

SCPVTBT: A PROGRAM IN R TO SIMULTANEOUSLY COMPARE THE PREDICTIVE VALUES OF TWO BINARY DIAGNOSTIC TESTS IN PAIRED DESIGN

José Antonio Roldán Nofuentes^{*1}, Juan de Dios Luna del Castillo*, Miguel Angel Montero Alonso** & Saad Bouh Ould Sidaty***

*Biostatistics, Department of Statistics, School of Medicine, University of Granada, Avenida de Madrid s/n, 18071, Granada, Spain

**Department of Statistics, School of Social Science, Melilla, University of Granada, Spain

***School of Medicine, University of Nouakchott, Mauritania

RESUMEN

La comparación de los valores predictivos positivos y negativos de dos tests diagnósticos binarios es un tópico importante en el estudio de los métodos estadísticos para el diagnóstico médico. El programa SCPVTBT (Simultaneously Comparison of Predictive Values of Two Binary Tests) es un programa escrito en R que permite comparar simultáneamente los valores predictivos positivos y negativos de dos tests diagnósticos binarios en diseños apareados. Este programa resuelve el test de hipótesis global para comparar simultáneamente los valores predictivos al error α , y si el test global es significativo las causas de la significación se investigan comparando los valores predictivos positivos y los valores predictivos negativos de forma independiente y aplicando a continuación un método de comparaciones múltiples. El programa SCPVTBT está disponible libremente en la URL <http://www.ugr.es/~bioest/software.htm>.

KEY WORDS: Positive and negative predictive values, Comparison of two binary diagnostic tests, R software.

ABSTRACT

Comparison of the positive and negative predictive values of two binary diagnostic tests is an important topic in the statistical methods for medical diagnosis. The program SCPVTBT (Simultaneously Comparison of Predictive Values of Two Binary Tests) is a program written in R to simultaneously compare the positive and negative predictive values of two binary diagnostic tests in paired design. This program solves the global hypothesis test of simultaneously comparison of predictive values to an error rate of α , and if global hypothesis test is significant, the causes of the significance are carried out by comparing the positive predictive values and the negative predictive values independently and subsequently applying a method of multiple comparisons. The program SCPVTBT is freely available for download from the URL <http://www.ugr.es/~bioest/software.htm>.

PALABRAS CLAVE: Valores predictivos positivo y negativo, Comparación de dos tests diagnósticos binarios, Software R.

1. INTRODUCTION

A diagnostic test is a medical test that is applied to a individual in order to obtain a provisional diagnosis regarding the presence or absence of a disease. If the test outcome is binary, the accuracy of the test is measured in terms of sensitivity and specificity. Sensitivity is the probability of a positive test result given that the patient is diseased and specificity is the probability of a negative test result given that the patient is non-diseased. Another way to describe the diagnostic value of a diagnostic test are the positive and negative predictive values. Positive predictive value (*PPV*) and negative predictive value (*NPV*) represent the clinical accuracy of a binary diagnostic test when it is applied to a cohort of individuals. *PPV* is the probability of an individual being diseased when the test result is positive and *NPV* is the probability of an individual not being diseased when the test result is negative. A topic of interest in the study of statistical methods for medical diagnosis is to compare the predictive values (*PVs*) of two binary diagnostic tests with respect to the same gold standard. When two diagnostic tests and the gold standard are applied to the same random sample of subjects (paired design), the estimation and comparison of the positive (negative) predictive values has been the object of several studies [1, 2, 3]. In all these studies the *PPVs* and the *NPVs* of the two diagnostic tests are compared with one another independently, that is, two hypothesis tests are resolved, one for comparing the two *PPVs* and the other for comparing the two *NPVs*. Roldán Nofuentes et al [4] have shown that the comparison of *PPVs* and of *NPVs* of two binary diagnostic tests independently may produce errors. In order to solve this problem, Roldán Nofuentes et al [4] have proposed a method based on the chi-squared distribution, which allows the simultaneous comparison of *PPVs* and *NPVs* of two binary diagnostic tests in paired designs. In this

¹ jaroldan@ugr.es

report, we summarize the method of Roldán Nofuentes et al [4] and propose a program written in R [5] to solve this inference problem.

2. THE STATISTICAL METHOD

The results obtained by applying the two diagnostic tests and the gold standard to all individuals in a random sample are shown in Table 1.

	Test 1 positive		Test 1 negative		Total
	Test 2 positive	Test 2 negative	Test 2 positive	Test 2 negative	
Standard positive	s_{11}	s_{10}	s_{01}	s_{00}	s
Standard negative	r_{11}	r_{10}	r_{01}	r_{00}	r
Total	n_{11}	n_{10}	n_{01}	n_{00}	n

Table 1. Results for the study of coronary artery disease.

The simultaneously comparison of the PVs of two binary tests consist in solve the hypothesis test

$$H_0 : (PPV_1 = PPV_2 \text{ and } NPV_1 = NPV_2)$$

vs

$$H_1 : (PPV_1 \neq PPV_2 \text{ and/or } NPV_1 \neq NPV_2),$$

where PPV_i and NPV_i are the positive and negative predictive values of i th binary test ($i = 1, 2$). The maximum likelihood estimators of the predictive values are

$$\widehat{PPV}_1 = \frac{s_{10} + s_{11}}{s_{10} + s_{11} + r_{10} + r_{11}} \quad \text{and} \quad \widehat{NPV}_1 = \frac{r_{00} + r_{01}}{s_{00} + s_{01} + r_{00} + r_{01}}$$

for test 1, and

$$\widehat{PPV}_2 = \frac{s_{01} + s_{11}}{s_{00} + s_{11} + r_{01} + r_{11}} \quad \text{and} \quad \widehat{NPV}_2 = \frac{r_{00} + r_{10}}{s_{00} + s_{10} + r_{00} + r_{10}}$$

for test 2, and applying the delta method, the estimated variances-covariances of the estimators of the PVs are:

$$\begin{aligned} \widehat{V}(\widehat{PPV}_1) &\approx \frac{(s_{10} + s_{11})(r_{01} + r_{11})}{(s_{10} + s_{11} + r_{10} + r_{11})^2}, \quad \widehat{V}(\widehat{PPV}_2) \approx \frac{(s_{01} + s_{11})(r_{01} + r_{11})}{(s_{01} + s_{11} + r_{01} + r_{11})^2} \\ \widehat{V}(\widehat{NPV}_1) &\approx \frac{(s_{00} + s_{01})(r_{00} + r_{01})}{(s_{00} + s_{01} + r_{00} + r_{01})^2}, \quad \widehat{V}(\widehat{NPV}_2) \approx \frac{(s_{00} + s_{10})(r_{00} + r_{10})}{(s_{00} + s_{01} + r_{00} + r_{10})^2} \\ \widehat{Cov}(\widehat{PPV}_1, \widehat{PPV}_2) &\approx \frac{s_{01}s_{10}r_{11} + s_{11}\{r_{01}(r_{10} + r_{11}) + r_{11}(s_{01} + s_{10} + s_{11} + r_{10} + r_{11})\}}{(s_{01} + s_{11} + r_{01} + r_{11})^2(s_{10} + s_{11} + r_{10} + r_{11})^2} \\ \widehat{Cov}(\widehat{PPV}_1, \widehat{NPV}_2) &\approx \frac{s_{00}(s_{10} + s_{11})r_{10} + s_{10}r_{10}(s_{10} + s_{11} + r_{00} + r_{10}) + s_{10}(r_{00} + r_{10})r_{11}}{(s_{00} + s_{10} + r_{00} + r_{10})^2(s_{10} + s_{11} + r_{10} + r_{11})^2} \\ \widehat{Cov}(\widehat{PPV}_2, \widehat{NPV}_1) &\approx \frac{s_{00}(s_{01} + s_{11})r_{01} + s_{01}r_{01}(s_{01} + s_{11} + r_{00} + r_{01}) + s_{01}(r_{00} + r_{01})r_{11}}{(s_{00} + s_{01} + r_{00} + r_{01})^2(s_{01} + s_{11} + r_{01} + r_{11})^2} \\ \widehat{Cov}(\widehat{NPV}_1, \widehat{NPV}_2) &\approx \frac{s_{00}(r_{00} + r_{01})r_{10} + r_{00}\{r_{00}^2 + s_{01}s_{10} + s_{00}(s_{01} + s_{10} + r_{00} + r_{01})\}}{(s_{00} + s_{01} + r_{00} + r_{01})^2(s_{00} + s_{10} + r_{00} + r_{10})^2} \\ \widehat{Cov}(\widehat{PPV}_1, \widehat{NPV}_1) &\approx 0 \quad \text{and} \quad \widehat{Cov}(\widehat{PPV}_2, \widehat{NPV}_2) \approx 0 \end{aligned}$$

The contrast statistics for the hypothesis test $H_0 : (PPV_1 = PPV_2 \text{ and } NPV_1 = NPV_2)$ is

$$Q^2 = \hat{\boldsymbol{\eta}}^T \boldsymbol{\Phi}^T (\boldsymbol{\Phi} \hat{\boldsymbol{\Sigma}} \boldsymbol{\Phi}^T)^{-1} \boldsymbol{\Phi} \hat{\boldsymbol{\eta}},$$

where

$$\hat{\eta} = (PPV_1, \widehat{PPV}_2, \widehat{NPV}_1, \widehat{NPV}_2),$$

$\hat{\Sigma}$ is the estimated variance-covariance matrix of $\hat{\eta}$ and ϕ is the design matrix, i.e.

$$\phi = \begin{pmatrix} 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & -1 \end{pmatrix}.$$

The statistic Q^2 is distributed asymptotically according to a central chi-square distribution with two degrees of freedom if H_0 is true. To apply this method it is necessary that all predictive values can be estimated and that matrix $\phi \hat{\Sigma} \phi^T$ is non-singular. Therefore, the method cannot be apply if the are many observed frequencies are equal to zero. If this global hypothesis test is significant to the an error rate of α , the investigation of the causes of the significance is carried out by comparing the positive predictive values and the negative predictive values independently (our program uses the method of Wang et al [3]) and subsequently applying a method of multiple comparisons (our program uses the Hochberg method [6]) to the same error rate of α . Simulation experiments performed have shown that samples of between 300 and 500 subjects are required in order for the power of the global hypothesis test to be high (over 80%).

3. THE PROGRAM SCPVTBT

The program SCPVTBT (Simultaneously Comparison of Predictive Values of Two Binary Tests) is freely available for download from the <http://www.ugr.es/~bioest/software.htm>. This program calculates the values of the estimators of the PVs and their corresponding standard errors, solves firstly the global hypothesis test to an error rate of α , and if this hypothesis test is significant the causes of the significance are investigated by applying the method of Wang et al [3] and Hochberg's method [6] to the same error rate of α . For $\alpha = 5\%$, the command to run the SCPVTBT program is

$$\text{SCPVTBT}(s_{11}, s_{10}, s_{01}, s_{00}, r_{11}, r_{10}, r_{01}, r_{00}),$$

and

$$\text{SCPVTBT}(s_{11}, s_{10}, s_{01}, s_{00}, r_{11}, r_{10}, r_{01}, r_{00}, \alpha)$$

for another value of α . This program is very easy to use and it shows on the screen the results obtained by applying the said method, without the need to have an advanced knowledge of Statistics. The results are also saved as a text file called "Results.txt" in the same folder where the SCPVTBT program is running. The use of this program requires the installation of the free "R" software [5].

4. EXAMPLE

For a practical application, one can take the example used by Roldán Nofuentes et al [4] for the diagnosis of coronary artery disease. In Table 2 results obtained on applying two diagnostic tests are shown (Test 1: echocardiography with dobutamine, and Test 2: myocardial perfusion scintigraphy) and the gold standard (coronary angiography) on a sample of 548 males aged over 50. By applying the SCPVTBT program with $\alpha = 5\%$, the estimations of positive and negative predictive values for Test 1 are 82.84% and 87.50% respectively, and those for Test 2 are 81.54% and 84.99% respectively. The statistic for the global hypothesis test is $Q^2 = 4.2517$ (p -value = 0.1193), which is why the null hypothesis is not rejected to the error $\alpha = 5\%$. As a result, the homogeneity of the two positive predictive values and the two negative predictive values of the echocardiography with dobutamine and of the myocardial perfusion scintigraphy is not rejected.

	Test 1 positive		Test 1 negative		Total
	Test 2 positive	Test 2 negative	Test 2 positive	Test 2 negative	
Standard positive	152	17	7	36	212
Standard negative	25	10	11	290	336
Total	177	27	18	326	548

Table 2. Results for the study of coronary artery disease.

If the method of Wang et al [3] is applied in this example, the statistic for $H_0 : PPV_1 = PPV_2$ is $\chi^2 = 0.4508$ (p -value = 0.5020) and the statistic for $H_0 : NPV_1 = NPV_2$ is $\chi^2 = 4.2461$ (p -value = 0.0393). Thus, to error $\alpha = 5\%$ the equality of the positive predictive values for both diagnostic tests is not rejected, but the equality of the negative predictive values is rejected. In this example one can see that comparing the positive predictive values and comparing the negative predictive values of two diagnostic tests independently may produce erroneous conclusions.

Acknowledgements: This research was supported by the Spanish Ministry of Science, Grant Number MTM2009-08886. The authors would like to thank Prof. María del Mar Rueda and Prof. Carlos Bouza for her invitation to write this article for this book.

REFERENCES

- [1] BENNETT, B.M. (1972): On comparison of sensitivity, specificity and predictive value of a number of diagnostic procedures. **Biometrics**, 28, 793-800.
- [2] LEISENRING, W., ALONZO, T. & PEPE, M.S. (2000): Comparisons of predictive values of binary medical diagnostic tests for paired designs. **Biometrics**, 56, 345-351.
- [3] WANG, W., DAVIS, C.S. & SOONG, S.J. (2006): Comparison of predictive values of two diagnostic tests from the same sample of subjects using weighted least squares. **Statistics in Medicine**, 25, 2215-2229.
- [4] ROLDÁN NOFUENTES, J.A., LUNA DEL CASTILLO, J.D. & MONTERO ALONSO, M.A. (2012): Global hypothesis test to simultaneously compare the predictive values of two binary diagnostic tests. **Computational Statistics and Data Analysis, Special issue "Computational Statistics for Clinical Research"**, 56, 1161-1173.
- [5] R DEVELOPMENT CORE TEAM. (2011): R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, 2011, **URL** <http://www.R-project.org>.
- [6] HOCHBERG, Y. (1988): A sharper Bonferroni procedure for multiple tests of significance. **Biometrika**, 75, 800-802.