Türkiye Klinikleri Biyoistatistik Dergisi
*Turkiye Klinikleri Journal of Biostatistics* Turkiye Klinikleri J Biostat. 2023;15(1):52-66

ORİJİNAL ARAŞTIRMA ORIGINAL RESEARCH DOI: 10.5336/biostatic.2022-94692

# A Simulation Study for the Performances of Two Dependent Correlation Coefficients Comparison Tests with a Real Data Application on Acute Appendicitis

# İki Bağımlı Korelasyon Katsayısı Karşılaştırma Testlerinin Performansları için Simülasyon Çalışması ve Akut Apandisit Verileri Üzerinde Bir Uygulama

Deniz SIĞIRLI[a], Pınar TAŞAR[b]

[a]Department of Biostatistics, Bursa Uludağ University Faculty of Medicine, Bursa, Türkiye
[b]Department of General Surgery, Bursa Uludağ University Faculty of Medicine, Bursa, Türkiye

**ABSTRACT Objective:** Although it is frequently encountered in many studies to examine the relationships between different features of the individuals by using correlation coefficient, it is a situation that can be ignored to statistically test whether there is a difference between the correlation coefficients obtained. In this study, it is aimed to compare the performances of statistical tests proposed for the comparison of dependent overlapping correlation coefficients, in terms of their Type I error rates, within the framework of a wide simulation scenario. **Material and Methods:** The 6 test procedures were compared with a simulation study, conducted at 5 different intercorrelation magnitudes, with 5 different null correlation coefficient magnitudes, at 6 different sample sizes. **Results:** Pearson and Filon's z (PF) test performed poorly compared to other 5 procedures in most cases. For small intercorrelation magnitudes Steiger's modification of Dunn and Clark's z (SM) test, Meng, Rosenthal, and Rubin's z (MRR) test, Rosenthal, and Rubin's z test, Hittner, May, and Silver's modification of Dunn and Clark's z (HMS) test and ZOU's approach for overlapping correlations (ZA) procedures outperformed PF test and Hendrickson, Stanley, and Hills' modification of Williams' t test (HSHM) especially in small to moderate sample sizes. For larger intercorrelation coefficients, HSHM test gave better results in small to moderate sample sizes and ZA procedure maintained its superiority at the 0.7 intercorrelation level. **Conclusion:** Tests' performances in terms of Type I error are affected from the magnitude of null correlation, magnitude of intercorrelation and sample size, in different ways. It will be helpful to consider these issues when selecting the appropriate statistical test.

**Keywords:** Comparing correlation coefficients; dependent correlations; overlapping variables

**ÖZET Amaç:** Pek çok çalışmada, birimlerin farklı özellikleri arasındaki ilişkilerin araştırılması sık karşılaşılan bir durum olmakla birlikte elde edilen 2 korelasyon katsayısı arasında istatistiksel olarak anlamlı fark olup olmadığının test edilmesi genellikle göz ardı edilen bir durumdur. Bu çalışmada, bağımlı örtüşen korelasyon katsayılarının karşılaştırılması için kullanılan istatistiksel testlerin performanslarının Tip I hata düzeyi bakımından, geniş bir simülasyon senaryosu çerçevesinde karşılaştırılması amaçlanmıştır. **Gereç ve Yöntemler:** Altı farklı test prosedürü, ortak olmayan değişkenler arasındaki korelasyon katsayısının 5 farklı düzeyi için 5 farklı yokluk hipotezi korelasyon katsayısı düzeyi için ve 6 farklı örneklem büyüklüğünde karşılaştırılmıştır. **Bulgular:** Pearson ve Filon'un z (PF) testi pek çok durumda diğer 5 testten daha kötü performans göstermiştir. Ortak olmayan değişkenler arasındaki korelasyon katsayısının düşük düzeyleri için ve özellikle küçük-orta örneklem büyüklüklerinde; "Steiger's modification of Dunn and Clark's z" (SM), "Meng, Rosenthal, and Rubin's z test", "Hittner, May, and Silver's modification of Dunn and Clark's z (HMS) test" ve "ZOU's approach for overlapping correlations" (ZA) testleri, PF ve "Hendrickson, Stanley, and Hills' modification of Williams' t (HSHM)" testlerinden daha iyi performans vermiştir. Ortak olmayan değişkenler arasındaki korelasyon katsayısının yüksek düzeylerinde ise küçük-orta örneklem büyüklüklerinde HSHM testi daha iyi sonuçlar vermiş ve ZA prosedürü 0,7 iner-korelasyon düzeyinde üstünlüğünü korumuştur. **Sonuç:** Testlerin Tip I hata oranları bakımından performansları, korelasyon katsayılarının büyüklükleri ve örneklem büyüklüklerindeki değişimlerden farklı şekillerde etkilenmektedirler. Uygun istatistiksel testi seçerken bu noktalara dikkat edilmesi faydalı olacaktır.

**Anahtar kelimeler:** Korelasyon katsayılarının karşılaştırılması; bağımlı korelasyonlar; örtüşen değişkenler

**Correspondence:** Deniz SIĞIRLI
Department of Biostatistics, Bursa Uludağ University Faculty of Medicine, Bursa, Türkiye
**E-mail:** sigirli@uludag.edu.tr

Peer review under responsibility of Turkiye Klinikleri Journal of Biostatistics.

***Received:*** 02 Dec 2022 ***Received in revised form:*** 08 Dec 2022 ***Accepted:*** 12 Dec 2022 ***Available online:*** 02 Jan 2023

To explore the existence of a possible relationship between 2 variables is one of the main tasks in almost all quantitative research, including those in the social and behavioral sciences. Whether these 2 variables, which are measured from the individuals of the same sample, are related to each other, and if so, the degree and the direction of this relationship is usually examined with the Pearson product-moment correlation coefficient.[1,2] Many studies involve evaluating the magnitudes of 2 such correlation coefficients in a comparative manner. At this point, instead of just looking at the correlation coefficients or using various exploratory data analysis methods, it is important to test whether there is a statistically significant difference between the correlation coefficients and to check whether the observed difference between the coefficients occurs by chance, in order to interpret which variables have a stronger correlation.

In the topic of comparing 2 correlation coefficients, there are broadly 2 main situations. One situation is that; it can be aimed to compare 2 correlation coefficients from independent samples.[3] In the second situation, the difference of 2 correlation coefficients measured within a single sample might be the matter of interest, which yields dependent correlation coefficients. For this second situation, correlations will have a correlation matrix of their own.[4] The most common situation when comparing 2 dependent correlation coefficients is that the correlations share a common variable. We would like to correlate X with Y and X with Z, and would like to investigate whether the correlation between X and Y is larger/smaller than the correlation between X and Z.[5] Or for instance, we would like to calculate the correlations between the percentage of giving correct answer in one task and the activity of 2 brain regions (e.g. parietal and occipital). For this example, if the correlation of percentage of correct answers and the activity in one region was statistically significant and the correlation of percentage of correct answers and the activity in the other region was not statistically significant, this could lead to a fallacy if the researcher stops the analysis at this stage because that does not mean that the 2 correlations differ.[6] Comparing these 2 situations requires the usage of a statistical test for the hypothesis test of the difference between 2 correlation coefficients.

When there is a common variable in the calculation of correlation coefficients, this situation is usually referred to as "overlapping" correlations.[7] Many tests have been proposed to examine this issue in the literature.[5,7-15] However, performing a statistical test to compare the magnitudes of the correlation coefficients is often overlooked in researches. In this study, it is aimed to compare the performances of the statistical tests proposed for the comparison of 2 dependent overlapping correlation coefficients, in terms of Type I error levels, within the framework of a wide simulation scenario, and also to demonstrate the usage of these statistical tests on real clinical data.

## MATERIAL AND METHODS

We compared 6 test procedures for the case of comparing 2 dependent correlation coefficients obtained from the same sample which has one variable in common (overlapping). Let the 2 correlations that are being compared would be $\rho_{12}$ and $\rho_{13}$. Their related sample statistics can be given as $r_{12}$ and $r_{13}$, for a sample with a size of $n$. For these notations the test statistics for testing the hypothesis of $H_0: \rho_{12} = \rho_{13}$ vs. $H_1: \rho_{12} \neq \rho_{13}$, which performances were compared in the present study can be given as follows.

### I. PEARSON AND FILON'S Z TEST

The first test statistic was provided by the Pearson and Filon which follows standard normal distribution, and was given as in equation-1.[4,16]

$$z = \frac{\sqrt{n}(r_{12} - r_{13})}{\sqrt{(1 - r_{12}^2)^2 + (1 - r_{13}^2)^2 - 2k}} \qquad (1)$$

In equation-1, the term $k$ was given as follows.

$$k = r_{23}(1 - r_{12}^2 - r_{13}^2) - \frac{1}{2}(r_{12}.r_{13})(1 - r_{12}^2 - r_{13}^2 - r_{23}^2) \tag{2}$$

## II. HENDRICKSON, STANLEY, AND HILLS' MODIFICATION OF WILLIAMS' T TEST

After Hotelling (1931) proposed a test statistic for comparing 2 dependent correlations that have a variable in common, Williams (1959) performed a modification on the Hotelling's test.[8,9] This test provided by Hendrickson and Collins was as a modification of the test statistics provided by Williams.[9,13] The test statistics follows t distribution with n-3 degree of freedom (equation-3).[13,15,17]

$$t = \frac{\sqrt{(n-3)(1+r_{23})}\,(r_{12} - r_{13})}{\sqrt{2|R| + \frac{(r_{12} - r_{13})^2(1 - r_{23})^3}{4\,(n-1)}}} \tag{3}$$

where

$$|R| = 1 + 2r_{12}r_{13}r_{23} - r_{12}^2 - r_{13}^2 - r_{23}^2. \tag{4}$$

## III. STEIGER'S MODIFICATION OF DUNN AND CLARK'S Z

Steiger's test statistics for comparing 2 dependent overlapping correlation coefficients is given in equation-5.[4]

$$z = \frac{\sqrt{n-3}(z_{12} - z_{13})}{\sqrt{2 - 2e}} \tag{5}$$

where $z_{12}$ and $z_{13}$ are the Fisher's z transformations of $r_{12}$ and $r_{13}$, which the Fisher's z transformation for converting $r$ to $z$ was defined as in equation-6,[18]

$$z = \frac{1}{2}[\ln(1+r) - \ln(1-r)] \tag{6}$$

and

$$e = \frac{r_{23}(1 - 2\bar{r}^2) - (\frac{1}{2})\bar{r}^2(1 - 2\bar{r}^2 - r_{23}^2)}{(1 - \bar{r}^2)^2} \tag{7}$$

In equation-7, $\bar{r}$ was given as in equation-8.

$$\bar{r} = \frac{r_{12} + r_{13}}{2} \tag{8}$$

## IV. MENG, ROSENTHAL, AND RUBIN'S Z TEST

The test statistics is given in equation-9.[14]

$$z = \frac{\sqrt{n-3}(z_{12} - z_{13})}{\sqrt{2(1 - r_{23})f}} \tag{9}$$

where $z_{12}$ and $z_{13}$ are the Fisher's z transformations of $r_{12}$ and $r_{13}$. Fisher's z transformation for converting $r$ to $z$ was given before in equation-6.[17] In equation-9, $f$ was given as in equation-10,

$$f = \frac{1 - g\overline{r^2}}{1 - \overline{r^2}} \tag{10}$$

and $g$ and $\overline{r^2}$ were given as follows.

$$g = \frac{1 - r_{23}}{2(1 - \overline{r^2})} \tag{11}$$

$$\overline{r^2} = \frac{r_{12}^2 + r_{13}^2}{2} \tag{12}$$

## V. HITTNER, MAY, AND SILVER'S MODIFICATION OF DUNN AND CLARK'S Z TEST

The test statistics was proposed by Hittner et al. which was based on the back transforming the average z values which was first proposed by Silver and Dunlap, and is given as in equation-13.[15,19]

$$z = \frac{\sqrt{n-3}(z_{12} - z_{13})}{\sqrt{2 - 2h}} \tag{13}$$

In equation-13, $z_{12}$ and $z_{13}$ are the Fisher's z transformations of $r_{12}$ and $r_{13}$, which the Fisher's z transformation for converting $r$ to $z$ was given before in equation-6.[18] The term $h$ in equation-13 can be given as follows,

$$h = \frac{r_{23}\left(1 - 2\bar{r}_z^2\right) - (\frac{1}{2})\bar{r}_z^2\left(1 - 2\bar{r}_z^2 - r_{23}^2\right)}{\left(1 - \bar{r}_z^2\right)^2} \tag{14}$$

where

$$\bar{r}_z = \frac{\exp\,(2\bar{z} - 1)}{\exp\,(2\bar{z} + 1)} \tag{15}$$

and

$$\bar{z} = \frac{z_{12} + z_{13}}{2}\ . \tag{16}$$

## VI. ZOU'S APPROACH FOR OVERLAPPING CORRELATIONS

This approach of Zou calculates confidence intervals for a difference between 2 dependent overlapping correlation coefficients.[5] The null hypothesis of the 2 correlations are equal is being rejected if the confidence interval does not include zero. A lower and upper limits of the proposed confidence interval were defined as in equation-17 and -18, respectively.[5]

$$LL = r_{12} - r_{13} - \sqrt{(r_{12} - l_1)^2 + (u_2 - r_{13})^2 - 2j(r_{12} - l_1)(u_2 - r_{13})} \tag{17}$$

$$UL = r_{12} - r_{13} + \sqrt{(u_1 - r_{12})^2 + (r_{13} - l_2)^2 - 2j(u_1 - r_{12})(r_{13} - l_2)} \tag{18}$$

where $l$ and $u$ can be given as defined in equation-19 and -20.

$$l = \frac{\exp(2.l') - 1}{\exp(2.l') + 1} \tag{19}$$

$$u = \frac{\exp(2.u') - 1}{\exp(2.u') + 1} \tag{20}$$

and $l'$ and $u'$ were given as follows.

$$l', u' = z - z_{\frac{\alpha}{2}}\sqrt{1/(n-3)}\,, z + z_{\frac{\alpha}{2}}\sqrt{1/(n-3)} \tag{21}$$

The term $j$ in equation-17 and -18 can be given as in equation-22.

$$j = \frac{\left(r_{23} - \frac{1}{2} r_{12} r_{13}\right)(1 - r_{12}^2 - r_{13}^2 - r_{23}^2) + r_{23}^3}{(1 - r_{12}^2)(1 - r_{13}^2)} \quad (22)$$

## SIMULATION SCENARIOS

We obtained Type I error rates for these statistical tests for a wide range of simulation scenarios. Type I error rates were evaluated based on the stringent criterion of Bradley, in which they would range between 0.045 and 0.055.[20]

For the situation of comparing 2 overlapping Pearson correlation coefficients, $\rho_{12}$ and $\rho_{13}$, there are totally 3 correlation coefficients which one of them is the intercorrelation coefficient ($\rho_{23}$). We evaluated 5 different magnitudes of the correlation coefficients tested in the null hypothesis as: $\rho_{12} = \rho_{13} = 0.1, 0.3, 0.5, 0.7, 0.9$. For the intercorrelation coefficient we evaluated again 5 different magnitudes as: $\rho_{23} = 0.1, 0.3, 0.5, 0.7, 0.9$.

The data generating process for obtaining correlated data was performed based on previous studies.[7,15] For the simulation study, correlated data were generated by constructing the $XU$ matrix, where $U_{3\times3}$ is the root matrix of the correlation matrix and $X_{n\times3}$ is the matrix which is constructed by generating independent $3.n$ data from standard normal distribution. $R$ is the correlation matrix which is given as in equation-23. $U$ matrix was obtained by calculating the Cholesky decomposition of the correlation matrix $R$, which $R = UU^T$ and $U^T$ is the transpose of $U$.[21]

$$R = \begin{bmatrix} 1 & \rho_{12} & \rho_{13} \\ \rho_{12} & 1 & \rho_{23} \\ \rho_{13} & \rho_{23} & 1 \end{bmatrix} \quad (23)$$

The sample sizes were taken to be 10, 20, 30, 50, 100 and 200 in all scenarios. The Type I error rates were calculated for α=0.05. The repetition number was 1,000. gmat v. 0.2.2 and cocor v. 1.1-3 R packages were used in the processes of data generating and implementing the statistical tests.[22,23] In real data example normality of the data was tested with the Shapiro-Wilk test. Relationships between the variables were evaluated with the Pearson correlation coefficient. This study was conducted in accordance with the principles of the Declaration of Helsinki.

# RESULTS

Type I error rates were obtained for these tests: PF test for overlapping correlations; HSHM; SM test for overlapping correlations; MRR; HMS. For the ZA, proportion of the number of times that the confidence interval does not include 0 was calculated under 1,000 repetitions.

In each table, results were given for different magnitudes of the null hypothesis correlation coefficients ($\rho_{12} = \rho_{13}$) and for different sample sizes. In Table 1, Table 2, Table 3, Table 4, Table 5 we gave the simulation results for 5 different intercorrelations magnitudes, i.e. for $\rho_{23} = 0.1, 0.3, 0.5, 0.7$ and $0.9$, respectively (Table 1, Table 2, Table 3, Table 4, Table 5).

**TABLE 1:** Simulation results for the scenario where intercorrelation coefficient is 0.1 ($\rho_{23} = 0.1$).

| Null hypothesis correlation coefficients | n | Method | | | | | |
|---|---|---|---|---|---|---|---|
| | | PF | HSHM | SM | MRR | HMS | ZA |
| $\rho_{12} = \rho_{13} = 0.1$ | 10 | 0.136 | 0.062 | **0.055** | **0.046** | **0.052** | 0.064 |
| | 20 | 0.084 | **0.050** | 0.044 | 0.042 | 0.044 | **0.055** |
| | 30 | 0.062 | **0.048** | **0.047** | **0.045** | **0.047** | **0.048** |
| | 50 | 0.065 | **0.049** | **0.049** | **0.049** | **0.049** | **0.050** |
| | 100 | **0.049** | 0.043 | 0.043 | 0.042 | 0.043 | 0.043 |
| | 200 | **0.050** | **0.049** | **0.048** | **0.047** | **0.048** | **0.048** |
| | | | | | | | |
| $\rho_{12} = \rho_{13} = 0.3$ | 10 | 0.117 | 0.063 | **0.055** | 0.043 | **0.055** | 0.063 |
| | 20 | 0.078 | 0.056 | **0.049** | **0.045** | **0.049** | **0.051** |
| | 30 | 0.064 | 0.057 | **0.048** | 0.044 | **0.047** | **0.050** |
| | 50 | 0.061 | **0.054** | **0.049** | **0.047** | **0.048** | **0.051** |
| | 100 | **0.052** | **0.051** | **0.046** | **0.045** | **0.046** | **0.047** |
| | 200 | **0.053** | 0.056 | **0.049** | **0.049** | **0.049** | **0.049** |
| | | | | | | | |
| $\rho_{12} = \rho_{13} = 0.5$ | 10 | 0.077 | 0.064 | **0.055** | **0.051** | **0.051** | 0.057 |
| | 20 | 0.066 | 0.068 | **0.048** | **0.046** | **0.046** | **0.052** |
| | 30 | **0.051** | 0.066 | **0.045** | 0.044 | **0.045** | **0.048** |
| | 50 | 0.057 | 0.072 | 0.056 | **0.055** | **0.055** | 0.056 |
| | 100 | **0.048** | 0.060 | **0.045** | 0.043 | 0.044 | **0.045** |
| | 200 | **0.051** | 0.075 | **0.051** | **0.051** | **0.051** | **0.051** |
| | | | | | | | |
| $\rho_{12} = \rho_{13} = 0.7$ | 10 | 0.036 | 0.174 | **0.045** | **0.048** | 0.056 | **0.048** |
| | 20 | 0.038 | 0.203 | 0.041 | 0.044 | 0.044 | 0.044 |
| | 30 | 0.033 | 0.215 | 0.041 | **0.045** | **0.046** | **0.045** |
| | 50 | 0.057 | 0.248 | 0.061 | 0.061 | 0.061 | 0.061 |
| | 100 | 0.039 | 0.197 | 0.040 | 0.042 | 0.042 | 0.041 |
| | 200 | 0.058 | 0.224 | 0.058 | 0.058 | 0.058 | 0.058 |
| | | | | | | | |
| $\rho_{12} = \rho_{13} = 0.9$ | 10 | - | - | - | - | - | - |
| | 20 | - | - | - | - | - | - |
| | 30 | - | - | - | - | - | - |
| | 50 | - | - | - | - | - | - |
| | 100 | - | - | - | - | - | - |
| | 200 | - | - | - | - | - | - |

- Results for this scenario could not be computed since the correlation matrices were not positive definite.
PF: Pearson and Filon's z test; HSHM: Hendrickson, Stanley, and Hills' modification of Williams' t test; SM: Steiger's modifications of Dunn and Clark's z test for overlapping correlations; MRR: Meng, Rosenthal, and Rubin's z test; HMS: Hittner, May, and Silver's modification of Dunn and Clark's z test; ZA: ZOU's approach for overlapping correlations.

Type I error rates meeting the Bradley's criterion are given in bold.

**TABLE 2:** Simulation results for the scenario where intercorrelation coefficient is 0.3 ($\rho_{23} = 0.3$).

| Null hypothesis correlation coefficients | n | Method | | | | | |
|---|---|---|---|---|---|---|---|
| | | PF | HSHM | SM | MRR | HMS | ZA |
| $\rho_{12} = \rho_{13} = 0.1$ | 10 | 0.121 | 0.061 | **0.054** | 0.044 | **0.053** | 0.063 |
| | 20 | 0.083 | **0.051** | 0.044 | 0.040 | 0.043 | **0.048** |
| | 30 | 0.057 | **0.048** | **0.046** | 0.042 | **0.046** | **0.048** |
| | 50 | 0.063 | **0.050** | **0.049** | **0.046** | **0.049** | **0.050** |
| | 100 | **0.048** | **0.045** | 0.044 | 0.044 | 0.044 | **0.045** |
| | 200 | **0.048** | **0.048** | **0.047** | **0.046** | **0.047** | **0.047** |
| | | | | | | | |
| $\rho_{12} = \rho_{13} = 0.3$ | 10 | 0.112 | 0.067 | **0.051** | 0.040 | **0.046** | 0.064 |
| | 20 | 0.074 | **0.055** | **0.051** | **0.047** | **0.051** | **0.053** |
| | 30 | 0.066 | **0.049** | **0.047** | **0.046** | **0.047** | **0.047** |
| | 50 | 0.061 | **0.054** | **0.052** | **0.050** | **0.052** | **0.052** |
| | 100 | **0.050** | **0.050** | **0.048** | **0.046** | **0.048** | **0.049** |
| | 200 | **0.050** | **0.051** | **0.048** | **0.048** | **0.048** | **0.048** |
| | | | | | | | |
| $\rho_{12} = \rho_{13} = 0.5$ | 10 | 0.073 | 0.065 | **0.051** | **0.046** | **0.048** | 0.060 |
| | 20 | 0.064 | 0.057 | **0.050** | **0.046** | **0.048** | **0.053** |
| | 30 | 0.057 | 0.060 | **0.049** | **0.047** | **0.048** | **0.053** |
| | 50 | 0.056 | 0.062 | **0.054** | **0.053** | **0.054** | 0.056 |
| | 100 | **0.046** | **0.051** | **0.045** | **0.045** | **0.045** | **0.045** |
| | 200 | 0.044 | **0.053** | 0.044 | 0.044 | 0.044 | 0.044 |
| | | | | | | | |
| $\rho_{12} = \rho_{13} = 0.7$ | 10 | 0.033 | 0.076 | **0.051** | **0.049** | **0.052** | **0.052** |
| | 20 | 0.040 | 0.091 | **0.050** | **0.050** | **0.050** | **0.051** |
| | 30 | 0.037 | 0.088 | **0.047** | **0.047** | **0.048** | **0.047** |
| | 50 | **0.047** | 0.105 | **0.055** | **0.055** | 0.056 | 0.056 |
| | 100 | 0.043 | 0.082 | 0.043 | 0.043 | 0.043 | 0.043 |
| | 200 | 0.057 | 0.104 | 0.057 | 0.057 | 0.057 | 0.058 |
| | | | | | | | |
| $\rho_{12} = \rho_{13} = 0.9$ | 10 | - | - | - | - | - | - |
| | 20 | - | - | - | - | - | - |
| | 30 | - | - | - | - | - | - |
| | 50 | - | - | - | - | - | - |
| | 100 | - | - | - | - | - | - |
| | 200 | - | - | - | - | - | - |

- Results for this scenario could not be computed since the correlation matrices were not positive definite.
PF: Pearson and Filon's z test; HSHM: Hendrickson, Stanley, and Hills' modification of Williams' t test; SM: Steiger's modifications of Dunn and Clark's z test for overlapping correlations; MRR: Meng, Rosenthal, and Rubin's z test; HMS: Hittner, May, and Silver's modification of Dunn and Clark's z test; ZA: ZOU's approach for overlapping correlations.

Type I error rates meeting the Bradley's criterion are given in bold.

**TABLE 3:** Simulation results for the scenario where intercorrelation coefficient is 0.5 ($\rho_{23} = 0.5$).

| Null hypothesis correlation coefficients | n | Method | | | | | |
|---|---|---|---|---|---|---|---|
| | | PF | HSHM | SM | MRR | HMS | ZA |
| $\rho_{12} = \rho_{13} = 0.1$ | 10 | 0.108 | 0.063 | **0.049** | 0.038 | **0.048** | 0.059 |
| | 20 | 0.074 | **0.055** | **0.049** | 0.039 | **0.048** | **0.055** |
| | 30 | 0.056 | **0.047** | 0.042 | 0.040 | 0.042 | **0.048** |
| | 50 | 0.061 | **0.046** | **0.045** | 0.042 | **0.045** | **0.046** |
| | 100 | **0.048** | **0.045** | **0.045** | **0.045** | **0.045** | **0.045** |
| | 200 | **0.048** | **0.045** | 0.044 | 0.044 | 0.044 | **0.045** |
| | | | | | | | |
| $\rho_{12} = \rho_{13} = 0.3$ | 10 | 0.098 | 0.065 | **0.046** | 0.039 | 0.040 | 0.058 |
| | 20 | 0.067 | 0.060 | **0.052** | **0.047** | **0.051** | 0.060 |
| | 30 | 0.058 | **0.046** | **0.046** | **0.045** | **0.045** | **0.046** |
| | 50 | 0.062 | **0.054** | **0.052** | **0.051** | **0.052** | **0.052** |
| | 100 | **0.049** | **0.048** | **0.045** | 0.044 | **0.045** | **0.047** |
| | 200 | **0.049** | **0.047** | **0.046** | **0.045** | **0.046** | **0.046** |
| | | | | | | | |
| $\rho_{12} = \rho_{13} = 0.5$ | 10 | 0.066 | 0.062 | **0.052** | **0.046** | 0.043 | 0.059 |
| | 20 | 0.063 | 0.059 | **0.053** | **0.045** | **0.049** | 0.056 |
| | 30 | 0.057 | 0.057 | **0.046** | **0.046** | **0.045** | **0.051** |
| | 50 | 0.057 | **0.055** | **0.054** | **0.052** | **0.053** | **0.055** |
| | 100 | 0.044 | **0.046** | 0.041 | 0.041 | 0.041 | 0.044 |
| | 200 | **0.045** | **0.049** | 0.044 | 0.044 | 0.044 | **0.045** |
| | | | | | | | |
| $\rho_{12} = \rho_{13} = 0.7$ | 10 | 0.029 | 0.059 | **0.050** | **0.049** | **0.050** | 0.056 |
| | 20 | 0.039 | 0.066 | 0.058 | 0.056 | 0.056 | 0.061 |
| | 30 | 0.039 | 0.064 | **0.051** | **0.049** | **0.048** | **0.051** |
| | 50 | **0.045** | 0.062 | **0.053** | **0.053** | **0.051** | **0.053** |
| | 100 | 0.037 | **0.050** | 0.039 | 0.039 | 0.039 | 0.040 |
| | 200 | **0.047** | 0.060 | **0.048** | **0.048** | **0.048** | **0.048** |
| | | | | | | | |
| $\rho_{12} = \rho_{13} = 0.9$ | 10 | - | - | - | - | - | - |
| | 20 | - | - | - | - | - | - |
| | 30 | - | - | - | - | - | - |
| | 50 | - | - | - | - | - | - |
| | 100 | - | - | - | - | - | - |
| | 200 | - | - | - | - | - | - |

- Results for this scenario could not be computed since the correlation matrices were not positive definite.
PF: Pearson and Filon's z test; HSHM: Hendrickson, Stanley, and Hills' modification of Williams' t test; SM: Steiger's modifications of Dunn and Clark's z test for overlapping correlations; MRR: Meng, Rosenthal, and Rubin's z test; HMS: Hittner, May, and Silver's modification of Dunn and Clark's z test; ZA: ZOU's approach for overlapping correlations.

Type I error rates meeting the Bradley's criterion are given in bold.

**TABLE 4:** Simulation results for the scenario where intercorrelation coefficient is 0.7 ($\rho_{23} = 0.7$).

| Null hypothesis correlation coefficients | n | Method | | | | | |
|---|---|---|---|---|---|---|---|
| | | **PF** | **HSHM** | **SM** | **MRR** | **HMS** | **ZA** |
| $\rho_{12} = \rho_{13} = 0.1$ | 10 | 0.094 | 0.058 | 0.042 | 0.034 | 0.038 | **0.045** |
| | 20 | 0.063 | **0.050** | 0.042 | 0.037 | 0.041 | **0.047** |
| | 30 | 0.057 | **0.050** | 0.044 | 0.042 | 0.043 | **0.048** |
| | 50 | **0.053** | **0.046** | 0.042 | 0.039 | 0.041 | **0.045** |
| | 100 | **0.049** | **0.046** | **0.046** | **0.046** | **0.046** | **0.046** |
| | 200 | **0.045** | **0.044** | 0.044 | 0.044 | 0.044 | 0.044 |
| | | | | | | | |
| $\rho_{12} = \rho_{13} = 0.3$ | 10 | 0.084 | 0.059 | 0.042 | 0.037 | 0.038 | **0.049** |
| | 20 | 0.066 | **0.053** | **0.047** | 0.042 | **0.046** | **0.053** |
| | 30 | **0.050** | 0.038 | 0.037 | 0.037 | 0.037 | 0.038 |
| | 50 | **0.053** | **0.050** | 0.043 | 0.043 | 0.043 | **0.048** |
| | 100 | **0.051** | **0.048** | **0.047** | **0.046** | **0.046** | **0.047** |
| | 200 | **0.046** | 0.044 | 0.044 | 0.043 | 0.044 | 0.044 |
| | | | | | | | |
| $\rho_{12} = \rho_{13} = 0.5$ | 10 | **0.046** | 0.057 | 0.044 | 0.034 | 0.038 | **0.052** |
| | 20 | 0.057 | **0.055** | **0.049** | 0.043 | **0.045** | **0.054** |
| | 30 | **0.049** | **0.048** | 0.044 | 0.043 | 0.044 | **0.047** |
| | 50 | **0.050** | **0.051** | **0.047** | **0.047** | **0.047** | **0.050** |
| | 100 | 0.043 | 0.043 | 0.043 | 0.043 | 0.043 | 0.043 |
| | 200 | **0.047** | **0.048** | **0.047** | **0.047** | **0.047** | **0.047** |
| | | | | | | | |
| $\rho_{12} = \rho_{13} = 0.7$ | 10 | 0.027 | **0.046** | 0.040 | 0.036 | 0.036 | **0.047** |
| | 20 | 0.035 | 0.062 | 0.061 | 0.059 | 0.058 | 0.062 |
| | 30 | 0.038 | 0.056 | **0.050** | **0.048** | **0.047** | **0.055** |
| | 50 | 0.040 | **0.051** | **0.051** | **0.048** | **0.046** | **0.051** |
| | 100 | 0.042 | **0.047** | 0.044 | 0.044 | 0.044 | 0.044 |
| | 200 | **0.048** | **0.051** | **0.048** | **0.048** | **0.048** | **0.049** |
| | | | | | | | |
| $\rho_{12} = \rho_{13} = 0.9$ | 10 | 0.004 | 0.090 | **0.046** | **0.048** | 0.063 | **0.051** |
| | 20 | 0.009 | 0.091 | 0.044 | 0.044 | 0.056 | **0.047** |
| | 30 | 0.017 | 0.100 | **0.048** | **0.050** | 0.060 | **0.052** |
| | 50 | 0.041 | 0.124 | **0.054** | **0.055** | 0.059 | **0.055** |
| | 100 | 0.030 | 0.088 | 0.036 | 0.036 | 0.039 | 0.036 |
| | 200 | **0.054** | 0.106 | 0.059 | 0.059 | 0.060 | 0.059 |

PF: Pearson and Filon's z test; HSHM: Hendrickson, Stanley, and Hills' modification of Williams' t test; SM: Steiger's modifications of Dunn and Clark's z test for overlapping correlations; MRR: Meng, Rosenthal, and Rubin's z test; HMS: Hittner, May, and Silver's modification of Dunn and Clark's z test; ZA: ZOU's approach for overlapping correlations.

Type I error rates meeting the Bradley's criterion are given in bold.

**TABLE 5:** Simulation results for the scenario where intercorrelation coefficient is 0.9 ($\rho_{23} = 0.9$).

| Null hypothesis correlation coefficients | n | Method | | | | | |
|---|---|---|---|---|---|---|---|
| | | PF | HSHM | SM | MRR | HMS | ZA |
| $\rho_{12} = \rho_{13} = 0.1$ | 10 | 0.067 | 0.059 | 0.032 | 0.030 | 0.031 | 0.027 |
| | 20 | 0.058 | **0.052** | 0.040 | 0.038 | 0.040 | 0.042 |
| | 30 | **0.054** | **0.047** | 0.043 | 0.042 | 0.043 | 0.044 |
| | 50 | 0.044 | 0.042 | 0.040 | 0.040 | 0.040 | 0.041 |
| | 100 | **0.050** | **0.048** | **0.046** | **0.045** | **0.046** | **0.048** |
| | 200 | **0.050** | **0.050** | **0.050** | **0.050** | **0.050** | **0.050** |
| | | | | | | | |
| $\rho_{12} = \rho_{13} = 0.3$ | 10 | 0.057 | **0.055** | 0.038 | 0.035 | 0.037 | 0.029 |
| | 20 | **0.055** | **0.055** | 0.041 | 0.039 | 0.040 | 0.038 |
| | 30 | **0.050** | **0.047** | 0.044 | 0.043 | 0.044 | 0.040 |
| | 50 | **0.048** | 0.044 | 0.041 | 0.040 | 0.041 | 0.038 |
| | 100 | **0.052** | **0.051** | **0.051** | **0.051** | **0.051** | **0.050** |
| | 200 | 0.043 | 0.043 | 0.043 | 0.043 | 0.043 | 0.043 |
| | | | | | | | |
| $\rho_{12} = \rho_{13} = 0.5$ | 10 | 0.037 | 0.061 | 0.038 | 0.037 | 0.035 | 0.028 |
| | 20 | **0.049** | **0.048** | 0.040 | 0.039 | 0.039 | 0.033 |
| | 30 | 0.043 | 0.044 | 0.040 | 0.039 | 0.040 | 0.033 |
| | 50 | 0.041 | **0.045** | 0.040 | 0.040 | 0.040 | 0.037 |
| | 100 | 0.053 | **0.054** | **0.052** | **0.052** | **0.052** | **0.051** |
| | 200 | **0.048** | **0.048** | **0.048** | **0.048** | **0.048** | **0.048** |
| | | | | | | | |
| $\rho_{12} = \rho_{13} = 0.7$ | 10 | 0.017 | **0.054** | 0.043 | 0.041 | 0.038 | 0.026 |
| | 20 | 0.036 | 0.058 | **0.049** | **0.048** | **0.048** | 0.038 |
| | 30 | 0.031 | **0.047** | 0.044 | 0.044 | 0.044 | 0.034 |
| | 50 | 0.039 | **0.047** | **0.045** | **0.045** | **0.045** | 0.041 |
| | 100 | 0.044 | **0.047** | **0.047** | **0.047** | **0.047** | **0.046** |
| | 200 | **0.046** | **0.049** | **0.048** | **0.048** | **0.048** | **0.047** |
| | | | | | | | |
| $\rho_{12} = \rho_{13} = 0.9$ | 10 | 0.001 | **0.053** | **0.050** | **0.049** | 0.042 | **0.051** |
| | 20 | 0.007 | 0.060 | 0.060 | 0.060 | 0.056 | 0.061 |
| | 30 | 0.022 | 0.057 | 0.056 | 0.056 | **0.053** | 0.057 |
| | 50 | 0.033 | **0.050** | **0.049** | **0.049** | **0.048** | **0.050** |
| | 100 | 0.036 | **0.046** | 0.043 | 0.042 | 0.042 | 0.044 |
| | 200 | 0.038 | **0.048** | 0.043 | 0.043 | 0.042 | 0.043 |

PF: Pearson and Filon's z test; HSHM: Hendrickson, Stanley, and Hills' modification of Williams' t test; SM: Steiger's modifications of Dunn and Clark's z test for overlapping correlations; MRR: Meng, Rosenthal, and Rubin's z test; HMS: Hittner, May, and Silver's modification of Dunn and Clark's z test; ZA: ZOU's approach for overlapping correlations.

Type I error rates meeting the Bradley's criterion are given in bold.

## REAL DATA APPLICATION

Acute appendicitis is the most common indication of the acute abdominal surgery.[24] However, the incidence of negative laparotomy in patients who underwent appendectomy with a preliminary diagnosis of acute appendicitis is still around 20%.[25] Although this rate decreases with radiological evaluations, the effectiveness of various serum biomarkers in diagnosis is evaluated due to both the radiation exposure of the patient and the costs. Among these biomarkers, normal values of the acute phase reactant C-reactive protein (CRP), has been associated with a normal appendix and has been thought to reduce negative laparotomy.[26] Also it has been shown that an increase in CRP and leukocyte levels may also be important indicators of complicated appendicitis.[27] However it has also been reported that, the decrease in leukocytes values and CRP formation with increasing age, may affect the correct diagnosis of these biomarkers in acute appendicitis cases.[28]

In the present study data related to age, CRP and white blood cells (WBC) counts of the 49 patients, operated with the diagnosis of acute appendicitis were collected retrospectively from the records of Bursa Uludağ Faculty of Medicine General Surgery Department. The Ethical Committee of Uludağ University Medical Faculty Clinical Research approved this retrospective study (date: November 25, 2022, no: 2022-18/49). We examined the relationship between CRP and age, also between WBC and age for these patients. While there was a significant negative correlation between WBC and age (r=-0.323, p=0.023), no significant correlation was found between CRP and age (r=0.194, p=0.183). We conducted the aforementioned statistical tests, to test the hypothesis of $H_0: \rho_{12} = \rho_{13}$ vs. $H_1: \rho_{12} \neq \rho_{13}$. p values were close to each other except for the PF test, which all demonstrated statistically significant differences between 2 correlation coefficients (Table 6).

**TABLE 6:** Results of the statistical tests performed for comparing two dependent overlapping correlation coefficients.

| Method | Test-statistics | p value |
|---|---|---|
| PF | -2.731 | 0.006 |
| HSHM | -2.606 | 0.012 |
| SM | -2.489 | 0.013 |
| MRR | -2.451 | 0.014 |
| HMS | -2.488 | 0.013 |
| ZA* | [-0.865: -0.116] | |

*The confidence interval of the difference between the 2 correlations; PF: Pearson and Filon's z test; HSHM: Hendrickson, Stanley, and Hills' modification of Williams' t test; SM: Steiger's modifications of Dunn and Clark's z test for overlapping correlations; MRR: Meng, Rosenthal, and Rubin's z test; HMS: Hittner, May, and Silver's modification of Dunn and Clark's z test; ZA: ZOU's approach for overlapping correlations.

## DISCUSSION

For the scenario where the intercorrelation coefficient was 0.1; on the small to medium null hypothesis correlation coefficient magnitudes, the SM, MRR, HM, and ZA tests outperformed the PF and HSHM tests, except that the HSHM test gave results close to nominal level for the scenario where the 0 correlation coefficient equals to the intercorrelations coefficient. But it diverged from the nominal level and tended to be fairly liberal as the magnitude of the null correlation levels increased. The PF test performed well only for large sample sizes, on small to moderate null correlation coefficient magnitudes. The 6 procedures could not perform well for large null correlation coefficient (0.7), except MRR, HMS and ZA gave better results in small sample sizes for large null correlation coefficient.

For the scenario where the intercorrelation coefficient was 0.3; the results were similar to, but slightly better than the scenario where the intercorrelation correlation was 0.1. For small null correlation coefficient magnitudes, HSHM, SM, HMS and ZA procedures gave better results than PF and MRR tests, giving Type I error rates close to the nominal level. MRR test tended to give more conservative results for the small null correlation coefficient (0.1). At moderate null correlation magnitude and for small to moderate sample sizes at large null correlation magnitude; SM, MRR, HMS and ZA procedures, outperformed the PF and HSHM tests, whereas the magnitude of null correlations increased HSHM test diverged from the nominal level by getting more liberal. For this scenario where the intercorrelation coefficient was 0.3 SM, HMS and ZA gave results similar to nominal level, except in high null correlation and large sample sizes.

For the scenario where the intercorrelations coefficient was 0.5; PF test could not perform well except large sample sizes, regardless of the magnitude of the null correlation coefficient. For this scenario, while HSHM gave better results in small null hypothesis correlations, it diverged from the nominal level and tended to be more liberal in small to moderate sample sizes as the magnitude of the null correlation levels increased. On very small null correlation coefficient (0.1) while HSHM, SM, HMS and ZA procedures gave better results than the other tests in small to moderate sample sizes, the MRR test could not perform well and tended to be fairly conservative as the sample size decreased. For the null correlation coefficients of similar magnitude to the intercorrelation coefficient, SM, MRR and HMS tended to give better results than PF, HSHM and ZA procedures at small sample sizes; on the contrary ZA, HSHM and PF gave better results than them in large sample sizes.

For the scenario where the intercorrelations coefficient was 0.7; SM, MRR, HMS tests gave poorer results than the scenarios where the intercorrelation coefficient was smaller. The ZA approach gave better results than the other tests, regardless of the magnitude of the null correlation coefficient. While PF and HSHM tests gave better results than the SM, MRR and HMS tests on small to moderate null correlations, on the contrary SM and MRR tests gave better results than PF and HSHM tests on large null correlations coefficients. For very large null correlation coefficient (0.9), PF, HSHM and HMS tests couldn’t reach Bradley’s criterion in almost all sample sizes.

For the scenario where the intercorrelations coefficient was 0.9; HSHM test gave better results than the other tests, regardless of the magnitude of the null correlation coefficient. SM, MRR and HMS tests couldn’t reach Bradley’s criterion in most scenarios, except where the null correlation coefficient was 0.7 and also except large and very large sample sizes for small to moderate null hypothesis correlations. HSHM and partially PF outperformed the other procedures on smaller null hypothesis correlations with small to moderate sample sizes. HSHM, SM, MRR and HMS test gave better performances than the PF and ZA procedures for the 0.7 null hypothesis correlation coefficient. ZA procedure couldn’t reach Bradley’s criterion in almost all scenarios, for small to moderate sample sizes.

In general, for small intercorrelations SM, MRR, HMS and ZA procedures outperformed PF and HSHM tests especially in small to moderate sample sizes. Hittner et al. also found greater Type I error rates for the HSHM test than the SM and HMS tests across all sample sizes and for small to moderate intercorrelations.[15] May et al. stated in their study, which they compared four tests including HSHM and MMR tests that; for small intercorrelation magnitude (0.1) with a 0.7 null correlation magnitude, HSHM test gave inflated Type I error rates, while Type I error rates for the MMR test were close to nominal level for this scenerio.[7] This result was in agreement with ours, for small intercorrelations in all sample sizes. They also mentioned in their study that, inflated Type I error rates for HSHM were not associated with small sample, instead it appeared to be a function of the magnitudes of null correlation and intercorrelation coefficients.[7] For larger intercorrelation coefficients, the superiority of the SM, MRR, HMS and ZA procedures ended. While the ZA procedure maintained its superiority at the 0.7 intercorrelation level, it also diverged from the nominal level at the 0.9 intercorrelation level and for these 2 intercorrelation levels, HSHM test gave better results in small to moderate sample sizes. On the other hand, regardless of the sample size and at almost all null correlation magnitudes, the HSHM test gave more liberal results than the other tests, except PF. The worst performance was given by the PF test especially at small intercorrelations; giving fairly liberal results in small null correlations and fairly conservative results in higher null correlations, it failed to meet Bradley's criterion at all in small to moderate sample sizes. Also in real data example, p values were close to each other except for the PF test. This finding related to PF test, collaborates the study of Silver et al., where they found progressively more liberal results for PF test as the null correlations decreased in their study which they had compared the tests for dependent but nonoverlapping correlations.[29]

Although the Pearson correlation coefficient was considered in this study, it has been mentioned in different studies that the above test statistics regarding to the comparison of dependent correlations and procedures involving the Fisher's z transformations of Pearson product-moment correlations can also be applied to Spearman's correlation coefficient, provided that the sample size is equal or greater than 10 and that the Spearman's rho correlation coefficient is less than 0.9.[30,31] Future studies can also be performed to explore the performances of test procedures in case of comparing Spearman correlation coefficients, and for the cases where the data do not follow normal distribution.

## CONCLUSION

It is apparent from the simulation results that tests are affected from the magnitudes of null correlations, from the magnitudes of intercorrelations and from the sample size, in different ways. Therefore, it would be useful to decide on the test to be used according to these magnitudes and the sample size. But to make more generalizable comments, the test should also be compared in terms of their power.

***Source of Finance***

*During this study, no financial or spiritual support was received neither from any pharmaceutical company that has a direct connection with the research subject, nor from a company that provides or produces medical instruments and materials which may negatively affect the evaluation process of this study.*

***Conflict of Interest***

*No conflicts of interest between the authors and/or family members of the scientific and medical committee members or members of the potential conflicts of interest, counseling, expertise, working conditions, share holding and similar situations in any firm.*

***Authorship Contributions***

***Idea/Concept:*** *Deniz Sığırlı;* ***Design:*** *Deniz Sığırlı, Pınar Taşar;* ***Control/Supervision:*** *Deniz Sığırlı, Pınar Taşar;* ***Data Collection and/or Processing:*** *Pınar Taşar;* ***Analysis and/or Interpretation:*** *Deniz Sığırlı;* ***Literature Review:*** *Deniz Sığırlı, Pınar Taşar;* ***Writing the Article:*** *Deniz Sığırlı;* ***Critical Review:*** *Pınar Taşar;* ***References and Fundings:*** *Deniz Sığırlı;* ***Materials:*** *Pınar Taşar.*

## REFERENCES

1. Allen M. The SAGE Encyclopedia of Communication Research Methods. 1st ed. Thousand Oaks: SAGE Publications Inc.; 2017. [Crossref]
2. Pearson K. Notes on regression and inheritance in the case of two parents. Proceedings of the Royal Society of London. 1985;58:240-2. [Link]
3. Wilcox RR. Comparing pearson correlations: dealing with heteroscedasticity and nonnormality. Commun. Stat. B: Simul. Comput. 2009;38(10):2220-34. [Crossref]
4. Steiger JH. Tests for comparing elements of a correlation matrix. Psychol Bull. 1980;87(2):245-51. [Crossref]
5. Zou GY. Toward using confidence intervals to compare correlations. Psychol Methods. 2007;12(4):399-413. [Crossref] [PubMed]
6. Nieuwenhuis S, Forstmann BU, Wagenmakers EJ. Erroneous analyses of interactions in neuroscience: a problem of significance. Nat Neurosci. 2011;14(9):1105-7. [Crossref] [PubMed]
7. May K, Hittner JB. Tests for comparing dependent correlations revisited: A Monte Carlo study. J Exp Educ. 1997;65(3):257-69. [Crossref]
8. Hotelling H. The generalization of Student's ratio. Ann Math Stat. 1931;2(3):360-78. [Crossref]
9. Williams EJ. The comparison of regression variables. J R Stat Soc Series B Stat Methodol J R Stat Soc B. 1959;21(2):396-9. [Crossref]
10. Olkin I. Correlations revisited. In: Stanley JC, ed. Improving Experimental Design and Statistical Analysis. 1st ed. Chicago, IL: Rand McNally; 1967. p.102-28.
11. Dunn OJ, Clark VA. Correlation coefficients measured on the same individuals. J Am Stat Assoc. 1969;64(325):366-77. [Crossref]
12. Dunn OJ, Clark VA. Comparison of tests of the equality of dependent correlation coefficients. Am Stat Assoc. 1971;66(336):904-8. [Crossref]

13. Hendrickson GF, Collins JR. Olkin's new formula for the significance of r13 vs. r23 compared with Hotelling's method. Am Educ Res J. 1970;7(2):189-95. [Crossref]
14. Meng XL, Rosenthal R, Rubin DB. Comparing correlated correlation coefficients. Psychol Bull. 1992;111(1):172-5. [Crossref]
15. Hittner JB, May K, Silver NC. A Monte Carlo evaluation of tests for comparing dependent correlations. J Gen Psychol. 2003;130(2):149-68. [Crossref] [PubMed]
16. Pearson K, Filon LNG. Mathematical contributions to theory of evolution: IV. on the probable errors of frequency constants and on the influence of random selection and correlation. Proc. R. Soc. Lond. 1898;62(379):173-6. [Crossref]
17. Weaver B, Wuensch KL. SPSS and SAS programs for comparing Pearson correlations and OLS regression coefficients. Behav Res Methods. 2013;45(3):880-95. Erratum in: Behav Res Methods. 2013;45(3):896-7. [Crossref] [PubMed]
18. Fisher RA. On the 'probable error' of a coefficient of correlation deduced from a small sample. Metron. 1921;1:3-32. [Link]
19. Silver N, Dunlap WP. Averaging correlation coefficients: Should Fisher's z transformation be used? J Appl Psychol. 1987;72(1):146-8. [Crossref]
20. Bradley JV. Robustness. Br J Math Stat Psychol. 1978;31(2):144-51. [Crossref]
21. Rao CR. Linear Statistical Inference and Its Applications. 2nd ed. New York: John Wiley & Sons Inc.; 1973. [Crossref]
22. Córdoba I, Varando G, Bielza C, Larra-aga P. gmat: Simulation of graphical correlation matrices. R package version 0.2.2. CRAN, 2020.
23. Diedenhofen B, Musch J. cocor: a comprehensive solution for the statistical comparison of correlations. PLoS One. 2015;10(3):e0121945. Erratum in: PLoS One. 2015;10(6):e0131499. [Crossref] [PubMed] [PMC]
24. Brown-Forestiere R, Furiato A, Foresteire NP, Kashani JS, Waheed A. Acute appendicitis: clinical clues and conundrums related to the greatest misses. Cureus. 2020;12(5):e8051. [Crossref] [PubMed] [PMC]
25. Acharya A, Markar SR, Ni M, Hanna GB. Biomarkers of acute appendicitis: systematic review and cost-benefit trade-off analysis. Surg Endosc. 2017;31(3):1022-31. [Crossref] [PubMed] [PMC]
26. Asfar S, Safar H, Khoursheed M, Dashti H, al-Bader A. Would measurement of C-reactive protein reduce the rate of negative exploration for acute appendicitis? J R Coll Surg Edinb. 2000;45(1):21-4. [PubMed]
27. Beltrán MA, Almonacid J, Vicencio A, Gutiérrez J, Cruces KS, Cumsille MA. Predictive value of white blood cell count and C-reactive protein in children with appendicitis. J Pediatr Surg. 2007;42(7):1208-14. [Crossref] [PubMed]
28. Paajanen H, Mansikka A, Laato M, Kettunen J, Kostiainen S. Are serum inflammatory markers age dependent in acute appendicitis? J Am Coll Surg. 1997;184(3):303-8. [PubMed]
29. Silver C, Hitter JB, May K. Testing dependent correlations with non-overlapping variables: A Mote Carlo simulation. J Exp Educ. 2004;73(11):53-69. [Crossref]
30. Sheshkin DJ. Handbook of Parametric and Nonparametric Statistical Procedures. 3rd ed. Boca Raton FL: Chapman & Hall/CRC; 2004.
31. Zar JH. Biostatistical Analysis. 4th ed. Upper Saddle River NJ: Prentice Hall; 1999.