Comparison of Concordance Correlation Coefficient and Coefficient of Individual Agreement in Assessing Agreement

Huiman X. Barnhart[1], Michael Haber[2], Yuliya Lokhnygina[1] and Andrzej S. Kosinski[1]

[1]Department of Biostatistics and Bioinformatics

and Duke Clinical Research Institute

Duke University

PO Box 17969

Durham, NC 27715

[2]Department of Biostatistics

The Rollins School of Public Health

Emory University

Atlanta, GA 30322

mhaber@sph.emory.edu

Corresponding Author: Huiman X. Barnhart

huiman.barnhart@duke.edu

Tel: 919-668-8403

Fax: 919-668-7049

Summary

In method comparison and reliability studies, it is often important to assess agreement between multiple measurements made by different methods, devices, laboratories, observers, or instruments. For continuous data, the concordance correlation coefficient (CCC) is a popular index for assessing agreement between multiple methods on the same subject where none of the methods is treated as reference. Barnhart et al. (2007) proposed coefficient of individual agreement (CIA) to assess individual agreement between multiple methods for situations with and without a reference method extending the concept of individual bioequivalence from the FDA 2001 guidelines. In this paper, we propose a new CCC for assessing agreement between multiple methods where one of the methods is treated as reference. We compare the properties of the CCC and CIA and their dependency on the relative magnitude of between-subject variability and within-subject variability. The relationship between CCC and CIA as well as the impact of between-subject variability are presented algebraically and graphically. Several examples are presented to explain the interpretation of the CCC and CIA values.

Keywords: agreement; concordance correlation coefficient; method comparison; intraclass correlation coefficient; coefficient of individual agreement

1 Introduction

In method comparison and reliability studies, it is often important to assess agreement between measurements made by multiple methods, devices, laboratories, observers, or instruments. For continuous data, the concordance correlation coefficient (CCC) is the most popular index for assessing agreement. The CCC was originally developed by Lin (1989) for two methods ($J = 2$) each making a single reading on a subject. It was later extended to multiple J methods 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) where none of the methods is treated as reference. These extensions included the original CCC for $J = 2$ as a special case. Barnhart and Williamson (2001) used generalized estimating equations (GEE) approach to model the pairwise CCCs as a function of covariates. Chinchilli et al. (1996), King et al. (2007) extended the CCC for data with repeated measures comparing two methods. Quiroz (2005) extended the CCC for data with repeated measures comparing multiple methods by using the two way ANOVA model without interaction. Due to the assumptions on the ANOVA model, the CCC defined by Quiroz (2005) is the special case of CCC by Barnhart et al. (2005) for data with replications. For data with J methods without replications, the CCC corresponds to a version of intraclass correlation coefficient defined by a two-way ANOVA model with or without interaction term (McGraw and Wong, 1996; Barnhart et al., 2002; Carrasco and Jover. 2003). Recently, Barnhart et al. (2007) proposed a coefficient of individual agreement (CIA) to assess individual agreement between multiple measurements based on the concept of individual equivalence, extending the concept of individual bioequivalence in FDA 2001 guidelines for industry. There is a need to understand the similarities and difference between these two indices. As illustrated by Atkinson and Nevill (1997), an increase in the between-subject variability would imply a larger value of CCC even if the individual difference between measurements by the two

methods remains the same. Therefore, it is of particular interest to quantify the degree of impact of the between-subject variability on these two indices. The focus of this paper is to compare the properties of the CCC and the CIA, to investigate the relationship between them and to understand the impact of the between-subject variability on these two indices.

In section 2, we first present the CCC and CIA indices with parameters based on the general model $Y_{ijk} = \mu_{ij} + \epsilon_{ijk}$ for ith subject and jth method, where none of the methods are considered as a reference. We then propose a new CCC for multiple methods where one of the methods is treated as reference. The expressions for CCC and CIA are re-written under parameterization which allows to establish the relationship between CCC and CIA and to understand the degree of impact of various components. We pay special attention to the relative magnitude of between-subject variability and within-subject variability. The relationship between CCC and CIA as well as the impact of between-subject variability are presented algebraically and graphically. Several examples are presented in section 3 to aid in interpretation of the CCC and CIA values and related components. We conclude with recommendations on how to use the CCC and CIA in practice in section 4.

2 Comparison of CCC and CIA and Their Dependency on Between-subject Variability

We first consider the case with two methods ($J = 2$), and then extend the results to the case of J methods. We present the CCC for data without replication (Lin, 1989; Lin, et al. 2002, Barnhart et al., 2002) and the CCC and CIA for data with replications (Barnhart et al., 2005). This CCC treats all methods symmetrically and thus is appropriate for assessing agreement between methods without a reference. In section 2.3, we propose a new CCC for multiple methods where one of the methods is treated as reference and compare this CCC

to the CIA for multiple methods with reference.

2.1 The Case of Two Methods

Consider two readings Y_{i1}, Y_{i2} by two methods on subject i. The CCC is defined as (Lin, 1989)

$$\rho_c = 1 - \frac{E[(Y_{i1} - Y_{i2})^2]}{E[(Y_{i1} - Y_{i2})^2 | Y_{i1}, Y_{i2} \quad \text{are independent}]} = \frac{2\sigma_1\sigma_2\rho}{\sigma_1^2 + \sigma_2^2 + (\mu_1 - \mu_2)^2} = \rho\chi_a$$

where $\rho = Corr(Y_{i1}, Y_{i2})$ is referred to as the precision component and

$$\chi_a = \frac{2\sigma_1\sigma_2}{2\sigma_1\sigma_2 + (\mu_1 - \mu_2)^2 + (\sigma_1 - \sigma_2)^2}$$

is referred to as the accuracy component with $\mu_j = E(Y_{ij})$ and $\sigma_j^2 = Var(Y_{ij})$. We note that χ_a assesses location shift ($(\mu_1 - \mu_2)^2$) and scale shift ($(\sigma_1 - \sigma_2)^2$) relative to the scales (σ_j). For fixed values of location and scale shifts, χ_a is an increasing function of σ_j and we have $\chi_a \to 1$ and $\rho_c \to \rho$ as $\sigma_j \to \infty, j = 1, 2$. Intuitively as the measurement range increases, σ_j would increase, and the correlation ρ is also likely to increase even if the individual differences $Y_{i1} - Y_{i2}$ remain the same. Therefore, for data sets where the location and scale shifts stay the same, one would obtain high CCC value for data with large variability of σ_j (Atkinson and Nevill, 1997). In the extreme case, if σ_j is so much larger than the location and scale shifts, the CCC would be close to ρ which may be close to 1. Thus, when reporting the CCC for assessing agreement, estimated values for μ_j, σ_j and ρ should also be reported and one should not compare the CCC values among different data sets unless σ_j's are similar. Because of its dependency on σ_j the CCC should be viewed as an agreement index conditional on between-subject variability.

Assume that other than the subject's own characteristics and the measurement method, there are no other external factors influencing the reading value Y_{ij}. Then $\sigma_j^2 = Var(Y_{ij})$ contains both the variability of the subject's true value and the variability of the random

error (within-subject) contributed by method j. Thus, it is sensible to decompose σ_j^2 as the sum of these two sources of variabilities. Without true gold standard, one would never know the subject's true value. However, subject's true value by method j, μ_{ij}, may be estimated if replicated measurements are taken on the same subject by the same method. Let Y_{ijk} be the kth replicated measurement for the ith subject by the jth method. We write $Y_{ijk} = \mu_{ij} + \epsilon_{ijk}$ with the following minimal common assumptions: (1) μ_{ij} and ϵ_{ijk} are independent with means of $E(\mu_{ij}) = \mu_j$ and $E(\epsilon_{ijk}) = 0$ and between-subject and within-subject variances of $Var(\mu_{ij}) = \sigma_{Bj}^2$ and $Var(\epsilon_{ijk}) = \sigma_{Wj}^2$, respectively; (2) $\mu_{ij}'s$ are correlated with $Corr(\mu_{i1}, \mu_{i2}) = \rho_\mu$ and ϵ_{ijk}'s are uncorrelated. Thus we have $\sigma_j^2 = \sigma_{Bj}^2 + \sigma_{Wj}^2, Cov(Y_{i1}, Y_{i2}) = \rho\sigma_1\sigma_2 = \rho_\mu\sigma_{B1}\sigma_{B2}$ and

$$\rho = \frac{\sigma_{B1}\sigma_{B2}\rho_\mu}{\sqrt{\sigma_{B1}^2 + \sigma_{W1}^2}\sqrt{\sigma_{B2}^2 + \sigma_{W2}^2}} \leq \rho_\mu.$$

The CCC for data with replications can be written as

$$\begin{aligned} \rho_c &= 1 - \frac{E[(Y_{i1k} - Y_{i2k'})^2]}{E[(Y_{i1k} - Y_{i2k'})^2 | Y_{i1k}, Y_{i2k'} \quad \text{are independent}]} \\ &= \frac{2\sigma_{B1}\sigma_{B2}\rho_\mu}{2\sigma_{B1}\sigma_{B2} + (\mu_1 - \mu_2)^2 + (\sigma_{B1} - \sigma_{B2})^2 + \sigma_{W1}^2 + \sigma_{W2}^2} \end{aligned}$$

This is the total-CCC defined in Barnhart et al. (2005). Note that the above expression for CCC holds for the case without replications because we can mathematically write $Y_{ij} = \mu_{ij} + \epsilon_{ij}$ even though we can not estimate σ_{Bj} and σ_{Wj}. Barnhart et al. (2005) also defined inter-CCC at the level of μ_{ij}'s as

$$\begin{aligned} \rho_c(\mu) &= 1 - \frac{E[(\mu_{i1} - \mu_{i2})^2]}{E[(\mu_{i1} - \mu_{i2})^2 | \mu_{i1}, \mu_{i2} \quad \text{are independent}]} \\ &= \frac{2\sigma_{B1}\sigma_{B2}\rho_\mu}{2\sigma_{B1}\sigma_{B2} + (\mu_1 - \mu_2)^2 + (\sigma_{B1} - \sigma_{B2})^2} \end{aligned}$$

and they used intraclass correlation coefficient (ICC) to assess intra-method agreement, where ICC for method j is expressed as

$$\rho_j^I = \frac{\sigma_{Bj}^2}{\sigma_{Bj}^2 + \sigma_{Wj}^2}.$$

The (total) CCC is related to inter-CCC and ICC by

$$\frac{1}{\rho_c} = \frac{1}{\rho_c(\mu)} + \frac{1}{\gamma}$$

where

$$\frac{1}{\gamma} = \frac{\sigma_{W1}^2 + \sigma_{W2}^2}{2\rho_\mu \sigma_{B1}\sigma_{B2}} = \frac{1}{2\rho_\mu}(\frac{\sigma_{B1}}{\sigma_{B2}}\frac{1-\rho_1^I}{\rho_1^I} + \frac{\sigma_{B2}}{\sigma_{B1}}\frac{1-\rho_2^I}{\rho_2^I}),$$

is the weighted sum of the odds of $1-\rho_j^I$. For comparison purpose with CIA, we use σ_{Bj} (that is probably better than using σ_j) to represent the between-subject variability hereafter. In general, we have

- The CCC decreases as the location and scale shifts ($(\mu_1-\mu_2)^2$, $(\sigma_{B1}-\sigma_{B2})^2$) increase.
- The CCC decreases as the within-subject variability (σ_{Wj}) increases.
- The CCC increases as the between-subject variability (σ_{Bj}) increases.
- The CCC increases as the "true" correlation (ρ_μ) increases.

The inter-CCC has the same properties as the CCC except without dependency on the within-subject variability (σ_{Wj}).

The CIA index was introduced based on the concept of individual agreement (Barnhart et al, 2007, Haber and Barnhart, 2007), that was extended from the individual bioequivalence concept in 2001 FDA guidelines for industry. Specifically, individual agreement between methods is good only if variability of individual measurements from these methods is similar to the variability of replicated measurements within a method. The CIA for the case of two methods where neither of them is treated as reference is defined as

$$\begin{aligned}
\psi^N &= \frac{[E(Y_{i1k}-Y_{i1k'})^2 + E(Y_{i2k}-Y_{i2k'})^2)]/2}{E[(Y_{i1k}-Y_{i2k'})^2]} \\
&= \frac{\sigma_{W1}^2+\sigma_{W2}^2}{\sigma_D^2+(\mu_1-\mu_2)^2+\sigma_{W1}^2+\sigma_{W2}^2} \\
&= \frac{\sigma_{W1}^2+\sigma_{W2}^2}{2(1-\rho_\mu)\sigma_{B1}\sigma_{B2}+(\mu_1-\mu_2)^2+(\sigma_{B1}-\sigma_{B2})^2+\sigma_{W1}^2+\sigma_{W2}^2},
\end{aligned}$$

where $\sigma_D^2 = Var(\mu_{i1} - \mu_{i2}) = \sigma_{B1}^2 + \sigma_{B2}^2 - 2\sigma_{B1}\sigma_{B2}\rho_\mu$ is the subject-by-method interaction. The ψ^N ranges from 0 to 1 and we want to have high value of CIA^N to claim satisfactory individual agreement. Barnhart et al. (2007) suggested to have $\psi^N \geq 0.445$ for good individual agreement where the inter-method variability (or total between method variability) is within 125% (or 225%) of the within-subject variability. Haber and Barnhart (2007) suggested to have $\psi^N \geq 0.8$ for excellent individual agreement where the inter-method variability (or total between method variability) is within 25% (or 125%) of within-subject variability. We emphasize that the index CIA is used for assessing individual agreement only if the within-subject variability σ_{Wj}^2 is established to be acceptable. This can be examined by repeatability coefficient (Bland and Altman, 1999), $1.96 * \sqrt{2 * \sigma_{Wj}^2}$, to see whether it is less than or equal to an acceptable value within which the difference between any two readings by the same method should lie for 95% of subjects. In the special case that there are no location and scale shifts (i.e., $\mu_j = \mu, \sigma_{Bj} = \sigma_B$ for all j) and $\rho_\mu = 1$, the CIA is the ratio of the odds of the CCC over the odds of the intra ICC, i. e.,

$$\psi^N = \frac{\rho_c/(1-\rho_c)}{\rho^I/(1-\rho^I)}.$$

In general, we have

- The CIA decreases as the location and scale shifts ($(\mu_1 - \mu_2)^2$, $(\sigma_{B1} - \sigma_{B2})^2$) increase.
- The CIA increases as the within-subject variability (σ_{Wj}) increases.
- The CIA decreases as the between-subject variability (σ_{Bj}) increases.
- The CIA increases as the "true" correlation (ρ_μ) increases.

Thus, the CIA is similar to the CCC in its relationship with the location and scale shifts and the "true" correlation. However, the CIA differs from the CCC in its relationship with the within-subject variability (σ_{Wj}) and the between-subject variability (σ_{Bj}).

If $\rho_\mu = 0$, then the CCC is equal to zero and it is independent of between-subject variability. In this case, the CIA measures the location, scale shifts and the between-subject variability relative to the within-subject variability and we have $\psi^N \geq \rho_c$.

If general, the CCC and the CIA have the following relationship:

$$\psi^N = \frac{\rho_c}{1-\rho_c}\frac{1}{\gamma} \qquad \text{or} \qquad \rho_c = \frac{\gamma\psi^N}{1+\gamma\psi^N}, \quad \text{if} \quad \rho_c \neq 0, 1,$$

where $1/\gamma$ is defined previously and it is a weighted sum of the odds of $1-\rho_j^I$. The CIA is related to inter-CCC and intra ICC by

$$\psi^N = \frac{1}{1-\frac{1}{\rho_c(\mu)}+\frac{1}{\gamma}}.$$

Thus, we have $\psi^N \leq \rho_c$ if $\rho_c \geq 1-1/\gamma$ and $\psi^N \geq \rho_c$ if $\rho_c \leq 1-1/\gamma$.

For simplicity, we consider case of equal between-subject variabilities and equal within-subject variabilities, i.e., $\sigma_{Bj}^2 = \sigma_B^2, \sigma_{Wj}^2 = \sigma_W^2, j = 1, 2$, to assess the impact of between-subject variability on these two indices. Let $d = \sigma_B^2/\sigma_W^2$ denote the magnitude of between-subject variability relative to the within-subject variability. Then ICC $= \rho_1^I = \rho_2^I = \rho^I = d/(d+1), \gamma = \rho_\mu d$, and the CCC and the CIA are functions of d, ρ_μ and $(\mu_1-\mu_2)^2/(2\sigma_W^2)$ with the following expressions:

$$\rho_c = \frac{d\rho_\mu}{d+(\mu_1-\mu_2)^2/(2\sigma_W^2)+1}, \qquad \psi^N = \frac{1}{(1-\rho_\mu)d+(\mu_1-\mu_2)^2/(2\sigma_W^2)+1}.$$

The CCC and CIA are related by

$$\psi^N = \frac{\rho_c}{1-\rho_c}\frac{1}{\rho_\mu d} = \frac{1}{1-\rho_c}\frac{1}{d+(\mu_1-\mu_2)^2/(2\sigma_W^2)+1}, \quad \text{if} \quad \rho_c \neq 0, 1.$$

In practice, the within-subject variability and the true correlation (ρ_μ) may remain constant or may increase as the between-subject variability increases. As long as the within-subject variability is considered to be acceptable, the impact of between-subject variability can be assessed through the magnitude of d.

To understand the dependency of the CCC and the CIA on the magnitude (d) of the between-subject variability relative to the within-subject variability, we produce a series of graphs (Figures 1-3) that provide the values of CCC and the CIA for a fixed value of d as well as the trend of the CCC and CIA as d increases. As d increases, we examine combinations of the following: $\rho_\mu = 1, 0.8$, $\mu_1 \neq \mu_2$ or $\mu_1 = \mu_2$ (e.g., $(\mu_1 - \mu_2)^2 = 9$ or $(\mu_1 - \mu_2)^2 = 0$), σ_W^2 is a constant or is an increasing function of d (e.g, $\sigma_W^2 = 9/2$ or $\sigma_W^2 = d$). We first look at these two agreement indices as functions of d (Figure 1).

If $\rho_\mu = 1$, the CCC measures the location, scale shifts and within-subject variability relative to the between-subject variability, but the CIA measures the location and scale shifts relative to the within-subject variability and it is independent of the between-subject variability. If the within-subject variability remains constant as d increases, Figure 1(a) shows that the CCC is increasing to 1 while the CIA remains constant as d increases for fixed location shift and within-subject variability. If the within-subject variability increases as the between-subject variability increases, e.g., $\sigma_W^2 = d$ implying that $\sigma_B^2 = d^2$, thus both the CCC and CIA increase as d increases. As shown in Figure 1(b), the CCC increases faster than the CIA as d increases. If there is no location shift ($\mu_1 = \mu_2$), then CIA is equal to 1 which is always larger than the CCC (Figure 1(c)).

If $0 < \rho_\mu < 1$, then for constant within-subject variability and true correlation (ρ_μ), the CCC increases to ρ_μ and the CIA decreases to zero as $d \to \infty$ (Figure 1(d)) for fixed values of location shift and true correlation (ρ_μ). However, if both the within-subject variability and the true correlation (ρ_μ) is an increasing function of d, e.g., $\sigma_W^2 = d$ and $\rho_\mu = \rho^I = d/(d+1)$, then both the CCC and CIA increase as d increases (Figure 1(e)) with the CCC increases faster than the CIA, for the case of $\mu_1 \neq \mu_2$. If $\mu_1 = \mu_2$, then as d increases, the CCC increases to 1 and the CIA decreases to 0.5 (Figure 1(f)). We see that the CIA stabilizes faster than the CCC.

In summary, we find that the CIA is less dependent on the between-subject variability than the CCC from the perspective of their dependency on the relative magnitude of between-subject variability and within-subject variability.

To understand how the value of the CIA is related to the value of CCC, the curves of CIA as a function of CCC for fixed value of d are shown in Figures 2 and 3. We also super impose a 45 degree line in the figure to see which value is larger. For fixed value of d, we note that CCC is bounded by $d/(d+1)$. We should point out that these curves correspond to various values of ρ_μ because $\rho_\mu = \rho_c(d + (\mu_1 - \mu_2)^2/2\sigma_W^2 + 1)/d$. Figure 2 shows the curves when $\mu_1 \neq \mu_2$ for cases of a constant within-subject variability and non-constant within-subject variability for different values of d. Figure 3 shows the curves when $\mu_1 = \mu_2$.

In both Figures 2 and 3, the value of CIA is less than CCC for most of the situations. If d is very large, then CIA is close to zero for most values of the CCC and then CIA jumps up rapidly to approach to one as the CCC approaches 1. Because the curves are similar in Figures 2 and 3, we focus on Figure 3 where $\mu_1 = \mu_2$. In this case, the CCC and CIA have the following expressions:

$$\rho_c = \rho^I \rho_\mu = \frac{d\rho_\mu}{d+1}, \qquad \psi^N = \frac{1}{d(1-\rho_\mu)+1},$$

and the CIA is related to the CCC by

$$\psi^N = \frac{1}{(d+1)(1-\rho_c)}.$$

We note that

$$\begin{aligned}
\psi^N &\geq \rho_c \quad \text{if} \quad d \leq 3 \\
\psi^N &\geq \rho_c \quad \text{if} \quad d > 3 \quad \text{and} \quad -1 \leq \rho_c \leq 0.5 - 0.5\sqrt{\tfrac{d-3}{d+1}} \\
\psi^N &\leq \rho_c \quad \text{if} \quad d > 3 \quad \text{and} \quad 0.5 - 0.5\sqrt{\tfrac{d-3}{d+1}} \leq \rho_c \leq 0.5 + 0.5\sqrt{\tfrac{d-3}{d+1}} \\
\psi^N &\geq \rho_c \quad \text{if} \quad d > 3 \quad \text{and} \quad 0.5 + 0.5\sqrt{\tfrac{d-3}{d+1}} \leq \rho_c \leq d/(d+1)
\end{aligned}$$

In summary, if $d \leq 3$, then the CIA is always larger than or equal to the CCC. If $d > 3$, then the CIA is larger than the CCC for low value of CCC, but smaller than the CCC for

most value of CCC, and can be larger than the CCC for only extremely high value of CCC. If d is very large, i.e., $d \geq 100$ or $\rho^I \geq 0.99$, then CIA is close to zero for most values of the CCC and then CIA jumps up rapidly to approach to one as the CCC approaches 1.

It is interesting to note that the above observations in Figure 3 hold approximately in general if the difference of the population means, $\mu_1 - \mu_2$, is relatively small as compared to σ_B (e.g., Figure 2(b)). This is particularly true if as $\sigma_{Bj} \to \infty$, the value of $\mu_1 - \mu_2$ is fixed, and $\sigma^2_{Bj}/\sigma^2_{Wj} \to d$, $\sigma_{B1}/\sigma_{B2} \to 1$. This corresponds to the case that the between-subject variability is a multiple of the within-subject variability and the location shift is negligible relative to the between-subject variability (implying that the coefficient of variation for the difference of two readings for the same subject is close to zero).

In summary, the CCC value is usually larger (smaller) than the CIA value for large (small) relative magnitude of between-subject variability to within-subject variability.

2.2 The Case of Multiple Methods without Reference

Consider J readings $Y_{i1}, \ldots, Y_{iJ}$ by J methods on subject i. The CCC is defined as (Lin, 1989, 2000; King and Chinchilli, 2001; Barnhart et al., 2002)

$$\begin{aligned} \rho_c &= 1 - \frac{E[\sum_{j=1}^{J-1}\sum_{j'=j+1}^{J}(Y_{ij} - Y_{ij'})^2]}{E[\sum_{j=1}^{J-1}\sum_{j'=j+1}^{J}(Y_{ij} - Y_{ij'})^2 | Y_{i1}, \ldots, Y_{iJ} \quad \text{are independent}]} \\ &= \frac{2\sum_{j=1}^{J-1}\sum_{j'=j+1}^{J}\sigma_j\sigma_{j'}\rho_{jj'}}{(J-1)\sum_{j=1}^{J}\sigma_j^2 + \sum_{j=1}^{J-1}\sum_{j'=j+1}^{J}(\mu_j - \mu_{j'})^2} \end{aligned}$$

where $E(Y_{ij}) = \mu_j, Var(Y_{ij}) = \sigma_j^2, corr(Y_{ij}, Y_{ij'}) = \rho_{jj'}$. For data with replication, let Y_{ijk} be the kth replicated measurements for the ith subject by the jth method and write $Y_{ijk} = \mu_{ij} + \epsilon_{ijk}$ with the same common assumptions in section 2.1. The CCC for data with replications can be written as

$$\rho_c = \frac{2\sum_{j=1}^{J-1}\sum_{j'=j+1}^{J}\sigma_{Bj}\sigma_{Bj'}\rho_{\mu jj'}}{(J-1)\sum_{j=1}^{J}\sigma_{Bj}^2 + \sum_{j=1}^{J-1}\sum_{j'=j+1}^{J}(\mu_j - \mu_{j'})^2 + (J-1)\sum_{j=1}^{J}\sigma_{Wj}^2}$$

$$= \frac{2\sum_{j=1}^{J-1}\sum_{j'=j+1}^{J}\sigma_{Bj}\sigma_{Bj'}\rho_{\mu jj'}}{\sum_{j=1}^{J-1}\sum_{j'=j+1}^{J}[2\sigma_{Bj}\sigma_{Bj'}+(\mu_j-\mu_{j'})^2+(\sigma_{Bj}-\sigma_{Bj'})^2+\sigma_{Wj}^2+\sigma_{Wj'}^2]},$$

by noting that $\sum_{j=1}^{J-1}\sum_{j'=j+1}^{J}(\sigma_{Bj}^2+\sigma_{Bj'}^2)=(J-1)\sum_{j=1}^{J}\sigma_{Bj}^2$. This is the total-CCC defined in Barnhart et al. (2005). Again, this expression for CCC holds for the case without replications if we write $Y_{ij}=\mu_{ij}+\epsilon_{ij}$ even though we can not estimate σ_{Bj} and σ_{Wj} separately. The inter-CCC defined at the level of μ_{ij}'s is

$$\begin{aligned}\rho_c(\mu) &= 1-\frac{E[\sum_{j=1}^{J-1}\sum_{j'=j+1}^{J}(\mu_{ij}-\mu_{ij'})^2]}{E[\sum_{j=1}^{J-1}\sum_{j'=j+1}^{J}(\mu_{ij}-\mu_{ij'})^2|\mu_{i1},\ldots,\mu_{iJ} \quad \text{are independent}]}\\ &= \frac{2\sum_{j=1}^{J-1}\sum_{j'=j+1}^{J}\sigma_{Bj}\sigma_{Bj'}\rho_{\mu jj'}}{\sum_{j=1}^{J-1}\sum_{j'=j+1}^{J}[2\sigma_{Bj}\sigma_{Bj'}+(\mu_j-\mu_{j'})^2+(\sigma_{Bj}-\sigma_{Bj'})^2]}\end{aligned}$$

We use the ICCs, ρ_j^I, to assess the intra-method agreement for method j. Similar to what is shown in section 2.1, the (total) CCC is related with the inter-CCC and ICCs by

$$\frac{1}{\rho_c}=\frac{1}{\rho_c(\mu)}+\frac{1}{\gamma}$$

where

$$\frac{1}{\gamma}=\frac{(J-1)\sum_{j=1}^{J}\sigma_{Wj}^2}{2\sum_{j=1}^{J-1}\sum_{j'=j+1}^{J}\sigma_{Bj}\sigma_{Bj'}\rho_{\mu jj'}}=\sum_{j=1}^{J}\omega_j\frac{1-\rho_j^I}{\rho_j^I}$$

with $\omega_j=\sigma_{Bj}^2/(2\sum_{j=1}^{J-1}\sum_{j'=j+1}^{J}\sigma_{Bj}\sigma_{Bj'}\rho_{\mu jj'})$ is the weighted sum of the odds of $1-\rho_j^I$.

The CIA defined in Barnhart et al. (2007) for the case of J methods where none of them is treated as reference can be written as

$$\begin{aligned}\psi^N &= \frac{\sum_{j=1}^{J}E(Y_{ijk}-Y_{ijk'})^2/2}{\sum_{j=1}^{J-1}\sum_{j'=j+1}^{J}E[(Y_{ij}-Y_{ij'})^2]/(J-1)}\\ &= \frac{\sum_{j=1}^{J-1}\sum_{j'=j+1}^{J}(\sigma_{Wj}^2+\sigma_{Wj'}^2)}{\sum_{j=1}^{J-1}\sum_{j'=j+1}^{J}[2(1-\rho_{\mu jj'})\sigma_{Bj}\sigma_{Bj'}+(\mu_j-\mu_{j'})^2+(\sigma_{Bj}-\sigma_{Bj'})^2+\sigma_{Wj}^2+\sigma_{Wj'}^2]}.\end{aligned}$$

It can be shown that the CCC and CIA are related in the same way as in the case of two methods, i.e.,

$$\psi^N=\frac{\rho_c}{1-\rho_c}\frac{1}{\gamma} \qquad \text{or} \qquad \rho_c=\frac{\gamma\psi^N}{1+\gamma\psi^N}, \quad \text{if} \quad \rho_c\neq 0,1.$$

Based on the expressions for the CCC and the CIA shown above, the properties of the CCC and CIA for the case of two methods extend similarly to the case of J methods. Similarly, the

comparisons of the CCC and CIA in terms of between-subject variability can be extended from the case of the two methods to the case of J methods.

2.3 The Case of Multiple Methods with One Reference

We use the notation in section 2.2 and assume that the first $J-1$ methods are new and the Jth method is treated as reference. We propose the new CCC for the case of multiple methods with one reference as

$$\begin{aligned}\rho_c^R &= 1 - \frac{E[\sum_{j=1}^{J-1}(Y_{ij}-Y_{iJ})^2]}{E[\sum_{j=1}^{J-1}(Y_{ij}-Y_{iJ})^2|Y_{i1},\ldots,Y_{iJ} \quad \text{are independent}]} \\ &= \frac{2\sum_{j=1}^{J-1}\sigma_j\sigma_J\rho_{jJ}}{\sum_{j=1}^{J-1}[\sigma_j^2+\sigma_J^2+(\mu_j-\mu_J)^2]}.\end{aligned}$$

When there is one new method and one reference, the ρ_c^R is the same as the usual ρ_c for the case without reference in section 2.1. For data with replications, the ρ_c^R can be written as

$$\rho_c^R = \frac{2\sum_{j=1}^{J-1}\sigma_{Bj}\sigma_{BJ}\rho_{\mu jJ}}{\sum_{j=1}^{J-1}[2\sigma_{Bj}\sigma_{BJ}+(\mu_j-\mu_J)^2+(\sigma_{Bj}-\sigma_{BJ})^2+\sigma_{Wj}^2+\sigma_{WJ}^2]}.$$

We can also define inter-CCC at the level of μ_{ij}'s as

$$\begin{aligned}\rho_c(\mu)^R &= 1 - \frac{E[\sum_{j=1}^{J-1}(\mu_{ij}-\mu_{iJ})^2]}{E[\sum_{j=1}^{J-1}(\mu_{ij}-\mu_{iJ})^2|\mu_{i1},\ldots,\mu_{iJ} \quad \text{are independent}]} \\ &= \frac{2\sum_{j=1}^{J-1}\sigma_{Bj}\sigma_{BJ}\rho_{\mu jJ}}{\sum_{j=1}^{J-1}[2\sigma_{Bj}\sigma_{BJ}+(\mu_j-\mu_J)^2+(\sigma_{Bj}-\sigma_{BJ})^2]}.\end{aligned}$$

It can be shown that the CCC is related with the inter-CCC and ICCs by

$$\frac{1}{\rho_c^R} = \frac{1}{\rho_c(\mu)^R} + \frac{1}{\gamma^{R*}}$$

where

$$\frac{1}{\gamma^{R*}} = \frac{\sum_{j=1}^{J-1}(\sigma_{Wj}^2+\sigma_{WJ}^2)}{2\sum_{j=1}^{J-1}\sigma_{Bj}\sigma_{BJ}\rho_{\mu jJ}} = \sum_{j=1}^{J}\omega_j^R\frac{1-\rho_j^I}{\rho_j^I},$$

with $\omega_j^R = \sigma_{Bj}^2/(2\sum_{j=1}^{J-1}\sigma_{Bj}\sigma_{BJ}\rho_{\mu jJ}), j=1,\ldots,J-1$ and $\omega_J^R = (J-1)\sigma_{BJ}^2/(2\sum_{j=1}^{J-1}\sigma_{Bj}\sigma_{BJ}\rho_{\mu jJ})$ is the weighted sum of the odds of $1-\rho_j^I$.

The CIA defined in Barnhart et al. (2007) for the case of J methods where the Jth method is treated as reference can be written as

$$\begin{aligned}\psi^R &= \frac{E(Y_{iJk} - Y_{iJk'})^2/2}{\sum_{j=1}^{J-1} E[(Y_{ij} - Y_{iJ})^2]/(J-1)} \\ &= \frac{\sigma_{WJ}^2}{\sum_{j=1}^{J-1}[2(1-\rho_{\mu jJ})\sigma_{Bj}\sigma_{BJ} + (\mu_j - \mu_J)^2 + (\sigma_{Bj} - \sigma_{BJ})^2 + \sigma_{Wj}^2 + \sigma_{WJ}^2]}.\end{aligned}$$

The ψ^R in general ranges from 0 to 1 because in general we have $\sigma_{WJ}^2 \leq \sigma_{Wj}^2, j = 1, \ldots, J-1$. It is possible for ψ^R to be greater than 1 and in this case the new method is better than the reference method because the new method would have smaller within-subject variability and difference between the new method and the reference is smaller than the difference between the replications within the reference. Similar to ψ^N, we want to have high value of CIA^R to claim satisfactory individual agreement with $\psi^R \geq 0.445$ for good agreement and $\psi^R \geq 0.8$ for excellent agreement.

It can be shown that the CCC and CIA are related in general as

$$\psi^R = \frac{\rho_c^R}{1-\rho_c^R}\frac{1}{\gamma^R} \qquad \text{or} \qquad \rho_c^R = \frac{\gamma^R\psi^R}{1+\gamma^R\psi^R}, \quad \text{if} \quad \rho_c^R \neq 0, 1,$$

where

$$\frac{1}{\gamma^R} = \frac{(J-1)\sigma_{WJ}^2}{\sum_{j=1}^{J-1}\sigma_{Bj}\sigma_{BJ}\rho_{\mu jJ}}.$$

We note that γ^R is related to γ^{R*} defined earlier by

$$\gamma^R = \gamma^{R*}\frac{\sigma_{WJ}}{\sigma_{*R}^2},$$

where

$$\sigma_{*R}^2 = \frac{1}{2}\left(\frac{\sum_{j=1}^{J-1}\sigma_{Wj}^2}{J-1} + \sigma_{WJ}^2\right)$$

is the weighted average of within-subject variabilities. If $\sigma_{Wj}^2 = \sigma_{WJ}^2, j = 1, \ldots, J-1$, then $\gamma^R = \gamma^{R*}$.

These definitions and relationships show that the properties examined between the CCC and CIA for the case of two methods in section 2.1 can be extended to the case of multiple methods with one reference. Estimation for the new CCC and inter-CCC for multiple

methods with one reference can be obtained through the SAS procedure `MIXED` as described in Barnhart et al. (2005).

3 Examples

We present four examples to compare the point estimates of the CCC and CIA from four different data sets. The first data set comes from Eliasziw et al. (1994) to assess agreement between a manual goniometer and an electro-goniometer for measuring knee joint angles. It has a total of 29 subjects where 3 replications were taken by each of the two goniometers. The second data set comes from Haber et al. (2005) to assess agreement between two radiologists in grading the coronary artery calcium score. It has 12 patients where 2 replications were taken by each radiologist. The third data set comes from the Emory carotid stenosis study (Barnhart et al., 2005). We assess agreement between the two new methods, two dimensional and three dimensional magnetic resonance angiography (MRA-2D and MRA-3D), and between the two new methods versus the reference method, the invasive intra-arterial angiogram (IA). We consider the readings by the three observers using the same method as the replications of the method for illustration. This data set contains 55 subjects where 3 replications were taken by each method. The fourth data set was taken from the paper by Bland and Altman (1999) where two human observers and a semi-automatic blood pressure monitor (denoted as machine) made three replicated readings of systolic blood pressure on 85 subjects. We assess the pairwise agreement among the two observers and the machine where we consider the observer as a reference and not as a reference.

Table 1 displays the estimates of population means, within-subject variability, between-subject variability, intra-ICC and repeatability coefficient for each of the four data sets. For all four data sets, the location and scale shifts are small relative to the within-subject or

between-subject variability except the systolic blood pressure data where the reading by the machine has higher mean and higher within-subject variability than the ones by the two observers. The intra-ICCs are high (> 0.9) for all data sets except the carotid stenosis data where the intra-ICC is about 0.6. This is because the between-subject variability is relatively much higher than the within-subject variability except in the carotid stenosis data. We note that the repeatability coefficient is large (in the range of 60's) in practical sense for MRA-2D and MRA-3D in the carotid stenosis data. However, their intra-ICCs are moderate (about 0.6). This indicates that there is substantial within-subject variability when using the new method MRA-2D or MRA-3D.

Table 2 gives the estimates for the CCC and CIA for the four data examples where none of the methods is treated as reference. The CCC is estimated by using the SAS procedure `MIXED` that was described in Barnhart et al. (2005) for data with replications. The method of moment is used for estimating the CIA (Barnhart et al., 2007). To better understand and compare the estimates of the CCC and CIA, estimates of other parameters such as true correlation (ρ_μ) and $d = \sigma_B^2/\sigma_W^2$ are also displayed in table 2. Because the scale shift is small, the estimate of d is computed as the sum of σ_{Bj}^2 over the sum of σ_{Wj}^2, a weighted average of $\sigma_{Bj}^2/\sigma_{Wj}^2$ with weights $\sigma_{Wj}^2/\sum_j \sigma_{Wj}^2$. The estimated value of d ranges from 1.6 to 275 with the largest value coming from the calcium scoring data and the smallest value coming from the carotid stenosis data.

In the calcium scoring data, the CCC is 0.995 as compared to CIA of 0.754. The estimated true correlation ρ_μ is close to 1. With a very large value of $d = 275$ and $\rho_\mu \approx 1$, this corresponds to the situation observed in Figure 1(a) or 1(b), the CCC is close to one due to large between-subject variability and the CIA measures the location shift relative to the within-subject variability.

In the goniometer data, the CCC is 0.944 as compared to CIA of 0.287. The estimated

true correlation is 0.977. With a relatively large $d = 61.4$, small location shift and small (possibly constant) within-subject variability, this may correspond to the situation in Figure 1(d) where the CCC is large due to large between-subject variability and the CIA is small due to small within-subject variability.

In the systolic blood pressure data, we look at three pairwise comparisons. For assessing agreement between observers 1 and 2, the CCC is 0.973 and the CIA is 1.0. The estimated inter-CCC and true correlation are also equal to 1.0. This may correspond to the situation in Figure 1(c) where CIA is one and the CCC is close to one due to small location and scale shifts and medium value of d. However, for agreement between observer 1 versus machine or observer 2 versus machine, the CCC values are around 0.7 while the CIA values are small around 0.18. There are large location shift and within-subject variability shift in these two comparisons. Therefore the CCC value is not as large as the CCC comparing the two observers. However, the CCC is not very low despite a mean difference of 16 points between the observer (1 or 2) and machine – quite poor agreement from a practical point of view at the population level. This is because there is large between-subject variability. The increased within-subject variability should help to increase the value of CIA in the comparison of observer vs. machine. However, larger location shift, smaller estimated true correlation (0.83) and medium value of d make the CIA value small as observed in Figure 1(d).

For the above three examples, the CCC values are larger than the CIA values except the comparisons of two observers in the blood pressure example. In the carotid stenosis data that compares MRA-2D versus MRA-3D, the CCC value (about 0.58 for left and right arteries) is smaller than the CIA value (0.881 for left artery and 0.917 for right artery). This is mainly because there is considerable within-subject variability that decreases the CCC and increases the CIA. With estimated value of d less than 2, we expect the CCC is less than the CIA as seen in Figure 2.

Table 3 presents the estimates for three examples where one of the methods is treated as reference. We see similar patterns when comparing the CCC value to the CIA value. In general, the CCC and the CIA values where one method is treated as reference are less than the CCC and the CIA values without any reference, respectively. This is because the within-subject variability for the reference method is smaller than the within-subject variability for the new methods.

4 Discussion

We proposed a new CCC for assessing agreement for multiple methods with one reference. We compared the CCC and CIA indices for assessing agreement in their dependency on between-subject variability relative to the within-subject variability for fixed location and scale shifts. We also characterized the relationship between the CCC and the CIA. In general, we find that the CIA value is less dependent on the relative magnitude of between-subject variability to the within-subject variability than the CCC value.

In using the CCC and the CIA for assessing agreement in practice, one needs to consider the magnitude of the between-subject variability and within-subject variability. The first question is whether the within-subject variability is acceptable based on the subject matter for the considered measurement range. This can be examined by repeatability coefficient that provides the value within which any two readings by the same method would lie for 95% of subjects. If the answer is yes, then consider using the CIA index, especially when the between-subject variability is large relative to the within-subject variability. If the answer is no or not sure, then consider using both the CCC and the CIA indices with appropriate interpretations. We need to keep in mind that a high CCC value may be driven by large between-subject variability and a high CIA value may be driven by large within-subject

variability. If the between-subject variability varies greatly from different populations of subjects, the CCC values from these populations can not be compared, especially if one tries to find the population that gives the best agreement of the methods. However, the CIA values from these populations may be compared if the magnitude of the between-subject variability relative to the within-subject variability is similar across these populations and the between-subject variabilities in these populations are acceptable.

Acknowledgements

This research is supported by the National Institutes of Health Grant R01 MH70028

References

1. Atkinson, G. and Nevill, A. (1997). Comment on the use of concordance correlation to assess the agreement between two variables. *Biometrics* 53: 775-777.

2. Barnhart, H.X. and Williamson, J.M. (2001). Modeling Concordance Correlation via GEE to Evaluate Reproducibility. *Biometrics* 57: 931-940.

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

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

5. Barnhart, H.X., Haber, M., and Kosinski, A.S. (2007). Assessing Individual Agreement. *Journal of Biopharmaceutical Statistics* a special issue on agreement, accepted.

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

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

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

9. 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.

10. FDA. Guidance for industry: Statistical approaches to establishing bioequivalence, Food and Drug Administration, Center for Drug Evaluation and Research (CDER), January 2001, BP.

11. Haber, M. and Barnhart, H.X. (2007). A general approach to evaluating agreement between two observers or methods of measurement from quantitative data with replicated measurements. *Statistical Methods in Medical Research*, in press.

12. Haber, M., Barnhart, H.X., Song, J., and Gruden, J. (2005). Interobserver Variability: a New Approach in Evaluating Interobserver Agreement. *Journal of Data Sciences* 3: 69-83.

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

14. King, T.S., Chinchilli, V.M., Carrasco, J. (2007). A repeated measures concordance correlation coefficient, *Statistics in Medicine* in press.

15. Lin, L.I. (1989). A concordance correlation coefficient to evaluate reproducibility. *Biometrics* 45:225-268.

16. Lin, L.I. (2000). A note on the concordance correlation coefficient. *biometrics* 56: 324-325.

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

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

19. Quiroz, J. (2005). Assessment of equivalence using a concordance correlation coefficient in a repeated measurement design. *Journal of Biopharmaceutical Statistics* 15:913-928.

Table 1. Estimates of population means, within-subject and between-subject variabilities and intra-ICC for the four data examples.

	μ_j	σ^2_{Wj}	σ^2_{Bj}	Intra ICC (ρ^I_j)	Rep. Coef.
Goniometers (Eliasziw, et al., 1994)					
Manual Goniometer	1.437	0.736	53.8	0.986	2.37
Electro-goniometer	0.046	0.977	53.8	0.982	2.74
Calcium Scoring (Haber, et al., 2005)					
Radiologist A	35.833	7.667	1025.7	0.993	7.67
Radiologist B	36.125	0.125	1116.2	0.999	0.98
Carotid Stenosis (Barnhart et al., 2005)					
Left MRA-2D	43.7	576.7	966.5	0.626	66.5
Left MRA-3D	48.2	520.2	953.7	0.647	63.2
Left IA	38.0	139.7	1061.2	0.884	32.8
Right MRA-2D	45.9	568.5	887.7	0.610	66.0
Right MRA-3D	43.9	550.0	903.6	0.622	64.7
Right IA	33.8	88.0	965.2	0.916	26.0
Systolic Blood Pressure (Bland and Altman, 1999)					
Observer 1	127.4	37.4	936.0	0.962	17.0
Observer 2	127.3	38.0	917.1	0.960	17.0
Machine	143.0	83.1	983.2	0.922	25.3

Table 2. Estimates of CCC, CIA and other parameters from the four data examples.

	CCC	CIA	Inter-CCC	Intra-ICC	True Corr.	W. Average
	ρ_c	ψ^N	$\rho_c(\mu)$	mean ρ^I	ρ_μ	$d = \frac{\sigma_B^2}{\sigma_W^2}$
Goniometers (Eliasziw, et al., 1994)						
Manual vs. Electro Goniometer	0.944	0.287	0.959	0.984	0.977	61.4
Calcium Scoring (Haber, et al., 2005)						
Radiologist A vs. B	0.995	0.754	0.998	0.996	0.999	274.9
Carotid Stenosis (Barnhart et al., 2005)						
Left: MRA-2D vs. MRA-3D	0.589	0.881	0.919	0.637	0.929	1.75
Right: MRA-2D vs. MRA-3D	0.579	0.917	0.939	0.616	0.941	1.60
Systolic Blood Pressure (Bland and Altman, 1999)						
Observer 1 vs. 2	0.973	1.0	1.0	0.961	1.0	24.6
Observer 1 vs. Machine	0.701	0.178	0.740	0.942	0.834	15.9
Ovserver 2 vs. Machine	0.700	0.179	0.739	0.941	0.836	15.7

Table 3. Estimates of CCC and CIA for multiple methods with one reference.

	CCC ρ_c^R	CIA ψ^R	Inter-CCC $\rho_c(\mu)^R$
Goniometers (Eliasziw, et al., 1994)			
Manual* vs. Electro Goniometer	0.944	0.246	0.959
Carotid Stenosis (Barnhart et al., 2005)			
Left: MRA-2D and MRA-3D vs. IA*	0.514	0.209	0.683
Right: MRA-2D and MRA-3D vs. IA*	0.607	0.103	0.805
Systolic Blood Pressure (Bland and Altman, 1999)			
Observer 1* vs. Machine	0.701	0.117	0.740
Observer 2* vs. Machine	0.700	0.117	0.739

* The method is used as reference.

Figure 1. CCC and CIA as functions of d

Panel (a). $(\mu_1 - \mu_2)^2 = 2\sigma_W^2 = 9, \rho_\mu = 1$

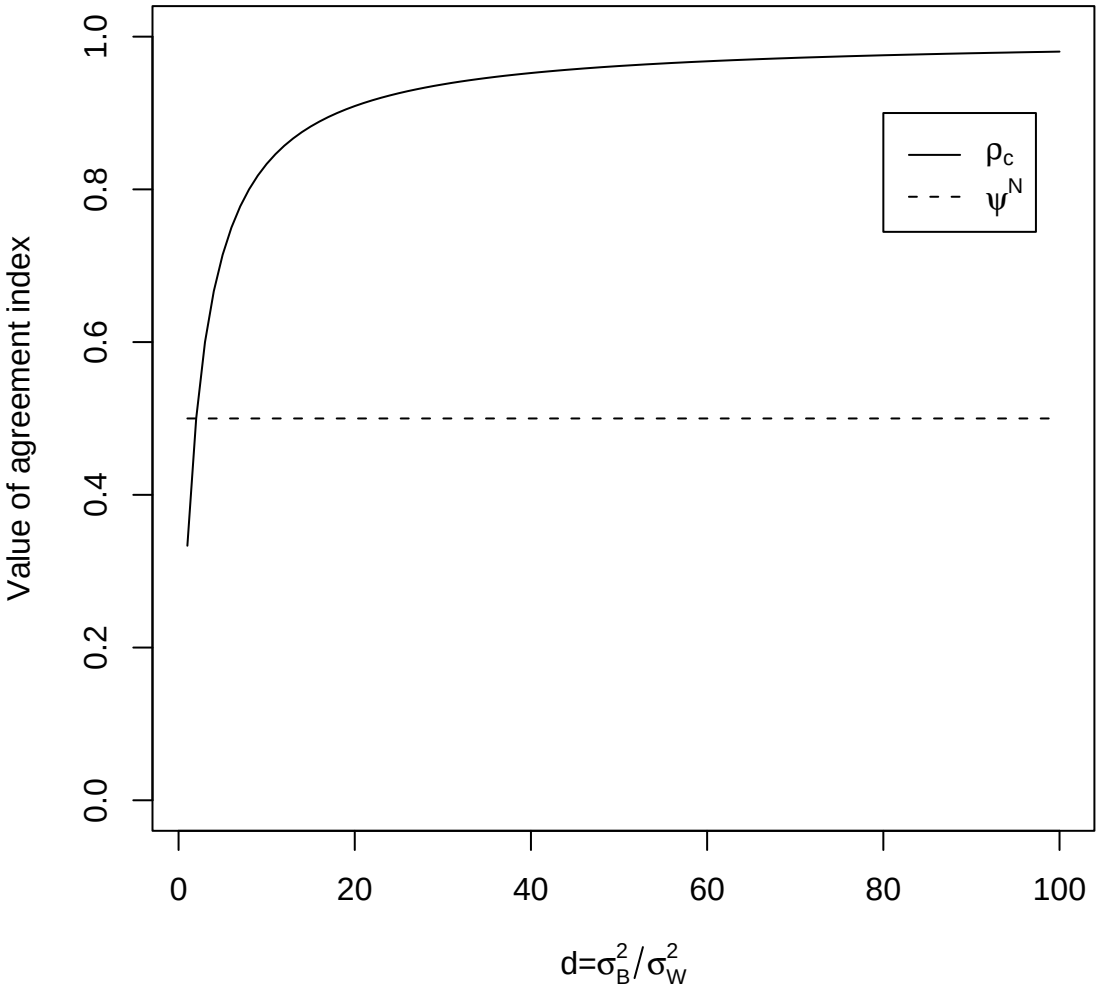

Figure 1. CCC and CIA as functions of d

Panel (b). $(\mu_1 - \mu_2)^2 = 9, \sigma_W^2 = d, \rho_\mu = 1$

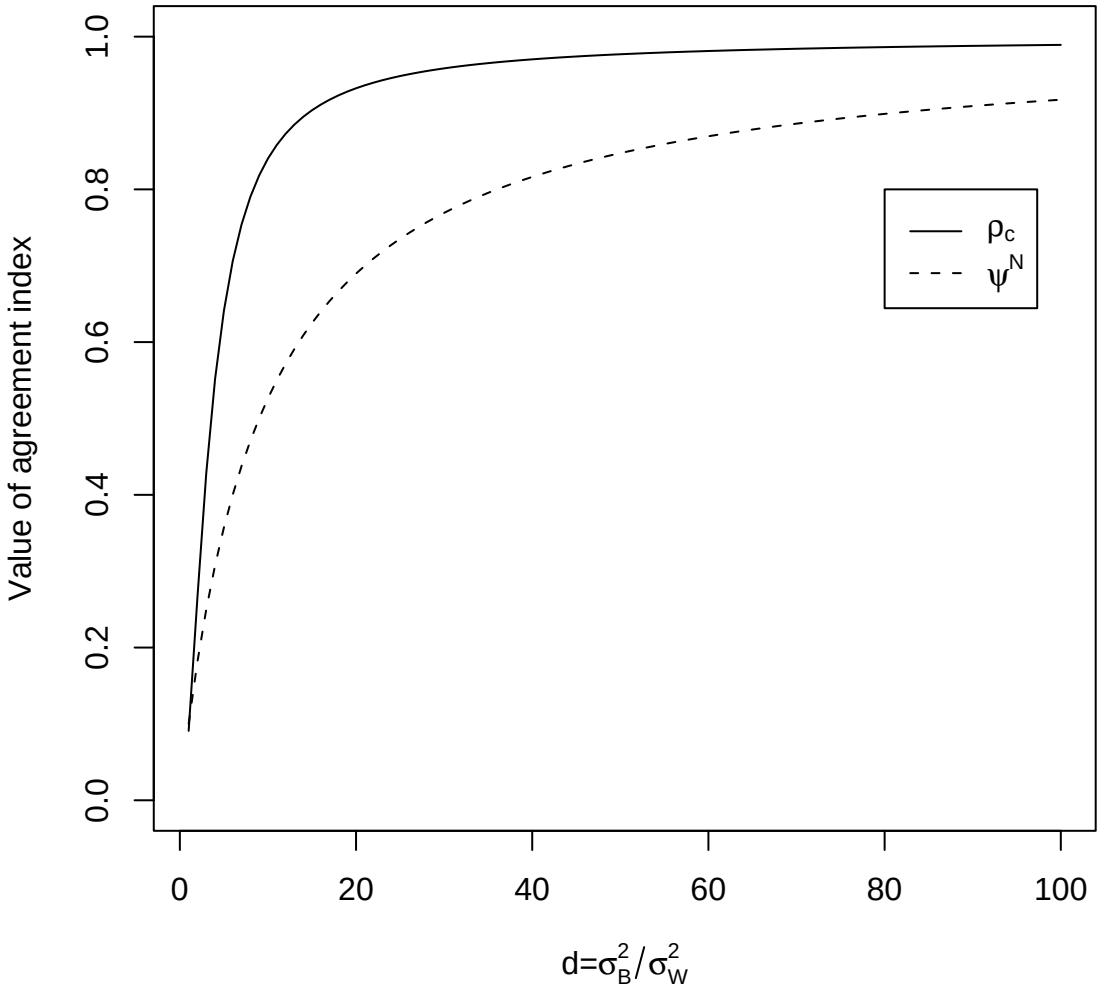

Figure 1. CCC and CIA as functions of d

Panel (c). $(\mu_1 - \mu_2) = 0, \sigma_W^2 = d, \rho_\mu = 1$

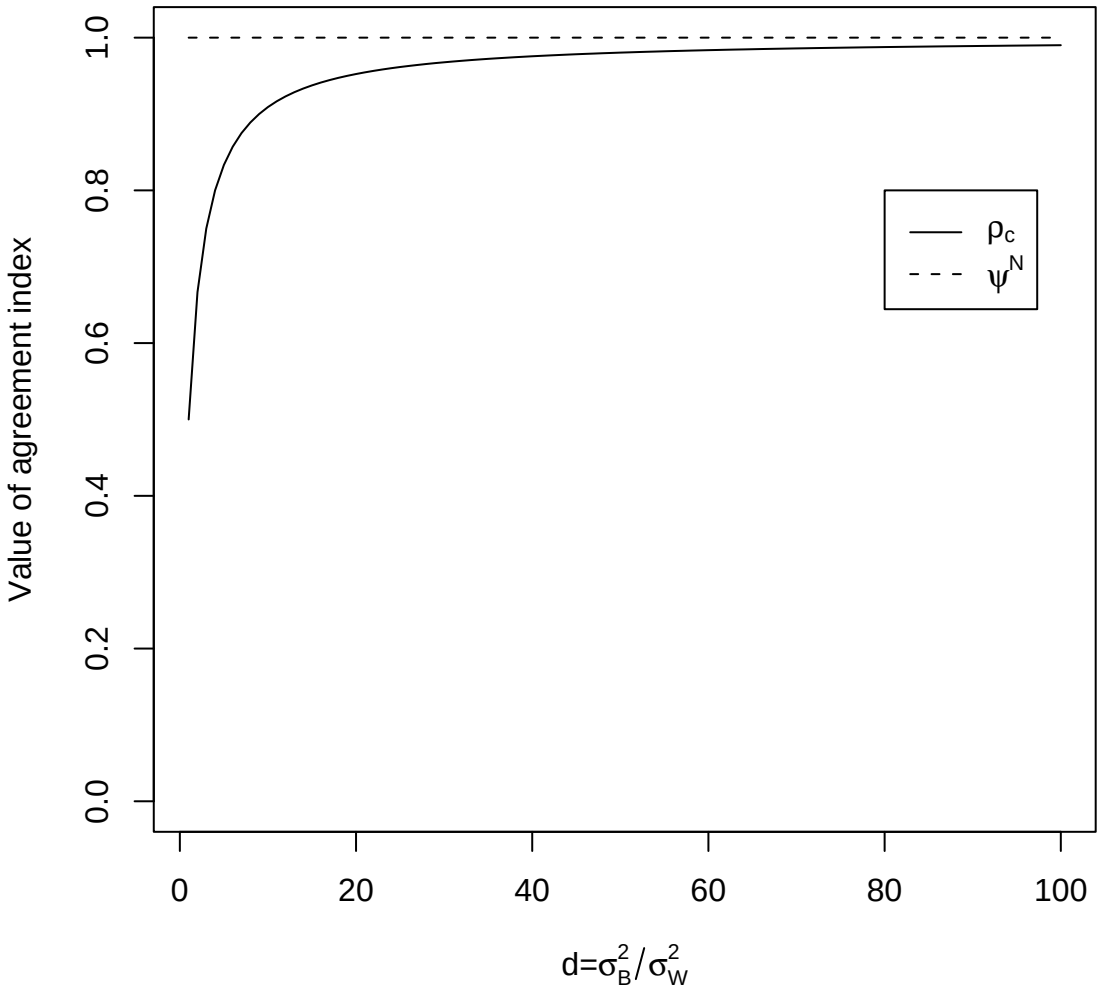

Figure 1. CCC and CIA as functions of d

Panel (d). $(\mu_1 - \mu_2)^2 = 2\sigma_W^2 = 9, \rho_\mu = 0.8$

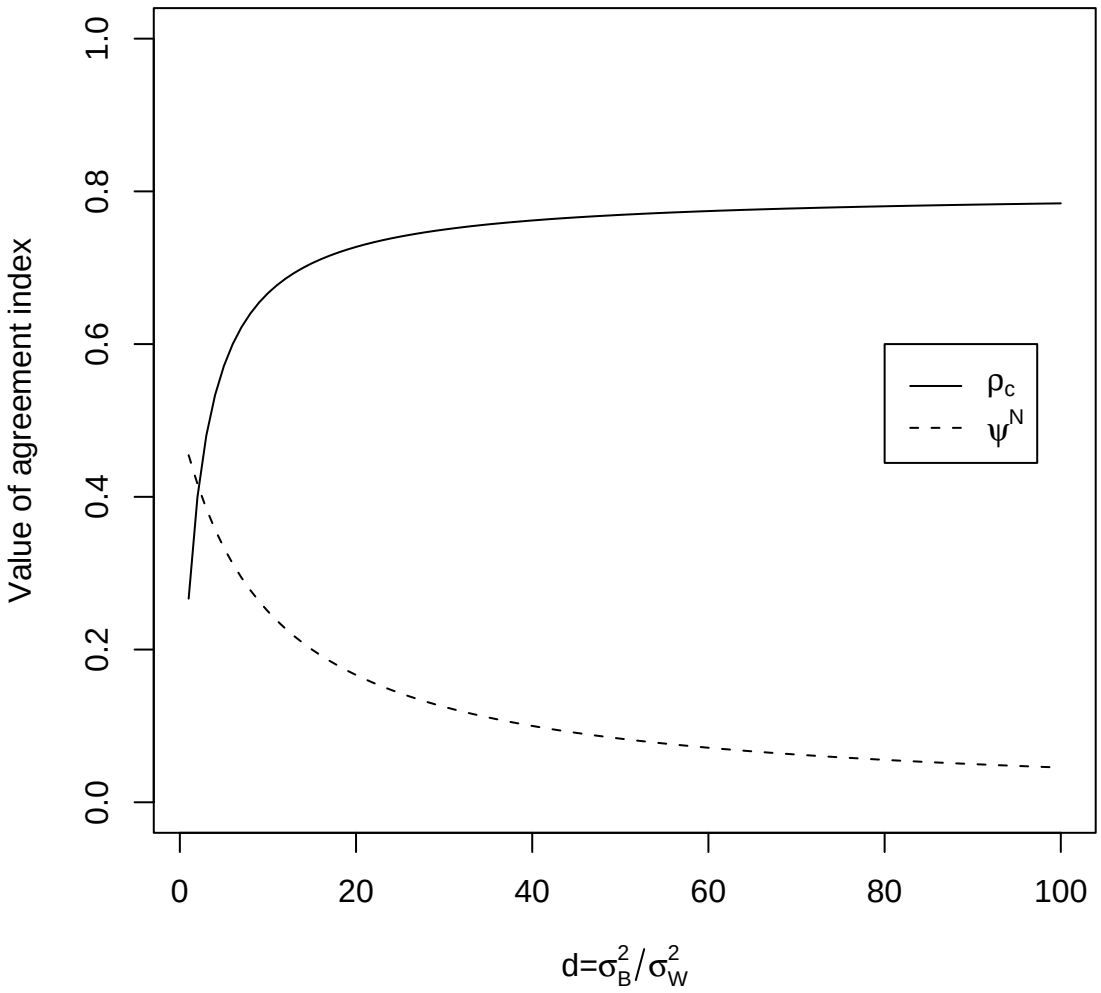

Figure 1. CCC and CIA as functions of d

Panel (e). $(\mu_1 - \mu_2)^2 = 9, \sigma_W^2 = d, \rho_\mu = d/(d+1)$

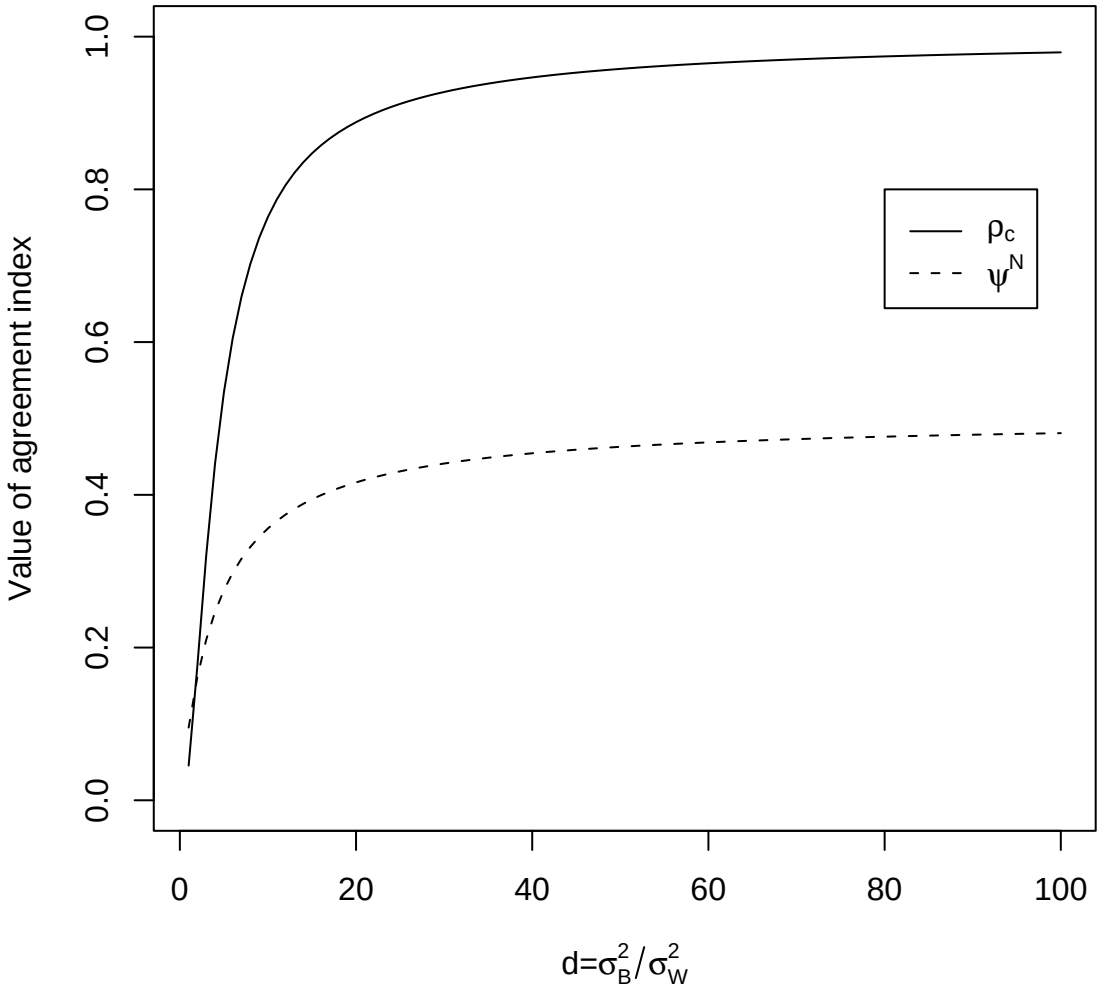

Figure 1. CCC and CIA as functions of d

Panel (f). $\mu_1 - \mu_2 = 0, \sigma_W^2 = d, \rho_\mu = d/(d+1)$

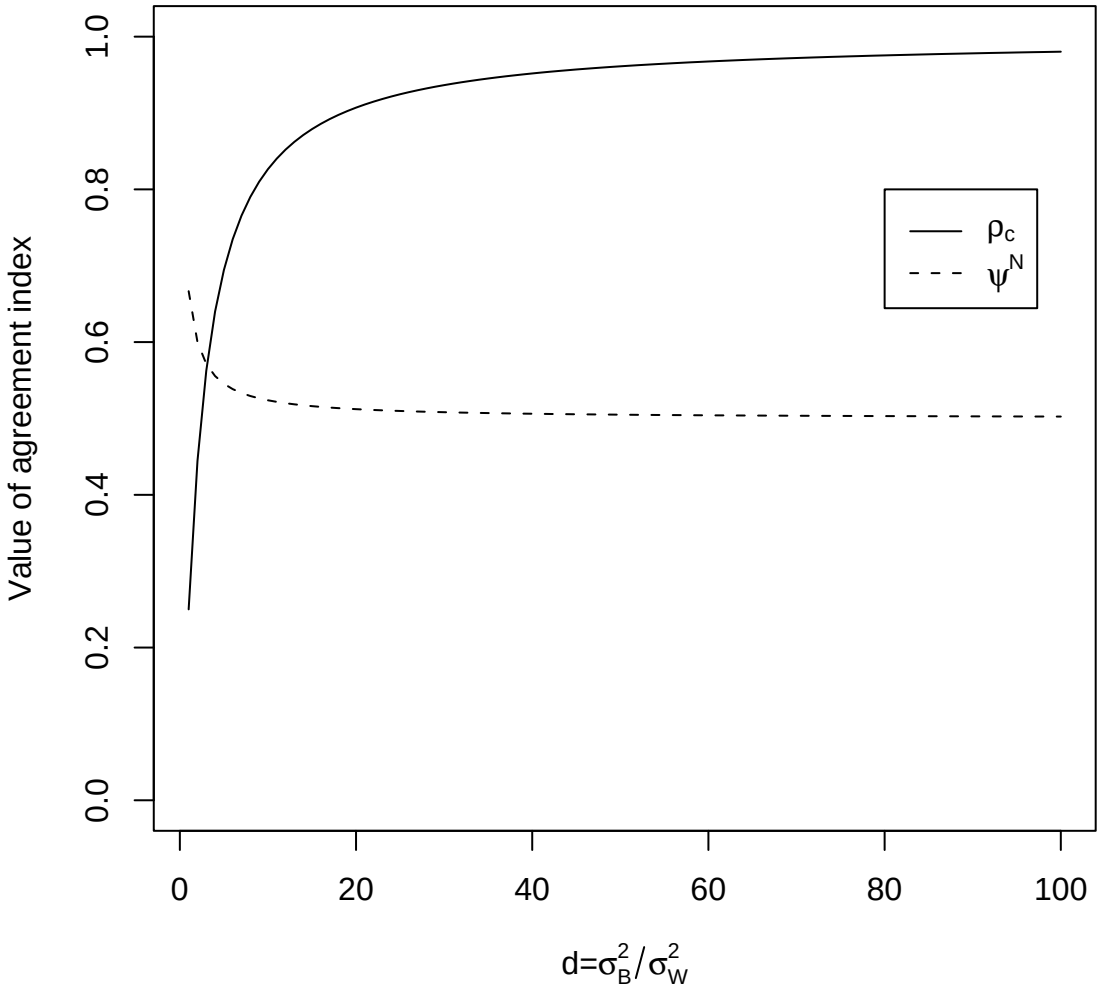

Figure 2. CIA as a function of CCC for fixed $d = \frac{\sigma_B^2}{\sigma_W^2}$

Panel (a), $(\mu_1 - \mu_2)^2 = 2\sigma_W^2 = 9$

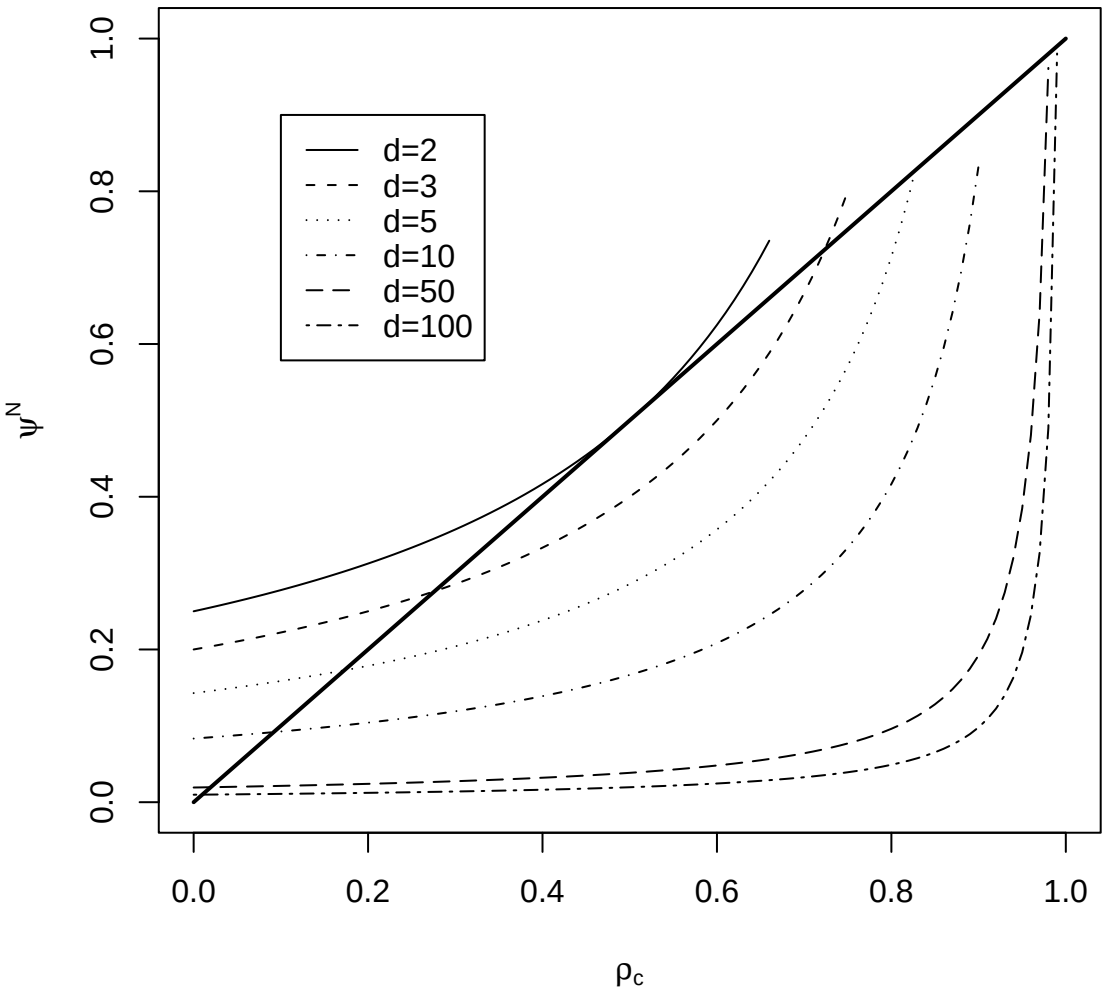

Figure 2. CIA as a function of CCC for fixed $d = \frac{\sigma_B^2}{\sigma_W^2}$

Panel (b), $(\mu_1 - \mu_2)^2 = 9, \sigma_W^2 = d$

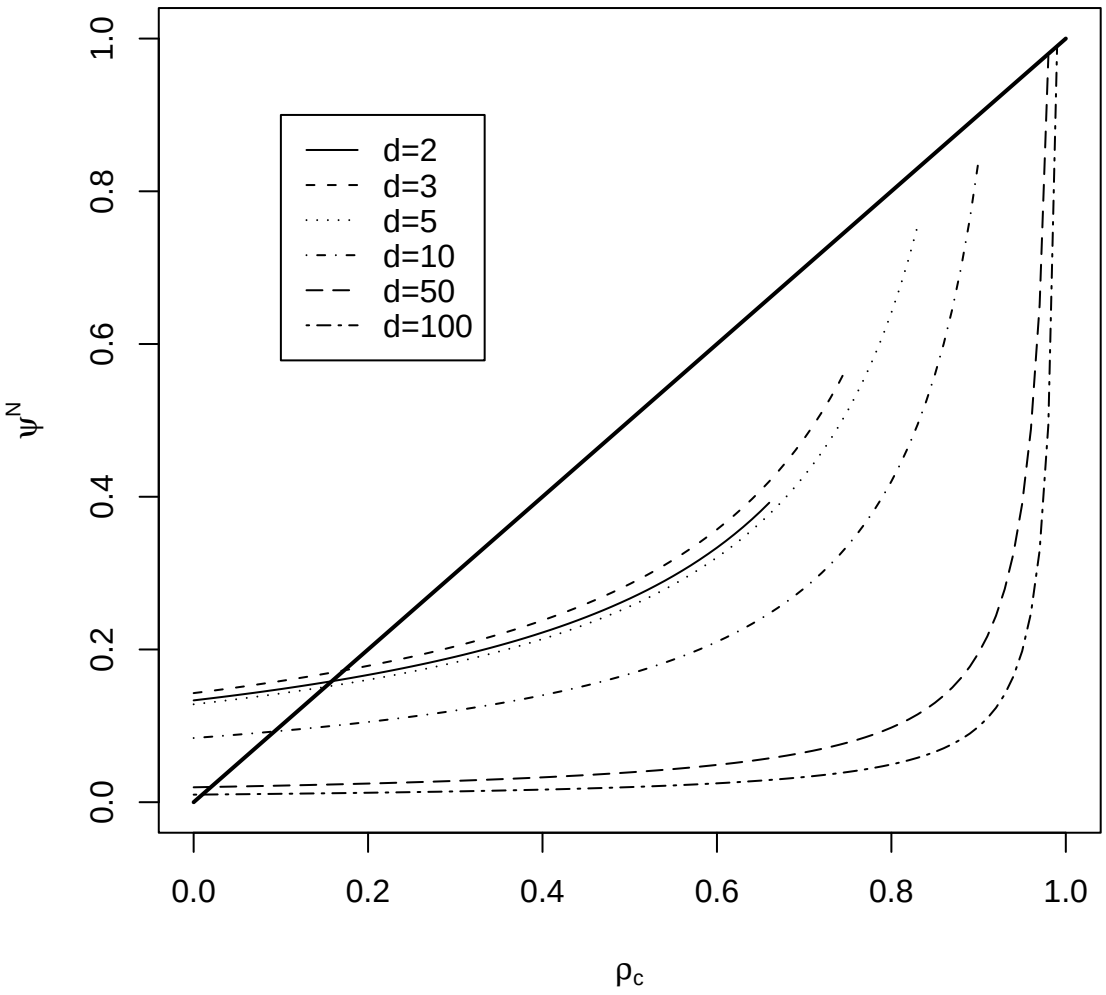

Figure 3. CIA as a function of CCC for fixed $d = \frac{\sigma_B^2}{\sigma_W^2}$ and $\mu_1 - \mu_2 = 0$