same observer on the same subject (*i.e.* Y_{ijk} vs $Y_{ijk'}$). The absolute agreement is assessing the agreement of an observed reading by one observer on one time to an observed reading by a different observer at a different time on the same subject (*i.e.* Y_{ijk} vs $Y_{ij'k'}$). This absolute agreement for data with repeated measurements is similar to total agreement for data with replicated reading in Barnhart et al. (2005). Both absolute and total agreement contain inter-observer and intra-observer agreement, while absolute agreement is for data with repeated measures and total agreement is for data with replications. The new indices, CCC_R 's, for inter-observer, intra-observer, and absolute agreement are denoted as CCC_R^{inter} , CCC_R^{intra} , and CCC_R^{abs} , respectively, and are described as below.

2.2.1 CCC_R^{inter} for inter-observer agreement

Lin (1989) developed the concordance correlation coefficient (CCC) to evaluate the agreement by scaling the mean square difference between two readings taken from two

fixed observers as

$$\rho_c = 1 - \frac{E[(Y_{i1} - Y_{i2})^2]}{E_I[(Y_{i1} - Y_{i2})^2]},$$

where E_I is the conditional expectation given independence of Y_{i1} and Y_{i2} . Let $Y_{i1k}, Y_{i2k}, \dots, Y_{iJk}$ be the measurements made by J fixed observers where each observer has only one reading for each subject, then Barnhart et al. (2005) extended Lin's *CCC* and proposed a total *CCC* to assess the agreement among multiple observers for data with replications. The total-*CCC* can be expressed as

$$\rho_c(Y) = 1 - \frac{\sum_{j=1}^{J-1} \sum_{j'=j+1}^J E_i[(Y_{ijk} - Y_{ij'k})^2 / J(J-1)]}{\sum_{j=1}^{J-1} \sum_{j'=j+1}^J E_I[(Y_{ijk} - Y_{ij'k})^2 / J(J-1)]},$$

where E_I is the conditional expectation given independence of Y_{i1k}, Y_{i2k}, \dots , and Y_{iJk} . By extending the total-*CCC* to assess the agreement between random observers at random times for data with repeated measurements, we define the inter-observer *CCC* for random observer and random time by replacing the summation $\sum_{j=1}^{J-1} \sum_{j'=j+1}^J$ with the expectation $E_{jj'k}$ for the new agreement index where $E_{jj'k}$ is the expectation respect to random observers j, j' and random time k . Specifically,

$$CCC_R^{inter} = 1 - \frac{E_{jj'k} E_i[(Y_{ijk} - Y_{ij'k})^2 | j, j', k]}{E_{jj'k} E_I[(Y_{ijk} - Y_{ij'k})^2 | j, j', k]}, \quad (2.1)$$

where E_I is the conditional expectation given independence of Y_{ijk} and $Y_{ij'k}$ conditional on observer and time. With $E_{jk} Var_i(Y_{ijk} | j, k) = E_{jk} \sigma_{jk}^2 = \sigma^2$ and $Var_{jk} E_i(Y_{ijk} | j, k) = Var_{jk} \mu_{jk} = \tau^2$, we can obtain

$$\begin{aligned} & E_{jj'k} E_i[(Y_{ijk} - Y_{ij'k})^2 | j, j', k] \\ &= 2\sigma^2 + E_{jj'k}(\mu_{jk} - \mu_{j'k})^2 - 2E_{jj'k} Cov_i(Y_{ijk}, Y_{ij'k} | j, j', k) \\ &= 2\sigma^2 + 2\tau^2 - 2Cov_{jj'k}(\mu_{jk}, \mu_{j'k}) - (E_{jk}(\mu_{jk}) - E_{j'k}(\mu_{j'k}))^2 - 2E_{jj'k} Cov_i(Y_{ijk}, Y_{ij'k} | j, j', k), \end{aligned}$$

and

$$E_{jj'k} E_I[(Y_{ijk} - Y_{ij'k})^2 | j, j', k] = 2\sigma^2 + 2\tau^2 - 2Cov_{jj'k}(\mu_{jk}, \mu_{j'k}) - (E_{jk}(\mu_{jk}) - E_{j'k}(\mu_{j'k}))^2,$$

Here we assume that $E_{jk}(\mu_{jk}) = E_{j'k}(\mu_{j'k})$, where $E_{jk}(\mu_{jk}) = E_j E_k(\mu_{jk} | j)$. Then, the inter-observer agreement for random *CCC* can be expressed as

$$\begin{aligned} CCC_R^{inter} &= 1 - \frac{2\sigma^2 + 2\tau^2 - 2Cov_{jj'k}(\mu_{jk}, \mu_{j'k}) - 2E_{jj'k} Cov_i(Y_{ijk}, Y_{ij'k} | j, j', k)}{2\sigma^2 + 2\tau^2 - 2Cov_{jj'k}(\mu_{jk}, \mu_{j'k})} \\ &= \frac{E_{jj'k} Cov_i(Y_{ijk}, Y_{ij'k} | j, j', k)}{\sigma^2 + \tau^2 - Cov_{jj'k}(\mu_{jk}, \mu_{j'k})}. \end{aligned} \quad (2.2)$$

2.2.2 CCC_R^{intra} for intra-observer agreement

The intra-observer CCC for random observers and random time points is obtained similarly by considering the expected mean square difference for measurements taken on the same subject by the same observer at different time points. We consider two situations for the intra-observer agreement: (1) subjects' true values do not change over time; (2) subjects' true values change over time.

If subjects' true values do not change over time by the same observer with the assumption of $E_{jk}(\mu_{jk}) = E_{jk'}(\mu_{jk'})$, the CCC_R^{intra} is defined as

$$CCC_R^{intra} = 1 - \frac{E_{jkk'} E_i([Y_{ijk} - Y_{ijk'}]^2 | j, k, k')}{E_{jkk'} E_I([Y_{ijk} - Y_{ijk'}]^2 | j, k, k')} \quad (2.3)$$

$$\begin{aligned} &= 1 - \frac{2\sigma^2 + 2\tau^2 - 2Cov_{jkk'}(\mu_{jk}, \mu_{jk'}) - 2E_{jkk'} Cov_i(Y_{ijk}, Y_{ijk'} | j, k, k')}{2\sigma^2 + 2\tau^2 - 2Cov_{jkk'}(\mu_{jk}, \mu_{jk'})} \\ &= \frac{E_{jkk'} Cov_i(Y_{ijk}, Y_{ijk'} | j, k, k')}{\sigma^2 + \tau^2 - Cov_{jkk'}(\mu_{jk}, \mu_{jk'})}, \end{aligned} \quad (2.4)$$

where E_I is the conditional expectation given independence of Y_{ijk} and $Y_{ijk'}$ conditional on subject and time.

However, if subjects' true values change over time, the mean square difference of $Y_{ijk} - Y_{ijk'}$ should be replaced by $(Y_{ijk} - \mu_{jk}) - (Y_{ijk'} - \mu_{jk'})$ so that the intra-observer agreement is not affected by the subjects' readings taken from time. In this case, the CCC_R^{intra} can be expressed as

$$CCC_R^{intra} = 1 - \frac{E_{jkk'} E_i([(Y_{ijk} - \mu_{jk}) - (Y_{ijk'} - \mu_{jk'})]^2 | j, k, k')}{E_{jkk'} E_I([(Y_{ijk} - \mu_{jk}) - (Y_{ijk'} - \mu_{jk'})]^2 | j, k, k')} \quad (2.5)$$

$$\begin{aligned} &= 1 - \frac{2\sigma^2 - 2E_{jkk'} Cov_i(Y_{ijk}, Y_{ijk'} | j, k, k')}{2\sigma^2} \\ &= \frac{E_{jkk'} Cov_i(Y_{ijk}, Y_{ijk'} | j, k, k')}{\sigma^2} \end{aligned} \quad (2.6)$$

where E_I is the conditional expectation given independence of Y_{ijk} and $Y_{ijk'}$ conditional on subject and time.

2.2.3 CCC_R^{abs} for absolute agreement

The absolute CCC for random observers and random time points is obtained similarly by considering the expected mean square difference for measurements taken on the same subject by different observers at different time points. Here, we also consider two situations for the absolute agreement: (1) subjects' true values do not change over time; (2) subjects' true values change over time.

If subjects' true values do not change over time by different observers with the assumption of $E_{jk}(\mu_{jk}) = E_{j'k'}(\mu_{j'k'})$, the CCC_R^{abs} is defined as

$$CCC_R^{abs} = 1 - \frac{E_{jj'kk'} E_i([Y_{ijk} - Y_{ij'k'}]^2 | j, j', k, k')}{E_{jj'kk'} E_I([Y_{ijk} - Y_{ij'k'}]^2 | j, j', k, k')} \quad (2.7)$$

$$= 1 - \frac{2\sigma^2 + 2\tau^2 - 2Cov_{jj'kk'}(\mu_{jk}, \mu_{j'k'}) - 2E_{jj'kk'} Cov_i(Y_{ijk}, Y_{ij'k'} | j, j', k, k')}{2\sigma^2 + 2\tau^2 - 2Cov_{jj'kk'}(\mu_{jk}, \mu_{j'k'})}$$

$$= \frac{E_{jj'kk'} Cov_i(Y_{ijk}, Y_{ij'k'} | j, j', k, k')}{\sigma^2 + \tau^2 - Cov_{jj'kk'}(\mu_{jk}, \mu_{j'k'})}, \quad (2.8)$$

where E_I is the conditional expectation given independence of Y_{ijk} and $Y_{ij'k'}$ conditional on subject and time.

However, if subjects' true values may change over time by different observers, the mean square difference of $Y_{ijk} - Y_{ij'k'}$ should be replaced by $(Y_{ijk} - \mu_{jk}) - (Y_{ij'k'} - \mu_{j'k'})$ so that the absolute agreement is not affected by the the subjects' readings taken from observers and time. Thus, the CCC_R^{abs} can be expressed as

$$CCC_R^{abs} = 1 - \frac{E_{jj'kk'} E_i([(Y_{ijk} - \mu_{jk}) - (Y_{ij'k'} - \mu_{j'k'})]^2 | j, j', k, k')}{E_{jj'kk'} E_I([(Y_{ijk} - \mu_{jk}) - (Y_{ij'k'} - \mu_{j'k'})]^2 | j, j', k, k')} \quad (2.9)$$

$$= 1 - \frac{2\sigma^2 - 2E_{jj'kk'} Cov_i(Y_{ijk}, Y_{ij'k'} | j, j', k, k')}{2\sigma^2}$$

$$= \frac{E_{jj'kk'} Cov_i(Y_{ijk}, Y_{ij'k'} | j, j', k, k')}{\sigma^2} \quad (2.10)$$

where E_I is the conditional expectation given independence of Y_{ijk} and $Y_{ij'k'}$ conditional on subject and time.

2.3 Estimation and Inference

To estimate the CCC for random observers and random time points, we used the method of moments approach for each component of CCC_R for inter-observer, intra-observer, and absolute agreement indices. In this section, we present the estimation and inference for CCC_R^{inter} , while CCC_R^{intra} and CCC_R^{abs} can be done in a similar fashion. Specifically, the CCC_R^{inter} is estimated by

$$\widehat{CCC}_R^{inter} = \frac{E_{jj'k} \widehat{Cov}_i(Y_{ijk}, Y_{ij'k} | j, j', k)}{\hat{\tau}_c^2 + \hat{\sigma}^2}, \quad (2.11)$$

where $\tau_c^2 = \tau^2 - Cov_{jj'k}(\mu_{jk}, \mu_{j'k})$. Note here that $\sigma^2 = E_{jk} Var_i(Y_{ijk} | j, k) = E_{jk} \sigma_{jk}^2$. Thus, σ_{jk}^2 can be estimated by the sample variance conditional on observer j and time k with $\hat{\sigma}_{jk}^2 = \sum_{i=1}^N (Y_{ijk} - \bar{Y}_{.jk})^2 / (N - 1)$, and a natural unbiased estimator of σ^2 can be estimated by an average of the sample conditional variance with

$$\hat{\sigma}^2 = \frac{\sum_{j=1}^J \sum_{k=1}^K \hat{\sigma}_{jk}^2}{JK}.$$

The covariance conditional on subject j and time k is $Cov_i(Y_{ijk}, Y_{ij'k} | j, j', k)$, where the notation Cov_i is the conditional covariance of Y_{ijk} and $Y_{ij'k}$ with respect to subjects. Then $Cov_i(Y_{ijk}, Y_{ij'k} | j, j', k)$ can be estimated by the sample conditional covariance with

$$\widehat{Cov}_i(Y_{ijk}, Y_{ij'k} | j, j', k) = \sum_{i=1}^N (Y_{ijk} - \bar{Y}_{.jk})(Y_{ij'k} - \bar{Y}_{.j'k}) / (N - 1),$$

and a natural unbiased estimator of the numerator of CCC_R^{inter} can be estimated by an average of the sample conditional covariance with

$$E_{jj'k} \widehat{Cov}_i(Y_{ijk}, Y_{ij'k} | j, j', k) = \sum_{j=1}^{J-1} \sum_{j'=j'}^J \sum_{k=1}^K \widehat{Cov}_i(Y_{ijk}, Y_{ij'k} | j, j', k) / K \binom{J}{2}.$$

Since

$$\begin{aligned} & E_{jj'k} E_i([\bar{Y}_{.jk} - \bar{Y}_{.j'k}]^2 | j, j', k) \\ &= E_{jj'k} Var_i([\bar{Y}_{.jk} - \bar{Y}_{.j'k}] | j, j', k) + E_{jj'k} (\mu_{jk} - \mu_{j'k})^2 \\ &= \frac{1}{N} E_{jj'k} (\sigma_{jk}^2 + \sigma_{j'k}^2 - 2Cov_i(Y_{ijk}, Y_{ij'k} | j, j', k)) + E_{jj'k} (\mu_{jk} - \mu_{j'k})^2 \\ &= \frac{2}{N} (\sigma^2 - E_{jj'k} Cov_i(Y_{ijk}, Y_{ij'k} | j, j', k)) + 2\tau^2 - 2Cov_{jj'k}(\mu_{jk}, \mu_{j'k}) \\ &= \frac{2}{N} (\sigma^2 - E_{jj'k} Cov_i(Y_{ijk}, Y_{ij'k} | j, j', k)) + 2\tau_c^2, \end{aligned}$$

with the assumption of $E_{jk}(\mu_{jk}) = E_{j'k}(\mu_{j'k})$, then the expression of τ_c^2 from the above equation can be written as

$$\tau_c^2 = \frac{1}{2} E_{jj'k} E_i([\bar{Y}_{jk} - \bar{Y}_{j'k}]^2 | j, j', k) - \frac{1}{N} [\sigma^2 - E_{jj'k} \text{Cov}_i(Y_{ijk}, Y_{ij'k} | j, j', k)].$$

Thus, a natural unbiased estimator of τ_c^2 can be estimated by plugging in the estimates of $\hat{\sigma}^2$ and $\widehat{E_{jj'k} \text{Cov}_i(Y_{ijk}, Y_{ij'k} | j, j', k)}$ as

$$\hat{\tau}_c^2 = \frac{\sum_{j=1}^{J-1} \sum_{j'=j'}^J \sum_{k=1}^K (\bar{Y}_{jk} - \bar{Y}_{j'k})^2}{2K \binom{J}{2}} - \frac{1}{N} [\hat{\sigma}^2 - \widehat{E_{jj'k} \text{Cov}_i(Y_{ijk}, Y_{ij'k} | j, j', k)}].$$

In order to deal with the inference of CCC_R and easily construct confidence intervals for inter-observer, intra-observer, and absolute agreement, we consider the bootstrap method for the three agreement indices. Efron and Tibshirani (1993) described several different approaches of confidence intervals based on the use of bootstrap method without asymptotically normality assumptions or other complex properties for the point estimator. The percentile method and bootstrap- t method are useful for constructing such confidence intervals. Some complex methods are based on the percentile version which may improve the coverage probability, such as *bias-corrected and accelerated* method (BC_α) and *approximate bootstrap confidence* method (ABC). However, in our simulation studies, we found that the percentile method, BC_α method, and ABC method did not perform well. Below we describe the bootstrap- t method for constructing a 95% bootstrap confidence interval. We make a minor modification for the bootstrap- t method since this method presented in Efron's book is usually suitable only for some simple statistics, such as the sample mean and sample median. We present the modified bootstrap- t method for the confidence interval of CCC_R^{inter} because of its better performance in our simulation studies; the same method for constructing the confidence interval can be used for CCC_R^{intra} and CCC_R^{abs} , respectively.

Let $(k^{1b}, k^{2b}, \dots, k^{Kb})$, $(j^{1b}, j^{2b}, \dots, j^{Jb})$, and $(i^{1b}, i^{2b}, \dots, i^{Nb})$ be the b^{th} bootstrap sample drawn with replacement from indices of times $(1, \dots, K)$, observers $(1, \dots, J)$, and subjects $(1, \dots, N)$, respectively. The B independent bootstrap data sets can be obtained by $\{Y_{i^{jb}j^{jb}k^{kb}}, i = 1, \dots, N; j = 1, \dots, J; k = 1, \dots, K\}$, where $b = 1, \dots, B$. The CCC_R^{inter} for each bootstrap sample is estimated by

$$\widehat{CCC}_R^{inter}(b) = \frac{\widehat{E_{jj'k} \text{Cov}_i(Y_{i^{jb}j^{jb}k^{kb}}, Y_{i^{jb}j'^{jb}k^{kb}} | j, j', k)}}{\hat{\tau}_{cb}^2 + \hat{\sigma}_b^2}, \quad (2.12)$$

where $E_{jj'k} \widehat{Cov}_i(Y_{ib_jjb_kkb}, Y_{ib_{j'}b_{k'}kb}|j, j', k)$, $\hat{\tau}_{cb}^2$, and $\hat{\sigma}_b^2$ are the estimators for the b^{th} bootstrap sample. The $100(1-2\alpha)\%$ bootstrap- t confidence interval of CCC_R^{inter} is defined as

$$(\widehat{CCC}_R^{inter} - \hat{t}^{(1-\alpha)} \cdot \hat{se}, \widehat{CCC}_R^{inter} - \hat{t}^{(\alpha)} \cdot \hat{se}), \quad (2.13)$$

where $\hat{t}^{(\alpha)}$ refers to the α th percentile of $(\widehat{CCC}_R^{inter}(b) - \widehat{CCC}_R^{inter})/\hat{se}(b)$, where \hat{se} is the estimated standard error of \widehat{CCC}_R^{inter} , and $\hat{se}(b)$ is the estimated standard error of $\widehat{CCC}_R^{inter}(b)$ for the bootstrap sample $Y_{ib_jjb_kkb}$. However, unlike the examples in Efron's book, we do not find a formula for the estimated standard error of \widehat{CCC}_R^{inter} . Thus, we approximate \hat{se} by $\hat{se}(b)$ to obtain the the $100(1-2\alpha)\%$ confidence interval of CCC_R^{inter} as

$$(\widehat{CCC}_R^{inter} - \hat{t}^{(1-\alpha)} \cdot \hat{se}(b), \widehat{CCC}_R^{inter} - \hat{t}^{(\alpha)} \cdot \hat{se}(b)). \quad (2.14)$$

2.4 Simulations

To evaluate the performance of CCC_R^{abs} , CCC_R^{inter} , and CCC_R^{intra} for random observers and random time, we carried out simulations based on 1000 Monte Carlo data sets. We consider that the number of observers drawn from the observer population is 2, 4, 6, respectively ($J = 2, 4, 6$), and three time points ($K = 3$) for each observer are randomly drawn from the time population. The sample size of subjects for each combination of the number of observers and time points is 20, 50, 100, 150, 200, 250, 300, and 500, respectively.

To illustrate the process of generating the data set, consider the situation with two random observers ($J = 2$) and three random time points ($K = 3$). Assume that the observer population has normal distribution with mean μ and variance σ_μ^2 . Two observers' overall mean parameters μ_1 and μ_2 were randomly selected from the observer population $N(\mu, \sigma_\mu^2)$. For a given observer j , we generated μ_{jk} at random time k with $E_k(\mu_{jk}|j) = \mu_j$ and $Var_k(\mu_{jk}|j) = \sigma_{\mu_j}^2$, where $\sigma_{\mu_j}^2$ was generated from a gamma distribution with $\sigma_{\mu_j}^2 \sim Gamma(\sigma_0^2, 1)$, $j = 1, 2$. We will use the following assumptions for the correlations of μ_{jk} 's (1) $Corr(\mu_{jk}, \mu_{jk'}) = \rho_1$, (2) $Corr(\mu_{jk}, \mu_{j'k}) = \rho_2$, and (3) $Corr(\mu_{jk}, \mu_{j'k'}) = \rho_1\rho_2$, where ρ_1 is the correlation coefficient within observers and ρ_2 is the correlation coefficient between observers. Then, the realization $(\mu_{11}, \mu_{12}, \mu_{13}, \mu_{21}, \mu_{22}, \mu_{23})^T$ of the true measurements by two observers at three time points was generated from a multivariate normal distribution

with mean $(\mu_1, \mu_1, \mu_1, \mu_2, \mu_2, \mu_2)^T$ and variance-covariance matrix

$$\Sigma_{\mu}^{6 \times 6} = \begin{pmatrix} \sigma_{\mu_1}^2 & \rho_1 \sigma_{\mu_1}^2 & \rho_1 \sigma_{\mu_1}^2 & \rho_2 \sigma_{\mu_1} \sigma_{\mu_2} & \rho_1 \rho_2 \sigma_{\mu_1} \sigma_{\mu_2} & \rho_1 \rho_2 \sigma_{\mu_1} \sigma_{\mu_2} \\ \rho_1 \sigma_{\mu_1}^2 & \sigma_{\mu_1}^2 & \rho_1 \sigma_{\mu_1}^2 & \rho_1 \rho_2 \sigma_{\mu_1} \sigma_{\mu_2} & \rho_2 \sigma_{\mu_1} \sigma_{\mu_2} & \rho_1 \rho_2 \sigma_{\mu_1} \sigma_{\mu_2} \\ \rho_1 \sigma_{\mu_1}^2 & \rho_1 \sigma_{\mu_1}^2 & \sigma_{\mu_1}^2 & \rho_1 \rho_2 \sigma_{\mu_1} \sigma_{\mu_2} & \rho_1 \rho_2 \sigma_{\mu_1} \sigma_{\mu_2} & \rho_2 \sigma_{\mu_1} \sigma_{\mu_2} \\ \rho_2 \sigma_{\mu_1} \sigma_{\mu_2} & \rho_1 \rho_2 \sigma_{\mu_1} \sigma_{\mu_2} & \rho_1 \rho_2 \sigma_{\mu_1} \sigma_{\mu_2} & \sigma_{\mu_2}^2 & \rho_1 \sigma_{\mu_2}^2 & \rho_1 \sigma_{\mu_2}^2 \\ \rho_1 \rho_2 \sigma_{\mu_1} \sigma_{\mu_2} & \rho_2 \sigma_{\mu_1} \sigma_{\mu_2} & \rho_1 \rho_2 \sigma_{\mu_1} \sigma_{\mu_2} & \rho_1 \sigma_{\mu_2}^2 & \sigma_{\mu_2}^2 & \rho_1 \sigma_{\mu_2}^2 \\ \rho_1 \rho_2 \sigma_{\mu_1} \sigma_{\mu_2} & \rho_1 \rho_2 \sigma_{\mu_1} \sigma_{\mu_2} & \rho_2 \sigma_{\mu_1} \sigma_{\mu_2} & \rho \sigma_{\mu_2}^2 & \rho_1 \sigma_{\mu_2}^2 & \sigma_{\mu_2}^2 \end{pmatrix}.$$

Finally, the observed data $(Y_{i11}, Y_{i12}, Y_{i13}, Y_{i21}, Y_{i22}, Y_{i23})^T$ for each subject i was generated from a multivariate normal distribution with mean $(\mu_{11}, \mu_{12}, \mu_{13}, \mu_{21}, \mu_{22}, \mu_{23})^T$ and variance covariance matrix

$$\Sigma^{6 \times 6} = \begin{pmatrix} \sigma_{11}^2 & \rho_1 \sigma_{11} \sigma_{12} & \rho_1 \sigma_{11} \sigma_{13} & \rho_2 \sigma_{11} \sigma_{21} & \rho_1 \rho_2 \sigma_{11} \sigma_{22} & \rho_1 \rho_2 \sigma_{11} \sigma_{23} \\ \rho_1 \sigma_{12} \sigma_{11} & \sigma_{12}^2 & \rho_1 \sigma_{12}^2 \sigma_{13} & \rho_1 \rho_2 \sigma_{12} \sigma_{21} & \rho_2 \sigma_{12} \sigma_{22} & \rho_1 \rho_2 \sigma_{12} \sigma_{23} \\ \rho_1 \sigma_{13} \sigma_{11} & \rho_1 \sigma_{13} \sigma_{12} & \sigma_{13}^2 & \rho_1 \rho_2 \sigma_{13} \sigma_{21} & \rho_1 \rho_2 \sigma_{13} \sigma_{22} & \rho_2 \sigma_{13} \sigma_{23} \\ \rho_2 \sigma_{11} \sigma_{21} & \rho_1 \rho_2 \sigma_{11} \sigma_{22} & \rho_1 \rho_2 \sigma_{11} \sigma_{23} & \sigma_{21}^2 & \rho_1 \sigma_{21} \sigma_{22} & \rho_1 \sigma_{21} \sigma_{23} \\ \rho_1 \rho_2 \sigma_{12} \sigma_{21} & \rho_2 \sigma_{12} \sigma_{22} & \rho_1 \rho_2 \sigma_{12} \sigma_{23} & \rho_1 \sigma_{22} \sigma_{21} & \sigma_{22}^2 & \rho_1 \sigma_{22} \sigma_{23} \\ \rho_1 \rho_2 \sigma_{13} \sigma_{21} & \rho_1 \rho_2 \sigma_{13} \sigma_{22} & \rho_2 \sigma_{13} \sigma_{23} & \rho \sigma_{23} \sigma_{21} & \rho_1 \sigma_{23} \sigma_{22} & \sigma_{23}^2 \end{pmatrix},$$

where σ_{jk}^2 was generated from a gamma distribution with $\sigma_{jk}^2 \sim \text{Gamma}(\sigma_T^2, 1)$, $j = 1, 2, k = 1, 2, 3$. With these assumptions, we have $\sigma^2 = \sigma_T^2$ and $\tau_c^2 = \sigma_{\mu}^2 + \sigma_0^2$. Table 2.1 shows the true values of μ , σ_{μ}^2 , σ_0^2 , σ_T^2 , ρ_1 , and ρ_2 used for data generation as well as the corresponding true values of CCC_R . The true values of CCC_R^{abs} and CCC_R^{inter} provide situations with strong, moderate, and mild agreement between observers, and the true values of CCC_R^{intra} represents strong agreement within observers.

We evaluate the performance of CCC_R for inter-observer, intra-observer, and absolute agreement based on 1000 Monte Carlo data sets by calculating (1) the average estimates of CCC_R , (2) the standard deviation of the estimates of CCC_R , and (3) the coverage probability for a modified 95% bootstrap- t confidence interval. The results are shown in Table 2.3-2.4 for $J = 2, 4, 6$, respectively, which indicate that the bias of the point estimate of CCC_R decreases as the sample size and the number of observers increase. The estimate is very close to the true value for a large sample size and a large number of observers. The coverage probability of CCC_R is more close to 0.95 when the sample size and the number

Table 2.1: Three scenarios of the true values for the proposed CCC_R

	μ	σ_μ^2	σ_0^2	σ_T^2	ρ_1	ρ_2	CCC_R^{abs}	CCC_R^{inter}	CCC_R^{intra}
Scenario 1	10	1	10	3	0.95	0.90	0.764	0.805	0.937
Scenario 2	10	11	10	5	0.95	0.90	0.477	0.502	0.962
Scenario 3	10	57	10	7	0.95	0.90	0.191	0.201	0.985

of observers increase. Overall, the point estimate is approximate to be unbiased and the coverage probability of CCC_R tends to near 0.95.

Table 2.2: Results of CCC_R^{abs} based on 1000 Monte Carlo data sets

CCC_R^{abs}	N	J=2			J=4			J=6		
		Mean	SD	Coverage	Mean	SD	Coverage	Mean	SD	Coverage
0.764	20	0.773	0.098	0.851	0.778	0.075	0.837	0.778	0.075	0.835
	50	0.770	0.055	0.894	0.770	0.044	0.894	0.773	0.039	0.892
	100	0.769	0.038	0.936	0.767	0.030	0.925	0.768	0.028	0.938
	150	0.768	0.030	0.945	0.767	0.024	0.933	0.767	0.024	0.952
	200	0.766	0.025	0.959	0.766	0.021	0.936	0.766	0.020	0.951
	250	0.765	0.022	0.967	0.766	0.019	0.942	0.765	0.017	0.948
	300	0.765	0.021	0.976	0.765	0.017	0.937	0.765	0.016	0.954
	500	0.765	0.016	0.973	0.765	0.013	0.955	0.764	0.012	0.946
0.477	20	0.463	0.174	0.854	0.469	0.132	0.869	0.471	0.106	0.881
	50	0.475	0.108	0.893	0.479	0.072	0.916	0.475	0.065	0.925
	100	0.476	0.074	0.916	0.478	0.052	0.935	0.479	0.045	0.942
	150	0.476	0.061	0.932	0.478	0.042	0.939	0.478	0.036	0.931
	200	0.476	0.054	0.951	0.478	0.037	0.955	0.478	0.032	0.950
	250	0.476	0.046	0.944	0.477	0.032	0.947	0.478	0.029	0.941
	300	0.476	0.042	0.948	0.477	0.029	0.952	0.478	0.025	0.945
	500	0.476	0.034	0.960	0.477	0.023	0.947	0.477	0.020	0.948
0.191	20	0.183	0.220	0.861	0.187	0.115	0.863	0.188	0.094	0.850
	50	0.186	0.137	0.909	0.187	0.072	0.914	0.188	0.057	0.914
	100	0.188	0.095	0.926	0.192	0.053	0.926	0.193	0.041	0.928
	150	0.188	0.078	0.927	0.192	0.042	0.948	0.189	0.032	0.950
	200	0.190	0.068	0.938	0.192	0.038	0.945	0.190	0.029	0.942
	250	0.190	0.062	0.944	0.189	0.034	0.946	0.190	0.027	0.931
	300	0.192	0.054	0.939	0.191	0.030	0.942	0.191	0.025	0.948
	500	0.191	0.042	0.954	0.191	0.023	0.946	0.191	0.018	0.949

Table 2.3: Results of CCC_R^{inter} based on 1000 Monte Carlo data sets

CCC_R^{inter}	N	J=2			J=4			J=6		
		Mean	SD	Coverage	Mean	SD	Coverage	Mean	SD	Coverage
0.805	20	0.824	0.089	0.827	0.820	0.068	0.845	0.824	0.061	0.826
	50	0.812	0.050	0.896	0.814	0.037	0.898	0.812	0.035	0.891
	100	0.809	0.033	0.916	0.809	0.024	0.925	0.809	0.024	0.918
	150	0.809	0.027	0.925	0.807	0.021	0.943	0.807	0.019	0.932
	200	0.808	0.024	0.930	0.807	0.018	0.937	0.807	0.016	0.942
	250	0.807	0.020	0.937	0.806	0.016	0.948	0.805	0.015	0.947
	300	0.806	0.019	0.947	0.805	0.015	0.943	0.805	0.013	0.938
	500	0.806	0.014	0.951	0.805	0.011	0.961	0.805	0.010	0.949
0.502	20	0.498	0.174	0.858	0.502	0.119	0.869	0.503	0.106	0.873
	50	0.499	0.110	0.889	0.500	0.070	0.921	0.500	0.063	0.909
	100	0.499	0.074	0.937	0.502	0.050	0.945	0.503	0.042	0.940
	150	0.500	0.059	0.932	0.504	0.041	0.940	0.501	0.034	0.942
	200	0.501	0.052	0.938	0.504	0.036	0.953	0.502	0.032	0.951
	250	0.501	0.045	0.942	0.500	0.031	0.951	0.503	0.028	0.958
	300	0.503	0.043	0.953	0.500	0.031	0.958	0.503	0.025	0.951
	500	0.503	0.033	0.955	0.503	0.022	0.955	0.502	0.020	0.949
0.201	20	0.199	0.218	0.862	0.199	0.120	0.862	0.202	0.094	0.865
	50	0.197	0.135	0.918	0.198	0.072	0.911	0.199	0.057	0.922
	100	0.199	0.095	0.937	0.200	0.054	0.945	0.199	0.041	0.927
	150	0.199	0.079	0.936	0.201	0.042	0.942	0.199	0.032	0.943
	200	0.199	0.068	0.955	0.201	0.037	0.953	0.200	0.029	0.935
	250	0.200	0.060	0.948	0.201	0.032	0.941	0.202	0.025	0.950
	300	0.199	0.055	0.943	0.201	0.030	0.948	0.201	0.023	0.948
	500	0.200	0.041	0.942	0.201	0.023	0.949	0.201	0.018	0.947

Table 2.4: Results of CCC_R^{intra} based on 1000 Monte Carlo data sets

CCC_R^{intra}	N	J=2			J=4			J=6		
		Mean	SD	Coverage	Mean	SD	Coverage	Mean	SD	Coverage
0.937	20	0.933	0.026	0.925	0.933	0.022	0.917	0.935	0.020	0.918
	50	0.936	0.014	0.933	0.936	0.012	0.945	0.936	0.012	0.945
	100	0.937	0.010	0.954	0.937	0.009	0.953	0.937	0.008	0.956
	150	0.937	0.008	0.955	0.937	0.007	0.946	0.937	0.007	0.942
	200	0.937	0.006	0.951	0.937	0.006	0.961	0.937	0.006	0.940
	250	0.937	0.006	0.963	0.937	0.006	0.961	0.937	0.006	0.958
	300	0.937	0.006	0.946	0.937	0.005	0.947	0.937	0.005	0.967
	500	0.937	0.004	0.958	0.937	0.004	0.946	0.937	0.004	0.953
0.962	20	0.960	0.013	0.941	0.961	0.011	0.929	0.962	0.010	0.916
	50	0.961	0.008	0.951	0.962	0.007	0.952	0.962	0.006	0.963
	100	0.962	0.005	0.970	0.962	0.004	0.961	0.962	0.004	0.971
	150	0.962	0.004	0.965	0.962	0.004	0.966	0.962	0.003	0.968
	200	0.962	0.004	0.965	0.962	0.003	0.954	0.962	0.003	0.963
	250	0.962	0.003	0.975	0.962	0.003	0.948	0.962	0.003	0.951
	300	0.962	0.003	0.962	0.962	0.003	0.962	0.962	0.002	0.954
	500	0.962	0.002	0.967	0.962	0.002	0.948	0.962	0.002	0.952
0.985	20	0.986	0.005	0.937	0.986	0.004	0.931	0.986	0.003	0.933
	50	0.985	0.003	0.957	0.986	0.002	0.956	0.985	0.002	0.958
	100	0.985	0.002	0.972	0.986	0.002	0.969	0.985	0.001	0.970
	150	0.985	0.002	0.962	0.985	0.001	0.961	0.985	0.001	0.963
	200	0.985	0.001	0.961	0.985	0.001	0.965	0.985	0.001	0.961
	250	0.985	0.001	0.964	0.985	0.001	0.961	0.985	0.001	0.960
	300	0.985	0.001	0.967	0.985	0.001	0.966	0.985	0.001	0.952
	500	0.985	0.001	0.967	0.985	0.001	0.956	0.985	0.001	0.949

2.5 Examples

2.5.1 Vertebral Body Data

We first use data from a vertebral body wedging study of pediatric patients. The data was measured in thoracic and lumbar area of vertebral body at the thoracolumbar junction for patients with ages between 0 and 17 years. Anterior and posterior vertebral body heights were measure by two experienced pediatric radiologists on six locations (T10, T11, T12, L1, L2, L3) of the vertebral body, providing a combination of the 12 locations measured for each patient. In all 100 patients measured by two radiologists, 20 patients were measured twice by the same radiologists and the remaining 80 patients were measured only once by each of the two radiologists. For illustration of the proposed methodology, we use this data set to determine the degree of the inter-observer, intra-observer, and absolute agreement between two radiologists in two situations below. In both situations, we assumed that two radiologists ($J = 2$) were randomly selected from the pediatric radiologists population. However, in the first situation, the time points are the two times when the same radiologists make measurements for the 20 patients. In the second situation, we consider the 12 locations as the "time points" where each of the two radiologists make measurements for all 100 patients. Because the radiologists made measurements twice in the first 20 patients, only the first measurement is used in this situation.

Situation 1: Two random time points

In the first situation, we assume that the two time points ($K=2$) are randomly selected from the time population. For illustration, the data was restricted to the 17 subjects ($N = 17$) without missing data by two radiologists at two time points. The descriptive statistics for 12 locations are shown in Tables 2.5-2.7, which indicated that differences are small between readings by the same reader at the same location (Table 2.5), between readers at the same time and the same location (Table 2.6), and between readers at different time and at the same location (Table 2.7). Figures 2.1-2.4 showed the scatter plots for the comparisons between two readers measured at the same time. The scatter plots showed good agreement between radiologists and the measurements by the two readers scattered close to the 45° line. Table 2.8 showed the point estimates of CCC_{RS} with the corresponding modified

95% bootstrap- t confidence intervals for inter-rater, intra-rater, and absolute agreement at each location. The results indicate similar agreement between two radiologists across 12 locations. In order to construct the common estimates of CCC_R across locations in this situation, we take the average of the 12 estimates of CCC_R as \widehat{CCC}_R to be the estimate for the common estimate of CCC_R . We use the modified bootstrap- t method to establish the 95% confidence interval for the common estimates of inter-observer, intra-observer and absolute agreement, respectively. Specifically, for assessing the inter-observer agreement, a total 200 independent bootstrap samples ($B = 200$) were drawn with replacement from indices of two time points, two radiologists, and 17 subjects, respectively. For each bootstrap samples, $\widehat{CCC}_{R1}^{inter}(b)$, $\widehat{CCC}_{R2}^{inter}(b)$, ..., $\widehat{CCC}_{R12}^{inter}(b)$ are the 12 estimates for the inter-observer agreement for each bootstrap b , $b = 1, \dots, 200$. We then take the average of the 12 estimates as $\widehat{CCC}_R^{inter}(b)$ to obtain the estimate for the common estimate for each bootstrap b . Then, the final 95% confidence interval is obtained by

$$(\widehat{CCC}_R^{inter} - \tilde{t}^{(1-\alpha)} \cdot \tilde{se}(b), \widehat{CCC}_R^{inter} - \tilde{t}^{(\alpha)} \cdot \tilde{se}(b)), \quad (2.15)$$

where $\tilde{t}^{(\alpha)}$ is the α th percentile of $(\widehat{CCC}_R^{inter}(b) - \widehat{CCC}_R^{inter})/\tilde{se}(b)$, and $\tilde{se}(b)$ is the estimated standard error of $\widehat{CCC}_R^{inter}(b)$; the confidence intervals for the common estimates of \widehat{CCC}_R^{intra} and \widehat{CCC}_R^{abs} can be obtained similarly and the results are shown in Table 2.8.

Situation 2: Random time points as random locations

In the second situation, We considered all 100 patients measured by two radiologists ($J = 2$) on the 12 locations ($K = 12$), where the locations are treated as repeated measurements randomly selected from the location population. For illustration, the data was restricted to the 90 subjects ($N = 90$) without missing data by two radiologists at 12 locations. Figures 2.5-2.6 showed the scatter plots for the comparisons between two radiologists for the measurements at the 12 locations. The scatter plots showed that overall measurements of two radiologists scattered more closely to the 45° line on all levels, respectively. Table 2.8 showed the point estimates of CCC_R s with the corresponding modified 95% bootstrap- t confidence interval for inter-rater, intra-rater, and absolute agreement. The results indicate similar agreement between two radiologists. Here, since the subjects' true values may change over "time", for calculating the estimates of CCC_R^{abs} and CCC_R^{intra} , the mean square

difference of $Y_{ijk} - Y_{ij'k'}$ and $Y_{ijk} - Y_{ijk'}$ should be replaced by $(Y_{ijk} - \mu_{jk}) - (Y_{ij'k'} - \mu_{j'k'})$ and $(Y_{ijk} - \mu_{jk}) - (Y_{ijk'} - \mu_{jk'})$, respectively in order to exclude the situation that the agreement may affect by the subjects' readings from time.

Note that the common CCC_R in Situation 1 assesses agreement between radiologists aggregated over the locations where the locations are treated as fixed, while the CCC_R in Situation 2 assesses the agreement between radiologists over all possible locations where locations are treated as random. Thus, the estimates of CCC_R in Situation 2 are lower than that of the common CCC_R in Situation 1 because random locations can cause lower agreement between radiologists. In both situations for the study of the vertebral body data, the estimates of CCC_R^{inter} in Table 2.8 are almost the same as the estimates of $\rho_{c,rm}$ proposed by King et al. (2007) for assessing the agreement of CCC when radiologists and time are treated as fixed.

Table 2.5: Description statistics for readings between times by readers and locations in Situation 1 for vertebral body data

Location		Reader 1			Reader 2		
Anterior(A)/Posterior(P)		1 st Read	2 nd Read	Diff	1 st Read	2 nd Read	Diff
T10 A	mean(sd)	17.971(5.212)	18.518(4.866)	-0.547(1.546)	18.765(5.640)	18.000(4.988)	0.765(1.322)
	min(max)	8.4(28.4)	9.0(25.0)	-2.6(3.4)	8.7(27.0)	9.3(25.0)	-1.4(2.9)
T10 P	mean(sd)	18.276(5.210)	18.553(4.516)	-0.276(1.602)	19.559(5.898)	18.388(4.674)	1.171(1.987)
	min(max)	9.0(28.6)	9.0(24.6)	-2.2(4.8)	9.3(27.0)	9.4(24.3)	-1.9(6.4)
T11 A	mean(sd)	18.800(5.492)	19.129(5.367)	-0.329(1.750)	19.459(6.189)	18.959(5.465)	0.500(1.470)
	min(max)	8.1(27.6)	8.7(27.3)	-2.5(5.0)	8.7(28.2)	8.8(27.1)	-2.6(3.4)
T11 P	mean(sd)	19.324(5.669)	19.647(5.398)	-0.324(1.649)	20.400(6.055)	19.441(5.614)	0.959(2.038)
	min(max)	8.4(28.3)	9.0(27.3)	-2.9(4.1)	9.3(28.6)	9.1(28.6)	-2.0(7.5)
T12 A	mean(sd)	20.394(5.780)	20.012(5.691)	0.382(1.364)	20.276(5.859)	19.759(5.386)	0.518(1.204)
	min(max)	9.8(28.5)	10.5(29.7)	-1.6(3.3)	10.2(28.1)	10.6(27.6)	-0.7(3.7)
T12 P	mean(sd)	20.712(5.752)	20.694(5.637)	0.018(1.160)	20.976(6.182)	20.418(5.748)	0.559(1.971)
	min(max)	10.4(29.0)	11.1(30.8)	-2.1(2.0)	10.1(28.9)	10.8(29.9)	-3.2(5.3)
L1 A	mean(sd)	21.400(6.284)	21.124(5.720)	0.276(1.186)	21.424(6.322)	20.888(5.431)	0.535(1.556)
	min(max)	9.9(31.8)	10.5(28.3)	-1.6(3.5)	10.2(31.0)	10.8(29.4)	-1.5(4.5)
L1 P	mean(sd)	21.988(6.683)	21.600(5.698)	0.388(1.488)	22.088(6.562)	21.800(5.801)	0.288(1.661)
	min(max)	10.1(32.2)	11.0(29.6)	-2.8(2.6)	10.1(31.9)	11.1(31.0)	-2.1(4.1)
L2 A	mean(sd)	22.694(6.569)	21.971(6.224)	0.724(2.063)	22.041(6.205)	21.924(6.071)	0.118(0.926)
	min(max)	10.1(33.1)	11.0(30.7)	-2.2(4.7)	10.7(30.5)	10.7(29.8)	-1.3(2.1)
L2 P	mean(sd)	23.224(6.717)	22.147(6.010)	1.076(2.238)	22.471(6.588)	22.471(6.323)	0.000(1.280)
	min(max)	9.6(33.1)	11.3(30.5)	-1.9(5.2)	10.1(32.2)	11.1(31.5)	-1.9(2.9)
L3 A	mean(sd)	22.759(6.866)	22.518(6.561)	0.241(1.738)	22.082(6.676)	21.988(6.382)	0.094(1.163)
	min(max)	9.6(32.7)	10.4(31.7)	-2.4(4.7)	10.1(32.0)	9.6(31.3)	-1.3(3.4)
L3 P	mean(sd)	22.988(6.924)	22.371(6.418)	0.617(1.414)	22.329(6.805)	22.041(6.176)	0.288(1.907)
	min(max)	9.6(33.5)	10.4(31.8)	-2.6(3.4)	10.5(33.6)	10.1(31.3)	-2.2(5.0)

Table 2.6: Description statistics for readings between readers by times and locations in Situation 1 for vertebral body data

Location		1 st Read			2 nd Read		
Anterior(A)/Posterior(P)		Reader 1	Reader 2	Diff	Reader 1	Reader 2	Diff
T10 A	mean(sd)	17.971(5.212)	18.765(5.640)	-0.794(1.682)	18.518(4.866)	18.000(4.988)	0.518(0.993)
	min(max)	8.4(28.4)	8.7(27.0)	-3.2(1.5)	9.0(25.0)	9.3(25.0)	-1.5(2.2)
T10 P	mean(sd)	18.276(5.210)	19.559(5.898)	-1.282(1.878)	18.553(4.516)	18.388(4.674)	0.165(1.013)
	min(max)	9.0(28.6)	9.3(27.0)	-5.4(1.6)	9.0(24.6)	9.4(24.3)	-1.6(1.9)
T11 A	mean(sd)	18.800(5.492)	19.459(6.189)	-0.658(1.678)	19.129(5.367)	18.959(5.465)	0.171(1.484)
	min(max)	8.1(27.6)	8.7(28.2)	-3.8(2.0)	8.7(27.3)	8.8(27.1)	-2.6(2.9)
T11 P	mean(sd)	19.324(5.669)	20.400(6.055)	-1.076(1.720)	19.647(5.398)	19.441(5.614)	0.206(1.614)
	min(max)	8.4(28.3)	9.3(28.6)	-4.4(2.4)	9.0(27.3)	9.1(28.6)	-4.2(3.4)
T12 A	mean(sd)	20.394(5.780)	20.276(5.859)	0.117(1.097)	20.012(5.691)	19.759(5.386)	0.253(1.219)
	min(max)	9.8(28.5)	10.2(28.1)	-2.9(1.8)	10.5(29.7)	10.6(27.6)	-2.2(2.3)
T12 P	mean(sd)	20.712(5.752)	20.976(6.182)	-0.265(1.412)	20.694(5.637)	20.418(5.748)	0.276(1.468)
	min(max)	10.4(29.0)	10.1(28.9)	-3.8(1.8)	11.1(30.8)	10.8(29.9)	-3.4(2.9)
L1 A	mean(sd)	21.400(6.284)	21.424(6.322)	-0.023(1.264)	21.124(5.720)	20.888(5.431)	0.235(1.340)
	min(max)	9.9(31.8)	10.2(31.0)	-2.7(1.7)	10.5(28.3)	10.8(29.4)	-1.2(4.7)
L1 P	mean(sd)	21.988(6.683)	22.088(6.562)	-0.100(1.352)	21.600(5.698)	21.800(5.801)	-0.200(1.337)
	min(max)	10.1(32.2)	10.1(31.9)	-2.6(2.2)	11.0(29.6)	11.1(31.0)	-2.1(3.7)
L2 A	mean(sd)	22.694(6.569)	22.041(6.205)	0.653(2.089)	21.971(6.224)	21.924(6.071)	0.047(1.693)
	min(max)	10.1(33.1)	10.7(30.5)	-3.9(3.3)	11.0(30.7)	10.7(29.8)	-3.1(2.5)
L2 P	mean(sd)	23.224(6.717)	22.471(6.588)	0.753(1.692)	22.147(6.010)	22.471(6.323)	-0.324(1.538)
	min(max)	9.6(33.1)	10.1(32.2)	-2.5(3.4)	11.3(30.5)	11.1(31.5)	-2.8(2.3)
L3 A	mean(sd)	22.759(6.866)	22.082(6.676)	0.676(1.541)	22.518(6.561)	21.988(6.382)	0.529(1.244)
	min(max)	9.6(32.7)	10.1(32.0)	-2.4(3.2)	10.4(31.7)	9.6(31.3)	-2.3(3.1)
L3 P	mean(sd)	22.988(6.924)	22.329(6.805)	0.659(1.457)	22.371(6.418)	22.041(6.176)	0.329(1.529)
	min(max)	9.6(33.5)	10.5(33.6)	-2.9(3.1)	10.4(31.8)	10.1(31.3)	-2.5(3.1)

Table 2.7: Description statistics for readings between readers and times by locations in Situation 1 for vertebral body data

Location		Reader 1	Reader 2		Reader 1	Reader 2	
Anterior(A)/Posterior(P)		1 st Read	2 nd Read	Diff	2 nd Read	1 st Read	Diff
T10 A	mean(sd)	17.971(5.212)	18.000(4.988)	-0.029(1.391)	18.518(4.866)	18.765(5.640)	-0.247(1.277)
	min(max)	8.4(28.4)	9.3(25.0)	-1.9(3.4)	9.0(25.0)	8.7(27.0)	-2.7(2.2)
T10 P	mean(sd)	18.276(5.210)	18.388(4.674)	-0.112(1.514)	18.553(4.516)	19.559(5.898)	-1.006(1.945)
	min(max)	9.0(28.6)	9.4(24.3)	-2.4(4.6)	9.0(24.6)	9.3(27.0)	-4.7(2.6)
T11 A	mean(sd)	18.800(5.492)	18.959(5.465)	-0.159(1.827)	19.129(5.367)	19.459(6.189)	-0.329(1.705)
	min(max)	8.1(27.6)	8.8(27.1)	-3.5(2.5)	8.7(27.3)	8.7(28.2)	-3.0(3.0)
T11 P	mean(sd)	19.324(5.669)	19.441(5.614)	-0.118(1.998)	19.647(5.398)	20.400(6.055)	-0.753(1.383)
	min(max)	8.4(28.3)	9.1(28.6)	-4.2(4.2)	9.0(27.3)	9.3(28.6)	-4.1(1.4)
T12 A	mean(sd)	20.394(5.780)	19.759(5.386)	0.635(0.923)	20.012(5.691)	20.276(5.859)	-0.265(1.275)
	min(max)	9.8(28.5)	10.6(27.6)	-0.9(2.3)	10.5(29.7)	10.2(28.1)	-2.6(2.0)
T12 P	mean(sd)	20.712(5.752)	20.418(5.748)	0.294(1.003)	20.694(5.637)	20.976(6.182)	-0.282(1.539)
	min(max)	10.4(29.0)	10.8(29.9)	-1.4(2.5)	11.1(30.8)	10.1(28.9)	-3.8(1.9)
L1 A	mean(sd)	21.400(6.284)	20.888(5.431)	0.512(1.495)	21.124(5.720)	21.424(6.322)	-0.300(1.267)
	min(max)	9.9(31.8)	10.8(29.4)	-1.9(4.5)	10.5(28.3)	10.2(31.0)	-3.1(1.5)
L1 P	mean(sd)	21.988(6.683)	21.800(5.801)	0.188(1.778)	21.600(5.698)	22.088(6.562)	-0.488(1.336)
	min(max)	10.1(32.2)	11.1(31.0)	-2.7(4.5)	11.0(29.6)	10.1(31.9)	-3.0(1.3)
L2 A	mean(sd)	22.694(6.569)	21.924(6.071)	0.771(2.127)	21.971(6.224)	22.041(6.205)	-0.071(1.528)
	min(max)	10.1(33.1)	10.7(29.8)	-4.1(4.1)	11.0(30.7)	10.7(30.5)	-2.9(2.3)
L2 P	mean(sd)	23.224(6.717)	22.471(6.323)	0.753(1.910)	22.147(6.010)	22.471(6.588)	-0.324(1.487)
	min(max)	9.6(33.1)	11.1(31.5)	-2.3(4.7)	11.3(30.5)	10.1(32.2)	-3.0(1.8)
L3 A	mean(sd)	22.759(6.866)	21.988(6.382)	0.770(1.768)	22.518(6.561)	22.082(6.676)	0.435(0.927)
	min(max)	9.6(32.7)	9.6(31.3)	-1.7(4.1)	10.4(31.7)	10.1(32.0)	-1.9(2.0)
L3 P	mean(sd)	22.988(6.924)	22.041(6.176)	0.947(2.058)	22.371(6.418)	22.329(6.805)	0.041(0.945)
	min(max)	9.6(33.5)	10.1(31.3)	-2.6(4.9)	10.4(31.8)	10.5(33.6)	-1.9(1.4)

Table 2.8: Results of the estimates and improved 95% bootstrap- t confidence intervals of CCC_R and the estimates of $\rho_{c,rm}$ for two situations for vertebral body data

Location	A/P	CCC_R^{abs} (95% C.I.)	CCC_R^{inter} (95% C.I.)	CCC_R^{intra} (95% C.I.)	$\rho_{c,rm}$ (95% C.I.)
Situation 1 (N=17, random radiologists and random time points)					
T10	Anterior	0.966 (0.939, 1.000)	0.965 (0.940, 1.000)	0.966 (0.938, 1.000)	0.965 (0.940, 1.000)
	Posterior	0.936 (0.889, 1.000)	0.956 (0.927, 1.000)	0.937 (0.889, 1.000)	0.956 (0.929, 1.000)
T11	Anterior	0.950 (0.920, 0.999)	0.961 (0.940, 1.000)	0.959 (0.927, 1.000)	0.961 (0.941, 1.000)
	Posterior	0.952 (0.916, 1.000)	0.957 (0.929, 1.000)	0.947 (0.952, 1.000)	0.957 (0.929, 1.000)
T12	Anterior	0.980 (0.965, 1.000)	0.979 (0.964, 1.000)	0.983 (0.958, 1.000)	0.979 (0.964, 1.000)
	Posterior	0.975 (0.955, 1.000)	0.970 (0.945, 1.000)	0.975 (0.936, 1.000)	0.968 (0.948, 1.000)
L1	Anterior	0.973 (0.949, 1.000)	0.976 (0.956, 1.000)	0.973 (0.952, 1.000)	0.976 (0.958, 1.000)
	Posterior	0.968 (0.946, 1.000)	0.976 (0.959, 1.000)	0.968 (0.950, 0.990)	0.976 (0.963, 1.000)
L2	Anterior	0.955 (0.920, 1.000)	0.953 (0.921, 1.000)	0.968 (0.940, 1.000)	0.951 (0.926, 1.000)
	Posterior	0.964 (0.941, 1.000)	0.968 (0.946, 1.000)	0.960 (0.926, 1.000)	0.964 (0.947, 1.000)
L3	Anterior	0.973 (0.955, 1.000)	0.974 (0.954, 1.000)	0.975 (0.955, 1.000)	0.974 (0.957, 1.000)
	Posterior	0.968 (0.947, 0.999)	0.972 (0.951, 1.000)	0.968 (0.943, 1.000)	0.971 (0.952, 1.000)
Overall		0.963 (0.939, 0.993)	0.967 (0.941, 1.000)	0.965 (0.936, 0.991)	0.967 (0.970, 0.993)
Situation 2 (N=90, random radiologists and random "time points"^a)					
		0.939 (0.888, 0.986)	0.937 (0.835, 1.000)	0.957 (0.944, 0.971)	0.937 (0.879, 0.992)

^a "time points" represent the locations.

2.5.2 Image Data

In the second example, we use the image data of 2D-echocardiogram from a Duke Cardiovascular Core Lab for illustration. The purpose of the study is to evaluate the pulmonary arterial hypertension measures by 2D-echocardiogram. To assess the agreement between sonographers who measure the 2D-echocardiogram images, two sonographers make measurement twice on 10 patients. Here we assumed that two sonographers ($J = 2$) are randomly selected from the sonographer population, and two time points ($K = 2$) are randomly selected from the time population. The variables of interests for assessing agreement are Visual Effect Fraction (VISEF), Left Ventricular Ejection Fraction (LVEF), and Right Atrium Volume (RAV). There are 10, 9, and 8 subjects with complete data for variables VISEF, LVEF, and RAV respectively. For illustration, we use this data set to determine the degree of the inter-observer, intra-observer, and absolute agreement between two sonographers for measurements on each of three variables. The descriptive statistics for three variables are shown in Tables 2.9-2.11, which indicated that differences are small between readings by the same reader at the same location (Table 2.9), between readers at the same time and the same location (Table 2.10), and between readers at different time and at the same location (Table 2.11). Figures 2.5-2.6 showed the scatter plots for measurements between two sonographers. Table 2.12 showed the point estimates with the corresponding improved 95% bootstrap- t confidence intervals based on CCC_R for inter-rater, intra-rater, and absolute agreement, respectively. The results indicated the inter-observer and absolute agreement between two sonographers are lower than intra-observer agreement by the same sonographer on these three variables. There is a need to investigate the reason for the disagreement of these two sonographers. In addition, Table 2.12 shows the estimates of $\rho_{c,rm}$ proposed by King et al. (2007) for CCC when sonographers and time are treated as fixed effects. The estimates of CCC_R^{inter} are almost the same as the estimate of $\rho_{c,rm}$ but their corresponding confidence intervals are different.

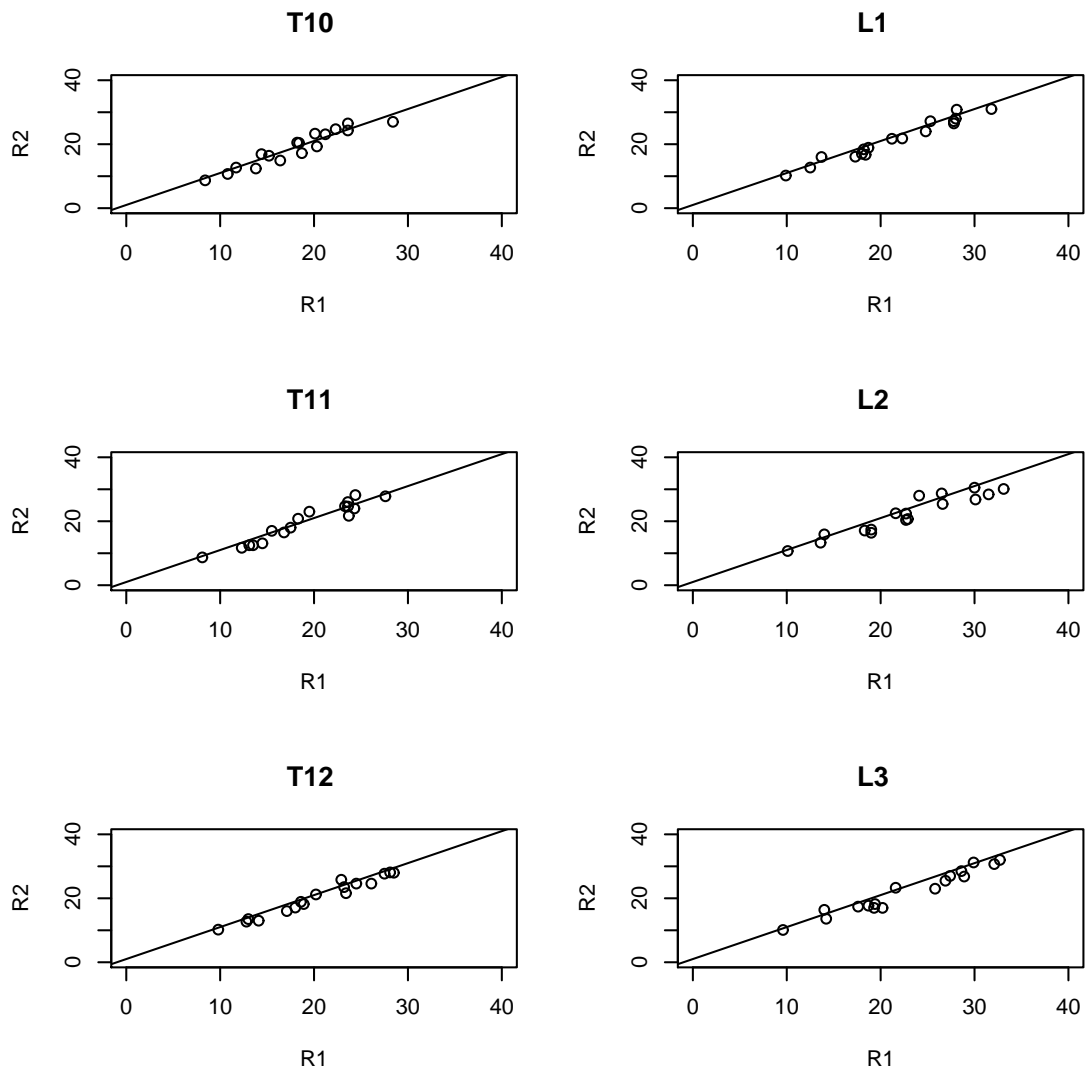


Figure 2.1: Scatter plots between readers for the Anterior at the first time point in Situation 1 for vertebral body data

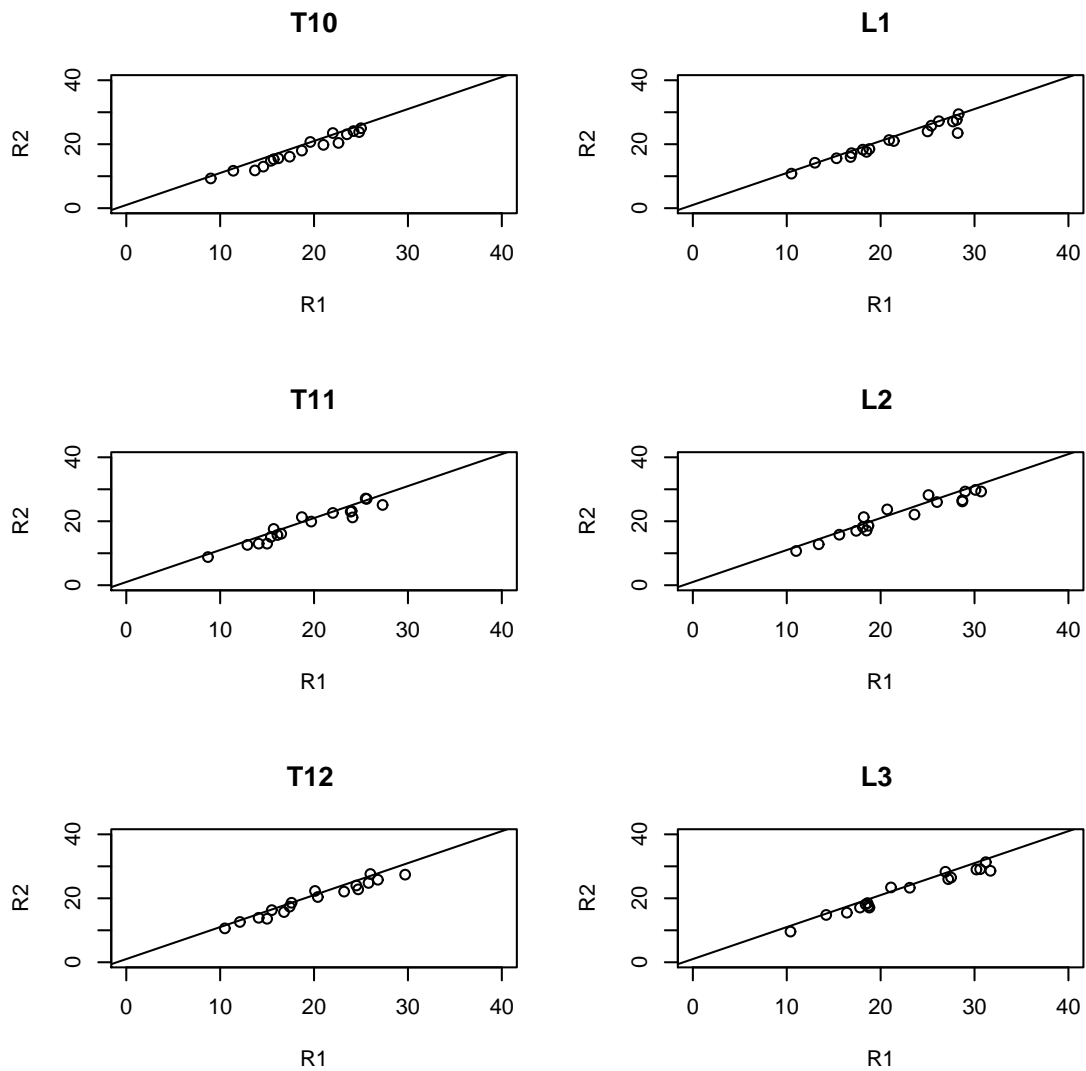


Figure 2.2: Scatter plots between readers for the Anterior at the second time point in Situation 1 for vertebral body data

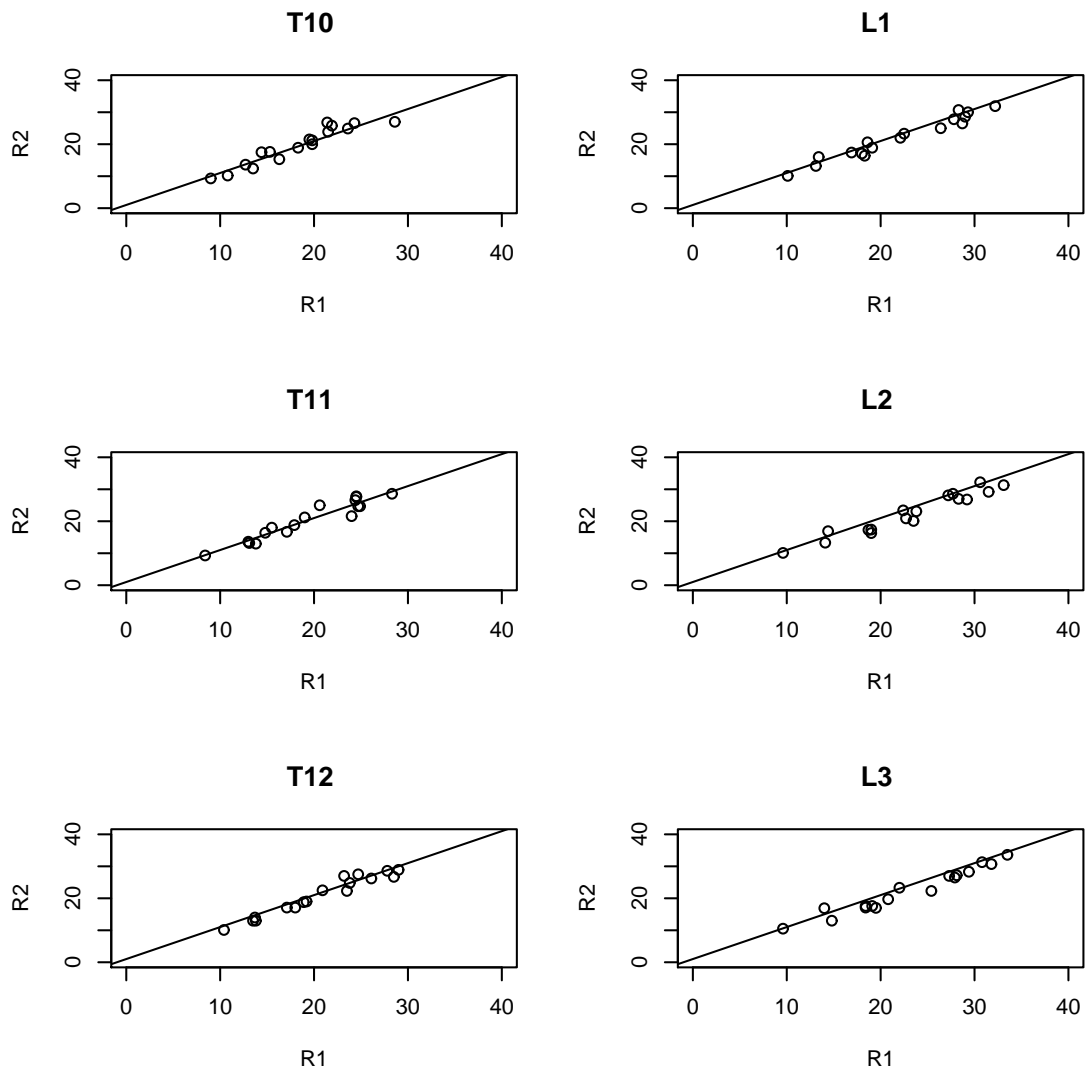


Figure 2.3: Scatter plots between readers for the Posterior at the first time point in Situation 1 for vertebral body data

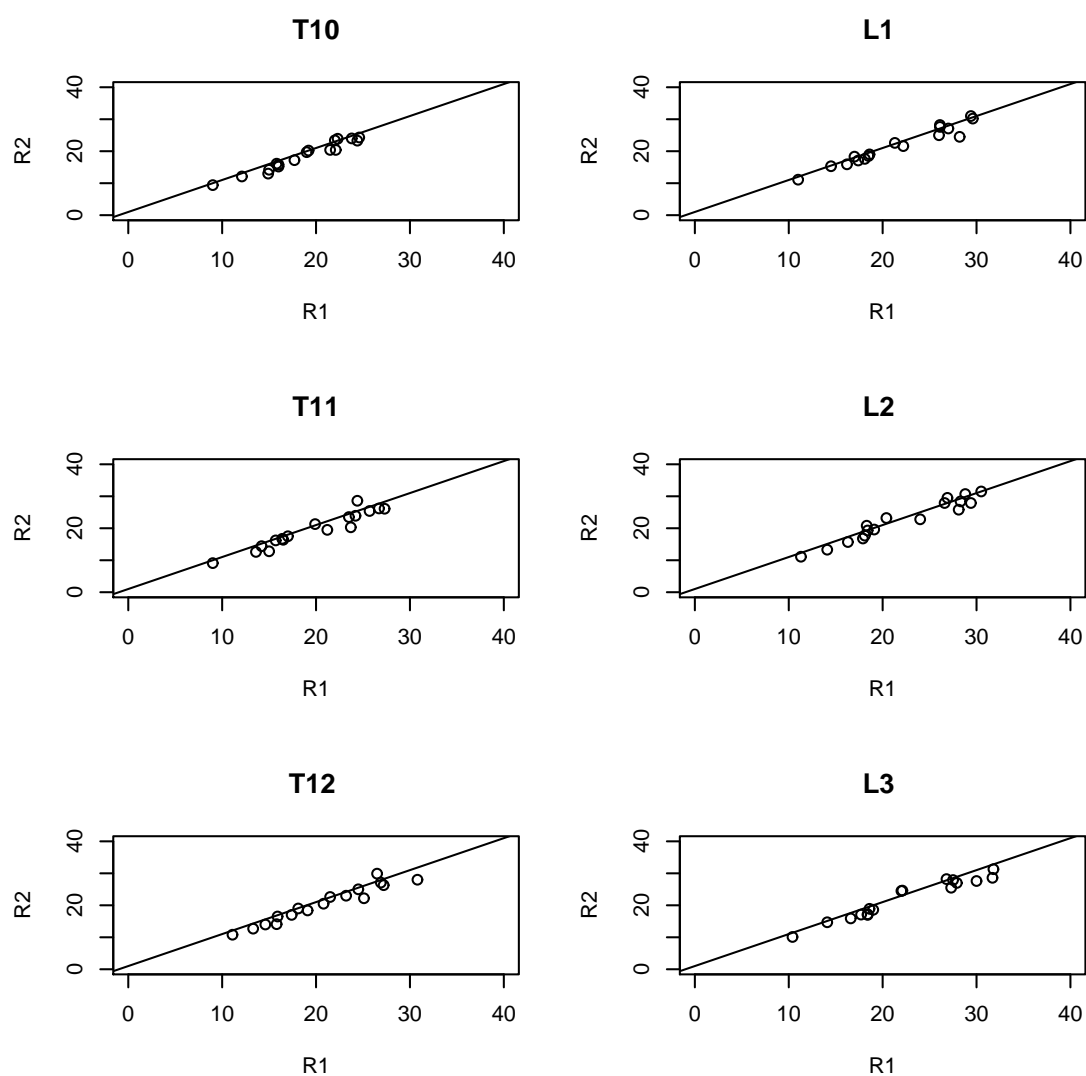


Figure 2.4: Scatter plots between readers for the Posterior at the second time point in Situation 1

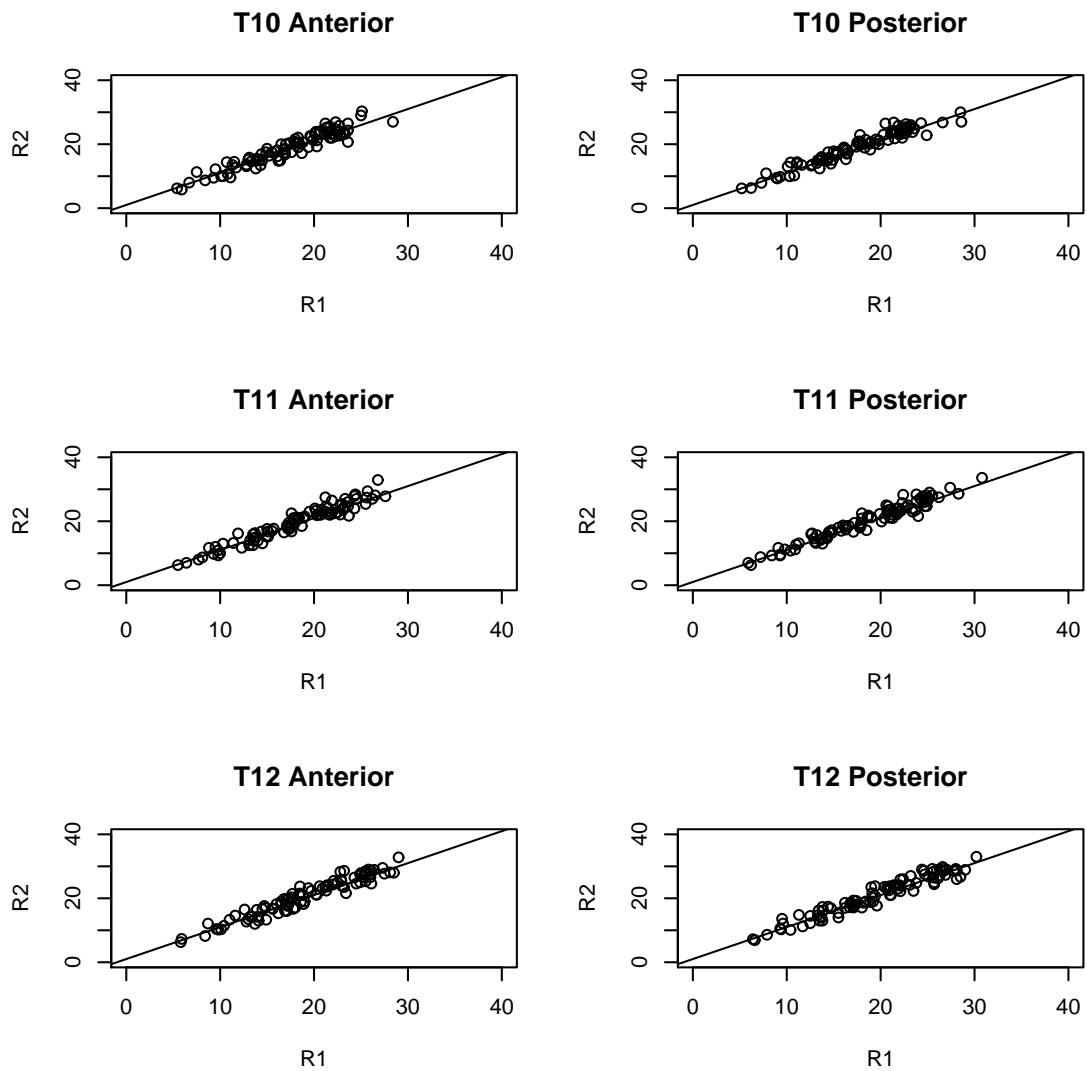


Figure 2.5: Scatter plots between readers of all T locations in Situation 2 for vertebral body data

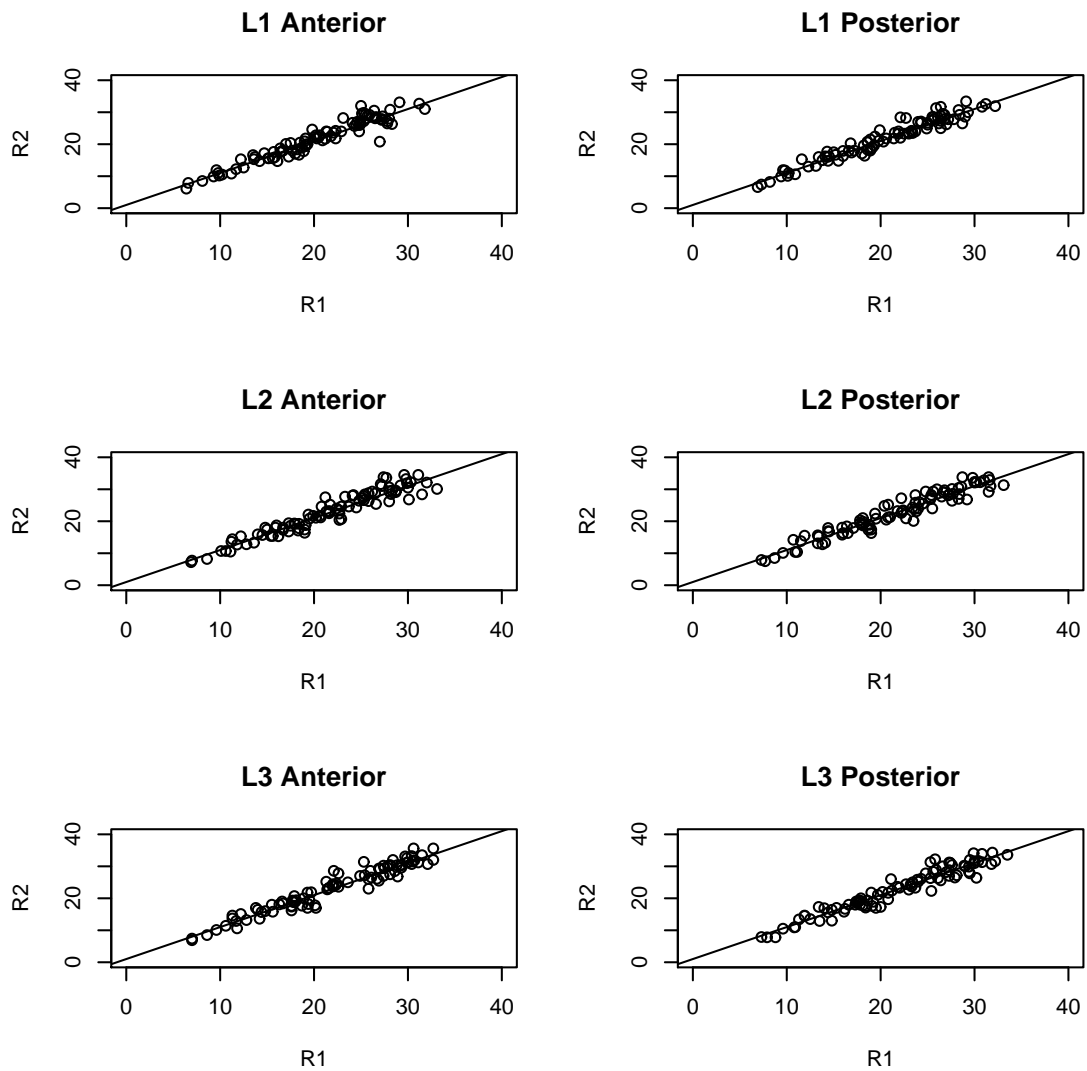


Figure 2.6: Scatter plots between readers for all L locations in Situation 2 for vertebral body data

Table 2.9: Description statistics for readings between times by readers and variables for image data

Variable	Reader 1				Reader 2		
		1 st Read	2 nd Read	Diff	1 st Read	2 nd Read	Diff
VISEF	mean(sd)	0.565 (0.111)	0.575 (0.103)	-0.010(0.032)	0.604 (0.084)	0.619(0.096)	-0.015(0.040)
	min(max)	0.35(0.70)	0.40(0.70)	-0.05(0.05)	0.45(0.75)	0.42(0.75)	-0.11(0.03)
LVEF	mean(sd)	58.881(11.106)	58.103(11.003)	0.779(2.585)	61.425(9.447)	63.342(10.278)	-1.917(4.938)
	min(max)	39.927(71.927)	41.187(71.813)	-2.630(5.11)	44.407(75.590)	42.820(75.747)	-12.367(3.39)
RAV	mean(sd)	28.025(14.138)	27.905(8.632)	0.120(8.429)	38.345(22.280)	41.666(24.589)	-3.322(9.720)
	min(max)	10.253(52.563)	16.603(43.480)	-14.14(14.973)	14.570 (77.350)	13.990(79.900)	-19.16(11.280)

Table 2.10: Description statistics for readings between readers by times and variables for image data

Variable		1 st Read			2 nd Read		
		Reader 1	Reader 2	Diff	Reader 1	Reader 2	Diff
VISEF	mean(sd)	0.565 (0.111)	0.604 (0.084)	-0.039(0.070)	0.575 (0.103)	0.619(0.096)	-0.044(0.088)
	min(max)	0.35(0.70)	0.45(0.75)	-0.15 (0.08)	0.40(0.70)	0.42(0.75)	-0.20(0.09)
LVEF	mean(sd)	58.881(11.106)	61.425(9.447)	-2.544(6.148)	58.103(11.003)	63.342(10.278)	-5.239(7.942)
	min(max)	39.927(71.927)	44.407(75.590)	-11.107(6.524)	41.187(71.813)	42.820(75.747)	-15.130(6.004)
RAV	mean(sd)	28.025(14.138)	38.345(22.280)	-10.320(9.670)	27.905(8.632)	41.666(24.589)	-13.761(16.769)
	min(max)	10.253(52.563)	14.570 (77.350)	-24.787(2.626)	16.603(43.480)	13.990(79.900)	-38.440(5.147)

Table 2.11: Description statistics for readings between readers and times by variables for image data

Variable		Reader 1	Reader 2	Diff	Reader 1	Reader 2	Diff
		1 st Read	2 nd Read		2 nd Read	1 st Read	
VISEF	mean(sd)	0.565 (0.111)	0.619(0.096)	-0.054(0.081)	0.575 (0.103)	0.604 (0.084)	-0.029(0.081)
	min(max)	0.35(0.70)	0.42(0.75)	-0.15 (0.09)	0.40(0.70)	0.45(0.75)	-0.20(0.08)
LVEF	mean(sd)	58.881(11.106)	63.342(10.278)	-4.460(8.115)	58.103(11.003)	61.425(9.447)	-3.323(7.377)
	min(max)	39.927(71.927)	42.820(75.747)	-16.773(4.954)	41.187(71.813)	44.407(75.590)	-14.973(7.574)
RAV	mean(sd)	28.025(14.138)	41.666(24.589)	-13.641(13.501)	27.905(8.632)	38.345(22.280)	-10.440(14.889)
	min(max)	10.253(52.563)	13.990(79.900)	-36.14(8.370)	16.603(43.480)	14.570 (77.350)	-39.76(3.804)

Table 2.12: Results of the estimates and improved 95% bootstrap- t confidence intervals of CCC_R and the estimates of $\rho_{c,rm}$ for image data

Variable	CCC_R^{abs} (95% C.I.)	CCC_R^{inter} (95% C.I.)	CCC_R^{intra} (95% C.I.)	$\rho_{c,rm}$ (95% C.I.)
VISEF	0.611 (0.320, 1.000)	0.624 (0.339, 1.000)	0.859 (0.743, 1.000)	0.624 (0.411, 1.000)
LVEF	0.679 (0.395, 1.000)	0.770 (0.616, 1.000)	0.869 (0.727, 1.000)	0.770 (0.620, 1.000)
RAV	0.583 (0.333, 1.000)	0.600 (0.413, 1.000)	0.727 (0.491, 1.000)	0.600 (0.414, 1.000)

Chapter 3

Assessing agreement with intraclass correlation coefficient and concordance correlation coefficient with repeated measures

3.1 Introduction

The intraclass correlation coefficient (ICC) and the concordance correlation coefficient (CCC) are two popular indices for assessing agreement between continuous measurements taken from different observers. ICC and CCC are usually used for data without and with replications and the comparison between these two indices based on this data structure under a general model is reported by Chen and Barnhart (2008). However, we can not use the methodology if repeated measurements rather than replications are collected. Several authors have studied the agreement indices of ICC and CCC for data with repeated measurements. Vangeneugden et al. (2005) investigated the ICC by linear mixed models with serial correlation for inter-observer, intra-observer, and absolute agreement, where both observer and time are treated as random effects. King et al. (2007) proposed a CCC for assessing inter-observer agreement for a response with repeated measurement, where both observer and time are treated as fixed effects. Chen and Barnhart (2009) also

proposed a *CCC* for assessing inter-observer, intra-observer, and absolute agreement for data with repeated measurements where observers and times are treated as random since researchers may assess agreement among many observers who take measurements at different time points.

In addition to treat both observer and time as random effects in defining *CCC*, there are situations that researchers may be interested in the case of random observer and fixed time, or fixed observer and random time. For example, a study is designed to assess the agreement among subjects' measurements taken by observers (e.g. nurse) for two shifts. Researches can conduct and assess the agreement study by randomly selecting several observers from the observer population and obtaining measurements from the chosen observers at two time points. In this study, observer is random and time is fixed because different observers may produce different measurements at the specific time. Another example is to assess agreement among a fixed number of observers for blood pressure measurements in a longitudinal study. Patients may have no scheduled visits to take measurements by these observers. In this example, observer is fixed and time is random because same observers may produce different measurements at different times. Therefore, any combinations of fixed or random effects for observer and time for agreement study of *CCC* or *ICC* can happen depending on the goals of researchers.

We propose new *CCCs* and *ICCs* for the remaining combinations of random or fixed effects for observer and time. We summarize and compare *CCCs* and *ICCs* between combinations of random or fixed effects for data with repeated measurements and illustrate the methodology with an example from image study.

3.2 Methodology

Consider that there are N randomly selected subjects where measurements are taken by J observers at K time points. Two factors, observer and time, can be treated either as random or fixed. If the factor is treated as random, the levels of this factor is treated as random sample from the corresponding population. If the factor is treated as fixed, the levels of this factor are the finite number of levels for this factor. In this section, we consider *CCCs* and *ICCs* for assessing inter-observer, intra-observer, and absolute agreement under four combinations of random or fixed effects of observer and time. (1) random observers

Table 3.1: Existing methods (Yes, No) for four cases comparing observer with repeated measures under different approaches and indices

Case	Observer	Time	Indices		
			Source	<i>CCC</i>	<i>ICC</i>
1	Random	Random		Yes	Yes
			inter-observer	<i>–yes</i>	<i>–yes</i>
			intra-observer	<i>–yes</i>	<i>–yes</i>
			absolute	<i>–yes</i>	<i>–yes</i>
2	Random	Fixed		No	No
3	Fixed	Random		No	No
4	Fixed	Fixed		Yes	No
			inter-observer	<i>–yes</i>	<i>–no</i>
			intra-observer	<i>–no</i>	<i>–no</i>
			absolute	<i>–no</i>	<i>–no</i>

- (1) random observer and random time; (2) random observer and fixed time;
(3) fixed observer and random time; (4) fixed observer and fixed time.

and random times. (2) random observers and fixed times. (3) fixed observers and random times. (4) fixed observers and fixed times.

Table 3.1 shows whether there exist *CCC* or *ICC* index for these four cases. New *CCC*s for case (1) have been proposed in Chapter 2. Vangeneugden et al. (2005) has proposed *ICC*s for case (1), and King et al. (2007) has proposed *CCC* for case (4) for assessing inter-observer agreement. In this paper, we propose new *CCC* and *ICC* indices for cases (2) through (4). We consider inter-observer, intra-observe, and absolute agreement for both *CCC* and *ICC*. The inter-observer agreement is assessing the agreement of an observed reading by one rater to an observed reading by a different rater on the same subject at the same time (*i.e.* Y_{ijk} vs $Y_{ij'k}$). The intra-observer agreement is assessing the agreement of an observed reading on one time to an observed reading on another time by the same observer on the same subject (*i.e.* Y_{ijk} vs $Y_{ijk'}$). The absolute agreement is assessing the agreement of an observed reading by one observer on one time to an observed reading by a different observer at a different time on the same subject (*i.e.* Y_{ijk} vs $Y_{ij'k'}$).

The definitions of *CCC*s (existing and new) for these four cases are summarized

in Table 3.2. Details about these new *CCC*s are presented in Section 3.2.1 to Section 3.2.3, respectively. For the definition of *ICC*, let $Y_{ijk} = \mu + \mu_i^P + \mu_k^T + \mu_j^O + \mu_{ik}^{PT} + \mu_{ij}^{PO} + \mu_{kj}^{TO} + e_{ijk}^{POT}$ be the three-way mixed-effects model, where μ_i^P is the random effect of subject i with the variance component σ_P^2 of factor P , etc. If the factor in the mixed-effects model is fixed, then there is a constraint for this factor. Following McGraw and Wong (1996), *ICC* is defined as the ratio of covariance divided by total variance due to either random or fixed effects. The definitions of *ICC*s (existing and new) for these four cases are summarized in Table 3.3. Details about these new *ICC* are presented in Section 3.2.4.

Table 3.2: Definitions and results of CCC for inter-observer, intra-observer, and absolute agreement under four cases

Case	CCC	
	<i>inter-observer</i>	<i>intra-observer</i>
(1)	$1 - \frac{E_{jj'k} E_i([Y_{ijk} - Y_{ij'k}]^2 j, j', k)}{E_{jj'k} E_I([Y_{ijk} - Y_{ij'k}]^2 j, j', k)}$ $= \frac{E_{jj'k} Cov_i(Y_{ijk}, Y_{ij'k} j, j', k)}{\sigma^2 + \tau^2 - Cov_{jj'k}(\mu_{jk}, \mu_{j'k})}$	$1 - \frac{E_{jkk'} E_i([Y_{ijk} - Y_{ijk'}]^2 j, k, k')}{E_{jkk'} E_I([Y_{ijk} - Y_{ijk'}]^2 j, k, k')}$ $= \frac{E_{jkk'} Cov_i(Y_{ijk}, Y_{ijk'} j, k, k')}{\sigma^2 + \tau^2 - Cov_{jkk'}(\mu_{jk}, \mu_{jk'})}$
(2)	$1 - \frac{\sum_{k=1}^K E_{jj'} E_i([Y_{ijk} - Y_{ij'k}]^2 j, j', k)^a}{\sum_{k=1}^K E_{jj'} E_I([Y_{ijk} - Y_{ij'k}]^2 j, j', k)}$ $= \frac{\sum_{k=1}^K E_{jj'} Cov_i(Y_{ijk}, Y_{ij'k} j, j', k)}{\sum_{k=1}^K \{\sigma_k^2 + \tau_k^2 - Cov_{jj'}(\mu_{jk}, \mu_{j'k})\}}$	$1 - \frac{\sum_{k=1}^{K-1} \sum_{k'=k+1}^K E_j E_i([Y_{ijk} - Y_{ijk'}]^2 j, k, k')^a}{\sum_{k=1}^{K-1} \sum_{k'=k+1}^K E_j E_I([Y_{ijk} - Y_{ijk'}]^2 j, k, k')}$ $= \frac{\sum_{k=1}^{K-1} \sum_{k'=k+1}^K 2E_j Cov_i(Y_{ijk}, Y_{ijk'} j, k, k')}{\sum_{k=1}^{K-1} \sum_{k'=k+1}^K \{\sigma_k^2 + \sigma_{k'}^2 + \tau_k^2 + \tau_{k'}^2 - 2Cov_j(\mu_{jk}, \mu_{jk'})\}}$
(3)	$1 - \frac{\sum_{j=1}^{J-1} \sum_{j'=j+1}^J E_k E_i([Y_{ijk} - Y_{ij'k}]^2 j, j', k)^a}{\sum_{j=1}^{J-1} \sum_{j'=j+1}^J E_k E_I([Y_{ijk} - Y_{ij'k}]^2 j, j', k)}$ $= \frac{\sum_{j=1}^{J-1} \sum_{j'=j+1}^J 2E_k Cov_i(Y_{ijk}, Y_{ij'k} j, j', k)}{\sum_{j=1}^{J-1} \sum_{j'=j+1}^J \{\sigma_j^2 + \sigma_{j'}^2 + \tau_j^2 + \tau_{j'}^2 - 2Cov_k(\mu_{jk}, \mu_{j'k})\}}$	$1 - \frac{\sum_{j=1}^J E_{kk'} E_i([Y_{ijk} - Y_{ijk'}]^2 j, k, k')^a}{\sum_{j=1}^J E_{kk'} E_I([Y_{ijk} - Y_{ijk'}]^2 j, k, k')}$ $= \frac{\sum_{j=1}^J E_{kk'} Cov_i(Y_{ijk}, Y_{ijk'} j, k, k')}{\sum_{j=1}^J \{\sigma_j^2 + \tau_j^2 - Cov_{kk'}(\mu_{jk}, \mu_{jk'})\}}$
(4)	$1 - \frac{\sum_{k=1}^K \sum_{j=1}^{J-1} \sum_{j'=j+1}^J E_i([Y_{ijk} - Y_{ij'k}]^2 j, j', k)}{\sum_{k=1}^K \sum_{j=1}^{J-1} \sum_{j'=j+1}^J E_I([Y_{ijk} - Y_{ij'k}]^2 j, j', k)}$ $= \frac{\sum_{k=1}^K \sum_{j=1}^{J-1} \sum_{j'=j+1}^J 2Cov_i(Y_{ijk}, Y_{ij'k} j, j', k)}{\sum_{k=1}^K \sum_{j=1}^{J-1} \sum_{j'=j+1}^J (\sigma_{jk}^2 + \sigma_{j'k}^2 + (\mu_{jk} - \mu_{j'k})^2)}$	$1 - \frac{\sum_{k=1}^{K-1} \sum_{k'=k+1}^K \sum_{j=1}^J E_i([Y_{ijk} - Y_{ijk'}]^2 j, k, k')^a}{\sum_{k=1}^{K-1} \sum_{k'=k+1}^K \sum_{j=1}^J E_I([Y_{ijk} - Y_{ijk'}]^2 j, k, k')}$ $= \frac{\sum_{k=1}^{K-1} \sum_{k'=k+1}^K \sum_{j=1}^J 2Cov_i(Y_{ijk}, Y_{ijk'} j, k, k')}{\sum_{k=1}^{K-1} \sum_{k'=k+1}^K \sum_{j=1}^J \{\sigma_{jk}^2 + \sigma_{jk'}^2 + (\mu_{jk} - \mu_{jk'})^2\}}$

(1) random observer and random time; (2) random observer and fixed time;

(3) fixed observer and random time; (4) fixed observer and fixed time.

^a New CCC indices.

Case	CCC
	<i>absolute</i>
(1)	$1 - \frac{E_{jj'kk'} E_i([Y_{ijk} - Y_{ij'k'}]^2 j, j', k, k')}{E_{jj'kk'} E_I([Y_{ijk} - Y_{ij'k'}]^2 j, j', k, k')}$ $= \frac{E_{jj'kk'} Cov_i(Y_{ijk}, Y_{ij'k'} j, j', k, k')}{\sigma^2 + \tau^2 - Cov_{jj'kk'}(\mu_{jk}, \mu_{j'k'})}$
(2)	$1 - \frac{\sum_{k=1}^{K-1} \sum_{k'=k+1}^K E_{jj'} E_i([Y_{ijk} - Y_{ij'k'}]^2 j, j', k, k')^a}{\sum_{k=1}^{K-1} \sum_{k'=k+1}^K E_{jj'} E_I([Y_{ijk} - Y_{ij'k'}]^2 j, j', k, k')}$ $= \frac{\sum_{k=1}^{K-1} \sum_{k'=k+1}^K 2E_{jj'} Cov_i(Y_{ijk}, Y_{ij'k'} j, j', k, k')}{\sum_{k=1}^{K-1} \sum_{k'=k+1}^K \{\sigma_k^2 + \sigma_{k'}^2 + \tau_k^2 + \tau_{k'}^2 - 2Cov_{jj'}(\mu_{jk}, \mu_{j'k'})\}}$
(3)	$1 - \frac{\sum_{j=1}^{J-1} \sum_{j'=j+1}^J E_{kk'} E_i([Y_{ijk} - Y_{ij'k'}]^2 j, j', k, k')^a}{\sum_{j=1}^{J-1} \sum_{j'=j+1}^J E_{kk'} E_I([Y_{ijk} - Y_{ij'k'}]^2 j, j', k, k')}$ $= \frac{\sum_{j=1}^{J-1} \sum_{j'=j+1}^J 2E_{kk'} Cov_i(Y_{ijk}, Y_{ij'k'} j, j', k, k')}{\sum_{j=1}^{J-1} \sum_{j'=j+1}^J \{\sigma_j^2 + \sigma_{j'}^2 + \tau_j^2 + \tau_{j'}^2 - 2Cov_{kk'}(\mu_{jk}, \mu_{j'k'})\}}$
(4)	$1 - \frac{\sum_{k=1}^{K-1} \sum_{k'=K+1}^K \sum_{j=1}^{J-1} \sum_{j'=j+1}^J E_i([Y_{ijk} - Y_{ij'k'}]^2 j, j', k, k')^a}{\sum_{k=1}^{K-1} \sum_{k'=K+1}^K \sum_{j=1}^{J-1} \sum_{j'=j+1}^J E_I([Y_{ijk} - Y_{ij'k'}]^2 j, j', k, k')}$ $= \frac{\sum_{k=1}^{K-1} \sum_{k'=K+1}^K \sum_{j=1}^{J-1} \sum_{j'=j+1}^J 2Cov_i(Y_{ijk}, Y_{ij'k'} j, j', k, k')}{\sum_{k=1}^{K-1} \sum_{k'=K+1}^K \sum_{j=1}^{J-1} \sum_{j'=j+1}^J \{\sigma_{jk}^2 + \sigma_{j'k'}^2 + (\mu_{jk} - \mu_{j'k'})^2\}}$

(1) random observer and random time; (2) random observer and fixed time;

(3) fixed observer and random time; (4) fixed observer and fixed time.

^a New CCC indices.

Table 3.3: Definitions of *ICC* for inter-observer, intra-observer, and absolute agreement under four cases based on the corresponding ANOVA models^a

Case	ICC		
	<i>inter-observer</i>	<i>intra-observer</i>	<i>absolute</i>
(1)	$\frac{\sigma_P^2 + \sigma_T^2 + \sigma_{PT}^2}{\sigma_P^2 + \sigma_O^2 + \sigma_T^2 + \sigma_{PO}^2 + \sigma_{PT}^2 + \sigma_{OT}^2 + \sigma_E^2}$	$\frac{\sigma_P^2 + \sigma_O^2 + \sigma_{PO}^2}{\sigma_P^2 + \sigma_O^2 + \sigma_T^2 + \sigma_{PO}^2 + \sigma_{PT}^2 + \sigma_{OT}^2 + \sigma_E^2}$	$\frac{\sigma_P^2}{\sigma_P^2 + \sigma_O^2 + \sigma_T^2 + \sigma_{PO}^2 + \sigma_{PT}^2 + \sigma_{OT}^2 + \sigma_E^2}$
(2)	$\frac{\sigma_P^2 + \sigma_{Pt}^2}{\sigma_P^2 + \sigma_O^2 + \phi_t^2 + \sigma_{PO}^2 + \sigma_{Pt}^2 + \sigma_{Ot}^2 + \sigma_E^2}^b$	$\frac{\sigma_P^2 + \sigma_O^2 + \sigma_{PO}^2 - \frac{1}{K-1}\sigma_{Pt}^2 - \frac{1}{K-1}\sigma_{Ot}^2}{\sigma_P^2 + \sigma_O^2 + \phi_t^2 + \sigma_{PO}^2 + \sigma_{Pt}^2 + \sigma_{Ot}^2 + \sigma_E^2}^b$	$\frac{\sigma_P^2 - \frac{1}{K-1}\sigma_{Pt}^2}{\sigma_P^2 + \sigma_O^2 + \phi_t^2 + \sigma_{PO}^2 + \sigma_{Pt}^2 + \sigma_{Ot}^2 + \sigma_E^2}^b$
(3)	$\frac{\sigma_P^2 + \sigma_T^2 + \sigma_{PT}^2 - \frac{1}{J-1}\sigma_{Po}^2 - \frac{1}{J-1}\sigma_{oT}^2}{\sigma_P^2 + \phi_o^2 + \sigma_T^2 + \sigma_{Po}^2 + \sigma_{PT}^2 + \sigma_{oT}^2 + \sigma_E^2}^b$	$\frac{\sigma_P^2 + \sigma_{Po}^2}{\sigma_P^2 + \phi_o^2 + \sigma_T^2 + \sigma_{Po}^2 + \sigma_{PT}^2 + \sigma_{oT}^2 + \sigma_E^2}^b$	$\frac{\sigma_P^2 - \frac{1}{J-1}\sigma_{Po}^2}{\sigma_P^2 + \phi_o^2 + \sigma_T^2 + \sigma_{Po}^2 + \sigma_{PT}^2 + \sigma_{oT}^2 + \sigma_E^2}^b$
(4)	$\frac{\sigma_P^2 + \sigma_{Pt}^2 - \frac{1}{J-1}\sigma_{Po}^2}{\sigma_P^2 + \phi_o^2 + \phi_t^2 + \sigma_{Po}^2 + \sigma_{Pt}^2 + \phi_{ot}^2 + \sigma_E^2}^b$	$\frac{\sigma_P^2 + \sigma_{Po}^2 - \frac{1}{K-1}\sigma_{Pt}^2}{\sigma_P^2 + \phi_o^2 + \phi_t^2 + \sigma_{Po}^2 + \sigma_{Pt}^2 + \phi_{ot}^2 + \sigma_E^2}^b$	$\frac{\sigma_P^2 - \frac{1}{K-1}\sigma_{Pt}^2 - \frac{1}{J-1}\sigma_{Po}^2}{\sigma_P^2 + \phi_o^2 + \phi_t^2 + \sigma_{Po}^2 + \sigma_{Pt}^2 + \phi_{ot}^2 + \sigma_E^2}^b$

(1) random observer and random time; (2) random observer and fixed time;

(3) fixed observer and random time; (4) fixed observer and fixed time.

^a See Section 3.2.4 for different ANOVA models for cases (1) through (4).

^b New *ICC* indices and see Section 3.2.4 for the definitions of ϕ terms.

3.2.1 CCC for random observers and fixed times

Lin (1989) developed the concordance correlation coefficient (CCC) to evaluate the agreement by scaling the mean square difference between two readings taken from two fixed observers as

$$\rho_c = 1 - \frac{E[(Y_{i1} - Y_{i2})^2]}{E_I[(Y_{i1} - Y_{i2})^2]},$$

where E_I is the conditional expectation given independence of Y_{i1} and Y_{i2} . Here we let $Y_{i1k}, Y_{i2k}, \dots, Y_{iJk}$ be the measurements made by J fixed observers where each observer has only one reading for each subject, then Barnhart et al. (2005) extended Lin's CCC and proposed a total CCC to assess the agreement among multiple observers for data with replications. The total-CCC can be expressed as

$$\rho_c(Y) = 1 - \frac{\sum_{j=1}^{J-1} \sum_{j'=j+1}^J E_i[(Y_{ijk} - Y_{ij'k})^2 / J(J-1)]}{\sum_{j=1}^{J-1} \sum_{j'=j+1}^J E_I[(Y_{ijk} - Y_{ij'k})^2 / J(J-1)]},$$

where E_I is the conditional expectation given independence of Y_{i1k}, Y_{i2k}, \dots , and Y_{iJk} . By extending the total-CCC to assess the agreement between random observers at random times for data with repeated measurements, we define the inter-observer CCC for random observer and fixed time by replacing the summation $\sum_{j=1}^{J-1} \sum_{j'=j+1}^J$ with the expectation $E_{jj'}$ in the new agreement index where $E_{jj'}$ is the expectation respect to random observers j, j' . We also use the summation $\sum_{k=1}^K$ to sum the elements with functions of fixed time effect k . Specifically,

$$CCC_{Ot}^{inter} = 1 - \frac{\sum_{k=1}^K E_{jj'} E_i((Y_{ijk} - Y_{ij'k})^2 | j, j', k)}{\sum_{k=1}^K E_{jj'} E_I((Y_{ijk} - Y_{ij'k})^2 | j, j', k)}, \quad (3.1)$$

where E_I is the conditional expectation given independence of Y_{ijk} and $Y_{ij'k}$ conditional on observer and time. With the notations of $E_j Var_i(Y_{ijk} | j, k) = E_j(\sigma_{jk}^2 | k) = \sigma_k^2$ and $Var_j E_i(Y_{ijk} | j, k) = Var_j(\mu_{jk} | k) = \tau_k^2$, we obtain

$$\begin{aligned} & E_{jj'} E_i((Y_{ijk} - Y_{ij'k})^2 | j, j', k) \\ &= 2\sigma_k^2 + E_{jj'}((\mu_{jk} - \mu_{j'k})^2 | k) - 2E_{jj'} Cov_i(Y_{ijk}, Y_{ij'k} | j, j', k) \\ &= 2\sigma_k^2 + 2\tau_k^2 - 2Cov_{jj'}(\mu_{jk}, \mu_{j'k}) + (E_j(\mu_{jk}) - E_{j'}(\mu_{j'k}))^2 - 2E_{jj'} Cov_i(Y_{ijk}, Y_{ij'k} | j, j', k), \end{aligned}$$

and $E_{jj'} E_I((Y_{ijk} - Y_{ij'k})^2 | j, j', k) = 2\sigma_k^2 + 2\tau_k^2 - 2Cov_{jj'}(\mu_{jk}, \mu_{j'k}) + (E_j(\mu_{jk}) - E_{j'}(\mu_{j'k}))^2$. Here we assume that $E_j(\mu_{jk}) = E_{j'}(\mu_{j'k})$, where $E_j(\mu_{jk}) = E_j(\mu_{jk} | k)$. Then, the inter-

observer agreement of CCC can be expressed as

$$\begin{aligned}
 CCC_{Ot}^{inter} &= 1 - \frac{\sum_{k=1}^K \{2\sigma_k^2 + 2\tau_k^2 - 2Cov_{jj'}(\mu_{jk}, \mu_{j'k}) - 2E_{jj'}Cov_i(Y_{ijk}, Y_{ij'k}|j, j', k)\}}{\sum_{k=1}^K \{2\sigma_k^2 + 2\tau_k^2 - 2Cov_{jj'}(\mu_{jk}, \mu_{j'k})\}} \\
 &= \frac{\sum_{k=1}^K E_{jj'}Cov_i(Y_{ijk}, Y_{ij'k}|j, j', k)}{\sum_{k=1}^K \{\sigma_k^2 + \tau_k^2 - Cov_{jj'}(\mu_{jk}, \mu_{j'k})\}} \quad (3.2)
 \end{aligned}$$

The intra-observer CCC for random observer and fixed time is obtained similarly by considering the expected mean square difference for measurements taken on the same subject by the same observer at different time points. We consider two situations for the intra-observer agreement: (1) subjects' true values do not change over time; (2) subjects' true values change over time.

(1) Subjects' true values not changing over time

If subjects' true values do not change over time by the same observer with the assumption of $E_j(\mu_{jk}) = E_j(\mu_{jk'})$, the CCC_{Ot}^{intra} is defined as

$$\begin{aligned}
 CCC_{Ot}^{intra} &= 1 - \frac{\sum_{k=1}^{K-1} \sum_{k'=k+1}^K E_j E_i([Y_{ijk} - Y_{ijk'}]^2 | j, k, k')}{\sum_{k=1}^{K-1} \sum_{k'=k+1}^K E_j E_i([Y_{ijk} - Y_{ijk'}]^2 | j, k, k')} \quad (3.3)
 \end{aligned}$$

$$\begin{aligned}
 &= 1 - \frac{\sum_{k \neq k'} \{\sigma_k^2 + \sigma_{k'}^2 + \tau_k^2 + \tau_{k'}^2 - 2Cov_j(\mu_{jk}, \mu_{jk'}) - 2E_j Cov_i(Y_{ijk}, Y_{ijk'} | j, k, k')\}}{\sum_{k \neq k'} \{\sigma_k^2 + \sigma_{k'}^2 + \tau_k^2 + \tau_{k'}^2 - 2Cov_j(\mu_{jk}, \mu_{jk'})\}} \\
 &= \frac{\sum_{k=1}^{K-1} \sum_{k'=k+1}^K 2E_j Cov_i(Y_{ijk}, Y_{ijk'} | j, k, k')}{\sum_{k=1}^{K-1} \sum_{k'=k+1}^K \{\sigma_k^2 + \sigma_{k'}^2 + \tau_k^2 + \tau_{k'}^2 - 2Cov_j(\mu_{jk}, \mu_{jk'})\}}, \quad (3.4)
 \end{aligned}$$

where E_I is the conditional expectation given independence of Y_{ijk} and $Y_{ijk'}$ conditional on subject and time.

(2) Subjects' true values changing over time

However, if subjects' true values change over time, the mean square difference of $Y_{ijk} - Y_{ijk'}$ should be replaced by $(Y_{ijk} - \mu_{jk}) - (Y_{ijk'} - \mu_{jk'})$ so that the intra-observer agreement is not affected by the subjects' readings from time. In this case, the CCC_{Ot}^{intra} can be expressed as

$$\begin{aligned}
CCC_{Ot}^{intra} &= 1 - \frac{\sum_{k=1}^{K-1} \sum_{k'=k+1}^K E_j E_i [(Y_{ijk} - \mu_{jk}) - (Y_{ijk'} - \mu_{jk'})]^2 |j, k, k')}{\sum_{k=1}^{K-1} \sum_{k'=k+1}^K E_j E_i [(Y_{ijk} - \mu_{jk}) - (Y_{ijk'} - \mu_{jk'})]^2 |j, k, k')} \quad (3.5) \\
&= 1 - \frac{\sum_{k=1}^{K-1} \sum_{k'=k+1}^K \{\sigma_k^2 + \sigma_{k'}^2 - 2E_j Cov_i(Y_{ijk}, Y_{ijk'} | j, k, k')\}}{\sum_{k=1}^{K-1} \sum_{k'=k+1}^K \{\sigma_k^2 + \sigma_{k'}^2\}} \\
&= \frac{\sum_{k=1}^{K-1} \sum_{k'=k+1}^K 2E_j Cov_i(Y_{ijk}, Y_{ijk'} | j, k, k')}{\sum_{k=1}^{K-1} \sum_{k'=k+1}^K (\sigma_k^2 + \sigma_{k'}^2)}, \quad (3.6)
\end{aligned}$$

where E_I is the conditional expectation given independence of Y_{ijk} and $Y_{ijk'}$ conditional on subject and time.

The absolute CCC for random observers and fixed time is obtained similarly by considering the expected mean square difference for measurements taken on the same subject by different observers at different time points. Here, we also consider two situations for the absolute agreement: (1) subjects' true values do not change over time; (2) subjects' true values change over time.

(1) Subjects' true values not changing over time

If subjects' true values do not change over time by the same observer with the assumption of $E_j(\mu_{jk}) = E_{j'}(\mu_{j'k'})$, the CCC_{Ot}^{abs} is defined as

$$\begin{aligned}
CCC_{Ot}^{abs} &= 1 - \frac{\sum_{k=1}^{K-1} \sum_{k'=k+1}^K E_{jj'} E_i [(Y_{ijk} - Y_{ij'k'})^2 | j, j', k, k')]}{\sum_{k=1}^{K-1} \sum_{k'=k+1}^K E_{jj'} E_i [(Y_{ijk} - Y_{ij'k'})^2 | j, j', k, k')]} \quad (3.7)
\end{aligned}$$

$$\begin{aligned}
&= 1 - \frac{\sum_{k \neq k'} \{\sigma_k^2 + \sigma_{k'}^2 + \tau_k^2 + \tau_{k'}^2 - 2Cov_{jj'}(\mu_{jk}, \mu_{j'k'}) - 2E_{jj'} Cov_i(Y_{ijk}, Y_{ij'k'} | j, j', k, k')\}}{\sum_{k \neq k'} \{\sigma_k^2 + \sigma_{k'}^2 + \tau_k^2 + \tau_{k'}^2 - 2Cov_{jj'}(\mu_{jk}, \mu_{j'k'})\}} \\
&= \frac{\sum_{k=1}^{K-1} \sum_{k'=k+1}^K 2E_{jj'} Cov_i(Y_{ijk}, Y_{ij'k'} | j, j', k, k')}{\sum_{k=1}^{K-1} \sum_{k'=k+1}^K \{\sigma_k^2 + \sigma_{k'}^2 + \tau_k^2 + \tau_{k'}^2 - 2Cov_{jj'}(\mu_{jk}, \mu_{j'k'})\}}, \quad (3.8)
\end{aligned}$$

where E_I is the conditional expectation given independence of Y_{ijk} and $Y_{ij'k'}$ conditional on subject and time.

(2) Subjects' true values changing over time

However, if subjects' true values may change over time by different observers, the mean

square difference of $Y_{ijk} - Y_{ij'k'}$ should be replaced by $(Y_{ijk} - \mu_{jk}) - (Y_{ij'k'} - \mu_{j'k'})$ so that the absolute agreement is not affected by the subjects' readings from observers and time. Thus, the CCC_R^{abs} can be expressed as

CCC_{Ot}^{abs}

$$= 1 - \frac{\sum_{k=1}^{K-1} \sum_{k'=k+1}^K E_{jj'} E_i([(Y_{ijk} - \mu_{jk}) - (Y_{ij'k'} - \mu_{j'k'})]^2 | j, j', k, k')}{\sum_{k=1}^{K-1} \sum_{k'=k+1}^K E_{jj'} E_i([(Y_{ijk} - \mu_{jk}) - (Y_{ij'k'} - \mu_{j'k'})]^2 | j, j', k, k')} \quad (3.9)$$

$$= 1 - \frac{\sum_{k=1}^{K-1} \sum_{k'=k+1}^K \{\sigma_k^2 + \sigma_{k'}^2 - 2E_{jj'} Cov_i(Y_{ijk}, Y_{ij'k'} | j, j', k, k')\}}{\sum_{k=1}^{K-1} \sum_{k'=k+1}^K \{\sigma_k^2 + \sigma_{k'}^2\}} \\ = \frac{\sum_{k=1}^{K-1} \sum_{k'=k+1}^K 2E_{jj'} Cov_i(Y_{ijk}, Y_{ij'k'} | j, j', k, k')}{\sum_{k=1}^{K-1} \sum_{k'=k+1}^K (\sigma_k^2 + \sigma_{k'}^2)}, \quad (3.10)$$

where E_I is the conditional expectation given independence of Y_{ijk} and $Y_{ij'k'}$ conditional on subject and time.

3.2.2 CCC for fixed observers and random times

We define the inter-observer CCC for fixed observer and random time by using E_k with respect to the random time k , where E_k is the expectation respect to random time k . We also use the summation $\sum_{j=1}^{J-1} \sum_{j'=j+1}^J$ for the fixed observers j and j' . Specifically,

$$CCC_{oT}^{inter} = 1 - \frac{\sum_{j=1}^{J-1} \sum_{j'=j+1}^J E_k E_i([Y_{ijk} - Y_{ij'k}]^2 | j, j', k)}{\sum_{j=1}^{J-1} \sum_{j'=j+1}^J E_k E_i([Y_{ijk} - Y_{ij'k}]^2 | j, j', k)}$$

where E_I is the conditional expectation given independence of Y_{ijk} and $Y_{ij'k}$ conditional on observer and time. With the notations of $E_k Var_i(Y_{ijk} | j, k) = E_k(\sigma_{jk}^2 | j) = \sigma_j^2$ and $Var_k E_i(Y_{ijk} | j, k) = Var_k(\mu_{jk} | j) = \tau_j^2$, we can obtain

$$E_k E_i((Y_{ijk} - Y_{ij'k})^2 | j, j', k) \\ = \sigma_j^2 + \sigma_{j'}^2 + E_k(\mu_{jk} - \mu_{j'k})^2 - 2E_k Cov_i(Y_{ijk}, Y_{ij'k} | j, j', k) \\ = \sigma_j^2 + \sigma_{j'}^2 + \tau_j^2 + \tau_{j'}^2 - 2Cov_k(\mu_{jk}, \mu_{j'k}) + (E_k(\mu_{jk}) - E_k(\mu_{j'k}))^2 \\ - 2E_k Cov_i(Y_{ijk}, Y_{ij'k} | j, j', k),$$

and

$$E_k E_i([Y_{ijk} - Y_{ij'k}]^2 | j, j', k) = \sigma_j^2 + \sigma_{j'}^2 + \tau_j^2 + \tau_{j'}^2 - 2Cov_k(\mu_{jk}, \mu_{j'k}) + (E_k(\mu_{jk}) - E_k(\mu_{j'k}))^2,$$

Here we assume that $E_k(\mu_{jk}) = E_k(\mu_{j'k})$, where $E_k(\mu_{jk}) = E_k(\mu_{jk}|j)$. Then, the inter-observer agreement for CCC can be expressed as

$$CCC_{oT}^{inter} = 1 - \frac{\sum_{j=1}^{J-1} \sum_{j'=j+1}^J E_k E_i([Y_{ijk} - Y_{ij'k}]^2 | j, j', k)}{\sum_{j=1}^{J-1} \sum_{j'=j+1}^J E_k E_I([Y_{ijk} - Y_{ij'k}]^2 | j, j', k)} \quad (3.11)$$

$$= 1 - \frac{\sum_{j \neq j'} \{\sigma_j^2 + \sigma_{j'}^2 + \tau_j^2 + \tau_{j'}^2 - 2Cov_k(\mu_{jk}, \mu_{j'k}) - 2E_k Cov_i(Y_{ijk}, Y_{ij'k} | j, j', k)\}}{\sum_{j \neq j'} \{\sigma_j^2 + \sigma_{j'}^2 + \tau_j^2 + \tau_{j'}^2 - 2Cov_k(\mu_{jk}, \mu_{j'k})\}} \quad (3.12)$$

$$= \frac{\sum_{j=1}^{J-1} \sum_{j'=j+1}^J 2E_k Cov_i(Y_{ijk}, Y_{ij'k} | j, j', k)}{\sum_{j=1}^{J-1} \sum_{j'=j+1}^J \{\sigma_j^2 + \sigma_{j'}^2 + \tau_j^2 + \tau_{j'}^2 - 2Cov_k(\mu_{jk}, \mu_{j'k})\}}$$

The intra-observer CCC for fixed observers and random time is obtained similarly by considering the expected mean square difference for measurements taken on the same subject by the same observer at different time points. We consider two situations for the intra-observer agreement: (1) subjects' true values do not change over time; (2) subjects' true values change over time.

(1) Subjects' true values not changing over time

If subjects' true values do not change over time by the same observer with the assumption of $E_k(\mu_{jk}) = E_{k'}(\mu_{jk'})$, the CCC_{oT}^{intra} is defined as

$$CCC_{oT}^{intra} = 1 - \frac{\sum_{j=1}^J E_{kk'} E_i([Y_{ijk} - Y_{ijk'}]^2 | j, k, k')}{\sum_{j=1}^J E_{kk'} E_I([Y_{ijk} - Y_{ijk'}]^2 | j, k, k')} \quad (3.13)$$

$$= 1 - \frac{\sum_{j=1}^J \{2\sigma_j^2 + 2\tau_j^2 - 2Cov_{kk'}(\mu_{jk}, \mu_{jk'}) - 2E_{kk'} Cov_i(Y_{ijk}, Y_{ijk'} | j, k, k')\}}{\sum_{j=1}^J \{2\sigma_j^2 + 2\tau_j^2 - 2Cov_{kk'}(\mu_{jk}, \mu_{jk'})\}} \quad (3.14)$$

$$= \frac{\sum_{j=1}^J E_{kk'} Cov_i(Y_{ijk}, Y_{ijk'} | j, k, k')}{\sum_{j=1}^J \{\sigma_j^2 + \tau_j^2 - Cov_{kk'}(\mu_{jk}, \mu_{jk'})\}}$$

where E_I is the conditional expectation given independence of Y_{ijk} and $Y_{ijk'}$ conditional on subject and time.

(2) Subjects' true values changing over time

However, if subjects' true values change over time, the mean square difference of $Y_{ijk} - Y_{ijk'}$

should be replaced by $(Y_{ijk} - \mu_{jk}) - (Y_{ijk'} - \mu_{jk'})$ so that the intra-observer agreement is not affected by the the subjects' readings from time. In this case, the CCC_{oT}^{intra} can be expressed as

$$CCC_{oT}^{intra} = 1 - \frac{\sum_{j=1}^J E_{kk'} E_i([(Y_{ijk} - \mu_{jk}) - (Y_{ijk'} - \mu_{jk'})]^2 | j, k, k')}{\sum_{j=1}^J E_{kk'} E_I([(Y_{ijk} - \mu_{jk}) - (Y_{ijk'} - \mu_{jk'})]^2 | j, k, k')} \quad (3.15)$$

$$\begin{aligned} &= 1 - \frac{\sum_{j=1}^J \{2\sigma_j^2 - 2E_{kk'} Cov_i(Y_{ijk}, Y_{ijk'} | j, k, k')\}}{\sum_{j=1}^J 2\sigma_j^2} \\ &= \frac{\sum_{j=1}^J E_{kk'} Cov_i(Y_{ijk}, Y_{ijk'} | j, k, k')}{\sum_{j=1}^J \sigma_j^2} \end{aligned} \quad (3.16)$$

where E_I is the conditional expectation given independence of Y_{ijk} and $Y_{ijk'}$ conditional on subject and time.

The absolute CCC for fixed observer and random time is obtained similarly by considering the expected mean square difference for measurements taken on the same subject by different observers at different time points. Here, we also consider two situations for the absolute agreement: (1) subjects' true values do not change over time; (2) subjects' true values change over time.

(1) Subjects' true values not changing over time

If subjects' true values do not change over time by different observers with the assumption of $E_k(\mu_{jk}) = E_{k'}(\mu_{j'k'})$, the CCC_{oT}^{abs} is defined as

$$CCC_{oT}^{abs} = 1 - \frac{\sum_{j \neq j'} E_{kk'} E_i([Y_{ijk} - Y_{ij'k'}]^2 | j, j', k, k')}{\sum_{j \neq j'} E_{kk'} E_I([Y_{ijk} - Y_{ij'k'}]^2 | j, j', k, k')} \quad (3.17)$$

$$\begin{aligned} &= 1 - \frac{\sum_{j \neq j'} \{\sigma_j^2 + \sigma_{j'}^2 + \tau_j^2 + \tau_{j'}^2 - 2Cov_{kk'}(\mu_{jk}, \mu_{j'k'}) - 2E_{kk'} Cov_i(Y_{ijk}, Y_{ij'k'} | j, j', k, k')\}}{\sum_{j \neq j'} \{\sigma_j^2 + \sigma_{j'}^2 + \tau_j^2 + \tau_{j'}^2 - 2Cov_{kk'}(\mu_{jk}, \mu_{j'k'})\}} \\ &= \frac{\sum_{j=1}^{J-1} \sum_{j'=j+1}^J 2E_{kk'} Cov_i(Y_{ijk}, Y_{ij'k'} | j, j', k, k')}{\sum_{j=1}^{J-1} \sum_{j'=j+1}^J \{\sigma_j^2 + \sigma_{j'}^2 + \tau_j^2 + \tau_{j'}^2 - 2Cov_{kk'}(\mu_{jk}, \mu_{j'k'})\}}, \end{aligned} \quad (3.18)$$

where E_I is the conditional expectation given independence of Y_{ijk} and $Y_{ij'k'}$ conditional on subject and time.

(2) *Subjects' true values changing over time*

However, if subjects' true values may change over time by different observers, the mean square difference of $Y_{ijk} - Y_{ij'k'}$ should be replaced by $(Y_{ijk} - \mu_{jk}) - (Y_{ij'k'} - \mu_{j'k'})$ so that the absolute agreement is not affected by the the subjects' readings from observers and time. Thus, the CCC_{oT}^{abs} can be expressed as

$$CCC_{oT}^{abs} = 1 - \frac{\sum_{j \neq j'} E_{kk'} E_i([(Y_{ijk} - \mu_{jk}) - (Y_{ij'k'} - \mu_{j'k'})]^2 | j, j', k, k')}{\sum_{j \neq j'} E_{kk'} E_i([(Y_{ijk} - \mu_{jk}) - (Y_{ij'k'} - \mu_{j'k'})]^2 | j, j', k, k')} \quad (3.19)$$

$$\begin{aligned} &= 1 - \frac{\sum_{j=1}^{J-1} \sum_{j'=j+1}^J \{\sigma_j^2 + \sigma_{j'}^2 - 2E_{kk'} Cov_i(Y_{ijk}, Y_{ij'k'} | j, j', k, k')\}}{\sum_{j=1}^{J-1} \sum_{j'=j+1}^J \{\sigma_j^2 + \sigma_{j'}^2\}} \\ &= \frac{\sum_{j=1}^{J-1} \sum_{j'=j+1}^J 2E_{kk'} Cov_i(Y_{ijk}, Y_{ij'k'} | j, j', k, k')}{\sum_{j=1}^{J-1} \sum_{j'=j+1}^J \{\sigma_j^2 + \sigma_{j'}^2\}} \end{aligned} \quad (3.20)$$

where E_I is the conditional expectation given independence of Y_{ijk} and $Y_{ij'k'}$ conditional on subject and time.

3.2.3 CCC for fixed observers and fixed times

We define the inter-observer CCC for fixed observer and fixed time by using the summation $\sum_{k=1}^K$ with respect to the fixed time k . We also use the summation $\sum_{j=1}^{J-1} \sum_{j'=j+1}^J$ for the fixed observers j and j' . Specifically,

$$CCC_{ot}^{inter} = 1 - \frac{\sum_{k=1}^K \sum_{j=1}^{J-1} \sum_{j'=j+1}^J E_i((Y_{ijk} - Y_{ij'k})^2 | j, j', k)}{\sum_{k=1}^K \sum_{j=1}^{J-1} \sum_{j'=j+1}^J E_i((Y_{ijk} - Y_{ij'k})^2 | j, j', k)} \quad (3.21)$$

where E_I is the conditional expectation given independence of Y_{ijk} and $Y_{ij'k}$ conditional on observer and time. King et al. (2007) proposed CCC for fixed observers and fixed times in four different types based on the matrix of weights between times. The CCC_{inter}^{ot} is one of these types with identity matrix of weights. With the notations of $Var_i(Y_{ijk} | j, k) = \sigma_{jk}^2$ and $E_i(Y_{ijk} | j, k) = \mu_{jk}$, we can obtain

$$E_i((Y_{ijk} - Y_{ij'k})^2 | j, j', k) = \sigma_{jk}^2 + \sigma_{j'k}^2 + (\mu_{jk} - \mu_{j'k})^2 - 2Cov_i(Y_{ijk}, Y_{ij'k} | j, j', k)$$

and

$$E_i((Y_{ijk} - Y_{ij'k})^2 | j, j', k) = \sigma_{jk}^2 + \sigma_{j'k}^2 + (\mu_{jk} - \mu_{j'k})^2.$$

Then, the inter-observer agreement for CCC can be expressed as

$$\begin{aligned}
 CCC_{ot}^{inter} &= 1 - \frac{\sum_{k=1}^K \sum_{j \neq j'} \{\sigma_{jk}^2 + \sigma_{j'k}^2 + (\mu_{jk} - \mu_{j'k})^2 - 2Cov_i(Y_{ijk}, Y_{ij'k} | j, j', k)\}}{\sum_{k=1}^K \sum_{j \neq j'} \{\sigma_{jk}^2 + \sigma_{j'k}^2 + (\mu_{jk} - \mu_{j'k})^2\}} \\
 &= \frac{\sum_{k=1}^K \sum_{j=1}^{J-1} \sum_{j'=j+1}^J 2Cov_i(Y_{ijk}, Y_{ij'k} | j, j', k)}{\sum_{k=1}^K \sum_{j=1}^{J-1} \sum_{j'=j+1}^J \{\sigma_{jk}^2 + \sigma_{j'k}^2 + (\mu_{jk} - \mu_{j'k})^2\}} \quad (3.22)
 \end{aligned}$$

The intra-observer CCC for fixed observer and fixed time is obtained similarly by considering the expected mean square difference for measurements taken on the same subject by the same observer at different time points. We consider two situations for the intra-observer agreement: (1) subjects' true values do not change over time; (2) subjects' true values change over time.

(1) Subjects' true values not changing over time

If subjects' true values do not change over time by the same observer, the CCC_{ot}^{intra} is defined as

$$CCC_{ot}^{intra} = 1 - \frac{\sum_{k=1}^{K-1} \sum_{k'=k+1}^K \sum_{j=1}^J E_i([Y_{ijk} - Y_{ijk'}]^2 | j, k, k')}{\sum_{k=1}^{K-1} \sum_{k'=k+1}^K \sum_{j=1}^J E_I([Y_{ijk} - Y_{ijk'}]^2 | j, k, k')} \quad (3.23)$$

$$\begin{aligned}
 &= 1 - \frac{\sum_{k \neq k'} \sum_{j=1}^J \{\sigma_{jk}^2 + \sigma_{jk'}^2 + (\mu_{jk} - \mu_{jk'})^2 - 2Cov_i(Y_{ijk}, Y_{ijk'} | j, j', k)\}}{\sum_{k \neq k'} \sum_{j=1}^J \{\sigma_{jk}^2 + \sigma_{jk'}^2 + (\mu_{jk} - \mu_{jk'})^2\}} \\
 &= \frac{\sum_{k=1}^{K-1} \sum_{k'=k+1}^K \sum_{j=1}^J 2Cov_i(Y_{ijk}, Y_{ijk'} | j, k, k')}{\sum_{k=1}^{K-1} \sum_{k'=k+1}^K \sum_{j=1}^J \{\sigma_{jk}^2 + \sigma_{jk'}^2 + (\mu_{jk} - \mu_{jk'})^2\}}, \quad (3.24)
 \end{aligned}$$

where E_I is the conditional expectation given independence of Y_{ijk} and $Y_{ijk'}$ conditional on subject and time.

(2) Subjects' true values changing over time

However, if subjects' true values change over time, the mean square difference of $Y_{ijk} - Y_{ijk'}$ should be replaced by $(Y_{ijk} - \mu_{jk}) - (Y_{ijk'} - \mu_{jk'})$ so that the intra-observer agreement is not affected by the the subjects' readings from time. In this case, the CCC_{ot}^{intra} can be expressed as

CCC_{ot}^{intra}

$$= 1 - \frac{\sum_{k=1}^{K-1} \sum_{k'=k+1}^K \sum_{j=1}^J E_i([(Y_{ijk} - \mu_{jk}) - (Y_{ijk'} - \mu_{jk'})]^2 | j, k, k')}{\sum_{k=1}^{K-1} \sum_{k'=k+1}^K \sum_{j=1}^J E_I([(Y_{ijk} - \mu_{jk}) - (Y_{ijk'} - \mu_{jk'})]^2 | j, k, k')} \quad (3.25)$$

$$= 1 - \frac{\sum_{k=1}^{K-1} \sum_{k'=k+1}^K \sum_{j=1}^J \{\sigma_{jk}^2 + \sigma_{jk'}^2 - 2Cov_i(Y_{ijk}, Y_{ijk'} | j, k, k')\}}{\sum_{k=1}^{K-1} \sum_{k'=k+1}^K \sum_{j=1}^J \{\sigma_{jk}^2 + \sigma_{jk'}^2\}} \quad (3.26)$$

$$= \frac{\sum_{k=1}^{K-1} \sum_{k'=k+1}^K \sum_{j=1}^J 2Cov_i(Y_{ijk}, Y_{ijk'} | j, k, k')}{\sum_{k=1}^{K-1} \sum_{k'=k+1}^K \sum_{j=1}^J \{\sigma_{jk}^2 + \sigma_{jk'}^2\}}$$

where E_I is the conditional expectation given independence of Y_{ijk} and $Y_{ijk'}$ conditional on subject and time.

The absolute CCC for fixed observer and fixed time is obtained similarly by considering the expected mean square difference for measurements taken on the same subject by different observers at different time points. Here, we also consider two situations for the absolute agreement: (1) subjects' true values do not change over time; (2) subjects' true values change over time.

(1) *Subjects' true values not changing over time*

If subjects' true values do not change over time by different observers, the CCC_{ot}^{abs} is defined as

$$CCC_{ot}^{abs} = 1 - \frac{\sum_{k=1}^{K-1} \sum_{k'=K+1}^K \sum_{j=1}^{J-1} \sum_{j'=j+1}^J E_i([Y_{ijk} - Y_{ij'k'}]^2 | j, j', k, k')}{\sum_{k=1}^{K-1} \sum_{k'=K+1}^K \sum_{j=1}^{J-1} \sum_{j'=j+1}^J E_I([Y_{ijk} - Y_{ij'k'}]^2 | j, j', k, k')} \quad (3.27)$$

$$= \frac{\sum_{k=1}^{K-1} \sum_{k'=K+1}^K \sum_{j=1}^{J-1} \sum_{j'=j+1}^J 2Cov_i(Y_{ijk}, Y_{ij'k'} | j, j', k, k')}{\sum_{k=1}^{K-1} \sum_{k'=K+1}^K \sum_{j=1}^{J-1} \sum_{j'=j+1}^J \{\sigma_{jk}^2 + \sigma_{j'k'}^2 + (\mu_{jk} - \mu_{j'k'})^2\}}, \quad (3.28)$$

where E_I is the conditional expectation given independence of Y_{ijk} and $Y_{ij'k'}$ conditional on subject and time.

(2) *Subjects' true values changing over time*

However, if subjects' true values may change over time by different observers, the mean square difference of $Y_{ijk} - Y_{ij'k'}$ should be replaced by $(Y_{ijk} - \mu_{jk}) - (Y_{ij'k'} - \mu_{j'k'})$ so that the absolute agreement is not affected by the the subjects' readings from observers and

time. Thus, the CCC_{ot}^{abs} can be expressed as

$$CCC_{ot}^{abs} = 1 - \frac{\sum_{k=1}^{K-1} \sum_{k'=K+1}^K \sum_{j=1}^{J-1} \sum_{j'=j+1}^J E_i([Y_{ijk} - Y_{ij'k'}]^2 | j, j', k, k')}{\sum_{k=1}^{K-1} \sum_{k'=K+1}^K \sum_{j=1}^{J-1} \sum_{j'=j+1}^J E_I([Y_{ijk} - Y_{ij'k'}]^2 | j, j', k, k')} \quad (3.29)$$

$$\begin{aligned} &= 1 - \frac{\sum_{k \neq k'} \sum_{j \neq j'} \{\sigma_{jk}^2 + \sigma_{j'k'}^2 - 2Cov_i(Y_{ijk}, Y_{ij'k'} | j, j', k, k')\}}{\sum_{k \neq k'} \sum_{k'=K+1}^K \sum_{j \neq j'} \{\sigma_{jk}^2 + \sigma_{j'k'}^2\}} \\ &= \frac{\sum_{k=1}^{K-1} \sum_{k'=K+1}^K \sum_{j=1}^{J-1} 2Cov_i(Y_{ijk}, Y_{ij'k'} | j, j', k, k')}{\sum_{k=1}^{K-1} \sum_{k'=K+1}^K \sum_{j=1}^{J-1} \sum_{j'=j+1}^J \{\sigma_{jk}^2 + \sigma_{j'k'}^2\}} \end{aligned} \quad (3.30)$$

where E_I is the conditional expectation given independence of Y_{ijk} and $Y_{ij'k'}$ conditional on subject and time.

3.2.4 ICC for repeated measurements

For data with replications, CCC reduces to ICC if the ANOVA assumptions used to define the ICC are correct. There are no corresponding ICC s for repeated measurements under different combinations of random or fixed observer and time effects. ICC s are usually defined by ANOVA types of models. Following Vangeneugden et al. (2005), we will use the linear mixed model with a form, e.g.

$$Y_{ijk} = \mu + \mu_i^P + \mu_k^T + \mu_j^O + \mu_{ik}^{PT} + \mu_{ij}^{PO} + \mu_{kj}^{TO} + e_{ijk}^{POT},$$

to define the new ICC s. Definitions of ICC s for the four combinations are shown in Section 2.4.1 through 2.4.4. We use letter P to denote subject effects, O to denote observer effects, and T to denote time effects. Upper case is used for the random factor and lower case is used for the fixed factor.

ICC for random observers and random times

Vangeneugden et al. (2005) defined ICC for the case of random observers and random times. For completeness, we summarize their definitions for assessing agreement. A three-way random-effects model can be expressed as

$$Y_{ijk} = \mu + \mu_i^P + \mu_k^T + \mu_j^O + \mu_{ik}^{PT} + \mu_{ij}^{PO} + \mu_{kj}^{TO} + e_{ijk}^{POT},$$

$i = 1, \dots, N; j = 1, \dots, J; k = 1, \dots, k$. In case (1) with random observers and random times, μ is the overall mean, μ_i^P is the random subject effect with $N(0, \sigma_P^2)$, μ_k^T is the random

time effect with $N(0, \sigma_T^2)$, μ_j^O is the random observer effect with $N(0, \sigma_O^2)$, μ_{ik}^{PT} is the joint random effect of subject i and time k with $N(0, \sigma_{PT}^2)$, μ_{ij}^{PO} is the joint random effect of subject i and observer j with $N(0, \sigma_{PO}^2)$, μ_{kj}^{TO} is the joint random effect of time k and observer j with $N(0, \sigma_{TO}^2)$, and e_{ijk}^{POT} is the random error with $N(0, \sigma_E^2)$. The ICCs with these ANOVA assumptions for inter-observer, intra-observer, and absolute agreement are given by

$$ICC_R^{inter} = \frac{Cov(Y_{ijk}, Y_{ij'k})}{Var(Y_{ijk})} \quad (3.31)$$

$$= \frac{\sigma_P^2 + \sigma_T^2 + \sigma_{PT}^2}{\sigma_P^2 + \sigma_O^2 + \sigma_T^2 + \sigma_{PO}^2 + \sigma_{PT}^2 + \sigma_{OT}^2 + \sigma_E^2} \quad (3.32)$$

$$= Corr(Y_{ijk}, Y_{ij'k}). \quad (3.33)$$

$$ICC_R^{intra} = \frac{Cov(Y_{ijk}, Y_{ijk'})}{Var(Y_{ijk})} \quad (3.34)$$

$$= \frac{\sigma_P^2 + \sigma_O^2 + \sigma_{PO}^2}{\sigma_P^2 + \sigma_O^2 + \sigma_T^2 + \sigma_{PO}^2 + \sigma_{PT}^2 + \sigma_{OT}^2 + \sigma_E^2} \quad (3.35)$$

$$= Corr(Y_{ijk}, Y_{ijk'}). \quad (3.36)$$

$$ICC_R^{abs} = \frac{Cov(Y_{ijk}, Y_{ij'k'})}{Var(Y_{ijk})} \quad (3.37)$$

$$= \frac{\sigma_P^2}{\sigma_P^2 + \sigma_O^2 + \sigma_T^2 + \sigma_{PO}^2 + \sigma_{PT}^2 + \sigma_{OT}^2 + \sigma_E^2} \quad (3.38)$$

$$= Corr(Y_{ijk}, Y_{ij'k'}). \quad (3.39)$$

ICC_R^{inter} , ICC_R^{intra} , and ICC_R^{abs} are same as the reliability coefficient of interrater, test-retest, and ρ^2 for ICC based on the absolute value of a score defined in Vangeneugden et al. (2005).

ICC for random observers and fixed times

A three-way mixed-effects model can be expressed as

$$Y_{ijk} = \mu + \mu_i^P + \mu_k^t + \mu_j^O + \mu_{ik}^{Pt} + \mu_{ij}^{PO} + \mu_{kj}^{tO} + e_{ijk}^{POT},$$

$i = 1, \dots, N; j = 1, \dots, J; k = 1, \dots, K$. In case (2) with random observers and fixed times, μ is the overall mean, μ_i^P is the random subject effect with $N(0, \sigma_P^2)$, μ_k^t is the fixed effect of time k with $\sum_{k=1}^K \mu_k^t = 0$, μ_j^O is the random observer effect with $N(0, \sigma_O^2)$, μ_{ik}^{Pt} is the joint random effect of subject i and time k with $N(0, \sigma_{Pt}^2)$, μ_{ij}^{PO} is the joint random effect of subject i and observer j with $N(0, \sigma_{PO}^2)$, μ_{kj}^{tO} is the joint random effect of time k and observer j with $N(0, \sigma_{tO}^2)$, and e_{ijk}^{POT} is the random error with $N(0, \sigma_E^2)$. Since there are both random and fixed factors in this model, the variance due to the random factor is $\sigma_P^2 + \sigma_O^2 + \sigma_{PO}^2 + \sigma_{Pt}^2 + \sigma_{Ot}^2 + \sigma_E^2$. The variance due to the fixed factor is ϕ_t^2 . Then the total variation is $Var_{TOt} = \sigma_P^2 + \sigma_O^2 + \phi_t^2 + \sigma_{PO}^2 + \sigma_{Pt}^2 + \sigma_{Ot}^2 + \sigma_E^2$. The ICCs with these ANOVA assumptions for inter-observer, intra-observer, and absolute agreement are given by

$$ICC_{Ot}^{inter} = \frac{Cov(Y_{ijk}, Y_{ij'k})}{Var_{TOt}} \quad (3.40)$$

$$= \frac{\sigma_P^2 + \sigma_{Pt}^2}{\sigma_P^2 + \sigma_O^2 + \phi_t^2 + \sigma_{PO}^2 + \sigma_{Pt}^2 + \sigma_{Ot}^2 + \sigma_E^2} \quad (3.41)$$

$$ICC_{Ot}^{intra} = \frac{Cov(Y_{ijk}, Y_{ijk'})}{Var_{TOt}} \quad (3.42)$$

$$= \frac{\sigma_P^2 + \sigma_O^2 + \sigma_{PO}^2 - \frac{1}{K-1}\sigma_{Pt}^2 - \frac{1}{K-1}\sigma_{Ot}^2}{\sigma_P^2 + \sigma_O^2 + \phi_t^2 + \sigma_{PO}^2 + \sigma_{Pt}^2 + \sigma_{Ot}^2 + \sigma_E^2} \quad (3.43)$$

$$ICC_{Ot}^{abs} = \frac{Cov(Y_{ijk}, Y_{ij'k'})}{Var_{TOt}} \quad (3.44)$$

$$= \frac{\sigma_P^2 - \frac{1}{K-1}\sigma_{Pt}^2}{\sigma_P^2 + \sigma_O^2 + \phi_t^2 + \sigma_{PO}^2 + \sigma_{Pt}^2 + \sigma_{Ot}^2 + \sigma_E^2} \quad (3.45)$$

with the constraint of interaction effects as $\sum_{k=1}^K \mu_{ik} = 0; \sum_{k=1}^K \mu_{jk} = 0$, and the constraint of the negative correlation between these interaction effects as $Cov_i(\mu_{ik}^{PD}; \mu_{ik'}^{Pt}) = -\frac{\sigma_{Pt}^2}{K-1}; Cov(\mu_{jk}^{Ot}; \mu_{jk'}^{Ot}) = -\frac{\sigma_{Ot}^2}{K-1}$. The parameter $\phi_t^2 = \frac{\sum_{k=1}^K (\mu_k^t)^2}{K-1}$.

ICC for fixed observers and random times

A three-way random-effects model can be expressed as

$$Y_{ijk} = \mu + \mu_i^P + \mu_k^T + \mu_j^O + \mu_{ik}^{PT} + \mu_{ij}^{Po} + \mu_{kj}^{To} + e_{ijk}^{PoT},$$

$i = 1, \dots, N; j = 1, \dots, J; k = 1, \dots, k$. In case (3) with fixed observers and random times, μ is the overall mean, μ_i^P is the random subject effect with $N(0, \sigma_P^2)$, μ_k^T is the random time effect with $N(0, \sigma_T^2)$, μ_j^o is the fixed effect of observer J with $\sum_{j=1}^J \mu_j^o = 0$, μ_{ik}^{PT} is the joint random effect of subject i and time k with $N(0, \sigma_{PT}^2)$, μ_{ij}^{Po} is the joint random effect of subject i and observer j with $N(0, \sigma_{Po}^2)$, μ_{kj}^{To} is the joint random effect of time k and observer j with $N(0, \sigma_{To}^2)$, and e_{ijk}^{Pot} is the random error with $N(0, \sigma_E^2)$. Since there are both random and fixed factors in this model, the variance due to the random factor is $\sigma_P^2 + \sigma_T^2 + \sigma_{Po}^2 + \sigma_{PT}^2 + \sigma_{To}^2 + \sigma_E^2$. The variance due to the fixed factor is ϕ_o^2 . Then the total variation is $Var_{ToT} = \sigma_P^2 + \sigma_o^2 + \phi_o^2 + \sigma_{PT}^2 + \sigma_{To}^2 + \sigma_E^2$. The ICCs with these ANOVA assumptions for inter-observer, intra-observer, and absolute agreement are given by

$$ICC_{oT}^{inter} = \frac{Cov(Y_{ijk}, Y_{ij'k})}{Var_{ToT}} \quad (3.46)$$

$$= \frac{\sigma_P^2 + \sigma_T^2 + \sigma_{PT}^2 - \frac{1}{J-1}\sigma_{Po}^2 - \frac{1}{J-1}\sigma_{To}^2}{\sigma_P^2 + \phi_o^2 + \sigma_T^2 + \sigma_{Po}^2 + \sigma_{PT}^2 + \sigma_{To}^2 + \sigma_E^2} \quad (3.47)$$

$$ICC_{oT}^{intra} = \frac{Cov(Y_{ijk}, Y_{ijk'})}{Var_{ToT}} \quad (3.48)$$

$$= \frac{\sigma_P^2 + \sigma_{Po}^2}{\sigma_P^2 + \phi_o^2 + \sigma_T^2 + \sigma_{Po}^2 + \sigma_{PT}^2 + \sigma_{To}^2 + \sigma_E^2} \quad (3.49)$$

$$ICC_{oT}^{abs} = \frac{Cov(Y_{ijk}, Y_{ij'k'})}{Var_{ToT}} \quad (3.50)$$

$$= \frac{\sigma_P^2 - \frac{1}{J-1}\sigma_{Po}^2}{\sigma_P^2 + \phi_o^2 + \sigma_T^2 + \sigma_{Po}^2 + \sigma_{PT}^2 + \sigma_{To}^2 + \sigma_E^2} \quad (3.51)$$

with the constraint of interaction effects as $\sum_{j=1}^J \mu_{ij} = 0; \sum_{j=1}^J \mu_{jk} = 0$, and the constraint of the negative correlation between these interaction effects as $Cov_i(\mu_{ij}^{Po}; \mu_{ij'}^{Po}) = -\frac{\sigma_{Po}^2}{J-1}; Cov_j(\mu_{jk}^{To}; \mu_{j'k}^{To}) = -\frac{\sigma_{To}^2}{J-1}$. The parameter $\phi_o^2 = \frac{\sum_{j=1}^J (\mu_j^o)^2}{J-1}$.

ICC for fixed observers and fixed times

A three-way random-effects model can be expressed as

$$Y_{ijk} = \mu + \mu_i^P + \mu_k^t + \mu_j^o + \mu_{ik}^{Pt} + \mu_{ij}^{Po} + \mu_{kj}^{to} + e_{ijk}^{Pot},$$

$i = 1, \dots, N; j = 1, \dots, J; k = 1, \dots, K$. In case (4) with fixed observers and fixed times, μ is the overall mean, μ_i^P is the random subject effect with $N(0, \sigma_P^2)$, μ_k^t is the fixed effect of time k with $\sum_{k=1}^K \mu_k^t$, μ_j^o is the fixed effect of observer j with $\sum_{j=1}^J \mu_j^o$, μ_{ik}^{Pt} is the joint random effect of subject i and time k with $N(0, \sigma_{Pt}^2)$, μ_{ij}^{Po} is the joint random effect of subject i and observer j with $N(0, \sigma_{Po}^2)$, e_{ijk}^{Pot} is the random error with $N(0, \sigma_E^2)$. Since there are both random and fixed factors in this model, the variance due to the random factor is $\sigma_P^2 + \sigma_{Po}^2 + \sigma_{Pt}^2 + \sigma_E^2$. The variance due to the fixed factor is $\phi_o^2 + \phi_t^2 + \phi_{ot}^2$. Then the total variation is $Var_{Tot} = \sigma_P^2 + \phi_o^2 + \phi_t^2 + \sigma_{Po}^2 + \sigma_{Pt}^2 + \phi_{ot}^2 + \sigma_E^2$. The ICCs with these ANOVA assumptions for inter-observer, intra-observer, and absolute agreement are given by

$$ICC_{ot}^{inter} = \frac{Cov(Y_{ijk}, Y_{ij'k})}{Var_{Tot}} \quad (3.52)$$

$$= \frac{\sigma_P^2 + \sigma_{Pt}^2 - \frac{1}{J-1}\sigma_{Po}^2}{\sigma_P^2 + \phi_o^2 + \phi_t^2 + \sigma_{Po}^2 + \sigma_{Pt}^2 + \phi_{ot}^2 + \sigma_E^2} \quad (3.53)$$

$$ICC_{ot}^{intra} = \frac{Cov(Y_{ijk}, Y_{ijk'})}{Var_{Tot}} \quad (3.54)$$

$$= \frac{\sigma_P^2 + \sigma_{Po}^2 - \frac{1}{K-1}\sigma_{Pt}^2}{\sigma_P^2 + \phi_o^2 + \phi_t^2 + \sigma_{Po}^2 + \sigma_{Pt}^2 + \phi_{ot}^2 + \sigma_E^2} \quad (3.55)$$

$$ICC_{ot}^{abs} = \frac{Cov(Y_{ijk}, Y_{ij'k'})}{Var_{Tot}} \quad (3.56)$$

$$= \frac{\sigma_P^2 - \frac{1}{K-1}\sigma_{Pt}^2 - \frac{1}{J-1}\sigma_{Po}^2}{\sigma_P^2 + \phi_o^2 + \phi_t^2 + \sigma_{Po}^2 + \sigma_{Pt}^2 + \phi_{ot}^2 + \sigma_E^2} \quad (3.57)$$

with the constraint of interaction effects as $\sum_{j=1}^J \mu_{ij} = 0, \sum_{k=1}^K \mu_{ik} = 0$, and the constraint of the negative correlation between these interaction effects as $Cov_i(\mu_{ij}^{Po}, \mu_{ij'}^{Po}) = -\frac{1}{J-1}\sigma_{Po}^2; Cov_i(\mu_{ik}^{Pt}, \mu_{ik'}^{Pt}) = -\frac{1}{K-1}\sigma_{Pt}^2$. The parameters $\phi_o^2 = \frac{\sum_{j=1}^J (\mu_j^o)^2}{J-1}; \phi_t^2 = \frac{\sum_{k=1}^K (\mu_k^t)^2}{K-1}; \phi_{to}^2 = \frac{\sum_{k=1}^K \sum_{j=1}^J (\mu_{kj}^{to})^2}{(J-1)(K-1)}$.

3.3 Estimation and Inference

The point estimation and statistical inference of CCC for case (1) has been proposed by Chen and Barnhart (2009). Similar to their previous work, we present the estimation and inference of CCC for cases (2) through (4) in Section 3.3.1 and the estimation and inference of ICC for cases (1) through (4) in Section 3.3.2 with different ANOVA assumptions.

3.3.1 Estimation and Inference for CCC

Method of moments to estimate CCC and modified bootstrap- t method to construct the confidence interval are used for cases (2) through (4) below.

Random observers and fixed times

To estimate the CCC for random observers and fixed time points, we used the method of moments approach for each component of CCC_{Ot} for inter-observer intra-observer, and absolute agreement indices. In this section, we present the estimation and inference for CCC_{Ot}^{inter} , while CCC_{Ot}^{intra} and CCC_{Ot}^{abs} can be done in a similar fashion. Specifically, the CCC_{Ot}^{inter} is estimated by

$$\widehat{CCC}_{Ot}^{inter} = \frac{\sum_{k=1}^K E_{jj'} \widehat{Cov}_i(Y_{ijk}, Y_{ij'k} | j, j', k)}{\sum_{k=1}^K (\hat{\sigma}_k^2 + \hat{\tau}_{ck}^2)} \quad (3.58)$$

where $\tau_{ck}^2 = \tau_k^2 - Cov_{jj'}(\mu_{jk}, \mu_{j'k})$. Note here that $\sigma_k^2 = E_j Var_i(Y_{ijk} | j, k) = E_j(\sigma_{jk}^2 | k)$, which is a function of time point k . Thus, σ_{jk}^2 can be estimated by the sample variance conditional on subject j and time k with $\hat{\sigma}_{jk}^2 = \sum_{i=1}^N (Y_{ijk} - \bar{Y}_{.jk})^2 / (N - 1)$, and a natural unbiased estimator of σ_k^2 can be estimated by an average of the sample conditional variance with

$$\hat{\sigma}_k^2 = \frac{\sum_{j=1}^J \hat{\sigma}_{jk}^2}{J}.$$

The covariance conditional on subject j and time k is $Cov_i(Y_{ijk}, Y_{ij'k} | j, j', k)$, where the notation Cov_i is the conditional covariance of Y_{ijk} and $Y_{ij'k}$ respect to subjects. Then $Cov_i(Y_{ijk}, Y_{ij'k} | j, j', k)$ can be estimated by the sample conditional covariance with

$$\widehat{Cov}_i(Y_{ijk}, Y_{ij'k} | j, j', k) = \sum_{i=1}^N (Y_{ijk} - \bar{Y}_{.jk})(Y_{ij'k} - \bar{Y}_{.j'k}) / (N - 1),$$

and a natural unbiased estimator of the numerator of CCC_{Ot}^{inter} can be estimated by an average of the sample conditional covariance with

$$\widehat{E_{jj'}Cov_i}(Y_{ijk}, Y_{ij'k}|j, j', k) = \sum_{j=1}^{J-1} \sum_{j'=j'}^J \widehat{Cov_i}(Y_{ijk}, Y_{ij'k}|j, j', k) / \binom{J}{2}.$$

Since

$$\begin{aligned} & E_{jj'} E_i([\bar{Y}_{.jk} - \bar{Y}_{.j'k}]^2 | j, j', k) \\ &= E_{jj'} Var_i([\bar{Y}_{.jk} - \bar{Y}_{.j'k}] | j, j', k) + E_{jj'} (\mu_{jk} - \mu_{j'k})^2 \\ &= \frac{1}{N} E_{jj'} (\sigma_{jk}^2 + \sigma_{j'k}^2 - 2Cov_i(Y_{ijk}, Y_{ij'k} | j, j', k)) + E_{jj'} (\mu_{jk} - \mu_{j'k})^2 \\ &= \frac{2}{N} (\sigma_k^2 - E_{jj'} Cov_i(Y_{ijk}, Y_{ij'k} | j, j', k)) + 2\tau_k^2 - 2Cov_{jj'}(\mu_{jk}, \mu_{j'k}) \\ &= \frac{2}{N} (\sigma_k^2 - E_{jj'} Cov_i(Y_{ijk}, Y_{ij'k} | j, j', k)) + 2\tau_{ck}^2, \end{aligned}$$

with the assumption of $E_j(\mu_{jk}) = E_{j'}(\mu_{j'k})$, then the expression of τ_{ck}^2 form the above equation can be written as

$$\tau_{ck}^2 = \frac{1}{2} E_{jj'} E_i([\bar{Y}_{.jk} - \bar{Y}_{.j'k}]^2 | j, j', k) - \frac{1}{N} [\sigma_k^2 - E_{jj'} Cov_i(Y_{ijk}, Y_{ij'k} | j, j', k)].$$

Thus, a natural unbiased estimator of τ_{ck}^2 can be estimated by plugging in the estimates of $\hat{\sigma}_k^2$ and $\widehat{E_{jj'}Cov_i}(Y_{ijk}, Y_{ij'k}|j, j', k)$ as

$$\hat{\tau}_{ck}^2 = \frac{\sum_{j=1}^{J-1} \sum_{j'=j'}^J (\bar{Y}_{.jk} - \bar{Y}_{.j'k})^2}{2\binom{J}{2}} - \frac{1}{N} [\hat{\sigma}_k^2 - \widehat{E_{jj'}Cov_i}(Y_{ijk}, Y_{ij'k}|j, j', k)].$$

Therefore, we have $\widehat{CCC}_{Ot}^{inter}$

$$= \frac{\sum_{k=1}^K \widehat{E_{jj'}Cov_i}(Y_{ijk}, Y_{ij'k}|j, j', k)}{\sum_{k=1}^K (\hat{\sigma}_k^2 + \hat{\tau}_{ck}^2)} \quad (3.59)$$

$$= \frac{\sum_{k=1}^K \sum_{j=1}^{J-1} \sum_{j'=j'}^J \widehat{Cov_i}(Y_{ijk}, Y_{ij'k}|j, j', k) / \binom{J}{2}}{\sum_{k=1}^K \left\{ \frac{\sum_{j=1}^J \hat{\sigma}_{jk}^2}{J} + \frac{\sum_{j=1}^{J-1} \sum_{j'=j'}^J (\bar{Y}_{.jk} - \bar{Y}_{.j'k})^2}{2\binom{J}{2}} - \left[\frac{\sum_{j=1}^J \hat{\sigma}_{jk}^2}{NJ} - \frac{\sum_{j=1}^{J-1} \sum_{j'=j'}^J \widehat{Cov_i}(Y_{ijk}, Y_{ij'k}|j, j', k)}{N\binom{J}{2}} \right] \right\}} \quad (3.60)$$

In order to deal with the inference of CCC_{Ot} and easily construct confidence intervals for inter-observer, intra-observer, and absolute agreement when observers are treated as random factors and times are treated as fixed factors, We use the modified

bootstrap- t method for the confidence interval of CCC_{Ot}^{inter} presented in Chen and Barnhart (2009); the same method for constructing the confidence interval can be used for CCC_{Ot}^{intra} and CCC_{Ot}^{abs} , respectively. Let $(j^{1b}, j^{2b}, \dots, j^{Jb})$ and $(i^{1b}, i^{2b}, \dots, i^{Nb})$ be the b^{th} bootstrap sample drawn with replacement from indices of observers $(1, \dots, J)$ and subjects $(1, \dots, N)$, respectively. The B independent bootstrap data sets can be obtained by $\{Y_{i^{jb}j^{kb}}, i = 1, \dots, N; j = 1, \dots, J; k = 1, \dots, K\}$, where $b = 1, \dots, B$. The CCC_{Ot}^{inter} for each bootstrap sample is estimated by

$$\widehat{CCC}_{Ot}^{inter}(b) = \frac{\widehat{E_{jj'}Cov_i}(Y_{i^{jb}j^{kb}}, Y_{i^{jb}j'^{kb}}|j, j', k)}{\hat{\tau}_{ckb}^2 + \hat{\sigma}_{kb}^2}, \quad (3.61)$$

where $\widehat{E_{jj'}Cov_i}(Y_{i^{jb}j^{kb}}, Y_{i^{jb}j'^{kb}}|j, j', k)$, $\hat{\tau}_{ckb}^2$, and $\hat{\sigma}_{kb}^2$ are the estimators for the b^{th} bootstrap sample. The $100(1-2\alpha)\%$ bootstrap- t confidence interval of CCC_{Ot}^{inter} is defined as

$$(\widehat{CCC}_{Ot}^{inter} - \hat{t}_{Ot}^{(1-\alpha)} \cdot \hat{se}_{Ot}, \widehat{CCC}_{Ot}^{inter} - \hat{t}_{Ot}^{(\alpha)} \cdot \hat{se}_{Ot}), \quad (3.62)$$

where $\hat{t}_{Ot}^{(\alpha)}$ refers to the α th percentile of $(\widehat{CCC}_{Ot}^{inter}(b) - \widehat{CCC}_{Ot}^{inter})/\hat{se}_{Ot}(b)$, where \hat{se}_{Ot} is the estimated standard error of $\widehat{CCC}_{Ot}^{inter}$, and $\hat{se}_{Ot}(b)$ is the estimated standard error of $\widehat{CCC}_{Ot}^{inter}(b)$ for the bootstrap sample $Y_{i^{jb}j^{kb}}$. However, unlike the examples in Efron's book, we do not find a formula for the estimated standard error of $\widehat{CCC}_{Ot}^{inter}$. Thus, we approximate \hat{se}_{Ot} by $\hat{se}_{Ot}(b)$ to obtain the the $100(1-2\alpha)\%$ confidence interval of CCC_{Ot}^{inter} as

$$(\widehat{CCC}_{Ot}^{inter} - \hat{t}_{Ot}^{(1-\alpha)} \cdot \hat{se}_{Ot}(b), \widehat{CCC}_{Ot}^{inter} - \hat{t}_{Ot}^{(\alpha)} \cdot \hat{se}_{Ot}(b)). \quad (3.63)$$

Fixed observers and random times

To estimate the CCC for fixed observers and random time points, we used the method of moments approach for each component of CCC_{oT} for inter-observer intra-observer, and absolute agreement indices. In this section, we present the estimation and inference for CCC_{oT}^{inter} , while CCC_{oT}^{intra} and CCC_{oT}^{abs} can be done in a similar fashion. Specifically, the CCC_{oT}^{inter} is estimated by

$$\widehat{CCC}_{oT}^{inter} = \frac{\sum_{j=1}^{J-1} \sum_{j'=j+1}^J 2\widehat{E_kCov_i}(Y_{ijk}, Y_{ij'k}|j, j', k)}{\sum_{j=1}^{J-1} \sum_{j'=j+1}^J (\hat{\sigma}_j^2 + \hat{\sigma}_{j'}^2 + \hat{\tau}_{cj}^2 + \hat{\tau}_{cj'}^2)}$$

where $\tau_{cj}^2 = \tau_j^2 - Cov_k(\mu_{jk}, \mu_{j'k})$. Note here that $\sigma_j^2 = E_k Var_i(Y_{ijk}|j, k) = E_k(\sigma_{jk}^2|j)$, which is a function of observer j . Thus, σ_{jk}^2 can be estimated by the sample variance conditional on subject j and time k with $\hat{\sigma}_{jk}^2 = \sum_{i=1}^N (Y_{ijk} - \bar{Y}_{.jk})^2 / (N - 1)$, and a natural unbiased estimator of σ_j^2 can be estimated by an average of the sample conditional variance with

$$\hat{\sigma}_j^2 = \frac{\sum_{k=1}^K \hat{\sigma}_{jk}^2}{K}.$$

The covariance conditional on subject j and time k is $Cov_i(Y_{ijk}, Y_{ij'k}|j, j', k)$, where the notation Cov_i is the conditional covariance of Y_{ijk} and $Y_{ij'k}$ respect to subjects. Then $Cov_i(Y_{ijk}, Y_{ij'k}|j, j', k)$ can be estimated by the sample conditional covariance with

$$\widehat{Cov}_i(Y_{ijk}, Y_{ij'k}|j, j', k) = \sum_{i=1}^N (Y_{ijk} - \bar{Y}_{.jk})(Y_{ij'k} - \bar{Y}_{.j'k}) / (N - 1),$$

and a natural unbiased estimator of the numerator of CCC_{oT}^{inter} can be estimated by an average of the sample conditional covariance with

$$\widehat{E_k Cov}_i(Y_{ijk}, Y_{ij'k}|j, j', k) = \sum_{k=1}^K \widehat{Cov}_i(Y_{ijk}, Y_{ij'k}|j, j', k) / K.$$

Since

$$\begin{aligned} & E_k E_i([\bar{Y}_{.jk} - \bar{Y}_{.j'k}]^2 | j, j', k) \\ &= E_k Var_i([\bar{Y}_{.jk} - \bar{Y}_{.j'k}] | j, j', k) + E_k(\mu_{jk} - \mu_{j'k})^2 \\ &= \frac{1}{N} E_k(\sigma_{jk}^2 + \sigma_{j'k}^2 - 2Cov_i(Y_{ijk}, Y_{ij'k}|j, j', k)) + E_k(\mu_{jk} - \mu_{j'k})^2 \\ &= \frac{1}{N}(\sigma_j^2 + \sigma_{j'}^2 - 2E_k Cov_i(Y_{ijk}, Y_{ij'k}|j, j', k)) + \tau_j^2 + \tau_{j'}^2 - 2Cov_k(\mu_{jk}, \mu_{j'k}) \\ &= \frac{1}{N}(\sigma_j^2 + \sigma_{j'}^2 - 2E_k Cov_i(Y_{ijk}, Y_{ij'k}|j, j', k)) + \tau_{cj}^2 + \tau_{cj'}^2, \end{aligned}$$

with the assumption of $E_k(\mu_{jk}) = E_k(\mu_{j'k})$, then the expression of $\tau_{cj}^2 + \tau_{cj'}^2$ form the above equation can be written as

$$\tau_{cj}^2 + \tau_{cj'}^2 = E_k E_i([\bar{Y}_{.jk} - \bar{Y}_{.j'k}]^2 | j, j', k) - \frac{1}{N}[\sigma_j^2 + \sigma_{j'}^2 - 2E_k Cov_i(Y_{ijk}, Y_{ij'k}|j, j', k)].$$

Thus, a natural unbiased estimator of $\tau_{cj}^2 + \tau_{cj'}^2$ can be estimated by plugging in the estimates of $\hat{\sigma}_j^2$, $\hat{\sigma}_{j'}^2$ and $\widehat{E_k Cov}_i(Y_{ijk}, Y_{ij'k}|j, j', k)$ as

$$\hat{\tau}_{cj}^2 + \hat{\tau}_{cj'}^2 = \frac{\sum_{k=1}^K (\bar{Y}_{.jk} - \bar{Y}_{.j'k})^2}{K} - \frac{1}{N}[\hat{\sigma}_j^2 + \hat{\sigma}_{j'}^2 - 2\widehat{E_k Cov}_i(Y_{ijk}, Y_{ij'k}|j, j', k)].$$

Therefore, we have $\widehat{CCC}_{oT}^{inter}$

$$= \frac{\sum_{j=1}^{J-1} \sum_{j'=j+1}^J 2\widehat{E_k Cov}_i(Y_{ijk}, Y_{ij'k}|j, j', k)}{\sum_{j=1}^{J-1} \sum_{j'=j+1}^J (\hat{\sigma}_j^2 + \hat{\sigma}_{j'}^2 + \hat{\tau}_{cj}^2 + \hat{\tau}_{cj'}^2)} \quad (3.64)$$

$$= \frac{\sum_{j \neq j'} 2 \sum_{k=1}^K \widehat{Cov}_i(Y_{ijk}, Y_{ij'k}|j, j', k)/K}{\sum_{j \neq j'} \left\{ \frac{\sum_{k=1}^K (\hat{\sigma}_{jk}^2 + \hat{\sigma}_{j'k}^2)}{K} + \frac{\sum_{k=1}^K (\bar{Y}_{.jk} - \bar{Y}_{.j'k})^2}{K} - \left[\frac{\sum_{k=1}^K (\hat{\sigma}_{jk}^2 + \hat{\sigma}_{j'k}^2)}{NK} - \frac{\sum_{k=1}^K 2\widehat{Cov}_i(Y_{ijk}, Y_{ij'k}|j, j', k)}{NK} \right] \right\}} \quad (3.65)$$

In order to deal with the inference of CCC_{oT} and easily construct confidence intervals for inter-observer, intra-observer, and absolute agreement, when observers are treated fixed factors and times are treated as random factors, we consider the bootstrap method for the three agreement indices. Let $(k^{1b}, k^{2b}, \dots, k^{Kb})$ and $(i^{1b}, i^{2b}, \dots, i^{Nb})$ be the b^{th} bootstrap sample drawn with replacement from indices of times $(1, \dots, K)$ and subjects $(1, \dots, N)$, respectively. The B independent bootstrap data sets can be obtained by $\{Y_{i^{ib}j^{kb}k^{kb}}, i = 1, \dots, N; j = 1, \dots, J; k = 1, \dots, K\}$, where $b = 1, \dots, B$. The CCC_{oT}^{inter} for each bootstrap sample is estimated by

$$\widehat{CCC}_{oT}^{inter}(b) = \frac{\sum_{j=1}^{J-1} \sum_{j'=j+1}^J \widehat{E_k Cov}_i(Y_{i^{ib}j^{kb}k^{kb}}, Y_{i^{ib}j'^{kb}k^{kb}}|j, j', k)}{\sum_{j=1}^{J-1} \sum_{j'=j+1}^J (\hat{\tau}_{cjb}^2 + \hat{\tau}_{cj'b}^2 + \hat{\sigma}_{jb}^2 + \hat{\sigma}_{j'b}^2)}, \quad (3.66)$$

where $\widehat{E_k Cov}_i(Y_{i^{ib}j^{kb}k^{kb}}, Y_{i^{ib}j'^{kb}k^{kb}}|j, j', k)$, $\hat{\tau}_{cjb}^2$, $\hat{\tau}_{cj'b}^2$, $\hat{\sigma}_{jb}^2$, and $\hat{\sigma}_{j'b}^2$ are the estimators for the b^{th} bootstrap sample. The $100(1-2\alpha)\%$ bootstrap- t confidence interval of CCC_{oT}^{inter} is defined as

$$(\widehat{CCC}_{oT}^{inter} - \hat{t}_{oT}^{(1-\alpha)} \cdot \hat{se}_{oT}, \widehat{CCC}_{oT}^{inter} - \hat{t}_{oT}^{(\alpha)} \cdot \hat{se}_{oT}), \quad (3.67)$$

where $\hat{t}_{oT}^{(\alpha)}$ refers to the α th percentile of $(\widehat{CCC}_{oT}^{inter}(b) - \widehat{CCC}_{oT}^{inter})/\hat{se}_{oT}(b)$, where \hat{se}_{oT} is the estimated standard error of $\widehat{CCC}_{oT}^{inter}$, and $\hat{se}_{oT}(b)$ is the estimated standard error of $\widehat{CCC}_{oT}^{inter}(b)$ for the bootstrap sample $Y_{i^{ib}j^{kb}k^{kb}}$. However, unlike the examples in Efron's book, we do not find a formula for the estimated standard error of $\widehat{CCC}_{oT}^{inter}$. Thus, we approximate \hat{se}_{oT} by $\hat{se}_{oT}(b)$ to obtain the the $100(1-2\alpha)\%$ confidence interval of CCC_{oT}^{inter} as

$$(\widehat{CCC}_{oT}^{inter} - \hat{t}_{oT}^{(1-\alpha)} \cdot \hat{se}_{oT}(b), \widehat{CCC}_{oT}^{inter} - \hat{t}_{oT}^{(\alpha)} \cdot \hat{se}_{oT}(b)). \quad (3.68)$$

Fixed observers and fixed times

To estimate the CCC for fixed observers and fixed time points, we used the method of moments approach for each component of CCC_{ot} for inter-observer intra-observer, and absolute agreement indices. In this section, we present the estimation and inference for CCC_{ot}^{inter} , while CCC_{ot}^{intra} and CCC_{ot}^{abs} can be done in a similar fashion. Specifically, the CCC_{ot}^{inter} is estimated by

$$\widehat{CCC}_{ot}^{inter} = \frac{\sum_{k=1}^K \sum_{j=1}^{J-1} \sum_{j'=j+1}^J 2\hat{Cov}_i(Y_{ijk}, Y_{ij'k}|j, j', k)}{\sum_{k=1}^K \sum_{j=1}^{J-1} \sum_{j'=j+1}^J (\hat{\sigma}_{jk}^2 + \hat{\sigma}_{j'k}^2 + (\hat{\mu}_{jk}^2 - \hat{\mu}_{j'k}^2)^2)}. \quad (3.69)$$

Note here that $\sigma_{jk}^2 = Var_i(Y_{ijk}|j, k)$ and $\mu_{jk} = E_i(Y_{ijk}|j, k)$, which are functions of observer j and time k . Thus, σ_{jk}^2 can be estimated by the sample variance conditional on observer j and time k with $\hat{\sigma}_{jk}^2 = \sum_{i=1}^N (Y_{ijk} - \bar{Y}_{.jk})^2 / (N - 1)$. The unbiased estimate of μ_{jk} is $\sum_{i=1}^N Y_{ijk} / N$. The covariance conditional on subject j and time k is $Cov_i(Y_{ijk}, Y_{ij'k}|j, j', k)$, where the notation Cov_i is the conditional covariance of Y_{ijk} and $Y_{ij'k}$ respect to subjects. Then $Cov_i(Y_{ijk}, Y_{ij'k}|j, j', k)$ can be estimated by the sample conditional covariance with

$$\widehat{Cov}_i(Y_{ijk}, Y_{ij'k}|j, j', k) = \sum_{i=1}^N (Y_{ijk} - \bar{Y}_{.jk})(Y_{ij'k} - \bar{Y}_{.j'k}) / (N - 1).$$

Therefore, we have

$$\widehat{CCC}_{ot}^{inter} = \frac{\sum_{k=1}^K \sum_{j=1}^{J-1} \sum_{j'=j+1}^J 2\hat{Cov}_i(Y_{ijk}, Y_{ij'k}|j, j', k)}{\sum_{k=1}^K \sum_{j=1}^{J-1} \sum_{j'=j+1}^J (\hat{\sigma}_{jk}^2 + \hat{\sigma}_{j'k}^2 + (\hat{\mu}_{jk} - \hat{\mu}_{j'k})^2)} \quad (3.70)$$

$$= \frac{\sum_{k=1}^K \sum_{j=1}^{J-1} \sum_{j'=j+1}^J 2\hat{Cov}_i(Y_{ijk}, Y_{ij'k}|j, j', k)}{\sum_{k=1}^K \sum_{j=1}^{J-1} \sum_{j'=j+1}^J \hat{\sigma}_{jk}^2 + \hat{\sigma}_{j'k}^2 + (\bar{Y}_{.jk} - \bar{Y}_{.j'k})^2} \quad (3.71)$$

In order to deal with the inference of CCC_{ot} and easily construct confidence intervals for inter-observer, intra-observer, and absolute agreement, when both of observers and times are treated fixed factors, we consider the bootstrap method for the three agreement indices. Let $(i^{1b}, i^{2b}, \dots, i^{Nb})$ be the b^{th} bootstrap sample drawn with replacement from indices of subjects $(1, \dots, N)$. The B independent bootstrap data sets can be obtained by

$\{Y_{ibjk}, i = 1, \dots, N; j = 1, \dots, J; k = 1, \dots, K\}$, where $b = 1, \dots, B$. The CCC_{ot}^{inter} for each bootstrap sample is estimated by

$$\widehat{CCC}_{ot}^{inter}(b) = \frac{\sum_{k=1}^K \sum_{j=1}^{J-1} \sum_{j'=j+1}^J 2\hat{Cov}_i(Y_{ibjk}, Y_{ibj'k} | j, j', k)}{\sum_{k=1}^K \sum_{j=1}^{J-1} \sum_{j'=j+1}^J (\hat{\sigma}_{jkb}^2 + \hat{\sigma}_{j'kb}^2 + (\hat{\mu}_{jk} - \hat{\mu}_{j'k})^2)}, \quad (3.72)$$

where $\hat{Cov}_i(Y_{ibjk}, Y_{ibj'k} | j, j', k)$, $\hat{\sigma}_{jkb}^2$, $\hat{\sigma}_{j'kb}^2$, $\hat{\mu}_{jk}$, and $\hat{\mu}_{j'k}$ are the estimators for the b^{th} bootstrap sample. The $100(1-2\alpha)\%$ bootstrap- t confidence interval of CCC_{ot}^{inter} is defined as

$$(\widehat{CCC}_{ot}^{inter} - \hat{t}_{ot}^{(1-\alpha)} \cdot \hat{se}_{ot}, \widehat{CCC}_{ot}^{inter} - \hat{t}_{ot}^{(\alpha)} \cdot \hat{se}_{ot}), \quad (3.73)$$

where $\hat{t}_{ot}^{(\alpha)}$ refers to the α th percentile of $(\widehat{CCC}_{ot}^{inter}(b) - \widehat{CCC}_{ot}^{inter})/\hat{se}_{ot}(b)$, where \hat{se}_{ot} is the estimated standard error of $\widehat{CCC}_{ot}^{inter}$, and $\hat{se}_{ot}(b)$ is the estimated standard error of $\widehat{CCC}_{ot}^{inter}(b)$ for the bootstrap sample Y_{ibjk} . However, unlike the examples in Efron's book, we do not find a formula for the estimated standard error of $\widehat{CCC}_{ot}^{inter}$. Thus, we approximate \hat{se}_{ot} by $\hat{se}_{ot}(b)$ to obtain the the $100(1-2\alpha)\%$ confidence interval of CCC_{ot}^{inter} as

$$(\widehat{CCC}_{ot}^{inter} - \hat{t}_{ot}^{(1-\alpha)} \cdot \hat{se}_{ot}(b), \widehat{CCC}_{ot}^{inter} - \hat{t}_{ot}^{(\alpha)} \cdot \hat{se}_{ot}(b)). \quad (3.74)$$

3.3.2 Estimation and Inference for ICC

The definitions of ICCs for four cases are shown in Table 3.3. Fitting the ANOVA model usually produce an ANOVA table with means of the sums of squares (MS). The expected MS factor from different ANOVA tables are shown in Tables 3.4-3.7. The formulas for calculating ICC estimators for four cases based on the corresponding ANOVA tables for inter-observer, intra-observer, and absolute agreement are as follows:

The ICC for inter-observer agreement is estimated by

$$\begin{aligned} \widehat{ICC}^{inter} &= \frac{\frac{1}{JK}(MS_P - MS_{PO} - MS_{PD} + MS_E) + \frac{1}{NJ}(MS_T - MS_{PT} - MS_{OT} + MS_E) + \frac{1}{J}(MS_{PT} - MS_E)}{\frac{MS_P}{JK} + \frac{MS_O}{NK} + \frac{MS_T}{NJ} + w_1 \frac{MS_{PO}}{K} + w_2 \frac{MS_{PT}}{J} + w_3 \frac{MS_{OT}}{N} + w_4 MS_E} \end{aligned} \quad (3.75)$$

The ICC for intra-observer agreement is estimated by

Table 3.4: Mean square expectations for three-way ANOVA models based on $Y_{ijk} = \mu + \mu_i^P + \mu_k^T + \mu_j^O + \mu_{ik}^{PT} + \mu_{ij}^{PO} + \mu_{kj}^{TO} + e_{ijk}^{POT}$, where P , O , and T are random

Source	df	MS	EMS
Patient (P)	N-1	MS_P	$JK\sigma_P^2 + K\sigma_{PO}^2 + J\sigma_{PD}^2 + \sigma_e^2$
Observer (O)	J-1	MS_O	$NK\sigma_O^2 + K\sigma_{PO}^2 + N\sigma_{OT}^2 + \sigma_e^2$
Time (T)	K-1	MS_T	$NJ\sigma_T^2 + J\sigma_{PT}^2 + N\sigma_{OT}^2 + \sigma_e^2$
P \times O	(N-1)(J-1)	MS_{PO}	$K\sigma_{PO}^2 + \sigma_E^2$
P \times T	(N-1)(K-1)	MS_{PT}	$J\sigma_{PT}^2 + \sigma_E^2$
O \times T	(J-1)(K-1)	MS_{OT}	$N\sigma_{OT}^2 + \sigma_E^2$
Error	(N-1)(J-1)(K-1)	MS_E	σ_E^2

$$\begin{aligned}
& \widehat{ICC}^{inra} \\
&= \frac{\frac{1}{JK}(MS_P - MS_{PO} - MS_{PD} + MS_E) + \frac{1}{NK}(MS_O - MS_{PO} - MS_{OT} + MS_E) + \frac{1}{K}(MS_{PO} - MS_E)}{\frac{MS_P}{JK} + \frac{MS_O}{NK} + \frac{MS_T}{NJ} + w_1 \frac{MS_{PO}}{K} + w_2 \frac{MS_{PT}}{J} + w_3 \frac{MS_{OT}}{N} + w_4 MS_E}
\end{aligned} \tag{3.76}$$

The ICC for absolute agreement is estimated by

$$\widehat{ICC}^{abs} = \frac{\frac{1}{JK}(MS_P - MS_{PO} - MS_{PD} + MS_E)}{\frac{MS_P}{JK} + \frac{MS_O}{NK} + \frac{MS_T}{NJ} + w_1 \frac{MS_{PO}}{K} + w_2 \frac{MS_{PT}}{J} + w_3 \frac{MS_{OT}}{N} + w_4 MS_E} \tag{3.77}$$

where $w_1 = (1 - \frac{1}{J} - \frac{1}{N})$, $w_2 = (1 - \frac{1}{K} - \frac{1}{N})$, $w_3 = (1 - \frac{1}{J} - \frac{1}{K})$, and $w_4 = 1 + \frac{1}{NK} + \frac{1}{NJ} + \frac{1}{JK} - \frac{1}{N} - \frac{1}{J} - \frac{1}{K}$. Note here that MS_P = mean square for patients; MS_O = mean square for observers; MS_T = mean square for times; MS_{PO} = mean square for patients \times observers; MS_{PT} = mean square for patients \times times; MS_{OT} = mean square for observers \times times; MS_E = mean square error. Upper case for letters in MS are used for random factors. We use the method of moments to estimate the individual component used in the ICC definition. It turns out the expressions of ICC estimator under four cases are the same for inter-observer, intra-observer, and absolute agreement, respectively. These means that the point estimators are the same for four cases regardless whether observer and time are treated as random or fixed. However, the inference will be different. We construct the confidence intervals for these ICC s similarly by the modified bootstrap- t method.

Table 3.5: Mean square expectations for three-way ANOVA models based on $Y_{ijk} = \mu + \mu_i^P + \mu_k^t + \mu_j^O + \mu_{ik}^{Pt} + \mu_{ij}^{PO} + \mu_{kj}^{tO} + e_{ijk}^{POT}$, where P, O are random and t is fixed

Source	df	MS	EMS
Patient (P)	N-1	MS_P	$JK\sigma_P^2 + K\sigma_{PO}^2 + \sigma_e^2$
Observer (O)	J-1	MS_O	$NK\sigma_O^2 + K\sigma_{PO}^2 + \sigma_e^2$
Time (t)	K-1	MS_t	$NJ\phi_t^2 + J\frac{K}{K-1}\sigma_{Pt}^2 + N\frac{K}{K-1}\sigma_{Ot}^2 + \sigma_e^2$
P \times O	(N-1)(J-1)	MS_{PO}	$K\sigma_{PO}^2 + \sigma_e^2$
P \times t	(N-1)(K-1)	MS_{Pt}	$J\frac{K}{K-1}\sigma_{Pt}^2 + \sigma_e^2$
O \times t	(J-1)(K-1)	MS_{Ot}	$N\frac{K}{K-1}\sigma_{Ot}^2 + \sigma_e^2$
Error	(N-1)(J-1)(K-1)	MS_E	σ_e^2

Table 3.6: Mean square expectations for three-way ANOVA models based on $Y_{ijk} = \mu + \mu_i^P + \mu_k^T + \mu_j^o + \mu_{ik}^{PT} + \mu_{ij}^{Po} + \mu_{kj}^{To} + e_{ijk}^{PoT}$, where P, T are random and o is fixed

Source	df	MS	EMS
Patient (P)	N-1	MS_P	$JK\sigma_P^2 + J\sigma_{PD}^2 + \sigma_e^2$
Observer (o)	J-1	MS_o	$NK\phi_o^2 + K\frac{J}{J-1}\sigma_{Po}^2 + N\frac{J}{J-1}\sigma_{oT}^2 + \sigma_e^2$
Time (T)	K-1	MS_T	$NJ\sigma_T^2 + J\sigma_{PT}^2 + \sigma_e^2$
P \times o	(N-1)(J-1)	MS_{Po}	$K\frac{J}{J-1}\sigma_{Po}^2 + \sigma_e^2$
P \times T	(N-1)(K-1)	MS_{PT}	$J\sigma_{PT}^2 + \sigma_e^2$
o \times T	(J-1)(K-1)	MS_{oT}	$N\frac{J}{J-1}\sigma_{oT}^2 + \sigma_e^2$
Error	(N-1)(J-1)(K-1)	MS_E	σ_e^2

Table 3.7: Mean square expectations for three-way ANOVA models based on $Y_{ijk} = \mu + \mu_i^P + \mu_k^t + \mu_j^o + \mu_{ik}^{Pt} + \mu_{ij}^{Po} + \mu_{kj}^{to} + e_{ijk}^{Pot}$, where P is random, and o, t are fixed

Source	df	MS	EMS
Patient (P)	N-1	MS_P	$JK\sigma_P^2 + \sigma_e^2$
Observer (o)	J-1	MS_o	$NK\phi_o^2 + K\frac{J}{J-1}\sigma_{Po}^2 + N\frac{J}{J-1}\frac{K}{K-1}\sigma_{ot}^2 + \sigma_e^2$
Time (t)	K-1	MS_t	$NJ\phi_t^2 + J\frac{K}{K-1}\sigma_{Pt}^2 + N\frac{J}{J-1}\frac{K}{K-1}\sigma_{ot}^2 + \sigma_e^2$
P \times o	(N-1)(J-1)	MS_{Po}	$K\frac{J}{J-1}\sigma_{Po}^2 + \sigma_e^2$
P \times t	(N-1)(K-1)	MS_{Pt}	$J\frac{K}{K-1}\sigma_{Pt}^2 + \sigma_e^2$
o \times t	(J-1)(K-1)	MS_{ot}	$N\frac{J}{J-1}\frac{K}{K-1}\sigma_{ot}^2 + \sigma_e^2$
Error	(N-1)(J-1)(K-1)	MS_E	σ_e^2

3.4 Data Analysis

To demonstrate the use of the *CCC* and *ICC* indices for repeated measurements with either random or fixed observers and times, we use the image data discussed in Chen and Barnhart (2009) for illustration and comparison. The purpose of the image study is to evaluate the pulmonary arterial hypertension measures by 2D-echocardiogram. To assess the agreement between sonographers who measure the 2D-echocardiogram images, two sonographers make measurement twice on 10 patients. The variables of interests for assessing agreement are Visual Effect Fraction (VISEF), Left Ventricular Ejection Fraction (LVEF), and Right Atrium Volume (RAV). There are 10, 9, and 8 subjects with complete data for variables VISEF, LVEF, and RAV respectively. The point estimates and the corresponding 95% confidence interval based on the modified bootstrap-*t* method for *CCC* and *ICC* are present in Tables 3.8-3.9 for four cases. As indicated in the estimation and inference, the point estimates of *CCC*s are the same for all cases for each of the inter-observer, intra-observer, and absolute agreement. The point estimates of *ICC*s are also the same for all four cases as well. However, the confidence intervals are different with respect to whether the factor is treated as random or fixed. The point estimate of *ICC* is slightly larger than *CCC*. The confidence interval of *CCC* is slightly wider than *ICC* probably due to the less assumptions in *CCC*. The confidence intervals for both *ICC* and *CCC* in case (1) are slightly wider than case (4) probably due to the random effects of observer and time.

Table 3.8: Results of the estimates and improved 95% bootstrap- t confidence intervals of CCC under four cases for image data

Variable	CCC		
	CCC^{abs} (95% C.I.)	CCC^{inter} (95% C.I.)	CCC^{intra} (95% C.I.)
VISEF			
case ^a			
(1)	0.611 (0.381, 1.000)	0.624 (0.412, 1.000)	0.920 (0.861, 1.000)
(2)	0.611 (0.385, 1.000)	0.624 (0.409, 1.000)	0.920 (0.876, 1.000)
(3)	0.611 (0.380, 1.000)	0.624 (0.405, 1.000)	0.920 (0.885, 1.000)
(4)	0.612 (0.400, 1.000)	0.624 (0.418, 1.000)	0.941 (0.886, 1.000)
LVEF			
case ^a			
(1)	0.679 (0.452, 1.000)	0.721 (0.508, 1.000)	0.904 (0.800, 1.000)
(2)	0.679 (0.474, 1.000)	0.721 (0.567, 1.000)	0.904 (0.801, 1.000)
(3)	0.679 (0.473, 1.000)	0.721 (0.531, 1.000)	0.904 (0.802, 1.000)
(4)	0.679 (0.487, 1.000)	0.721 (0.564, 1.000)	0.906 (0.807, 1.000)
RAV			
case ^a			
(1)	0.583 (0.381, 1.000)	0.601 (0.401, 1.000)	0.900 (0.774, 1.000)
(2)	0.583 (0.490, 1.000)	0.601 (0.500, 0.836)	0.900 (0.800, 1.000)
(3)	0.583 (0.480, 0.970)	0.601 (0.510, 1.000)	0.900 (0.804, 1.000)
(4)	0.585 (0.491, 0.870)	0.601 (0.521, 0.934)	0.902 (0.806, 1.000)

^a (1) random observer and random time; (2) random observer and fixed time;

(3) fixed observer and random time; (4) fixed observer and fixed time

Table 3.9: Results of the estimates and improved 95% bootstrap- t confidence intervals of ICC under four cases for image data

Variable	ICC		
	ICC^{abs} (95% C.I.)	ICC^{inter} (95% C.I.)	ICC^{intra} (95% C.I.)
VISEF			
case ^a			
(1)	0.623 (0.384, 1.000)	0.643 (0.438, 1.000)	0.933 (0.865, 1.000)
(2)	0.623 (0.387, 1.000)	0.643 (0.439, 1.000)	0.933 (0.888, 1.000)
(3)	0.623 (0.389, 1.000)	0.643 (0.409, 1.000)	0.933 (0.889, 1.000)
(4)	0.623 (0.401, 1.000)	0.643 (0.442, 1.000)	0.933 (0.889, 1.000)
LVEF			
case ^a			
(1)	0.698 (0.496, 1.000)	0.729 (0.538, 1.000)	0.914 (0.828, 1.000)
(2)	0.698 (0.519, 1.000)	0.729 (0.577, 1.000)	0.914 (0.843, 1.000)
(3)	0.698 (0.536, 1.000)	0.729 (0.553, 1.000)	0.914 (0.845, 1.000)
(4)	0.698 (0.545, 1.000)	0.729 (0.578, 1.000)	0.914 (0.845, 1.000)
RAV			
case ^a			
(1)	0.598 (0.387, 1.000)	0.613 (0.401, 1.000)	0.900 (0.801, 1.000)
(2)	0.598 (0.492, 0.960)	0.613 (0.503, 1.000)	0.900 (0.848, 1.000)
(3)	0.598 (0.486, 1.000)	0.613 (0.521, 1.000)	0.900 (0.852, 1.000)
(4)	0.598 (0.494, 1.000)	0.613 (0.526, 1.000)	0.900 (0.855, 1.000)

^a (1) random observer and random time; (2) random observer and fixed time;

(3) fixed observer and random time; (4) fixed observer and fixed time

3.5 Discussion

In this paper, we have proposed new indices for *CCC* for assessing inter-observer, intra-observer, and absolute agreement under all four combinations for random or fixed effects of observer and time factors for data with repeated measurements. The point estimates of these *CCCs* regarding random or fixed effects are obtained by using the method of moments approach for each components of the *CCC* index. The statistical inference of this *CCC* is constructed by the modified bootstrap-*t* confidence interval. Since *CCC* can be reduced to *ICC* for data with replications if the correct ANOVA model holds, we considered a three-way ANOVA model with first and second order interaction effects to obtain the new *ICC* index. This index has different form for assessing inter-observer, intra-observer, and absolute agreement under different cases. The point estimate of this new *ICC* is obtained by estimating the variance components from the specific ANOVA models. The statistical inference of this *ICC* is constructed similarly by the modified bootstrap-*t* confidence interval. Note here that the new *ICC* proposed in this paper is considered for agreement measurements rather than consistency measurements. The consistency measurement is a kind of conditional correlation coefficient between two measurements when a factor is given.

The *ICC* usually needs ANOVA assumptions while *CCC* does not need. In this paper, the new *ICC* is based on a three-way ANOVA model with all interaction effects. In practice, we do not know whether a data set is from the three-way ANOVA model discussed in this paper, or from another three-way ANOVA model with different combinations of interaction effects. Thus, if there is need to choose between the new *ICC* and new *CCC*, we recommend to use the new *CCC* since it does not need ANOVA assumptions.

Bibliography

Barnhart, H. X. and Williamson, J. M. (2001). Modeling concordance correlation via GEE to evaluate reproducibility. *Biometrics* 57, 931-940.

Barnhart, H. X., Haber, M., and Song, J. (2002). Overall concordance correlation coefficient for evaluating agreement among multiple observers. *Biometrics* 58, 1020-1027.

Barnhart, H. X., Song, J., and Haber, M. (2005). Assessing intra, inter and total agreement with replicated readings. *Statistics in Medicine* 24, 1371-1384.

Barnhart, H. X., Haber, M. and Lin, L. I. (2007). An overview on assessing agreement with continuous measurements. *Journal of Biopharmaceutical Statistics* 17, 529-569.

Bartko, J. J. (1966). The intraclass correlation coefficient as a measure of reliability. *Psychological Reports* 19, 3-11.

Bartko, J. J. (1974). A note on the intraclass correlation coefficient as a measure of reliability. *Psychological Reports* 34, 418.

Bland, J. M. and Altman, D. G. (1999). Measuring agreement in method comparison studies. *Statistics Methods in Medical Research* 8, 135-160.

Chen, C.-C. and Barnhart, H.X. (2008). Comparison of ICC and CCC for assessing agreement for data without and with replications. *Computational Statistics and Data Analysis* 53(2), 554-564.

Carrasco, J. L. and Jover, L. (2003). Estimating the generalized concordance correlation coefficient through variance components. *Biometrics* 59, 849-858.

Chinchilli, V. M., Martel, J. K., Kumanyika, S., and Lloyd, T. (1996). A weighted concordance correlation coefficient for repeated measurement designs. *Biometrics* 52, 341-353.

Efron and Tibshirani (1993). *An introduction to the bootstrap*. New York : Chapman and Hall.

Eliasziw, M., Young, S. L., Woodbury, M. G., and Fryday-Field, K. (1994). Statistical methodology for the concurrent assessment of interrater and intrarater reliability: Using goniometric measurements as an example. *Physical Therapy* 74, 777-788.

Fay, M. P. (2005). Random marginal agreement coefficients: rethinking the adjustment for chance when measuring agreement. *Biometrics* 6, 171-180.

Fisher, R. A. (1925). *Statistical methods for research workers*. Edinburgh: Oliver and Boyd.

King, T. S. and Chinchilli, V. M. (2001). A generalized concordance correlation coefficient for continuous and categorical data. *Statistics in Medicine* 20, 2131-2147.

King, T. S., Chinchilli, V. M., Carrasco, J. (2007). A repeated measures concordance correlation coefficient. *Statistics in Medicine*, 26 3095-3113.

Lin, L. (1989). A concordance correlation coefficient to evaluate reproducibility. *Biometrics* 45, 255-268.

Lin, L., Hedayat, A. S., Sinha, B., and Yang, M. (2002). Statistical methods in assessing agreement: Models, issues and tools. *Journal of American Statistical Association* 97 (457), 257-270.

Lin, L., Hedayat, A. S., and Wenting, W. (2007). A unified approach for assessing agreement for continuous and categorical data. *Journal of Biopharmaceutical Statistics* 17, 629-652.

McGraw, K. O. and Wong, S. P. (1996). Forming inferences about some intraclass correlation coefficients. *Psychological Methods* 1, 30-46.

Müller, R. and Büttner, P. (1994). A critical discussion of intraclass correlation coefficients. *Statistics in Medicine* 13, 2465-2476.

Nickerson, C. A. (1997). Comment on "A Concordance Correlation Coefficient to Evaluate Reproducibility." *Biometrics* 53, 1503-1507.

Quiroz, J. (2005). Assessment of equivalence using a concordance correlation

coefficient in a repeated measurement design. *Journal of Biopharmaceutical Statistics*, 15, 913-928.

Shrout, P. E. and Fleiss, J. L. (1979). Intraclass correlations: Uses in assessing rater reliability. *Psychological Bulletin* 86, 420-428.

Vangeneugden, T., Laenen, A., Geys, H., Renard, D., and Molenberghs, G. (2005). Applying concepts of generalizability theory on clinical trial data to investigate sources of variation and their impact on reliability. *Biometrics* 61, 295-304.

Appendices

APPENDIX A

In Appendix A, we derive $E(\widehat{ICC})$ for data with replications mentioned in Section 2.2.2. The expected value of \widehat{ICC}_1 is

$$\begin{aligned} E(\widehat{ICC}_1) &\approx \frac{E(MS_\alpha - MS_e)}{E(MS_\alpha + (JK - 1)MS_e)} \\ &= \frac{\frac{J(K-1)}{JK-1} \sum_{j=1}^J \sigma_{Bj}^2 + (\frac{JK}{JK-1}) \sum_{j=1}^J \sum_{j'=1(j' \neq j)}^J \sigma_{Bj} \sigma_{Bj'} \rho_{\mu jj'} - \frac{1}{JK-1} (J \sum_{j=1}^J \mu_j^2 - (\sum_{j=1}^J \mu_j)^2)}{J(\sum_{j=1}^J \sigma_{Bj}^2 + \sum_{j=1}^J \sigma_{Wj}^2 + \sum_{j=1}^J \mu_j^2 - \frac{1}{J} (\sum_{j=1}^J \mu_j)^2)}. \end{aligned}$$

Similarly, we have

$$\begin{aligned} E(\widehat{ICC}_2) &\approx \frac{E(MS_\alpha - MS_e)}{E(MS_\alpha + (JK - 1)MS_e + J(MS_\beta - MS_e)/N)} \\ &= \frac{V_1 \sum_{j=1}^J \sigma_{Bj}^2 + V_2 \sum_{j=1}^J \sum_{j'=1(j' \neq j)}^J \sigma_{Bj} \sigma_{Bj'} \rho_{\mu jj'}}{U_1 \sum_{j=1}^J \sigma_{Bj}^2 + J \sum_{j=1}^J \sigma_{Wj}^2 + U_2 \sum_{j=1}^J \sum_{j'=1(j' \neq j)}^J \sigma_{Bj} \sigma_{Bj'} \rho_{\mu jj'} + U_3}, \end{aligned}$$

where $V_1 = 1 - \frac{(N-1)(J-1)}{(JK-1)N-J+1}$ and $V_2 = 1 + \frac{N-1}{(JK-1)N-J+1}$; $U_1 = 1 + \frac{J}{N} + \frac{(N-1)(J-1)}{N} \frac{(JK-1)N-J}{(JK-1)N-J+1}$, $U_2 = 1 - \frac{1}{N} \frac{J}{J-1} - \frac{N-1}{N} \frac{(JK-1)N-J}{(JK-1)N-J+1}$, and $U_3 = J(\sum_{j=1}^J \mu_j^2 - (\frac{1}{J-1}) \sum_{j=1}^J \sum_{j'=1(j' \neq j)}^J \mu_j \mu_{j'})$.

$$\begin{aligned} E(\widehat{ICC}_3) &\approx \frac{E(MS_\alpha - MS_\gamma)}{E(MS_\alpha + J(K-1)MS_e + (J-1)MS_\gamma + J(MS_\beta - MS_\gamma)/N)} \\ &= \frac{\frac{1}{J-1} \sum_{j=1}^J \sum_{j'=1(j' \neq j)}^J \sigma_{Bj} \sigma_{Bj'} \rho_{\mu jj'}}{\sum_{j=1}^J \sigma_{Bj}^2 + \sum_{j=1}^J \sigma_{Wj}^2 + (\sum_{j=1}^J \mu_j^2 - \frac{1}{J-1} \sum_{j=1}^J \sum_{j'=1(j' \neq j)}^J \mu_j \mu_{j'})} \\ &= CCC. \end{aligned}$$

For $J = 2$, the expressions are simplified as

$$E(\widehat{ICC}_1) = \frac{(K-1)(\sigma_{B1}^2 + \sigma_{B2}^2) + (2K)\sigma_{B1}\sigma_{B2}\rho_{\mu 12} - \frac{1}{2}(\mu_1 - \mu_2)^2}{(2K-1)(\sigma_{B1}^2 + \sigma_{B2}^2 + \sigma_{W1}^2 + \sigma_{W2}^2 + \frac{1}{2}(\mu_1 - \mu_2)^2)},$$

$$E(\widehat{ICC_2}) = \frac{V_1(\sigma_{B1}^2 + \sigma_{B2}^2) + V_2(2\sigma_{B1}\sigma_{B2}\rho_{\mu 12})}{U_1(\sigma_{B1}^2 + \sigma_{B2}^2) + ((4K - 2)N - 2)(\sigma_{W1}^2 + \sigma_{W2}^2) + U_2(2\sigma_{B1}\sigma_{B2}\rho_{\mu 12}) + U_3},$$

where $V_1 = 2N(K - 1)$ and $V_2 = 2(KN - 1)$; $U_1 = (2K - 1)(2N + 1) - 3$, $U_2 = 2 - 2K$, and $U_3 = ((4K - 2)N - 2)(\mu_1 - \mu_2)^2$, and

$$\begin{aligned} E(\widehat{ICC_3}) &= \frac{2\sigma_{B1}\sigma_{B2}\rho_{\mu 12}}{\sigma_{B1}^2 + \sigma_{B2}^2 + \sigma_{W1}^2 + \sigma_{W2}^2 + (\mu_1 - \mu_2)^2} \\ &= CCC. \end{aligned}$$

Considering $J > 2$ and let K go to infinity, we have

$$\lim_{K \rightarrow \infty} E(\widehat{ICC_1}) \approx \frac{\frac{1}{J} \sum_{j=1}^J \sigma_{Bj}^2 + \frac{1}{J} \sum_{j=1}^J \sum_{j'=1(j' \neq j)}^J \sigma_{Bj} \sigma_{Bj'} \rho_{\mu jj'}}{\sum_{j=1}^J \sigma_{Bj}^2 + \sum_{j=1}^J \sigma_{Wj}^2 + \sum_{j=1}^J \mu_j^2 - \frac{1}{J} (\sum_{j=1}^J \mu_j)^2},$$

and

$$\lim_{K \rightarrow \infty} E(\widehat{ICC_2}) \approx \frac{\sum_{j=1}^J \sigma_{Bj}^2 + \sum_{j=1}^J \sum_{j'=1(j' \neq j)}^J \sigma_{Bj} \sigma_{Bj'} \rho_{\mu jj'}}{U_{1K} \sum_{j=1}^J \sigma_{Bj}^2 + J \sum_{j=1}^J \sigma_{Wj}^2 + U_{2K} \sum_{j=1}^J \sum_{j'=1(j' \neq j)}^J \sigma_{Bj} \sigma_{Bj'} \rho_{\mu jj'} + U_{3K}},$$

where $U_{1K} = J + \frac{1}{N}$, $U_{2K} = \frac{-1}{N(J-1)}$, and $U_{3K} = J(\sum_{j=1}^J \mu_j^2 - \frac{1}{J-1} \sum_{j=1}^J \sum_{j'=1(j' \neq j)}^J \mu_j \mu_{j'})$.

In addition,

$$\lim_{K \rightarrow \infty} E(\widehat{ICC_3}) = CCC.$$