

Research Article

Open Access

Comparison of Dependent Pearson and Spearman Correlation Coefficients with and without Correction for Measurement Error

Bernard Rosner^{1,4*}, Wei Wang², Heather Eliassen^{3,4} and Eileen Hibert⁴

¹Department of Biostatistics, Harvard School of Public Health, Boston, MA 02115, USA ²Department of Medicine, Harvard Medical School, Brigham and Women's Hospital, Boston, MA 02115, USA ³Department of Nutrition and Epidemiology, Harvard School of Public Health, Boston, MA 02115, USA ⁴Channing Division of Network Medicine, Harvard Medical School, Boston, MA 02115, USA

Abstract

There already exist methods for comparing dependent Pearson correlation coefficients. However, each of the variables (X, Y) has associated random error; and a related question is after correcting for random error, which variable correlates most highly with the outcome variable Z. In this paper, we present methods for comparing dependent deattenuated correlation coefficients. This is a generalization of previous work for obtaining confidence limits for a single deattenuated correlation coefficient. In addition, we extend this work to the comparison of dependent Spearman correlation coefficients. The methods are illustrated with two examples. The first example concerns the comparison of nephrotoxicity of phenacetin and aspirin intake as measured by repeat biomarkers obtained from the same subjects. The second example is a comparison of the validity of different storage conditions for measuring HbA1c from dried blood specimens as compared to the gold standard of immediate processing. Results from using these methods indicate that phenacetin intake is more highly correlated with serum creatinine levels than aspirin intake and that short-term storage is preferable to long-term storage for assessment of HbA1c levels. We have available SAS software for comparing dependent deattenuated Pearson correlation and dependent Spearman correlations with and without deattentuation.

Keywords: Deattenuated correlation; Dependent correlations; Measurement error correction; Pearson correlation; Spearman correlation

Introduction

In 1967/68 a group of 1,256 women were recruited for the Swiss Analgesic Study [1,2]. Women were age 30-49 years and lived in the Basel, Switzerland area. The purpose of the study was to investigate whether there was an association between intake of phenacetincontaining analgesics and prevalence and incidence of kidney disease. To measure phenacetin intake a urinary metabolite (N-acetyl-Paminophenol (NAPAP)) was assayed from a urine sample in the clinic and from 2 additional urine samples collected at home on 2 separate days within 1 week of the clinic visit. Serum creatinine was used as a measure of kidney function. Women were seen at follow-up visits in 1969, 1970, 1971, 1972, 1975 and 1978. In the primary paper, mean NAPAP over the 3 replicates was categorized and related to both prevalence and incidence of kidney disease where an abnormal creatinine was defined as >1.5 mg/dl. In the present paper, we represent NAPAP and serum creatinine at the baseline clinic visit in 1968 as continuous variables and compute the correlation coefficient between these two measures. To distinguish phenacetin intake from general use of analgesics, urinary salicylates were also assessed in triplicate at baseline as an indication of intake of aspirin-containing analgesics. If we let X=NAPAP at the baseline clinic visit, Y=salicylates at the baseline clinic visit and Z=serum creatinine at the baseline clinic visit, then we wish to test the hypothesis $H_0: \rho_{xz} = \rho_{yz} \text{ vs } H_1: \rho_{xz} \neq \rho_{yz}$ (1)

where ρ_{xz} is the population correlation between X and Z and ρ_{yz} is defined similarly.

There already exist methods for comparing dependent correlation coefficients for normally distributed random variables obtained from the same subjects. In an excellent review paper on this subject [3] it is noted that in previous simulation studies, several methods are clearly inappropriate, while other methods preserve type I error under a variety of conditions. Among the latter are the method of Williams [4] which is an enhancement of a procedure initially proposed by Hotelling [5]. Hotelling's procedure is exact, but only under the condition that the sample and population standard deviations are the same for both X and Y (i.e., $s_x = \sigma_x$, $s_y = \sigma_y$). To test the hypotheses in (1) Williams proposes the test statistic

$$T_{1} = \left(R_{xz} - R_{yz}\right) \sqrt{\left[\frac{2(N-1)\left(1 + R_{xy}\right)}{\left[\frac{2(N-1)}{N-3}\right]}|R| + \overline{R}_{z}^{2}\left(1 - R_{xy}\right)^{3}} \sim t_{N-3} \text{ under } H_{0}$$

Where $R_{xz} = corr(x,z)$, R_{xy} and R_{yz} are defined similarly, N=sample size an

$$|R| = 1 - R_{xz}^2 - R_{yz}^2 - R_{xy}^2 + 2R_{xz}R_{yz}R_{xy}, \overline{R}_{z} = (R_{xz} + R_{yz})/2$$

and we reject H_0 if $|T_1| > t_{N-3,1-\alpha/2}$ and $t_{N-3,p} = p^{th}$ percentile of a t_{N-3} distribution.

In addition, Olkin and Finn [6] used a similar approach to compare more than 2 correlated correlation coefficients. Since in finite samples, the Fisher z transformation better approximates a normal distribution than the sample correlation coefficient, Dunn and Clark [7] proposed the test statistic

$$Z_{1}^{*} = \sqrt{N-3} \left(Z_{xz} - Z_{yz} \right) / \left(2 - 2s_{xz,yz} \right)^{1/2} \sim N(0,1) \text{ under } H_{0}$$

where

*Corresponding author: Bernard Rosner, Channing Division of Network Medicine, Harvard Medical School, Boston, MA 02115, USA, Tel: 617-525-2743; Fax: 617-731-1541; E-mail: stbar@channing.harvard.edu

Received April 01, 2015; Accepted May 22, 2015; Published May 29, 2015

Citation: Rosner B, Wang W, Eliassen H, Hibert E (2015) Comparison of Dependent Pearson and Spearman Correlation Coefficients with and without Correction for Measurement Error. J Biom Biostat 6: 226. doi:10.4172/2155-6180.1000226

Copyright: © 2015 Rosner B, et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Page 2 of 9

$$s_{xz,yz} = \frac{R_{xy} \left(1 - R_{xz}^2 - R_{yz}^2\right) - 0.5 \left(R_{xz}R_{yz}\right) \left(1 - R_{xz}^2 - R_{yz}^2 - R_{xy}^2\right)}{\left(1 - R_{xz}^2\right) \left(1 - R_{yz}^2\right)}$$

$$\begin{split} &Z_{\rm xz} {=} 0.5 \ ln \ [(1{+}\ {\rm R_{xz}})/\ (1{-}\ {\rm R_{xz}})], \ Z_{\rm yz} \ {\rm is \ defined \ similarly, \ and \ we \ reject} \\ &H_0 \ {\rm if \ }|Z_{\rm i}^*| {>} \ z_{\rm _{1,\underline{\alpha}}} \ , \ {\rm where \ } z_{\rm _p} {=} p^{\rm th} \ {\rm percentile \ of \ a \ N(0,1) \ distribution.} \end{split}$$

Furthermore, Meng, Rosenthal, and Rubin [8] propose a statistic that is asymptotically equivalent to the Dunn and Clark procedure, but with a simpler expression for $var(Z_{yz} - Z_{yz})$ given by:

$$Z_{xy} = \left(Z_{xz} - Z_{yz}\right) \sqrt{\frac{N-3}{2(1-R_{xy})h}} \sim N(0,1) \text{ under } H_0$$

where
$$h = \left(1 - f\overline{R}^2\right) / \left(1 - \overline{R}^2\right),$$

and
$$f = \min\left\{\left(1 - R_{xy}\right) / \left[2\left(1 - \overline{R}^2\right)\right], 1\right\} \text{ and } \overline{R}^2 = \left(R_{xz}^2 + R_{yz}^2\right) / 2$$

In addition Billion Remains and Comp[0]

In addition, Bilker, Brensinger, and Gur [9] propose a combined permutation-bootstrap approach to test the hypothesis in (1) based on the test statistic

 $Z_{xy} = (Z_{xz} - Z_{yz}) / \{var(Z_{xz} - Z_{yz})\}^{1/2}$

which is computationally intensive, but avoids the necessity of specifying the underlying distribution of Z_{xy} . Finally, the R package [10] cocor maintained by Diedenhofen [11] provides software to implement a variety of methods for comparing dependent Pearson correlation coefficients.

However, an assumption of [4-8] is that the joint distribution of (X, Y, Z) is multivariate normal. We relax this assumption and use the method of moments and the delta method to estimate $var(Z_{xz} - Z_{yz})$. Furthermore, in the Swiss Analgesic dataset, there is considerable intraindividual variation among replicate measures in both X and Y and the more important question is whether the true (underlying) mean value of X and Y at baseline, denoted by μ_x and μ_y are more highly correlated with Z. Thus, we wish to test the hypothesis

$$H_0: corr(\mu_X, Z) = corr(\mu_Y, Z) vs H_1: corr(\mu_X, Z) \neq corr(\mu_Y, Z).$$

This is an extension of previous work [12] that provides confidence limits for a single deattenuated correlation. Another issue is that it is clear that neither X nor Y is normally distributed. Hence, the Spearman correlation coefficient may be a more appropriate measure of association than the Pearson correlation coefficient in this dataset. Hence, we extend our methodology to the comparison of dependent Spearman correlation coefficients both with and without correction for measurement error. To our knowledge, there is no previous literature on the comparison of dependent Spearman correlation coefficients.

In this paper, we first describe the methodology and propose both an asymptotic test and an exact test. Second, we present a simulation study to assess the validity of the asymptotic test in finite samples and compare it to existing procedures for comparing dependent correlation coefficients. Finally, we describe two examples based on real data illustrating the use of these methods.

Comparison of Dependent Deattenuated Pearson Correlations

Hypothesis testing

We consider a classical measurement error model for X_i and Y_i of

the form

$$\begin{aligned} x_{ij} &= \alpha_x + \mu_{X_i} + e_{x_{ij}} & i = 1, \dots, n_1, j = 1, \dots, k \\ y_{ij} &= \alpha_y + \mu_{Y_i} + e_{y_{ij}} & i = 1, \dots, n_1, j = 1, \dots, k \end{aligned}$$

where

 $var(\mu_{x_i}) = \sigma_{x_x}^2, var(e_{x_y}) = \sigma_{ex}^2, var(\mu_{y_i}) = \sigma_{x_y}^2, var(e_{y_y}) = \sigma_{ey}^2, (\mu_{x_i}, e_{x_y}) and (\mu_{y_i}, e_{y_y}) are respectively independent, <math>corr(e_{x_y}, e_{y_y}) \neq 0$ if replicate measures of X and Y are obtained from the same subjects.

We define the true correlation between (*X* and *Z*) and (*Y* and *Z*) by

$$\rho_{xz,true} = corr(\mu_{X_i}, Z_i) \tag{3}$$

$$\rho_{yz,true} = corr(\mu_{Y_i}, Z_i)$$

Based on equations (2) and (3), it is straightforward to show that

$$\rho_{xz,true} = \frac{\rho_{xz}}{\sqrt{ICC_x}}$$

$$\rho_{yz,true} = \frac{\rho_{yz}}{\sqrt{ICC_y}}$$
(4)

where $\rho_{xz} = corr(X_{ij}, Z_i)\rho_{yz} = corr(Y_{ij}, Z_i)$

and

 $ICC_{x} = corr(X_{ij1}, X_{ij2}) = \sigma_{Ax}^{2} / (\sigma_{Ax}^{2} + \sigma_{ex}^{2}), ICC_{y} = corr(Y_{ij1}, Y_{ij2}) = \sigma_{Ay}^{2} / (\sigma_{Ay}^{2} + \sigma_{ey}^{2}) =$ the intra class correlations obtained from repeated measures of *X* and *Y* based on (2). We wish to test the hypothesis:

 $H_0: \rho_{xz,true} = \rho_{yz,true} \operatorname{vs} H_1: \rho_{xz,true} \neq \rho_{yz,true}$

Note that it is possible that $\rho_{xz} = \rho_{yz}$, while $\rho_{xz,true} > \rho_{yz,true}$ if the reproducibility of *X* is worse than that of *Y* (i.e., *ICC_x*<*ICC_y*). Similarly, it is possible that $\rho_{xz} < \rho_{yz}$, while $\rho_{xz,true} = \rho_{yz,true}$.

It is usually more efficient to base inferences for correlations based on Fisher's z statistic. Thus, we will test the hypothesis

$$H_0: Z_{xz,true} = Z_{yz,true} \text{ vs } H_1: Z_{xz,true} \neq Z_{yz,true}$$
 Where

$$Z_{xz,true} = \frac{1}{2} \ln \left[\left(1 + \rho_{xz,true} \right) / \left(1 - \rho_{xz,true} \right) \right] \text{ and } Z_{yz,true} \text{ is defined similarly.}$$

We will use the delta method to estimate

$$var(\hat{\Delta}_{z,true}) \equiv var(\hat{Z}_{xz,true} - \hat{Z}_{yz,true})$$

and obtain the test statistic

$$V_{xy,true} = \frac{\hat{\Delta}_{z,true}}{\sqrt{var(\hat{\Delta}_{z,true})}} \sim N(0,1) \text{ for large } N.$$
(5)

We have:

$$var(\hat{\Delta}_{z,true}) = var(\hat{Z}_{xz,true}) + var(\hat{Z}_{yz,true}) - 2cov(\hat{Z}_{xz,true}, \hat{Z}_{yz,true})$$
(6)
Where

 $var(\hat{Z}_{xz,true}), var(\hat{Z}_{yz,true})$ and $cov(\hat{Z}_{xz,true}, \hat{Z}_{yz,true})$ are derived in Appendices A, B and C of the supplementary materials, respectively.

Therefore, from Appendices A, B and C of the supplementary materials, we obtain $var(\hat{\Delta}_{z,true})$ and the test statistic $V_{xy,true}$ in (5), with asymptotic p-value

= 2 x $\left[1 - \Phi(|V_{xy,true}|)\right]$, where Φ =standard normal c.d.f.

Similarly, we can also test the hypothesis $H_0: \rho_{xz} = \rho_{yz}$ vs $H_1: \rho_{xz} \neq \rho_{yz}$ based on the large sample test statistic

$$V_{xy} = \frac{\Delta_z}{\sqrt{var(\hat{\Delta}_z)}} \sim N(0,1) \text{ for large N}$$
(7)

where $\hat{\Delta}_z = \hat{Z}_{xz} - \hat{Z}_{yz}$ and based on the delta method

$$var(\hat{\Delta}_{z}) = var(\hat{Z}_{xz} - \hat{Z}_{yz}) = var(\hat{Z}_{xz}) + var(\hat{Z}_{yz}) - 2 cov(\hat{Z}_{xz}, \hat{Z}_{yz})$$
(8)
$$= \frac{R_{xz}^{2}}{\left(1 - R_{xz}^{2}\right)^{2}} [var(\ln R_{xz})] + \frac{R_{yz}^{2}}{\left(1 - R_{yz}^{2}\right)^{2}} var[\ln(R_{yz})]$$
$$-2 \frac{R_{xz}R_{yz}}{\left(1 - R_{xz}^{2}\right)\left(1 - R_{yz}^{2}\right)} cov[\ln(R_{xz}), \ln(R_{yz})]$$

and

 $var[(\ln(R_{xz})], var[(\ln(R_{yz})]] and cov[\ln(R_{xz}), \ln(R_{yz})] are given in Appendices A, B and C, respectively of the supplementary materials.$

Interval estimation

It is also of interest to obtain confidence limits for $\Delta_{\rho,true} = \rho_{xz,true} - \rho_{yz,true}$. It is not possible to translate confidence limits for $\Delta_{Z,true}$ to corresponding confidence limits for $\Delta_{\rho,true}$. Instead, we will use $\Delta_{R,true} \equiv R_{xz} / \sqrt{\widehat{ICC_x}} - R_{yz} / \sqrt{\widehat{ICC_y}} \equiv R_{xz,true} - R_{yz,true}$ as a point estimate of $\Delta_{\rho,true}$. To obtain $\operatorname{var}(\Delta_{R,true})$ we use the formula:

$$var(\Delta_{R,true}) = var(R_{xz,true}) + var(R_{yz,true}) - 2cov(R_{xz,true}, R_{yz,true})$$
(9)

Since $R_{xz,true}$ is a ratio estimator, we first consider $var[ln(R_{xz,true})]$ and use the delta method to obtain

$$var(R_{xz,true}) \cong R_{xz,true}^2 var\left[\ln(R_{xz,true})\right]$$
(10)

Furthermore, from the delta method,

$$var\left[\ln\left(R_{x_{z},true}\right)\right] = var\left[\ln\left(R_{x_{z}}\right)\right] + \frac{1}{4}var\left[\ln\left(\widehat{ICC}_{x}\right)\right]$$
(11)

where $var[\ln(R_{\infty})]$ and $var[\ln(\widehat{ICC_x})]$ are obtained from Appendix A of the supplementary materials. Similarly,

$$var(R_{yz,true}) \cong R_{yz,true}^{2} var\left[\ln(R_{yz,true})\right]$$

$$= R_{yz,true}^{2} \left\{ var\left[\ln(R_{yz})\right] + \frac{1}{4} var\left[\ln(\widehat{ICC}_{y})\right] \right\}$$
(12)

where $var[\ln(R_{yz})]$ and $var[\ln(\widehat{ICC_y})]$ are obtained from Appendix B of the supplementary materials. Finally, from the delta method we have:

$$cov\left(R_{xz,true}, R_{yz,true}\right) \cong R_{xz,true}R_{yz,true}cov\left[\ln\left(R_{xz,true}\right), \ln\left(R_{yz,true}\right)\right]$$
(13)
$$= R_{xz,true}R_{yz,true}\left\{cov\left[\ln\left(R_{xz}\right), \ln(R_{yz})\right] + \frac{1}{4}cov\left[\ln\left(\widehat{ICC}_{x}\right), \ln\left(\widehat{ICC}_{y}\right)\right]\right\}$$

where $cov[\ln(R_{xz}),\ln(R_{yz})]$ and $cov[\ln(\widehat{ICC}_x),\ln(\widehat{ICC}_y)]$ are given in Appendix C of the supplementary materials.

Hence, if we combine (9)-(13) and assume asymptotic normality of $\hat{\Delta}_{R,true}$. we have an approximate $100\% \times (1-\alpha)CI$ for $\Delta_{\rho,true}$ given by

$$\hat{\Delta}_{R,true} \pm z_{1-\alpha/2} \sqrt{var(\hat{\Delta}_{R,true})}$$
(14)

Similarly, we can obtain confidence limits for $\Delta_{\rho} = \rho_{xz} - \rho_{yz}$ given by

$$\hat{\Delta}_{R} \pm z_{1-\alpha/2} \sqrt{var(\hat{\Delta}_{R})}$$
(15)

where

$$\hat{\Delta}_{R} = R_{xz} - R_{yz}$$

$$var(\hat{\Delta}_{R}) = var(R_{xz}) + var(R_{yz}) - 2 cov(R_{xz}, R_{yz})$$

$$= R_{xz}^{2} var[\ln(R_{xz})] + R_{yz}^{2} var[\ln(R_{yz})] - 2R_{xz}R_{yz} cov[\ln(R_{xz}), \ln(R_{yz})]$$

Small-Sample Inference

When sample size is small, permutation methods can be used to estimate levels of significance. The permutation distribution is generated by randomly shuffling the X, Y labels and computing the empirical distribution of $\hat{\Delta}_{z,true}$ and $\hat{\Delta}_{z}$ in (5) and (7), respectively. The rank of $\hat{\Delta}_{z,true}$ and $\hat{\Delta}_{z}$ based on the observed data with reference to their permutation distributions can be used to estimate exact two-sided p-values given by

$$p - value = 2 \operatorname{x} \left[\sum_{m=1}^{M} \frac{I\left(\hat{\Delta}_{z,true, permutation,m} \ge \left|\hat{\Delta}_{z,true, observed}\right|\right)}{M} \right]$$
(16)

And similarly for $\hat{\Delta}_z$, where M is the size of the permutation distribution and I(a)=1 if a is true,=0 if a is false.

Comparison of Dependent Spearman Correlations

Hypothesis testing

We can also consider the comparison of dependent Spearman correlations whereby we test the hypothesis

$$H_0: \rho_{xz,s} = \rho_{yz,s} \operatorname{vs} H_1: \rho_{xz,s} \neq \rho_{yz,s}$$
(17)

where $\rho_{xz,s}$ =population Spearman correlation between X and Z and $\rho_{yz,s}$ is defined similarly.

If the data are transformed to the probit scale then from Rosner and Glynn [13].

$$\rho_{xz,s} = \frac{6}{\pi} \sin^{-1} \left(\rho_{xz,probit} / 2 \right) \equiv \frac{6}{\pi} \sin^{-1} \left[\operatorname{corr} \left(H_{x_i}, H_{z_i} \right) / 2 \right]$$
(18)
$$\rho_{yz,s} = \frac{6}{\pi} \sin^{-1} \left(\rho_{yz,probit} / 2 \right) \equiv \frac{6}{\pi} \sin^{-1} \left[\operatorname{corr} \left(H_{y_i}, H_{z_i} \right) / 2 \right]$$

where $H_{x_i} \equiv probit \left[F_x(X_i) \right]$ and is estimated by $\Phi^{-1} \left[\frac{rank(X_i)}{N+1} \right]$, where $\Phi = c.d.f.$ of a N(0,1) distribution, $\rho_{xz,probit} =$ population correlation between the probit of X and the probit of Z and $\rho_{yz,probit}$ is defined similarly.

Thus, the hypothesis test in (17) is equivalent to the hypothesis test in (7) after transformation to the probit scale yielding a test statistic based

on
$$\hat{\Delta}_{z, probit} = Z_{xz, probit} - Z_{yz, probit}$$
 given by $V_{xy, probit} = \hat{\Delta}_{z, probit} / \sqrt{var(\hat{\Delta}_{z, probit})}$

Interval estimation

It is also of interest to obtain confidence limits for $\Delta_{\rho,s} = \rho_{xz,s} - \rho_{yz,s}$ which we estimate by $\Delta_{R,s} = R_{xz,s} - R_{yz,s}$. Based on the delta method and (18), we obtain:

$$var\left(R_{xz,s} - R_{yz,s}\right) = var\left(R_{xz,s}\right) + var\left(R_{yz,s}\right) - 2cov\left(R_{xz,s}, R_{yz,s}\right)$$
(19)

$$\approx \frac{9R_{xz,probit}^2}{\pi^2 \left(1 - \frac{R_{xz,probit}^2}{4}\right)} var \left[ln \left(R_{xz,probit}\right) \right] + \frac{9R_{yz,probit}^2}{\pi^2 \left(1 - \frac{R_{yz,probit}^2}{4}\right)} var \left[ln \left(R_{yz,probit}\right) \right]$$

Page 3 of 9

Page 4 of 9

$$-2\frac{9R_{xz,probit}R_{yz,probit}}{\pi^2 \sqrt{\left(1-\frac{R_{xz,probit}^2}{4}\right)\left(1-\frac{R_{yz,probit}^2}{4}\right)}}cov\left[ln\left(R_{xz,probit}\right),ln(R_{yz,probit})\right]$$

where

 $var[ln(R_{xz,probit})], var[ln(R_{yz,probit})]$ and $cov[ln(R_{xz,probit}), ln(R_{yz,probit})]$ are given in Appendices A, B and C, respectively, of the supplementary material.

Thus, a $100\% \times (1-\alpha)CI$ for $\Delta_{\rho,s}$ is given by

$$\Delta_{R,s} \pm z_{1-\alpha/2} \sqrt{var(R_{xz,s} - R_{yz,s})}$$
 (20)

Correction for measurement error

Rosner and Glynn [13] defined the deattenuated Spearman correlation by

$$\rho_{xz,s,true} = \left(6 / \pi\right) \sin^{-1} \left(\rho_{xz,probit,true} / 2\right)$$
(21)

where $\rho_{xz,probit,true}$ is the deattenuated correlation between probit scores for X and Z, respectively. Hence, the hypotheses $H_0: \rho_{xz,s,true} = \rho_{yz,s,true}$ vs $H_1: \rho_{xz,s,true} \neq \rho_{yz,s,true}$ are equivalent to the hypotheses

 $H_0: \rho_{xz, probit, true} = \rho_{yz, probit, true} \text{ vs } H_1: \rho_{xz, probit, true} \neq \rho_{yz, probit, true}$

Furthermore, the latter hypotheses can be tested using (5), where the Fisher z transformation is performed based on probit scores, yielding a test statistic based on $\hat{\Delta}_{z,probit,true} = Z_{xz,probit,true} - Z_{yz,probit,true}$ given by

$$V_{xy, probit, true} = \hat{\Delta}_{z, probit, true} / \sqrt{var(\hat{\Delta}_{z, probit, true})} \text{ and}$$
$$var(\hat{\Delta}_{z, true, probit, true}) \equiv var(\hat{Z}_{xz, probit, true} - \hat{Z}_{yz, probit, true})$$
(22)

$$= var(\hat{Z}_{xz, probit, true}) + var(\hat{Z}_{yz, probit, true}) - 2cov(\hat{Z}_{xz, probit, true}, \hat{Z}_{yz, probit, true})$$

and $var(\hat{Z}_{xz,probit,true}), var(\hat{Z}_{yz,probit,true})$ and $cov(\hat{Z}_{xz,probit,true}, \hat{Z}_{yz,probit,true})$ are obtained from equations (A3), (B1) and (C1), of Appendices A, B and C, respectively of the supplementary materials.

To obtain a $100\% \times (1-\alpha)CI$ for $\Delta_{R,s,true} = \rho_{xz,s,true} - \rho_{yz,s,true}$, we use a similar approach as in (19) and obtain:

$$var(R_{xz,s,true} - R_{yz,s,true}) = var(R_{xz,s,true}) + var(R_{yz,s,true}) - 2cov(R_{xz,s,true}, R_{yz,s,true})$$
(23)

$$\begin{split} & \cong \frac{9R_{zz,probit,rne}^2}{\pi^2 \left(1 - \frac{R_{zz,probit,rne}^2}{4}\right)} var \left[ln\left(R_{zz,probit,rne}\right)\right] + \frac{9R_{yz,probit,rne}^2}{\pi^2 \left(1 - \frac{R_{yz,probit,rne}^2}{4}\right)} var \left[ln\left(R_{yz,probit,rne}\right)\right] \\ & -2 \frac{9R_{zz,probit,true} R_{yz,probit,true}}{\pi^2 \sqrt{\left(1 - \frac{R_{zz,probit,true}^2}{4}\right)} Cov \left[ln\left(R_{zz,probit,true}\right), \ln(R_{yz,probit,true})\right]} \end{split}$$

where

$$var\left[ln(R_{xz,probit,true})\right], var\left[ln(R_{yz,probit,true})\right] \text{ and } cov\left[ln(R_{xz,probit,true}), ln(R_{yz,probit,true})\right]$$

are obtained from (11)-(13), respectively.

The corresponding large sample 100% \times (1- α)CI for $\Delta_{R,s,true}$ is given by

$$\hat{\Delta}_{R,s,true} \pm z_{1-\alpha/2} \sqrt{var(\hat{\Delta}_{R,s,true})}$$
(24)

where

$$\Delta_{R,s,true} = R_{xz,s,true} - R_{yz,s,true}$$

Simulations

We conducted simulation studies to assess the validity of the asymptotic procedure in (7) for comparing dependent correlations in finite samples and to compare this procedure with the procedures of Dunn and Clark [7] and Meng, Rosenthal and Rubin [8]. We simulated data for (X, Y, Z) from a multivariate normal distribution $N(\mu, \Sigma)$ with

$$\mu = (0,0,0) \text{ and } \sum = \begin{pmatrix} 1 & 0.732 & \rho_{xz} \\ 0.732 & 1 & \rho_{yz} \\ \rho_{xz} & \rho_{yz} & 1 \end{pmatrix}$$

We performed 4,000 simulations for each combination of N=100, 400 and

$$(\rho_{xz}, \rho_{yz}) = (0.305, 0.305), (0.500, 0.500) \text{ and } (0.305, 0.500)$$

We also simulated data from a log normal distribution (X*, Y*, Z*), where $X^* = exp(X)$, $Y^* = exp(Y)$ and $Z^* = exp(Z)$. The results are given in Table 1. We see that in the case of a normal distribution that all three procedures have appropriate type I error (designs 1-4), particularly in the case of n=400. Similarly, the power of the procedures (designs 5-6) is also comparable. However, in the case of a log normal distribution, type I error is more adequately preserved by the method of moments procedure, particularly for n=400. As the sample size increases, the inconsistent estimate of the variance provided by the Dunn and Clark and Meng procedures in the setting of a non-normal distribution yields variance estimates that are too low and type I errors that are too high. Conversely, as n increases the method of moments procedure yields consistent variance estimates and actual p-values that are close to nominal levels. Alternative approaches if non-normality is suspected would be to compare dependent Spearman correlations rather than Pearson correlations.

To assess the accuracy of the procedure in (5) to compare dependent deattenuated Pearson correlations we used the same data designs as in designs 1-6. We performed each set of simulations in two ways. First, we assumed that the intra class correlation between replicate measures of X (and Y) was known without error and set to 0.67. Thus, $var[\ln(\widehat{ICC_x})] = var[\ln(\widehat{ICC_y})] = cov[\ln(\widehat{ICC_x}),\ln(\widehat{ICC_y})] = 0$ in equations (A3), (B1) and (C2). Second, we generated replicate values of (X_1, X_2, Y_1, Y_2) from a multivariate normal distribution denoted by N (μ_R , Σ_R) where $\mu_{R=}(0,0,0,0)$ and

$$\sum_{R} = \begin{pmatrix} 1 & 0.67 & 0.60 & 0.50 \\ 0.67 & 1 & 0.50 & 0.60 \\ 0.60 & 0.50 & 1 & 0.67 \\ 0.50 & 0.60 & 0.67 & 1 \end{pmatrix}$$

and estimated ICC_x and ICC_y from the replicate data. For each set of simulations we estimated the empirical type I error of $V_{xy, true}$ and the bias and variance of $(\hat{Z}_{xz, true} - \hat{Z}_{yz, true})$ in (5). The results are given in Table 2.

We see in Table 2 that the median theoretical variance and the empirical variance of $\hat{Z}_{xz,true}, \hat{Z}_{yz,true}$ and $\hat{\Delta}_{z,true}$ are very similar both for the case where the ICC is assumed known (Type 1) and where it is estimated from the simulated data (Type 2). In addition, the empirical type I error is close to 0.05 (range from 0.039 to 0.061) for all parameter

Page 5 of 9

					т	ype I Error	Power			
Design	Distribution	ρ _{xz}	ρ _{yz}	N	Method of Moments*	Dunn and Clark	Meng	Method of Moments*	Dunn and Clark	Meng
1	Normal	0.305	0.305	100	0.061	0.054	0.052			
2	Normal	0.305	0.305	400	0.045	0.045	0.045			
3	Normal	0.5	0.5	100	0.063	0.054	0.052			
4	Normal	0.5	0.5	400	0.045	0.045	0.045			
5	Normal	0.305	0.5	100				0.866	0.842	0.839
6	Normal	0.305	0.5	400				0.987	1	1
7	Lognormal	0.305	0.305	100	0.102	0.162	0.16			
8	Lognormal	0.305	0.305	400	0.075	0.197	0.196			
9	Lognormal	0.5	0.5	100	0.116	0.247	0.244			
10	Lognormal	0.5	0.5	400	0.082	0.302	0.326			
11	Lognormal	0.305	0.5	100				0.495	0.581	0.577
12	Lognormal	0.305	0.5	400				0.812	0.919	0.919

*Based on equation 7

Table 1: Comparison of alternative methods for comparing dependent correlations.

					$\hat{Z}_{xz,true}$		$\hat{Z}_{_{yz,true}}$		$\hat{\Delta}_{z,true}$		Type I error
Туре	ICC _x (ICC _y)	ρ _{xz} (ρ _{yz})	N		Mean	Variance	Mean	Variance	Mean	Variance	
1	0.67	0.305	100	theoretical	0.391	0.0163	0.391	0.0162	0	0.0095	0.05
				empirical	0.395	0.0177	0.398	0.0178	-0.002	0.0103	0.061
	0.67	0.305	400	theoretical	0.391	0.0041	0.391	0.0041	0	0.0024	0.05
				empirical	0.392	0.0041	0.393	0.0042	-0.001	0.0023	0.045
	0.67	0.5	100	theoretical	0.71	0.0212	0.71	0.0212	0	0.0146	0.05
				empirical	0.72	0.0229	0.723	0.0233	-0.003	0.0161	0.058
	0.67	0.5	400	theoretical	0.71	0.0053	0.71	0.0053	0	0.0036	0.05
				empirical	0.712	0.0053	0.714	0.0055	-0.001	0.0036	0.044
2	0.67	0.305	100	theoretical	0.391	0.0166	0.391	0.0168	0	0.01	0.05
				empirical	0.395	0.0183	0.398	0.018	-0.003	0.0106	0.051
	0.67	0.305	400	theoretical	0.391	0.0042	0.391	0.0042	0	0.0025	0.05
				empirical	0.394	0.0042	0.393	0.0041	0.001	0.0026	0.048
	0.67	0.5	100	theoretical	0.71	0.0226	0.71	0.0222	0	0.0167	0.05
				empirical	0.724	0.027	0.728	0.0269	-0.004	0.0201	0.039
	0.67	0.5	400	theoretical	0.71	0.0056	0.71	0.0056	0	0.004	0.05
				empirical	0.715	0.0058	0.713	0.0057	0.002	0.0044	0.048

Theoretical variance estimates are medians over 4000 simulated samples; Type 1 simulations assume that ICCx, ICCy are known; Type 2 simulations estimates ICCx, ICCy, from the sample data; for both Type 1 and Type 2 simulations, the underlying ICCx, ICCy are provided in the 2nd column of the table. Simulation results are based on 4000 simulations.

Table 2: Simulation results for the test procedure in (5) and (6).

combinations considered. Furthermore, the bias in estimation of $\hat{\Delta}_{z,true}$ is close to 0 for all parameter combinations considered. The above variances are slightly larger when the ICC is estimated (Type 2) than when the ICC is assumed known (Type 1), particularly when $\rho_{xz} = \rho_{yz} = 0.5$.

In addition, we estimated the bias and variance of $\hat{\Delta}_{R,true}$ and the associated coverage probability based on (14). The results are given in Table 3. We note that the bias is close to 0 in all designs considered. Furthermore, the empirical type I error in designs where $\rho_{xz,true} = \rho_{yz,true}$ is close to 0.05 and the coverage probability both when $\rho_{xz,true} = \rho_{yz,true}$ and when $\rho_{xz,true} \neq \rho_{yz,true}$ is close to 95% both for N=100 and N=400.

Examples

Association between analgesic intake and level of kidney function

We use data from the Swiss Analgesic Study where in 1967/68

1,256 women provided a urine specimen in the clinic and 2 additional specimens at home on different days which were assayed for NAPAP, a metabolite indicating recent intake of phenacetin-containing analgesics and salicylates a metabolite indicating recent intake of aspirin-containing analgesics. In addition, serum creatinine, a marker of kidney function was also measured at the baseline clinic visit. We let X, Y and Z be the NAPAP, salicylates and serum creatinine at the baseline (1968) clinic visit. We wish to compare corr(X,Z) with corr(Y,Z). A total of 1168 women had complete data on X, Y and Z. Descriptive statistics for the (X,Y,Z) data are provided in Table 4.

We note that the distributions of each of X, Y and Z are skewed and not normally distributed. There is moderate correlation between NAPAP and salicylates (Pearson correlation=0.423) in part because some analgesics contain both phenacetin and aspirin. Pearson correlations between NAPAP and salicylates vs. serum creatinine were 0.190 and 0.068, respectively. The ICC for NAPAP and salicylates over the 3 determinations were 0.614 and 0.404, respectively, indicating moderate variability of intake on different days.

Page 6 of 9

							R	xz,true	R,	z,true	Â	R,true	Type I error	Coverage probability
ρ _{xz}	ρ _{vz}	N		Mean	Variance	Mean	Variance	Mean	Variance					
0.305	0.305	100	theoretical	0.373	0.012	0.373	0.0122	0	0.0073	0.05	95			
			empirical	0.37	0.0128	0.373	0.0131	-0.002	0.0075	0.051	94.6			
		400	theoretical	0.373	0.0031	0.373	0.0031	0	0.0018	0.05	95			
			empirical	0.373	0.003	0.372	0.0031	0.001	0.0019	0.048	95.1			
0.5	0.5	100	theoretical	0.611	0.0086	0.611	0.0087	0	0.0065	0.05	95			
			empirical	0.61	0.0094	0.612	0.0096	-0.002	0.0068	0.039	94.5			
		400	theoretical	0.611	0.0022	0.611	0.0022	0	0.0016	0.05	95			
			empirical	0.612	0.0022	0.611	0.0022	0.001	0.0017	0.048	95			
0.305	0.5	100	theoretical	0.373	0.012	0.611	0.0087	-0.238	0.0071		95			
			empirical	0.37	0.0132	0.612	0.0094	-0.242	0.0075		94.3			
		400	theoretical	0.373	0.0031	0.611	0.0022	-0.238	0.0018		95			
			empirical	0.373	0.003	0.611	0.0022	-0.238	0.0018		95			

ICCx and ICCy estimated from simulated data (type 2 simulations). Simulation results are based on 4000 simulations

Table 3: Simulation results for the estimation of $\Delta_{R, true}$ in (14).

					Pears	on (Spearman)	correlation	Pearson (Spearman)
	N	mean	sd	(10 th –90 th percentile)	Х	Y	Z	ICC
NAPAP (X) (o.d.)	1168	0.176	0.352	(0.000,0.675)	1	0.423	0.19	0.614
						(0.353)	(0.155)	(0.515)
salicylates (Y) (mg%)	1168	15.8	19.2	(10.0,35.0)		1	0.068	0.404
							(0.081)	(0.389)
serum creatinine (Z) (mg/dL)	1168	1.02	0.37	(0.69,1.50)			1	

NAPAP value at the baseline (1968) clinic visit.

Individual salicylate values were obtained in grouped form (1=0-19 mg%/2=20-49 mg%/3=50-99 mg%/4=100+mg%) at the baseline (1968) clinic visit, and assigned scores of 10 mg%, 35 mg%, 75 mg% and 100 mg%, respectively.

Spearman correlations are estimated based on (18)

N: Number of subjects with serum creatinine, NAPAP and salicylates, available at the baseline (1968) clinic visit

ICC: Intraclass Correlation among replicate urine values (based on urine specimens obtained at the baseline clinic visit and 2 additional urine specimens obtained at home on different days).

Table 4: Descriptive statistics for NAPAP, salicylates and serum creatinine at the baseline visit, Swiss Analgesic Study, n=1168.

Variable	observed correlation	Z transform	deattenuated correlation	Z transform
NAPAP, 1968 (o.d.)	0.19	0.192	0.243	0.248
salicylates, 1968 (mg%)	0.068	0.069	0.108	0.108
Difference between correlations	0.122	0.124	0.135	0.139
(95%CI)	(0.051, 0.193)		(0.035, 0.235)	
se		0.037		0.053
V statistic		3.344		2.64
p-value		0.001		0.008

The V statistic for the observed and deattenuated Pearson correlations is based on (7) and (5), respectively.

 Table 5: Observed and deattenuated Pearson Correlations between serum creatinine at baseline (1968) vs. each of NAPAP and salicylates at the baseline clinic visit (1968), Swiss Analgesic Study, n=1168.

Based on (7), the observed Pearson correlations were significantly different (V_{xy} =3.344, p=0.001) (Table 5). For comparison, we also analyzed these data using the Dunn and Clark [7] and Meng, et al. [8] procedures. Based on Dunn and Clark, we have Z_1^* =3.929, p-value=8.5 × 10⁻⁵. Based on Meng, et al, we have Z_{xy} =3.919, p-value=8.9 × 10⁻⁵. As in the simulations, the se's are inappropriately low for both of these procedures in the setting of these highly non-normal data, yielding p-values that are biased downward and type I errors that are too low. The deattenuated Pearson correlations were 0.243 and 0.108 for NAPAP (X) and salicylates (Y), respectively, vs. serum creatinine. Based on (5) there was a significant difference between these correlations, $V_{xy,true}$ =2.640, p=0.008.

Since NAPAP and salicylates were right-skewed, we also computed observed and deattenuated Spearman correlations which are presented in Table 6.

The observed Spearman correlations between Z and (X,Y) were 0.155 and 0.081, respectively, which were significantly different ($V_{xy,probit}$ =2.395, p=0.017). After correcting for deattentuation, the Spearman correlations between Z and (X,Y) were 0.212 and 0.127, respectively, which only showed a trend towards statistical significance ($V_{xy,probit,true}$ =1.849, p=0.064). Overall, the results from the comparison of correlation coefficients in Tables 5 and 6 were consistent with the

Page 7 of 9

Variable	observed Spearman correlation	observed probit correlation	Z transform	deattenuated Spearman correlation	deattenuated probit correlation	Z transform
NAPAP, 1968 (o.d.)	0.155	0.162	0.163	0.212	0.221	0.225
salicylates, 1968 (mg%)	0.081	0.085	0.085	0.127	0.133	0.134
Difference between correlations	0.074	0.077	0.078	0.085	0.088	0.091
(95% CI)	(0.013, 0.134)			(-0.005, 0.174)		
se			0.033			0.049
V statistic			2.395			1.849
p-value			0.017			0.064

The V statistic for the observed and deattenuated probit correlations is based on (7) and (5), respectively.

Table 6: Spearman Correlations between serum creatinine at baseline (1968) vs. each of NAPAP and salicylates at the baseline clinic visit (1968), Swiss Analgesic Study, n=168.

	observed Spearman correlation	observed probit correlation	Z transform	deattenuated Spearman correlation	deattenuated probit correlation	Z transform	ICCs
2 weeks of room temperature pre-shipping plus 4 weeks of freezer (X)	0.952	0.956	1.896	0.979	0.98	2.31	0.946
4 weeks of room temperature pre-shipping plus 12 weeks of freezer (Y)	0.711	0.727	0.922	0.776	0.79	1.072	0.835
difference	0.241	0.229	0.974	0.203	0.19	1.238	
se			0.348			0.827	
V statistic			2.797			1.497	
p-value (large sample)			0.005			0.134	
p-value (exact)			0.021			0.056	

The V statistics are based on (7) and (5) using the probit transformation.

Table 7: Spearman Correlations between HbA1c levels determined immediately vs. HbA1c levels determined after 2 delay periods, N= N_{ICC} =12.

results from the primary paper (Dubach, Levy and Muller, [1]) which showed that NAPAP and salicylates were significantly associated with serum creatinine at baseline, although these analyses were based on categorical definitions of NAPAP, salicylates and creatinine.

Comparison of storage conditions for HbA1c measurements in plasma

Recently dried blood spot assays of HbA1c have been used as biomarkers in many epidemiologic studies and simplified specimen handling would be cost-effective for these large scale community-based studies. A study was conducted at Brigham and Women's Hospital (BWH) to test the hypothesis that dried blood spot determinations for HbA1c are valid measurements with low-intensity storage conditions [14]. Blood samples were drawn into EDTA containing tubes and submitted for duplicate HPLC analysis of HbA1c (considered as the gold standard process). Blood for spotting was also drawn to identical EDTA tubes and then dropped randomly to blood-spot cards. After having been air-dried for at least 20 minutes these dried blood specimens were placed into single-sample, airtight bags with desiccant pouch. They were then stored at room temperature for 0, 2, or 4 weeks in the lab and were then shipped to Biosafe Laboratories, Inc. (Lake Forest IL) for immediate HbA1c analysis or placed in freezer storage [-80°C] for an additional 4 or 12 weeks. Dried blood spot determinations of HbA1c were performed in triplicate, blinded to storage conditions and protocol. Here we want to test the equality of the correlations between results from HPLC measures (Z) and results from dried blood spot measures after 2 weeks of room temperature pre-shipping followed by an additional 4 weeks of freezing after shipping (X), and the correlation between results from HPLC measures (Z) and results from dried blood spot measures after 4 weeks of room temperature pre-shipping followed by an additional 12 weeks of freezing after shipping (Y).

We refer to storage for 2 weeks at room temperature plus 4 additional weeks in the freezer as short-term storage (X), storage for 4 weeks at room temperature plus 12 additional weeks in the freezer as long-term storage (Y) and HPLC processing as Z. The observed and deattenuated Spearman correlations between Z vs. (X,Y) respectively are given in Table 7. We see that $\rho_{xz,s}=0.952$ while $\rho_{yz,s}=0.711$, which are significantly different both based on large-sample methods (p=0.005) and also more appropriate exact methods (p=0.021). After deattenuation, we obtain $\rho_{yz,s, true}$ =0.979 and $\rho_{yz,s, true}$ =0.776 (2-sided p-value=0.134 based on large sample methods and 0.056 based on exact methods). Thus, the length of storage has a borderline significant effect on the validity of the HbA1c assay with short-term storage preferable. Plots of $(\hat{\Delta}_{z, permutation})$ and $(\hat{\Delta}_{z, true, permutation})$ in equation 16 over the 4,096 elements of the permutation distribution are given in Figures 1 and 2, respectively. The distributions are somewhat skewed to the right, particularly for $(\hat{\Delta}_{z,true, permutation})$, which indicates the necessity of using exact methods in a small-sample setting.

Discussion

In this paper, we provide a method to compare two Pearson correlation coefficients (ρ_{xz}, ρ_{yz}) where the correlations are estimated from the same subjects. Unlike previous literature on the subject, we do not make the assumption that (X, Y, Z) are multivariate normal. It is shown in simulation studies that the method of moments estimator of the sampling variance of $(Z_{xz}-Z_{yz})$ is more appropriate than standard methods that assume normality, particularly in the setting of non-normal data such as in the analgesic intake data presented in the first example in this paper. We then extended our methods to provide for comparison of two deattenuated correlation coefficients ($\rho_{xz,true}, \rho_{yz,true}$) where the correlations are estimated from the same subjects. This method is generally applicable to compare the validity

Page 8 of 9





of two surrogate measures (X, Y) where validity is measured by the respective correlation with a third variable (Z) and both X and Y are subject to measurement error. Although the methods are asymptotic, we have shown in finite samples that the asymptotic properties of the test statistic and confidence limits are appropriate when N \geq 100. In addition, we provide a method for confidence interval estimation of both Δ_{ρ} and $\Delta_{\rho,true}$.

The methods were extended to the comparison of dependent Spearman correlations which may be more appropriate when (X, Y, Z) are not multivariate normal, which provide for the comparison of both ordinary and deattenuated Spearman correlations. To our knowledge, this is the first paper in the literature to discuss comparison of dependent Spearman correlations.

In this paper, we have considered deattenuated correlations of the form $corr(\mu_x Z)$ and $corr(\mu_y Z)$ which account for error in the estimation of X and Y. However, it is likely that Z is also measured with error and a useful extension would be to estimate and compare $corr(\mu_x \mu_y)$ vs

SAS macros for the comparison of dependent ordinary and deattenuated Pearson and Spearman correlations as discussed in this paper are available from the following website:

 $corr(\mu_{\mu},\mu_{z})$, where error in X, Y and Z are taken into account.

https://sites.google.com/a/channing.harvard.edu/bernardrosner/ channing.

Acknowledgment

We would like to thank Drs. Buxton and Seeman for providing the original data on HbAlc.

The research was supported by U.S. National Institute of Health grants (U54 CA155626-03, P30 CA006516-48, R01 CA050597-18 and P30 AG017265-099002). This work was conducted with support from Harvard Catalyst | The Harvard Clinical and Translational Science Center (National Center for Research Resources and the National Center for Advancing Translational Sciences, National Institutes of Health Award UL1 TR001102) and financial contributions from Harvard University and its affiliated academic healthcare centers. The content is solely the responsibility of the authors and does not necessarily represent the official views of Harvard Catalyst, Harvard University and its affiliated academic healthcare centers, of the National Institutes of Health.

References

- Dubach UC, Levy PS, Muller A (1971) Relationships between regular analgesic intake and urorenal disorders in a working female population of Switzerland. I. Initial results (1968). Am J Epidemiol 93: 425-434.
- 2. Dubach UC, Rosner B, Pfister E (1983) Epidemiologic study of abuse of

analgesics containing phenacetin. Renal morbidity and mortality (1968-1979). N Engl J Med 308: 357-362.

- 3. Steiger JH (1980) Tests for comparing elements of a correlation matrix. Psychological Bulletin 87: 245-251.
- Williams EJ (1959) The comparison of regression variables. Journal of the Royal Statistical Society Series B (Methodological) 21: 396-399.
- Hotelling H (1940) The selection of variates for use in prediction with comments on the general problem of nuisance parameters. Annals of Mathematical Statistics 11: 271-283.
- Olkin I, Finn J (1990) Testing Correlated Correlations. Psychological Bulletin 108: 330-333.
- Dunn OJ, Clark V (1969) Correlation Coefficients Measured on the Same Individuals. Journal of the American Statistical Association 64: 366-377.
- Meng XL, Rosenthal R, Rubin DB (1992) Comparing correlated correlation coefficients. Psychological Bulletin 111: 172-175.
- Bilker WB, Brensinger C, Gur RC (2004) A two factor ANOVA-like test for correlated correlations: CORANOVA. Multivariate Behavioral Research 39: 565-594.
- 10. RCoreTeam (2014) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.
- 11. Diedenhofen B (2014) cocor: Comparing correlations.
- Rosner B, Willett WC (1988) Interval estimates for correlation coefficients corrected for within-person variation: implications for study design and hypothesis testing. American journal of epidemiology 127: 377-386.
- Rosner B, Glynn RJ (2007) Interval estimation for rank correlation coefficients based on the probit transformation with extension to measurement error correction of correlated ranked data. Stat Med 26: 633-646.
- Buxton OM, Malarick K, Wang W, Seeman T (2009) Changes in dried blood spot Hb A1C with varied post collection conditions. Clin Chem 55: 1034-1036.

Page 9 of 9