Lin, Heydayat, Wu

Inter Intra, and Total CCC

Comparing this paper with Barnhart, Song & Haber (SIM 2005)

	This paper	Barnhart et al
Model	$Y_{ijl} = \mu + \alpha_i + \beta_j + \gamma_{ij} + e_{ijl}$	$Y_{ijl} = \mu_{ij} + e_{ijl}$
Assump-tions	Same interaction and error variances for all observers. Equal correlations for all pairs of observers Independence of all random effects	Independence of means and errors
		Under ANOVA model
Intra MSD	$E(Y_{ijl} - Y_{ijl'})^2$	$E(Y_{ijl} - Y_{ijl'})^2$
Intra CCC	$(\sigma_\alpha^2 + \sigma_\gamma^2)/(\sigma_\alpha^2 + \sigma_\gamma^2 + \sigma_e^2)$	$\delta_j^2/(\delta_j^2 + \sigma_{ej}^2)$ *
Inter MSD	$E(Y_{ij\bullet} - Y_{ij'\bullet})^2$	$E(\mu_{ij} - \mu_{ij'})^2$
Inter CCC	$\sigma_\alpha^2/(\sigma_\alpha^2 + \sigma_\beta^2 + \sigma_\gamma^2 + \sigma_e^2/m)$	$\sigma_\alpha^2/(\sigma_\alpha^2 + \sigma_\beta^2 + \sigma_\gamma^2)$
Total MSD	$E(Y_{ijl} - Y_{ij'l'})^2$	$E(Y_{ijl} - Y_{ij'l'})^2$
Total CCC	$\sigma_\alpha^2/(\sigma_\alpha^2 + \sigma_\beta^2 + \sigma_\gamma^2 + \sigma_e^2)$	$\sigma_\alpha^2/(\sigma_\alpha^2 + \sigma_\beta^2 + \sigma_\gamma^2 + \sigma_e^2$

* δ_j^2 is the between-subjects variability for observer *j*.

$\sigma_j^2 = \delta_j^2 + \sigma_{ej}^2$

In BSH, each observer has its own intra-CCC

Should inter-CCC be based on the observed ‘within’ means or on the true means?

Should the it depend on σ_e^2?

Should it depend on the number of replications

<u>Other issues</u>

Can we indeed get replicated observations?

Matched vs. unmatched replications

Agreement by chance = independence?

Example

Three methods for assessing carotid stenosis: angiogram (IA), 2D-MRA, 3D-MRA. The same three raters measured each patient using each of the three methods.

Inter CCC		LHW Left	BSH Left		LHW Right	BSH Right
Overall		0.668	0.763		0.742	0.848
IA vs 2D		0.675	0.755		0.761	0.846
IA vs 3D		0.556	0.624		0.689	0.765
2D vs 3D		0.773	0.925		0.778	0.943
Intra CCC						
Overall		0.702			0.691	
IA			0.882			0.915
2D			0.621			0.604
3D			0.614			0.616

Barnhart, Kosinski, Haber

ICC and CIA

$$CCC = 1 - \frac{E(Y_{i1} - Y_{i2})^2}{E\{(Y_{i1} - Y_{i2})^2 \mid \text{chance agreement}\}}$$

Does ‘chance agreement = independence’?

- Independence and (lack of) agreement are different things
- Observers are independent only if the measurements of at lest one of them is independent of the true value
- So, what is ‘agreement by chance’?

CIA is based on comparing the disagreement $MSD(Y_{1,}Y_2) = E(Y_{i1} - Y_{i2})^2$ to its expected value under 'disagreement by chance', i.e. when observers are equivalent, or interchangeable. It can be estimated from $MSD(Y_{jk}, Y_{jk'})$ based on replicated readings by observer *j*.

When observer 1 is the reference, use $MSD(Y_{1k}, Y_{1k'})$. When there is no reference, use the mean of $MSD(Y_{jk}, Y_{jk'})$ over all observers.

No modification of CCC when comparing to a reference

Comparison between CCC and CIA^N in terms of the ratio of the between-subjects and within-observers variabilities: $d = \frac{\sigma_B^2}{\sigma_W^2}$

$$CCC = \frac{\rho_\mu}{\frac{1}{2}(\mu_1 - \mu_2)^2 + \frac{1}{d} + 1}$$

$$CIA^N = \frac{1}{\frac{1}{2}(\mu_1 - \mu_2)^2 + (1 - \rho_\mu) \cdot d + 1}$$

Both increase with ρ_μ and decrease with $(\mu_1 - \mu_2)^2$

CCC increases when the between-subjects variability is higher

CIA increases when within subjects × observers variability is higher, i.e. it is more difficult to obtain a precise reading.

Biswas

Diagnostic Agreement

Imperfect reference test = R

New test = T

Observed two by two table:

	R+	**R-**	
T+	a	b	a+b
T-	c	d	c+d
	a+c	b+d	

PPA = Positive proportion of agreement = a/(a+c)

NPA = Negative percent agreement = d/(b+d)

D = unobserved disease status

Unobserved table:

	D=1			D=0		
	R=1	**R=0**		**R=1**	**R=0**	
T=1	a_1	b_1		a_0	b_0	
T=0	c_1	d_1		c_0	d_0	

$$Sn(R) = P(R = 1 \mid D = 1), \quad Sp(R) = P(R = 0 \mid D = 0)$$

$$Sn(T) = P(T = 1 \mid D = 1), \quad Sp(T) = P(T = 0 \mid D = 0)$$

We assume conditional independence of T and R given D.

Let $\pi = P(D)$ denote the prevalence of the disease.

Then $PPA = P(T = 1 \mid R = 1) =$

$P(T = 1 \mid R = 1,\ D = 1) \cdot P(D = 1 \mid R = 1) +$

$P(T = 1 \mid R = 1,\ D = 0) \cdot P(D = 0 \mid R = 1) =$

$[Sn(T) \cdot Sn(R) \cdot \pi + (1 - Sp(T)) \cdot (1 - Sp(R)) \cdot (1 - \pi)] / P(R = 1)$

Similarly $NPA = P(T = 0 \mid R = 0) =$

$[(1 - Sn(T)) \cdot (1 - Sn(R)) \cdot \pi + Sp(T) \cdot Sp(R) \cdot (1 - \pi)] / P(R = 0)$

In many cases π, $Sn(R)$, $Sp(R)$ are (approximately) known. PPA, NPA and $P(R = 1)$ can be estimated from the observed frequencies. Then one can estimate the sensitivity and specificity of the new test.

An artificial example

	R=1	**R=0**	
T=1	81	59	140
T=0	39	121	160
	120	180	300

Here PPA=0.675, NPA=0.672, P(R=1)=0.4

Suppose it is known that the prevalence is 1/3, and the sensitivity and specificity of the reference test are 0.80 and 0.85, respectively.

Then by solving the above equations we estimate that the new test has a sensitivity of 0.80 and a specificity of 0.70.