José Rafael Tovar, Jorge Alberto Achcar

Indexes to Measure Dependence between Clinical Diagnostic Tests: A Comparative Study

Revista Colombiana de Estadística, vol. 34, núm. 3, diciembre, 2011, pp. 433-450,

Universidad Nacional de Colombia

Colombia

Revista Colombiana de Estadística,

ISSN (Printed Version): 0120-1751

revcoles@unal.edu.co

Universidad Nacional de Colombia

Colombia

Revista Colombiana de Estadística
Diciembre 2011, volumen 34, no. 3, pp. 433 a 450

Indexes to Measure Dependence between Clinical Diagnostic Tests: A Comparative Study

Indices para medir dependencia entre pruebas para diagnóstico clínico: un estudio comparativo

José Rafael Tovar[1,a], Jorge Alberto Achcar[2,b]

[1]Departamento de Estatística, Instituto de Matemática Estatística e Computação Científica, Universidade Estadual de Campinas, Campinas, Brasil

[2]Departamento de Medicina Social FMRP, Faculdade de Saúde, Universidade de São Paulo, Riberão Preto, Brasil

Abstract

In many practical situations, clinical diagnostic procedures include two or more biological traits whose outcomes are expressed on a continuous scale and are then dichotomized using a cut point. As measurements are performed on the same individual there is a likely correlation between the continuous underlying traits that can go unnoticed when the parameter estimation is done with the resulting binary variables. In this paper, we compare the performance of two different indexes developed to evaluate the dependence between diagnostic clinical tests that assume binary structure in the results with the performance of the binary covariance and two copula dependence parameters.

***Key words*:** Copula, Farlie Gumbel Morgenstern distribution, Gumbel distribution.

Resumen

Muchos procedimientos de diagnóstico clínico médico exigen la evaluación de dos o mas rasgos biológicos que se ven alterados ante la presencia de fenómenos de enfermedad o infección, los cuales se expresan en una escala contínua de medición con posterior dicotomización usando de un valor límite o punto de corte. Dado que las mediciones son realizadas en el mismo indivíduo, los resultados probablemente presenten dependencia de algún tipo, lo cual puede ser ignorado en la etapa de análisis de datos dada la presentación binaria de los datos. En este estudio comparamos el comportamiento de dos parámetros de dependencia presentes en funciones de cópula con el de la covarianza binaria y dos índices creados para medir dependencia entre pruebas diagnósticas de respuesta dicótoma.

***Palabras clave*:** cópula, distribución Farlie, Gumbel.

[a]Professor. E-mail: rtovar34@hotmail.com
[b]Professor. E-mail: achacar@fmrp.usp.br

1. Introduction

The study of dependence between two clinical diagnostic tests has been a matter of interest in medical and statistical research. Many studies developed within clinical diagnostic tests framework have studied the conditional dependence between diagnostic tests using the binary covariance as dependence parameter (see for instance: (Thibodeau 1981), (Vacek 1985), (Torrance-Rynard & Walter 1997), (Enoe, Georgiadis & Johnson 2000) and (Dendukuri & Joseph 2001) among many others). Some authors as Georgiadis, Johnson & Gardner (2003), have used reparametrizations of the conditional correlation between binary tests to facilitate the prior specification in the implementation of a Bayesian estimation procedure. Bohning & Patilea (2008), consider one of the indexes used by Georgiadis et al. (2003) and developed another one to study the association between two diagnostic tests in designs where the individuals with negative outcome in both screening tests are not verified by "gold standard" (verification bias conditions).

Many diagnostic procedures include the measures of two or more biological traits directly observable or not, whose outcomes are initially expressed on a continuous scale and operationalized within a dichotomous representation using a cut point. It is possible that, there exists dependence between the two evaluated traits conditional on the true disease state and the same should be studied using indexes developed to study association between continuous variables, but as the data analyses is made with the binary data, the data analyst evaluates the conditional dependence hypothesis using indexes developed to binary variables.

In this paper, we use the indexes developed by Bohning & Patilea (2008) and we compare their performance with the performance of the binary covariance and the performance of the Farlie Gumbel Morgerstern (FGM) and Gumbel copula dependence parameters. The main goal is to evaluate the existing relationship among the five dependence parameters, where three of them (covariance and Böhning's indexes) are built to study dependence between binary variables and the other two (copula parameters) are developed to model dependence between continuous variables.

The paper is organized as follows: in Section 2, we introduce the estimation model formulation for two associated diagnostic tests, in Section 3, we present the comparative study among Böhning and Patilea's indexes, the binary covariance and the copula dependence parameters, in Section 4, we introduce two examples, one of them with simulated data and the other with published data. Finally, in Section 5, we present some conclusions on the results obtained.

2. Statistical Model with Two Dependent Screening Tests

Let us assume that we have a clinical diagnostic procedure that uses two screening tests whose performance we are interested to study and a reference procedure that classifies individuals as diseased and non-diseased without error called "gold

standard". Sometimes, the design of the study considers that those individuals with negative outcomes in both screening tests are not verified by the "gold standard" which is known as "verification bias". We assume that, the screening test outcomes are expressed on a continuous scale and they are exposed to a process of dichotomization using a cut point. We also assume that the test outcomes have a continuous dependent structure but the same can not be considered in the data analysis since they are presented in a binary form.

2.1. Modelling Dependence with Binary Structure

Let us denote by p the prevalence of a disease and by D a random variable related to the true disease status, where $D = 1$ denotes a diseased individual and $D = 0$ denotes a non-diseased individual. That is, $p = P(D = 1)$. Also, denote by T_1 and T_2, the two random variables associated to the test results, where $T_j = 1$, denotes a positive result and $T_j = 0$, denotes a negative result, $P(T_j = 1 \mid D = 1) = S_j$ is the sensitivity of the test j and $P(T_j = 0 \mid D = 0) = E_j$ is the specificity of the test j, for $j = 1, 2$. If we assume that, the dependence between tests can be modeled by the covariance (ψ parameter) using a Bernoulli distribution on the test outcome and we also assume the covariance is not necessarily the same in both populations ($\psi_D \neq \psi_{ND}$), we can use the Dendukuri's procedure to obtain the likelihood function contributions, as shown in Table 1.

Table 1: Likelihood contributions of all possible combinations of outcomes of T_1, T_2 and D assuming binary dependence structure (Values in brackets are unknown under verification bias. f_i: number of individuals for each combination of results)

				Contribution to likelihood
D	T_1	T_2	f_i	Binary dependence
1	1	1	a	$p[S_1S_2 + \psi_D]$
1	1	0	b	$p[S_1(1 - S_2) - \psi_D]$
1	0	1	c	$p[(1 - S_1)S_2 - \psi_D]$
1	0	0	$[d]$	$p[(1 - S_1)(1 - S_2) + \psi_D]$
0	1	1	e	$(1 - p)[(1 - E_1)(1 - E_2) + \psi_{ND}]$
0	1	0	f	$(1 - p)[(1 - E_1)E_2 - \psi_{ND}]$
0	0	1	g	$(1 - p)[E_1(1 - E_2) - \psi_{ND}]$
0	0	0	$[h]$	$(1 - p)[E_1E_2 + \psi_{ND}]$

2.2. Modelling Dependence using Copula Functions

Let us assume that the test outcomes are realizations of the random variables V_1 and V_2 measured in a positive continuous scale, that is, $V_1 > 0$ and $V_2 > 0$. Also, let us assume that two cut-off values ξ_1 and ξ_2 are chosen for each test in order to determine when an individual is classified as positive or negative. In this way we assume that an individual is classified as positive for test ν if $V_\nu > \xi_\nu$

that is, $T_\nu = 1$ if and only if $V_\nu > \xi_\nu$ for $\nu = 1, 2$. To measure the degree of the dependence structure between the random variables V_1 and V_2, let us consider the use of copula functions (For details about this topic, (Nelsen 1999) is a good reference). For specified univariate marginal distribution functions $F_1(v_1), \ldots, F_m(v_m)$, the function $C(F_1(v_1), \ldots, F_m(v_m))$ which is defined using a copula function C, results in a multivariate distribution function with univariate marginal distributions specified as $F_1(v_1), \ldots, F_m(v_m)$. Any multivariate distribution function F can be written in the form of a copula function, that is, if $F(v_1, \ldots, v_m)$ is a joint multivariate distribution function with univariate marginal distribution functions $F_1(v_1), \ldots, F_m(v_m)$, thus there exists a copula function $C(u_1, \ldots, u_m)$ such that, $F(v_1, \ldots, v_m) = C(F_1(v_1), \ldots, F_m(v_m))$. For the special case of bivariate distributions, we have $m = 2$. The approach to formulate a multivariate distribution using a copula is based on the idea that a simple transformation can be made of each marginal variable in such a way that each transformed marginal variable has an uniform distribution. Once this is done, the dependence structure can be expressed as a multivariate distribution on the obtained uniforms and a copula is precisely a multivariate distribution on marginally uniform random variables. In this way, there are many families of copulas which differ in the detail of the dependence they represent. In the bivariate case, let V_1 and V_2 be two random variables with continuous distribution functions F_1 and F_2. The probability integral transformation is applied separately for the two random variables to define $U = F_1(V_1)$ and $W = F_2(V_2)$ where U and W have uniform $(0, 1)$ distributions, but are usually dependent if V_1 and V_2 are dependent (V_1 and V_2 independent implies that U and W are independent). Specifying dependence between V_1 and V_2 is the same as specifying dependence between U and W, thus the problem reduces to specifying a bivariate distribution between two uniform variables, that is a copula. In this paper, we use two copula functions to study the dependence between two diagnostic tests namely: the Farlie Gumbel Morgerstern (FGM) copula and the Gumbel copula.

The FGM copula is defined by,

$$C(u, w) = uw[1 + \varphi(1 - u)(1 - w)] \tag{1}$$

where $u = F_1(v_1)$, $w = F_2(v_2)$ and $-1 \leq \varphi \leq 1$. As, φ measures the dependence between the two marginals, then, if $\varphi = 0$, we have independent random variables. We assume two dependence parameters φ_D and φ_{ND} with the same value for diseased and non-diseased individuals, respectively. From (1), the cumulative joint distribution and the joint survival distribution functions for the random variables V_1 and V_2 conditional on the diseased status (D subscript and superscript) are given respectively by,

$$F(v_1, v_2) = C(F_1(v_1), F_2(v_2)) = F_1(v_1)F_2(v_2) \\ [1 + \varphi(1 - F_1(v_1))(1 - F_2(v_2))] \tag{2}$$

$$S(v_1, v_2) = P(V_1 > v_1, V_2 > v_2) = 1 - F_1(v_1) - F_2(v_2) + F(v_1, v_2) \tag{3}$$

To obtain the likelihood function contributions within the diseased individuals group we have;

$$P(T_1 = 1, T_2 = 1 \mid D = 1) = P(V_1 > \xi_1, V_2 > \xi_2 \mid D = 1) = S_D(\xi_1, \xi_2),$$

$$P(T_1 = 1 \mid D = 1) = P(V_1 > \xi_1 \mid D = 1) = S_1,$$

$$P(T_2 = 1 \mid D = 1) = P(V_2 > \xi_2 \mid D = 1) = S_2,$$

$$F_1^D(\xi_1) = P(V_1 \leq \xi_1 \mid D = 1) = 1 - S_1,$$

and

$$F_2^D(\xi_2) = P(V_2 \leq \xi_2 \mid D = 1) = 1 - S_2$$

Using (2) we get,

$$\begin{aligned} F_D(\xi_1, \xi_2) &= F_1^D(\xi_1) F_2^D(\xi_2)[1 + \varphi(1 - F_1^D(\xi_1))(1 - F_2^D(\xi_2))] \\ &= (1 - S_1)(1 - S_2)(1 + \varphi_D S_1 S_2) \end{aligned}$$

and,

$$\begin{aligned} P(T_1 = 1, T_2 = 1 \mid D = 1) = S_D(\xi_1, \xi_2) = 1 - (1 - S_1) - (1 - S_2) + \\ (1 - S_1)(1 - S_2)(1 + \varphi_D S_1 S_2) \end{aligned}$$

That is,

$$P(T_1 = 1, T_2 = 1 | D = 1) = S_1 S_2 (1 + \varphi_D (1 - S_1)(1 - S_2)) \tag{4}$$

and

$$P(T_1 = 1, T_2 = 1, D = 1) = p S_1 S_2 (1 + \varphi_D (1 - S_1)(1 - S_2)) \tag{5}$$

Similarly, we get all likelihood contributions with diseased and non-diseased individuals (see Table 2).

The Gumbel copula, developed by Gumbel (1960) is defined as,

$$C(u, w) = u + w - 1 + (1 - u)(1 - w) \exp\{-\phi \ln(1 - u) \ln(1 - w)\} \tag{6}$$

In this model, the joint cumulative distribution function for the random variables V_1 and V_2 is given by,

$$\begin{aligned} F(v_1, v_2) = F_1(v_1) + F_2(v_2) - 1 + \\ (1 - F_1(v_1))(1 - F_2(v_2)) \exp\{-\phi \ln(1 - F_1(v_1)) \ln(1 - F_2(v_2))\} \end{aligned} \tag{7}$$

The dependence parameter of the Gumbel copula does not model positive linear correlations and when the two variables are independents, we have $\phi = 0$.

Employing the same arguments considered in the FGM copula to find the joint probabilities of all combinations with D, T_1 and T_2 and using (7) we obtain all the contributions for the likelihood function using the Gumbel copula (See Table 2).

Table 2: Likelihood contributions of all possible combinations of outcomes of T_1, T_2 and D when the dependence has the FGM copula or Gumbel copula structure. (Values in brackets are unknown under verification bias. f_i: number of individuals for each combination of results)

				Contribution to likelihood	
D	T_1	T_2	f_i	FGM copula	Gumbel copula
1	1	1	a	$pS_1S_2[1+\varphi_D(1-S_1)(1-S_2)]$	$pS_1S_2Q_1$
1	1	0	b	$pS_1(1-S_2)[1-\varphi_D(1-S_1)S_2]$	$pS_1[1-S_2Q_1]$
1	0	1	c	$p(1-S_1)S_2[1-\varphi_D S_1(1-S_2)]$	$pS_2[1-S_1Q_1]$
1	0	0	[d]	$p(1-S_1)(1-S_2)[1+\varphi_D S_1S_2]$	$p[1-S_1-S_2+S_1S_2Q_1]$
0	1	1	e	$(1-p)(1-E_1)(1-E_2)[1+\varphi_{ND}E_1E_2]$	$(1-p)(1-E_1)(1-E_2)Q_2$
0	1	0	f	$(1-p)(1-E_1)E_2[1-\varphi_{ND}E_1(1-E_2)]$	$(1-p)(1-E_1)[1-(1-E_2)Q_2]$
0	0	1	g	$(1-p)E_1(1-E_2)[1-\varphi_{ND}E_2(1-E_1)]$	$(1-p)(1-E_2)[1-(1-E_1)Q_2]$
0	0	0	[h]	$(1-p)E_1E_2[1+\varphi_{ND}(1-E_1)(1-E_2)]$	$(1-p)[E_1+E_2-1+(1-E_1)(1-E_2)Q_2]$

$Q_1 = \exp(-\phi_D \ln S_1 \ln S_2), \quad Q_2 = \exp(-\phi_{ND} \ln(1-E_1)\ln(1-E_2))$

3. Indexes Developed by Böhning and Patilea

Bohning & Patilea (2008), developed two association indexes to study the case of two dependent diagnostic tests in situations where it is not possible to verify the true disease status in individuals with negative outcome in both screening tests. The authors proposed computation of the indexes using the observed probabilities in the likelihood function. The Böhning and Patilea's indexes θ_i and α_i ($i = D$ denotes diseased individuals and $i = ND$ denotes non-diseased individuals) are defined as:

$$\begin{aligned} \theta_i &= \frac{P(T_1 = 1 \mid T_2 = 1, D = i)}{P(T_1 = 1, D = i)} \\ &= \frac{P(T_1 = 1, T_2 = 1, D = i)}{P(T_1 = 1, D = i)P(T_2 = 1, D = i)} \quad \theta_i \in (0, \infty) \end{aligned} \tag{8}$$

If $\theta_i = 1$ the tests results are independent; if $\theta_i < 1$ there is negative association between tests and if $\theta_i > 1$, the association between tests is positive.

$$\alpha_i = \frac{P(T_1 = 1, T_2 = 1, D = i)P(T_1 = 0, T_2 = 0, D = i)}{P(T_1 = 1, T_2 = 0, D = i)P(T_1 = 0, T_2 = 1, D = i)} \quad \alpha_i \in (0, \infty) \tag{9}$$

Thus, α_i is defined as the odds ratio in the ith diseased state, and when $\alpha_i = 1$ we have independence between tests; negative dependence is expressed by $\alpha_i < 1$ and positive dependence by $\alpha_i > 1$.

In spite of the fact that, both indexes measure dependence and they are within of the same range of values, they are different in nature. To establish the relationship between them, the authors considered a reparametrization given by: $a_i = \theta_i P(T_1 = 1 \mid D = i) = P(T_1 = 1 \mid T_2 = 1, D = i)$, $b_i = \theta_i P(T_2 = 1 \mid D = i) = P(T_2 = 1 \mid T_1 = 1, D = i)$ and $\eta_i = \frac{1}{\theta_i}$.

Let us rewrite the cell probabilities in the cross-tabulation as:

$$P(T_1 = 1, T_2 = 1, D = i) = \eta a_i b_i,$$

$$P(T_1 = 0, T_2 = 1 = \eta a_i(1 - b_i)),$$

$$P(T_1 = 1, T_2 = 0, D = i) = \eta(1 - a_i)b_i$$

and

$$P(T_1 = 0, T_2 = 0, D = i) = \eta(1 - a_i)(1 - b_i) + 1 - \eta$$

In this way, the α_i parameter can be expressed in terms of the parameter η_i, by,

$$\alpha_i = 1 + \frac{1 - \eta_i}{\eta_i(1 - a_i)(1 - b_i)} \tag{10}$$

The BP indexes were developed assuming that the tests have the same dependence within the disease and non-disease populations ($\theta_D = \theta_{ND}$ and $\alpha_D = \alpha_{ND}$) and they are useful when the design of the study does not include the verification with "gold standard" of those individuals with negative outcome in both screening tests. So, using the θ index, we can to estimate the unknown quantities of disease and non-disease individuals,

$$n_D = a + b + c + [d] \quad \text{and} \quad n_{ND} = e + f + g + [h]$$

as follows:

$$\begin{aligned} \widehat{n}_D &= \widehat{\theta}\left\{\frac{(a+b+1)(a+c+1)}{(a+1)} - 1\right\} = \widehat{\theta}q_1 \\ \widehat{n}_{ND} &= \widehat{\theta}\left\{\frac{(e+f+1)(e+g+1)}{(e+1)} - 1\right\} = \widehat{\theta}q_2 \end{aligned} \tag{11}$$

On the other hand, with the α index we can to estimate the unknown quantities d and h, as follows:

$$\begin{aligned} \widehat{d} &= \widehat{\alpha}\left\{\frac{(b+1)(c+1)}{(a+1)} - 1\right\} = \widehat{\alpha}(r_1 - 1) \\ \widehat{h} &= \widehat{\alpha}\left\{\frac{(f+1)(g+1)}{(e+1)} - 1\right\} = \widehat{\alpha}(r_2 - 1) \end{aligned} \tag{12}$$

where

$$\widehat{\theta} = \frac{n}{q_1 + q_2}, \qquad \widehat{\alpha} = \frac{u+2}{r_1 + r_2}$$

u is the quantity of individuals not verified by the "gold standard" and n is the total of participants in the screening study.

Assuming the three dependence structures for the two diagnostic tests, using the results showed in Tables 2 and 3 and the equations (8), (9), we obtained the analytic relationship between θ_i and α_i with ψ_i φ_i and ϕ_i.

- Diseased individuals population:

Binary Covariance

$$\theta_D = 1 + \frac{\psi_D}{S_1 S_2},$$

$$\alpha_D = \frac{[S_1 S_2 + \psi_D][(1-S_1)(1-S_2) + \psi_D]}{[S_1(1-S_2) - \psi_D][S_2(1-S_1) - \psi_D]},$$

$$\varphi_D = \psi_D [S_1 S_2 (1-S_1)(1-S_2)]^{-1},$$

$$\phi_D = -[\ln S_1 \ln S_2]^{-1} \ln[\psi_D S^{-1} S_2^{-1} + 1]$$

FGM copula

$$\theta_D = 1 + \varphi_D (1-S_1)(1-S_2);$$

$$\alpha_D = \frac{S_1 S_2 [1 + \varphi_D (1-S_1)(1-S_2)](1-S_1)(1-S_2)[1 + \varphi_D S_1 S_2]}{S_1(1-S_2)[1 - \varphi_D S_2 (1-S_1)] S_2 (1-S_1)[1 - \varphi_D S_1 (1-S_2)]}$$

Gumbel copula

$$\theta_D = \exp\{-\phi_D \ln S_1 \ln S_2\},$$

$$\alpha_D = \frac{\exp\{-\phi_D \ln S_1 \ln S_2\}[S_1 S_2 (1-S_1)(1-S_2)]}{[S_1 - S_1 S_2 \exp\{-\phi_D \ln S_1 \ln S_2\}][S_2 - S_1 S_2 \exp\{-\phi_D \ln S_1 \ln S_2\}]}$$

- Non-diseased individuals population:

Binary Covariance,

$$\theta_{ND} = 1 + \frac{\psi_{ND}}{(1-E_1)(1-E_2)},$$

$$\alpha_{ND} = \frac{[(1-E_1)(1-E_2) + \psi_{ND}][E_1 E_2 + \psi_{ND}]}{[(1-E_1)E_2 - \psi_{ND}][E_1(1-E_2) - \psi_{ND}]},$$

$$\varphi_{ND} = \psi_D [E_1 E 2 (1-E_1)(1-E_2)]^{-1},$$

$$\phi_{ND} = -[\ln(1-E_1) \ln(1-E_2)]^{-1} \ln[\psi_{ND} (1-E_1)^{-1} (1-E_2)^{-1} + 1]$$

FGM copula,

$$\theta_{ND} = 1 + \varphi_{ND} E_1 E_2,$$

$$\alpha_{ND} = \frac{(1-E_1)(1-E_2)[1 + \varphi_{ND} E_1 E_2] E_1 E_2 [1 + \varphi_{ND}(1-E_1)(1-E_2)]}{E_1(1-E_2)[1 - \varphi_{ND} E_2 (1-E_1)] E_2 (1-E_1)[1 - \varphi_{ND} E_1 (1-E_2)]}$$

Gumbel copula,

$$\theta_{ND} = \exp\{-\phi_{ND} \ln(1-E_1) \ln(1-E_2)\},$$

$$\alpha_{ND} = \frac{[(1-E_1)(1-E_2)\exp\{-k\}][E_1 + E_2 - 1 + (1-E_1)(1-E_2)\exp\{-k\}]}{[(1-E_1) - (1-E_1)(1-E_2)\exp\{-k\}][(1-E_2) - (1-E_1)(1-E_2)\exp\{-k\}]}$$

where, $k = \phi_{ND} \ln(1-E_1) \ln(1-E_2)\}$.

For two independent tests, we obtain $\lambda_i = 1$ and $\delta_i = 1$ when $\psi_i = 0$, $\varphi_i = 0$ and $\phi_i = 0$, regardless of the performance test values.

In all cases, the Böhning and Patilea's association indexes (BP indexes) are functions of the performance characteristics of the tests, and when the performance parameters are going to one or zero, the BP indexes could go to infinity or to be indeterminate. We have in Table 3, the limit values of the BP indexes when the test parameters are going to zero or one and the dependence coefficients are fixed at their extreme values.

TABLE 3: Limits of θ and α indexes when the performance test parameters (PTP), are going to zero or one and $\psi_i = \pm 1$, $\varphi_i = \pm 1$ and $\phi_i = 1$. (for diseased individuals, PTP are S_1 and S_2; for non-diseased individuals, PTP are E_1 and E_2)

				Limit values of θ and α				
Coefficient	Population	Limit values of PTP		$\psi_i = -1$	$\psi_i = 1$	$\varphi_i = -1$	$\varphi_i = 1$	$\phi_i = 1$
θ_D	Diseased individuals	0	0	$-\infty$	$+\infty$	0	2	0
		1	1	0	2	1	1	1
		0	1	$-\infty$	$+\infty$	1	1	1
		1	0	$-\infty$	$+\infty$	1	1	1
θ_{ND}	Non-diseased individuals	0	0	0	2	1	1	1
		1	1	$-\infty$	$+\infty$	0	2	0
		0	1	$-\infty$	$+\infty$	1	1	1
		1	0	$-\infty$	$+\infty$	1	1	1
α_D	Diseased individuals	0	0	0	2	0	2	0
		1	1	0	2	0	2	0
		0	1	1/2	$+\infty$	1/2	$+\infty$	0
		1	0	1/2	$+\infty$	1/2	$+\infty$	0
α_{ND}	Non-diseased individuals	0	0	0	2	0	2	0
		1	1	0	2	0	2	0
		0	1	1/2	$+\infty$	1/2	$+\infty$	0
		1	0	1/2	$+\infty$	1/2	$+\infty$	0

The relationship between covariance and copula parameters is not shown in Table 3, given that the covariance has zero as limit value in all combinations of extreme values made with the copula parameters and performance test parameters.

Observe that, under the hypothetical situation where we have two binary tests with the same perfect association within each individuals group ($\psi_D = -1$ and $\psi_{ND} = -1$ or $\psi_D = 1$ and $\psi_{ND} = 1$) if the tests have perfect negative association, we need for both tests to have absolutely perfect sensitivities ($S_1 = S_2 = 1$) and absolutely imperfect specificities ($E_1 = E_2 = 0$), to model association using θ_i; otherwise, we can not use it. If the tests have perfect positive association, we can model associations with θ_i values belonging to the interval $[2, \infty)$ provided that the performance parameters belong to the interval $(0, 1)$. In this way, values of θ_i very close to 2, will be related with ψ_i values close of zero.

On the other hand, if we have two tests, whose perfectly negative or positive dependence structure can be modeled using the FGM copula, we only can have agreement between the copula parameter and the BP indexes, when both sensitivities are equal to zero and both specificities are equal to one; then, under those conditions, we can only model associations when θ_i belongs to the interval $[0, 2]$. When the test parameters take values inside the within $(0, 1)$, the θ_i param-

eter would be indicating independence between tests while the FGM is indicating strong dependence between them.

For the Gumbel copula, we evaluated the extreme value 1 because this model is only applicable for positive dependence and 0 indicates independence. When the Gumbel dependence parameter indicates perfect positive dependence between two tests, both with absolutely imperfect sensitivities $S_1 = S_2 = 0$ or both with absolutely perfect specificities $E_1 = E_2 = 1$, the θ_i index takes the zero value indicating perfect negative association. When the tests have performance parameters belonging to interval $(0, 1)$, the θ_i parameter indicates independence between tests when those have perfect Gumbel dependence.

When the diagnostic tests have perfect FGM dependence, the θ index indicates independence and only when both tests have perfect specificities and absolutely imperfect sensitivity ($S_j = 0$), the index expresses a very weak association between tests ($\theta \in [0, 2]$).

Based on these facts, we observe that, the α_i parameter has a performance better than the θ_i parameter in their relations with the other parameters of association. For all combinations of sensitivity and specificity, α_i takes values within the range allowed by definition. When we have two binary tests negatively or positively associated with extreme values in their tests parameters, α_i takes values in the interval $[0, 2]$ for both populations, whereas, when the two tests have performance parameters belonging to the interval $(0, 1)$, the α_i parameter takes values in the interval $[1/2, \infty)$ for diseased and non-diseased individuals. For tests with dependence structure modeled by the FGM copula, the behaviour of α_i within groups of individuals is similar to that observed when we have two binary tests. When de dependence structure responds to the perfectly dependent Gumbel copula, in both populations, α_i indicates independence between tests regardless of the values of their performance parameters.

4. Examples

To illustrate the performance of the indexes, we show two examples, one of them with simulated data and the other one with a data set used by Bohning & Patilea (2008) to illustrate their methodology.

4.1. Example with Simulated Data

As a first example, we simulated 10000 pairs of observations with binary dependence structure and the same number of pairs of data for each copula structure (1000 diseased individuals and 9000 non-diseased individuals), considering the following conditions:

- Three dependence levels: weak (0.2), moderate (0.5) and strong (0.9),
- The dependence is the same in both populations ($\psi_D = \psi_{ND}$, $\varphi_D = \varphi_{ND}$ and $\phi_D = \phi_{ND}$)

- The specificities of the dependent tests are the same ($E_2 = E_3 = 0.95$) and the prevalence is relatively lower ($p = 0.10$)
- Stage 1: the dependent tests have the same relatively high sensitivities ($S_1 = S_2 = 0.85$)
- Stage 2: the dependent tests have the same relatively low sensitivities ($S_1 = S_2 = 0.45$)

We wrote a program in R to simulate pairs of variates with the different dependence forms. To simulate outcomes of the correlated binary variables Z_1, Z_2 we implemented the algorithm developed by Park, Park & Shin (1996) and to simulate the variables T_1, T_2 with FGM structure and the variables V_1,V_2 with Gumbel structure, we implemented algorithms introduced by Johnson (1987) as follows:

1. Binary data (ψ is the correlation coefficient)
   - Initialize p_1, p_2, $q_1 = 1 - p_1$, $q_2 = 1 - p_2$ and ψ_{12}
   - Let $\lambda_{11} = log\left\{1 + q_1 p_1^{-1}\right\}$, $\lambda_{22} = log\left\{1 + q_2 p_2^{-1}\right\}$ and $\lambda_{12} = \left\{1 + \psi_{12}\sqrt{\frac{q_1 q_2}{p_1 p_2}}\right\}$
   - Generate $X_1 \sim$ Poisson$(\lambda_{11} - \lambda_{12})$, $X_2 \sim$ Poisson$(\lambda_{22} - \alpha_{12})$ and $X_3 \sim$ Poisson(λ_{12})
   - Set $Y_1 = X_1 + X_3$ and $Y_2 = X_2 + X_3$
   - Set $Z_1 = 1$ if $Y_1 = 0$, else $Z_1 = 0$ and $Z_2 = 1$ if $Y_2 = 0$, else $Z_2 = 0$
   - Then, $Z_j \sim$ Bernoulli (p_j); $j = 1, 2$ and ψ_{12} is the correlation coefficient.
2. FGM data (φ is the dependence parameter)
   - Initialize φ
   - Generate variates $U_1 \sim U(0, 1)$, and $U_2 \sim U(0, 1)$
   - Set

$$\begin{aligned} T_1 &= U_1 \\ A &= \varphi(2U_1 - 1) - 1 \\ B &= [1 - 2\varphi(2U_1 - 1) + \varphi^2(2U_1 - 1)^2 + 4\varphi U_2(2U_1 - 1)]^{1/2} \\ T_2 &= 2U_2/(B - A) \end{aligned}$$

3. Gumbel data (ϕ is the dependence parameter)
   - Initialize ϕ
   - Generate $U_1 \sim U(0, 1)$, $U_2 \sim U(0, 1)$ and $U_3 \sim U(0, 1)$
   - Set $W_1 = -\ln(U_1)$ and $Y = -\ln(U_2)$
   - Compute $\beta = 1 + \phi W_1$ and $q = (\beta - \phi)/\beta$

- If $U_3 < q$, set $W_2 = \beta Y$ stop
- If $U_3 \geq q$, generate $U_4 \sim U(0,1)$, set $X_2 = \beta(Y - \ln U_4)$ and stop
- Let $V_1 = 1 - e^{-W_1}$ and $V_2 = 1 - e^{-W_2}$

As our data resulted from simulation, so we know all frequencies of individuals, but for the data analysis, we assume that we only have the total number of individuals with negative results in both tests.

The data analysis was made using the Bayesian paradigm, for that, we assumed that the screening tests have positive dependence ($P(\psi < 0) = P(\varphi < 0) = 0$) and we used the Beta(17,122), Beta(39.5, 39.5) and Beta(122, 17) as informative prior distributions for the weak, moderate and strong dependences respectively. To obtain the estimates we used a code in Winbugs software and we simulate $60,000$ Gibbs samples from the conditional distribution of each parameter. From these generated samples, we discarded the first $10,000$ samples to eliminate the effect of the initial values. The results obtained are showed in Table 4

Table 4: Simulated data with three different dependence structures in two scenarios and BP indexes estimates.

Scenary	f_i	$\psi = 0.2$	$\psi = 0.5$	$\psi = 0.9$	$\varphi = 0.2$	$\varphi = 0.5$	$\varphi = 0.9$	$\phi = 0.2$	$\phi = 0.5$	$\phi = 0.9$
	a	745	800	832	712	725	721	652	551	462
	b	116	68	13	131	137	133	152	175	190
	c	93	59	12	133	116	123	153	195	188
$S_1 = 0.85$	d	46	73	143	24	22	23	43	79	160
$S_2 = 0.85$	e	95	241	405	20	22	17	23	14	7
$E_1 = 0.95$	f	332	207	53	430	438	417	354	294	249
$E_2 = 0.95$	g	352	213	46	427	420	451	360	308	266
	h	8221	8339	8496	8123	8120	8115	8306	8384	8478
	θ	3.37	5.63	7.31	0.94	1.01	0.81	1.22	1.33	1.74
	$\hat{\alpha}$	6.67	44.5	1336	0.93	1.01	0.77	0.94	1.43	2.00
	a	238	350	433	198	209	201	275	334	406
	b	179	119	21	245	254	261	246	220	214
	c	214	105	33	239	239	238	223	246	209
$S_1 = 0.45$	d	369	426	513	298	298	300	256	200	171
$S_2 = 0.45$	e	95	241	405	20	22	17	19	17	15
$E_1 = 0.95$	f	332	207	53	430	438	417	353	329	255
$E_2 = 0.95$	g	352	213	46	427	420	451	365	320	264
	h	8221	8339	8496	8123	8120	8115	8263	8332	8466
	θ	3.59	6.88	10	0.94	0.94	0.81	1.23	1.34	1.76
	$\hat{\alpha}$	6.20	39.8	1130	0.93	0.93	0.78	1.28	1.41	1.99

The results presented in Table 4, confirm those shown in Table 3. When the data have a binary structure with linear dependence, both BP indexes tend to have high values. It is important to point out that, the sensibility has little effect on the index estimates but with low sensitivities the θ index shows a slight increase and the α index shows a opposite behaviour. If the data have a low or moderate FGM dependence both indexes express independence while for high FGM dependency both indexes indicate negative association in the data, that behaviour remains independent of the sensitivity of the tests. With low or moderate type Gumbel

dependences, the BP indexes indicate independence between tests, while with high Gumbel dependences the indexes express low dependence.

TABLE 5: Estimates of BP indexes and unknown quantities of diseased and non-diseased individuals within group with negative outcome in both screening tests, using data with binary and copula dependence. (n_D = 1,000; n_{ND} = 9,000 and $n_{(D+ND)}$ = 10,000)

Scenary	Dependence	BP Index	$\widehat{n}_D$	$\widehat{n}_{ND}$	$\widehat{d}$	$\widehat{h}$	$\widehat{n}_{(D+ND)}$
	$\psi = 0.2$	$\widehat{\theta} = 3.37$	3,264	6,728	2,310	5,949	9,991
		$\widehat{\alpha} = 6.67$	1,046	8,943	92	8,164	9,989
$S_1 = S_2 = 0.85$	$\psi = 0.5$	$\widehat{\theta} = 5.63$	5,247	4,747	4,320	4,086	9,994
$E_1 = E_2 = 0,95$		$\widehat{\alpha} = 44.5$	1,113	8,801	186	8,104	9,913
	$\psi = 0.9$	$\widehat{\theta} = 7.31$	6,266	3,728	5,409	3,224	9,994
		$\widehat{\alpha} = 1336$	1,149	7,518	292	7,014	8,666
	$\varphi = 0.2$	$\widehat{\theta} = 0.94$	940	9,043	-36	8,166	9,984
		$\widehat{\alpha} = 0.93$	998	9,002	22	8,125	10,000
$S_1 = S_2 = 0.85$	$\varphi = 0.5$	$\widehat{\theta} = 1.01$	1,010	8,967	32	8,087	9,977
$E_1 = E_2 = 0,95$		$\widehat{\alpha} = 1.01$	1,000	9,000	22	8,121	10,000
	$\varphi = 0.9$	$\widehat{\theta} = 0.81$	810	9,180	-167	8,295	9,990
		$\widehat{\alpha} = 0.77$	994	9,006	17	8,121	10,000
	$\phi = 0.2$	$\widehat{\theta} = 1.22$	1,222	8,759	239	8,022	9,981
		$\widehat{\alpha} = 0.94$	1,000	6,971	17	6,094	7,971
$S_1 = S_2 = 0.85$	$\phi = 0.5$	$\widehat{\theta} = 1.33$	1,328	8,665	341	7,999	9,992
$E_1 = E_2 = 0,95$		$\widehat{\alpha} = 1.42$	1,002	8,997	15	8,331	9,999
	$\phi = 0.9$	$\widehat{\theta} = 1.74$	1,731	8,250	745	7,716	9,981
		$\widehat{\alpha} = 2.00$	1,002	8,996	16	8,462	9,998
	$\psi = 0.2$	$\widehat{\theta} = 3.59$	2,841	7,167	2,210	6,388	10,008
		$\widehat{\alpha} = 6.20$	1,635	8,361	1,004	7,582	9,996
$S_1 = S_2 = 0.45$	$\psi = 0.5$	$\widehat{\theta} = 6.88$	4,194	5,801	3,620	5,140	9,995
$E_1 = E_2 = 0,95$		$\widehat{\alpha} = 39.8$	2,017	7,945	1,443	7,284	9,962
	$\psi = 0.9$	$\widehat{\theta} = 10.0$	4,886	5,100	4,596	4,399	9,986
		$\widehat{\alpha} = 1130$	2,435	6,438	1,948	5,934	8,872
	$\varphi = 0.2$	$\widehat{\theta} = 0.94$	939	9,062	248	8,185	10,001
		$\widehat{\alpha} = 0.93$	976	9,025	285	8,148	10,000
$S_1 = S_2 = 0.45$	$\varphi = 0.5$	$\widehat{\theta} = 0.94$	933	9,069	231	8,191	10,002
$E_1 = E_2 = 0,95$		$\widehat{\alpha} = 0.93$	971	9,027	269	8,149	9,998
	$\varphi = 0.9$	$\widehat{\theta} = 0.81$	816	9,180	116	8,295	9,996
		$\widehat{\alpha} = 0.78$	941	9,060	241	8,175	10,000
	$\phi = 0.2$	$\widehat{\theta} = 1.23$	1,160	8,831	416	8,094	9,990
		$\widehat{\alpha} = 1.28$	999	9,001	255	8,264	9,999
$S_1 = S_2 = 0.45$	$\phi = 0.5$	$\widehat{\theta} = 1.34$	1,289	8,730	489	8,064	10,000
$E_1 = E_2 = 0,95$		$\widehat{\alpha} = 1.41$	1,029	8,971	229	8,305	9,999
	$\phi = 0.9$	$\widehat{\theta} = 1.76$	1,652	8,345	823	7,811	10,000
		$\widehat{\alpha} = 1.99$	1,047	8,951	218	8,417	9,998

Using the index estimate, we computed the estimated unknown quantities of diseased and non-diseased individuals within group with negative outcome in both tests (d and h in Tables 1 and 2), using equations 11 and 12. We observed that, when the data have linear binary dependence the θ index overestimate d and n_D and underestimates h and n_{ND}, that effect is more evident when the covariance

level is increased. With weak and moderate linear binary dependences, the α index tends to overestimates all quantities but the observed bias is very little, if the dependence is strong, the observed behaviour is similar to the observed with the other index and it remains regardless of the test sensitivities. See Table 5.

When the diagnostic tests show a weak or strong FGM dependence structure and the sensitivities are higher than 0.5, the θ index takes a value lower than one indicating negative dependence which underestimates n_D and the estimate value for d is negative. For moderate FGM dependences the θ index expresses independence. If the test sensitivities are lower than 0,5, the n_D and d quantities are underestimated while n_{ND} and h are overestimated but the estimation biases are lower than those observed in data with linear binary dependence. For this type of dependences with high sensitivities the α index shows estimates very close to the true quantities but if the tests have sensitivities lower than 0.5, the behaviour is similar to that observed with binary data however the estimation biases are lower. See Table 5.

The obtained results using data within dependence type Gumbel, have a behaviour very similar with the at observed in binary linear dependent data, but in this case, the estimation bias is lower in all cases. Table 5

4.2. Example with Published Data

Bohning & Patilea (2008) used a published data set to illustrate the performance of their two indexes. The authors took the subset of data of serum cholesterol and body mass index as risk factors for cardiovascular disease considered in the Framingham Heart Study (Shurtleff 1974). In agreement with these authors, for that data, conditional on the disease status, the risk factors are positively and significantly associated as when measured by the Mantel-Haenszel odds ratio with the summary taken over disease status. With the estimate values of the indexes, they estimated the quantities of diseased and non-diseased individuals within group with negative outcome in both tests, using for that the equations 11 and 12.

We fit models assuming covariance, FGM dependence and Gumbel dependence and we obtained prevalence, performance test and dependence estimates under Bayesian paradigm. As we have six observed frequencies in the cross table and we have seven parameters to estimate, we have a non-identifiable model. So, in the same manner as Joseph, Gyorkos & Coupal (1995) we fitted models using Beta(a,b) as informative prior distributions on dependence parameters and Beta(1/2, 1/2) as non-informative prior distribution on prevalence and test parameters. Since we have no prior information on dependence parameters, we used the three prior distributions employed in the example with simulated data and we used the Deviance Information Criteria (DIC) to select the better fit. With our estimates, we computed the values of the BP indexes and we assumed that d and h are not known, so we use the estimators $\widehat{d}$ and $\widehat{h}$ to predicting the known n_D and n_{ND}. The results are showed in Table 6.

Table 6: Estimates of indexes and disease classes in a study with completely known disease status. (Observed frequencies: $a = 51$, $b = 19$, $c = 70$, $d = \mathbf{86}$, $e = 69$, $f = 20$, $g = 38$, $h = \mathbf{60}$, $n_{(D+ND)} = 413$)

Model	Index	$\widehat{n}_D$	$\widehat{n}_{ND}$	$\widehat{d}$	$\widehat{h}$	$\widehat{n}_{(D+ND)}$
Böhning*	$\widehat{\theta}$ =1,36	225	188	85	61	413
	$\widehat{\alpha}$ =3.79	242	171	104	44	414
FGM copula*	$\widehat{\theta}$ =1.32	219	183	79	56	402
	$\widehat{\alpha}$ =2.48	205	154	65	27	359
Covariance	$\widehat{\theta}_D$ =1.57	261	-	121	-	-
	$\widehat{\theta}_{ND}$ =1.08	-	202	-	75	463
	$\widehat{\alpha}_D$ =8.32	366	-	227	-	-
	$\widehat{\alpha}_{ND}$ =3.14	-	161	-	34	527
Gumbel copula	$\widehat{\theta}_D$ =0.45	75	-	(-65)	-	-
	$\widehat{\theta}_{ND}$ =0.78	-	129	-	(-11)	204
	$\widehat{\alpha}_D$ =0.23	146	-	6	-	-
	$\widehat{\alpha}_{ND}$ =0.27	-	130	-	3	276

*Models with homogeneous dependence

With this data set, the index values obtained after of fitting a model with Gumbel parameter show negative dependence between test outcomes and the estimates of the unknown quantities are negative which makes no sense, indicating the data do not fit well with the Gumbel copula. The model assuming binary dependence eliminates the assumption of homogeneity retained by the Böhning and FGM models. That model overestimates the numbers of individuals not verified by "gold standard" expressing that dependence between the tests do not have linear binary structure. The results obtained using the model with FGM dependence shows a better fit despite, it tends to underestimate both indexes a little and underestimate the unknown quantities. This implies tending a contradiction because in agreement with Bohning & Patilea (2008) when $\theta > 1$, the expected value of n_i, $(i = D, ND)$ will be below the true value of n and the amount of underestimation is determined by the value of θ; the higher value of θ, the higher the underestimation; therefore, if the index values obtained with FGM fitted model are lower than Böhning index, the estimate quantities for n_i should be higher than those observed with the Böhning model. In both models, the θ index estimate shows better behaviour than the other index.

5. Conclusions

In many clinical diagnostic procedures, it is necessary to use two or more (observable or not) biological traits expressed on a continuous scale in designs that includes verification with gold standard only for those individuals with at least one positive outcome in the screening tests. To obtain the diagnostic, those measures are dichotomized using a cut point, in this way, the final result is one of two values

(positive or negative). The continuous traits measured can be correlated in some way (not necessarily linear dependence) but when performing data analysis, can occur dependence is assumed with binary structure and not on the continuous structure. Given that the study planning has verification bias, some values in cross table are unknown so it is very complex to estimate the prevalence and performance test parameters using the maximum likelihood procedure. Many authors have considered the estimation problem using models with latent variables to complete the data set, others as Bohning & Patilea (2008) have developed reparametrizations using the observed incomplete data under binary structure assumption. In this paper, we studied the performance of models developed by Bohning & Patilea (2008) and we compared them with the performance of models that use covariance and copula functions to obtain information on the dependence between diagnostic tests.

Despite the covariance and the θ index have different parametric spaces, within the diseased population, we observed that, regardless of the population (diseased and non-diseased individuals) it exists a perfect linear relation between them, whenever there is the diagnostic tests have binary dependence structure, it is possible that the θ index to take values lower than zero and this range of values is not considered within the construction of the index. To have $\theta < 0$ indicates that we have at least one of the tests with sensitivity zero or at least one test without specificity and both situations are unacceptable in practical terms. The α index does not identify covariances lower than -0.5; and when tests with perfect sensitivities ($S_j = 1$) have a strong dependence expressed by a covariance close to unity, the α index takes values in a very constrained range [0,2] indicating very weak dependence; therefore, in cases where tests with perfect performance have perfect binary dependence structure, the BP index either may indicate values not allowed by construction or may underestimate the true dependence. It is obvious that tests with absolutely perfect or imperfect performance is a hypothetical situation very unlikely to occur in reality. In our simulation study with more realistic conditions, we observed that the BP indexes take values within range $(0, \infty)$, the relationship between covariance and α index grows more exponentially and the same do not show strong changes with the differences in the test sensitivities.

It is totally wrong to use the BP indexes with dichotomized data that initially have some of the two copula dependence structures studied, therefore, to use some of those indexes developed to binary data with dichotomized data, leads to erroneous conclusions regarding the dependence between tests which modify the final diagnostic result. In this work, we used two copula functions that model weak non linear dependences and the BP indexes failed to model them, whenever there are many other copula families which the BP indexes relationship could be studied. On the other hand, when the continuous traits are perfectly dependent with some of the copula structures studied, and we evaluate the dependence hypothesis using the dichotomized results and assume binary covariance as parameter, the estimation leads to conclude that test outcomes are independent from each other what directly affects the estimation of the test performance parameters and the prevalence.

Acknowledgments

The authors thank the editor and reviewers for their constructive suggestions. The first author was supported in part by a grant of the Coordenação de Aperfeiçõamento de Pessoal de Nivel Superior (CAPES); the second author was supported by grant of Conselho Nacional de Pesquisas (CNPq) Brazil. This paper is a part of the doctoral in statistics project of the first author.

[Recibido: enero de 2011 — Aceptado: junio de 2011]

References

Bohning, D. & Patilea, V. (2008), 'A capture-recapture approach for screening using two diagnostic tests with availability of disease status for the positives only', *Journal of the American Statistical Association* **103**, 212–221.

Dendukuri, N. & Joseph, L. (2001), 'Bayesian approaches to modelling the conditional dependence between multiple diagnostic tests', *Biometrics* **57**, 158–167.

Enoe, C., Georgiadis, M. P. & Johnson, W. O. (2000), 'Estimation of sensitivity and specificity of two diagnostic tests', *Preventive Veterinary Medicine* **45**, 61–81.

Georgiadis, M. P., Johnson, W. O. & Gardner, I. A. (2003), 'Correlation adjusted estimation of sensitivity and specificity of two diagnostic tests', *Journal of the Royal Statistical Society: Series C (Applied Statistics)* **52**, 63–76.

Gumbel, E. J. (1960), 'Bivariate exponential distributions', *Journal of the American Statistical Association* **55**, 698–707.

Johnson, M. E. (1987), *Multivariate Statistical Simulation*, Wiley and Sons, New York.

Joseph, L., Gyorkos, T. W. & Coupal, L. (1995), 'Bayesian estimation of disease prevalence and the parameters of diagnostic tests in the absence of a gold standard', *American Journal of Epidemiology* **141**, 263–272.

Nelsen, R. B. (1999), *An Introduction to Copulas*, Springer Verlag, New York.

Park, C. G., Park, T. & Shin, D. W. (1996), 'A simple method for generating correlated binary variates', *The American Statistician* **50**(4), 306–310.

Shurtleff, D. (1974), Some characteristics related to the incidence of cardiovascular disease and death: 18-year follow-up, *in* W. B. Kannel & T. Gordon, eds, 'An Epidemiological Investigation of Cardiovascular Disease. The Framinham Study', National Institute of Health, Washington, D. C.

Thibodeau, L. A. (1981), 'Evaluating diagnostic tests', *Biometrics* **37**, 801–804.

Torrance-Rynard, V. L. & Walter, S. D. (1997), 'Effects of dependent errors in the assessment of diagnostic tests performance', *Statistics in Medicine* **16**, 2157–2175.

Vacek, P. M. (1985), 'The effect of conditional dependence on the evaluation of diagnostic tests', *Biometrics* **41**, 959–968.