Statistics and Applications {ISSN 2454-7395 (online)}
Volume 19, No. 2, 2021 (New Series), pp 139-146

Joint Modeling of HIV and Tuberculosis through Copula-based Bivariate Binary Model

Atanu Pal[1] **and Arabinda Das**[2]
[1]*Department of Rheumatology and Nephrology, Institute of Post Graduate Medical Education and Research, Kolkata – 700 020, India*
[2]*Department of Statistics, Acharya Prafulla Chandra College, Kolkata – 700 131, India*

Received: 27 July 2020; Revised: 24 September 2020; Accepted: 08 December 2020

Abstract

Increasing spread of HIV facilitates the incidence of tuberculosis (TB) and HIV infected individuals co-infected with TB have much higher risk of developing active TB. In India, where TB infection is health burden, co-infection of TB with HIV increases the life-time risk by more than 5 times of developing TB. The objective of the study is to find the association between HIV and TB and the risk factors for infections of HIV and TB in Assam, north-east part of India. We used a joint bivariate binary model to accommodate the dependence between HIV and TB using a copula function. The maximum likelihood (ML) method has been used to estimate the model. We found a significant positive association between HIV and TB. The odds of developing TB in HIV infected person more than two-fold. We also found higher odds of HIV among younger people, who were men, military or paramilitary personnel, and heterosexual. These findings suggest that co-infection of HIV and TB is higher for this population and group interventions should be made to control the risk factors of the co-infection of HIV and TB in this part of India.

Key words: HIV; Tuberculosis; Bivariate binary outcomes; Copula function.

1. Introduction

While Human Immunodeficiency Virus (HIV) alone continues to be one of the deadliest diseases around the world, co-infection of tuberculosis (TB) with HIV is found to be the most leading cause of high mortality among people living with HIV (Corbett *et al.*, 2003; Lawn *et al.*, 2009). According to the UNAIDS report of 2019, about 37.9 million people globally were living with HIV at end of 2018 of which an estimated 10.0 million people developed TB disease, approximately 9% of all people living with HIV (UNAIDS fact sheet). Moreover, an estimated 49% of people living with HIV and TB are unaware of their co-infection and are therefore not receiving care. Increasing spread of HIV has become a major contributor in increasing the incidence of TB. Moreover, HIV infected individuals co-infected with TB have an annual risk of 5-15% of developing active TB, due to the reactivation of latent infection (Albalak *et al.*, 2007; Carvalho *et al.*,2001; Young, 2008; Mendelson, 2007). A HIV-positive person infected with TB has a 50 - 60% lifetime risk of developing TB as compared to an HIV-negative person who has only a 10% risk. Pulmonary tuberculosis showed no association with HIV and treatment failure. However co-infection of

Corresponding Author: Arabinda Das
Email: arabinda@apccollege.ac.in

pulmonary tuberculosis and HIV increases the probability of dying during treatment (Cabrera-Gaytán *et al*., 2016).

Thus, identification and treatment of TB as well as controlling the risk factors associated with TB are essential for successful HIV prevention in absence of effective vaccine. The objective of the present study was to find the association between HIV and TB in a survey among patients came to Assam Medical College and Hospital, Assam, north-east state of India and identify some of potential risk factors for HIV and TB infections. We used a joint response model with bivariate binary variables that measures the dependence between HIV and TB infections in assessing the socio-demographic, and biological risk factors of the two infections. Separate analysis of each infection ignores the dependence between the two infections and analyzing the data as if it were independent results in biased estimates.

The most common assumption for a bivariate binary mode to assume a bivariate normal distribution that guarantees normal marginals for the disturbances, and probit marginals for the binary dependent variables (Bhattacharya *et al*., 2006). Ghebremichael (2015) used a joint binary model to find the correlation between co-infection of HSV-2 and HIV-1. We consider a copula function based method to derive the joint model with the possibility of various marginal distributions and dependence structures for bivariate binary variables.

The copula approach of modelling gives wide flexibility in modelling for varied types of data. It has been widely used as a method to derive a joint bivariate distribution with non-normal marginal distributions and various dependence structures in recent years (see for details: Kolev and Paiva, 2009; Song, Li and Yuan, 2009 among others). Winkelmann (2012) proposed a bivariate probit model in modeling the effect of an endogenous binary regressor on a binary outcome variable with non-normal dependence among two variables using copula functions. There are number of copula families representing wide range of dependence structure among the random variables. A bivariate copula function, $C_\alpha(u_1, u_2)$ is a bivariate distribution function of the uniform [0,1] random variables and $\alpha \in \Omega$ measures the dependence between u_1 and u_2. Details of copula function can be found in Nelson (2006). The notion of statistical modeling by copula approach started with the pioneering work of Sklar (1959, 1973) where every multivariate distribution function can be uniquely constructed with a unique copula function which captures the dependence structure among the random variables and the marginal distribution functions of these random variables. Let $F(x_1, x_2; \lambda)$ be the joint distribution function of the random variables X_1 and X_2. Using Sklar's theorem, $F(x_1, x_2; \lambda)$ can be obtained with a unique copula function, $C_\alpha(u_1, u_2)$ such as $F(x_1, x_2; \lambda) = C_\alpha(F_1(x_1; \theta_1), F_2(x_2; \theta_2))$ where $F_i(x_i; \theta_i)$ is the distribution function of X_i, $\boldsymbol{\theta} = (\theta_1, \theta_2)'$ and $\boldsymbol{\lambda} = (\boldsymbol{\theta}, \alpha)'$.

The paper is organized as follows. We introduce a class of copula-based regression models for bivariate binary outcomes in Section 2 and discuss likelihood-based methods for model estimation. The model is estimated by maximum likelihood method. Section 3 introduces the data from the study population and summarizes the results from the proposed regression model applied to the data. Finally, Section 4, interprets the findings and concludes the paper.

2. Regression Model for Copula-based Bivariate Binary Outcomes

Consider two correlated binary outcomes X_i and Y_i obtained from each of *n* subjects, where X_i and Y_i are observed for the presence of HIV and TB respectively. The underlying

latent counterpart of X_i can be defined as $X_i^* = \boldsymbol{z}_{1i}^{'}\boldsymbol{\beta}_1 + \varepsilon_{1i}$ where $\boldsymbol{z}_{1i}$ denote explanatory variable, ε_{1i} is the random error and $\boldsymbol{\beta}_1$ is the corresponding parameter with

$$X_i = \begin{cases} 0 & if\ X_i^* \in (-\infty, \eta_1) \\ 1 & if\ X_i^* \in (\eta_1, \infty) \end{cases}$$

where η_1 is unknown threshold.

Similarly, the underlying latent counterpart of Y_i is presented as $Y_i^* = \boldsymbol{z}_{2i}^{'}\boldsymbol{\beta}_2 + \varepsilon_{2i}$ where $\boldsymbol{z}_{2i}$ denote explanatory variable, ε_{2i} is the random error and $\boldsymbol{\beta}_2$ is the corresponding parameter such that

$$Y_i = \begin{cases} 0 & if Y_i^* \in (-\infty, \gamma_1) \\ 1 & if Y_i^* \in (\gamma_1, \infty) \end{cases}$$

where γ_1 is unknown threshold.

Assuming $X_i^* \sim F_{X_i^*}$ and $Y_i^* \sim F_{Y_i^*}$ and the joint distribution function of (X_i^*, Y_i^*) as $F_{X_i^* Y_i^*}(x_i^*, y_i^*)$, the joint distribution of X_i and Y_i is then given by

$$\begin{aligned} &P(X_i = 0, Y_i = 0) = P(X_i^* \le \eta_1, Y_i^* \le \gamma_1) = F_{X_i^* Y_i^*}(\eta_1, \gamma_1) \\ &P(X_i = 0, Y_i = 1) = P(X_i^* \le \eta_1, \gamma_1 \le Y_i^*) = F_{X_i^*}(\eta_1) - F_{X_i^* Y_i^*}(\eta_1, \gamma_1) \\ &P(X_i = 1, Y_i = 0) = P(\eta_1 \le X_i^*, Y_i^* \le \gamma_1) \\ &= P(X_i^* \le \infty, Y_i^* \le \gamma_1) - P(X_i^* \le \eta_1, Y_i^* \le \gamma_1) \\ &= F_{Y_i^*}(\gamma_1) - F_{X_i^* Y_i^*}(\eta_1, \gamma_1) \end{aligned}$$

and

$$\begin{aligned} &P(X_i = 1, Y_i = 1) = P(\eta_1 \le X_i^*, \gamma_1 \le Y_i^*) \\ &= P(X_i^* \le \infty, Y_i^* \le \infty) - P(X_i^* \le \infty, Y_i^* \le \gamma_1) - P(X_i^* \le \eta_1, Y_i^* \le \infty) + P(X_i^* \le \eta_1, Y_i^* \le \gamma_1) \\ &= 1 - F_{Y_i^*}(\gamma_1) - F_{X_i^*}(\eta_1) - F_{X_i^* Y_i^*}(\eta_1, \gamma_1) \end{aligned}$$

The degree and type of dependence depends on the choice of copula. We consider the popular FGM copula, as this copula is comprehensive and allows for either negative or positive dependence. It also has a closed and simple analytic form. Moreover, the marginals are considered as standard logistic, i.e. logits for the binary dependent variables. The logit function is based on logistic distribution and its cdf turns out to be mathematically convenient. It is more popular in health sciences as the coefficients can be interpreted in terms of odds ratios. Therefore, the joint distribution function of X_i and Y_i with logit marginals and copula function is obtained as

$$F_{X_i^* Y_i^*}(x_i^*, y_i^*) = C_\alpha(\Lambda(x_i^*), \Lambda(y_i^*))$$

where $C_\alpha(.,.)$ is the copula function, $\Lambda(.)$ is the distribution function of a logistic random variable and is given by $\Lambda(z_i) = 1/(1 + \exp(-z_i))$.

Therefore, the joint density $f_{X_i, Y_i}(x, y) = P(X_i = x, Y_i = y)$ of X_i and Y_i under the copula is as follows:

$$P(X_i h = x, Y_i = y) = \begin{cases} C_\alpha\big(\Lambda(\zeta_1), \Lambda(\xi_1)\big) & if\ x = 0, y = 0 \\ C_\alpha(\Lambda(\zeta_1), 1) - C_\alpha\big(\Lambda(\zeta_1), \Lambda(\xi_1)\big) & if\ x = 0, y = 1 \\ C_\alpha\big(1, \Lambda(\xi_1)\big) - C_\alpha\big(\Lambda(\zeta_1), \Lambda(\xi_1)\big) & if\ x = 1, y = 0 \\ 1 - C_\alpha\big(1, \Lambda(\xi_1)\big) - C_\alpha(\Lambda(\zeta_1), 1) + C_\alpha\big(\Lambda(\zeta_1), \Lambda(\xi_1)\big) & if\ x = 1, y = 1 \end{cases}$$

where

$\zeta_1 = \eta_1 - \boldsymbol{z}_1^{'}\boldsymbol{\beta}_1$
$\xi_1 = \gamma_1 - \boldsymbol{z}_2^{'}\boldsymbol{\beta}_2$

Setting $\eta_1 = \gamma_1 = 0$ and under FGM copula, $C_\alpha(u_1, u_2) = u_1 u_2\{1 + \alpha(1 - u_1)(1 - u_2)\}$, the joint density $f_{X_i,Y_i}(x, y)$ becomes

$$P(X_i = 0, Y_i = 0) = \frac{1}{(1 + e^{-\zeta_1})}\frac{1}{(1 + e^{-\xi_1})}\left\{1 + \alpha\frac{1}{(1 + e^{\zeta_1})}\frac{1}{(1 + e^{\xi_1})}\right\}$$
$$P(X_i = 0, Y_i = 1) = \frac{1}{(1 + e^{-\zeta_1})}\left[1 - \frac{1}{(1 + e^{-\xi_1})}\left\{1 + \alpha\frac{1}{(1 + e^{\zeta_1})}\frac{1}{(1 + e^{\xi_1})}\right\}\right]$$
$$P(X_i = 1, Y_i = 0) = \frac{1}{(1 + e^{-\xi_1})}\left[1 - \frac{1}{(1 + e^{-\zeta_1})}\left\{1 + \alpha\frac{1}{(1 + e^{\zeta_1})}\frac{1}{(1 + e^{\xi_1})}\right\}\right]$$
$$P(X_i = 1, Y_i = 1) = 1 - \frac{1}{(1 + e^{-\zeta_1})} - \frac{1}{(1 + e^{-\xi_1})} + \frac{1}{(1 + e^{-\zeta_1})}\frac{1}{(1 + e^{-\xi_1})}\left\{1 + \alpha\frac{1}{(1 + e^{\zeta_1})}\frac{1}{(1 + e^{\xi_1})}\right\}$$

These joint probabilities described above depend on the selected copula as well as parameter vector $\boldsymbol{\lambda} = (\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \alpha)'$. If the true copula is assumed to belong to a parametric family, consistent and asymptotically normally distributed estimates of the parameter $\boldsymbol{\lambda}$ can be obtained through maximum likelihood method.

Assuming an independent sample of *n* observations on $(x_i, y_i, \boldsymbol{z}_{1i}, \boldsymbol{z}_{2i})$, the likelihood function is given by

$$L(\boldsymbol{\lambda}; x, y, \boldsymbol{z}_1, \boldsymbol{z}_2) = \prod_i\big\{P(X_i = 0, Y_i = 0|\boldsymbol{z}_{1i}, \boldsymbol{z}_{2i})^{(1-x_i)(1-y_i)} \times P(X_i = 0, Y_i = 1|\boldsymbol{z}_{1i}, \boldsymbol{z}_{2i})^{(1-x_i)y_i} \times P(X_i = 1, Y_i = 0|\boldsymbol{z}_{1i}, \boldsymbol{z}_{2i})^{x_i(1-y_i)} \times P(X_i = 1, Y_i = 1|\boldsymbol{z}_{1i}, \boldsymbol{z}_{2i})^{x_i y_i}\}$$

The corresponding log-likelihood function is then given by:

$$l(\boldsymbol{\lambda}; x, y, \boldsymbol{z}_1, \boldsymbol{z}_2) = \sum_{i=1}^{n}\{(1 - x_i)(1 - y_i)\log P(X_i = 0, Y_i = 0|\boldsymbol{z}_{1i}, \boldsymbol{z}_{2i}) + (1 - x_i)y_i \log P(X_i = 0, Y_i = 1|\boldsymbol{z}_{1i}, \boldsymbol{z}_{2i}) + x_i(1 - y_i)\log P(X_i = 1, Y_i = 0|\boldsymbol{z}_{1i}, \boldsymbol{z}_{2i}) + x_i y_i \log P(X_i = 1, Y_i = 1|\boldsymbol{z}_{1i}, \boldsymbol{z}_{2i})\}$$

The above presented log-likelihood function can be maximized using numerical optimization methods. These can use analytical first derivatives that have a relatively tractable form.

3. Data Analyses

This study was conducted at Department of TB & Chest Disease, Integrated Counselling and Testing Centre (ICTC) and Department of Medicine; Assam Medical College & Hospital, Dibrugarh from January 2007 to November 2008. Patients coming with suspected infection with HIV or Tuberculosis were sent to Department of TB and Chest Disease/RNTCP from ICTC. The study was approved by Assam Medical College & Hospital Ethics Committee and performed according to the Declaration of Helsinki, 1975. Written informed consent for participation in the study was obtained from participants. A total of 184 patients both male and female were screened for the study. Information was collected on different socio-demographic characteristics and sexual behaviors. Blood samples were drawn to test for HIV and Tuberculosis. HIV infection was determined using HIV enzyme-linked immunosorbent assay (ELISA), and reactive samples were confirmed using Wellcozyme HIV ELISA test. Western blot tests were used to confirm discordant ELISA test results. Presence of Tuberculosis were verified by acid fast bacilli (AFB) sputum smear, mycobacterial culture, histopathology, by clinical suspicion or radiological epidemiologic. Socio-demographic characteristics (age in years, gender, occupation), route of transmission, types of tuberculosis were considered as covariates in our analyses.

3.1. Results

One hundred and eighty four patients were recruited for the study, out of which 74% were male and 26% were female. The median age of the participants was 35 years (IQR = 25–54). Among the patients, 121 were HIV infected of which 35 were TB infected. The prevalence of TB among non-HIV infected patients was 11.11% compare to the prevalence of TB infection among HIV infected patients at 28.92% which is more than 2.6 times higher. The rate of HIV in TB positive was 83.33% compare to the rate of HIV among TB negative at 60.56%. The Pearson Chi-square test of interdependence between HIV and TB was highly significant indicating that patients who were infected by one of the infections were likely to be infected by the other as well (phi-coefficient = 0.201, p-value < 0.01). About 94% of the participants had heterosexual as route of transmission of HIV, and only 6% had homosexual as route of transmission of HIV. Eight-eight percent of HIV infected female were housewives, rest were commercial sexual workers. Among HIV infected male 74% were paramilitary & military personnel and private sector workers, 17% were drivers and rickshaw puller. Among TB infected patients 54% had pulmonary type of infection, and 46% had extra-pulmonary type of infection. Out of extra-pulmonary infection, lymphadenopathy were 26%, intestinal were 14% and meningitis were 6%. Table 1 showed male had more than twice odds (Odds Ratio (OR): 2.664, 95% Confidence Interval (CI): 0.845, 8.401) to have co-infection of HIV and TB. Urban patients had higher in-confection of HIV and TB (OR: 1.733, 95% CI: 0.672, 4.471) than their rural counterpart. Smokers showed almost 4 times odds (OR: 3.704, 95% CI: 1.513, 9.068) of developing co-infection. Drivers and paramilitary personnel showed seven times (OR: 7.200, 95% CI: 2.248, 23.063) and three times (OR: 2.965, 95% CI (0.964, 9.119) odds in favour of co-infection.

Table 1: Possible risk factors for tuberculosis (TB) in HIV and TB co-infected patients

		HIV + TB	HIV	OR (95% CI)
Gender				
	Female	4	22	1
	Male	31	64	2.664 (0.845, 8.401)
Age				
	<20	1	2	1
	20-40	29	57	0.983 (0.085, 11.295)
	40-60	4	24	3.000 (0.217, 41.353)
	>60	1	3	1.500 (0.055, 40.635)
Marital status				
	Married	11	25	1
	Single	24	61	1.118 (0.477, 2.622)
Residence				
	Rural	7	26	1
	Urban	28	60	1.733 (0.672, 4.471)
Smoking				
	No	8	45	1
	Yes	27	41	3.704 (1.513, 9.068)
Drug dependency				
	No	13	37	1
	Yes	22	49	1.278 (0.569, 2.865)
Occupation				
	Private	5	36	1
	Para-Military	14	34	2.965 (0.964, 9.119)
	Driver	16	16	7.200 (2.248, 23.063)
ROT				
Homosexual		1	10	1
Heterosexual		34	76	4.474 (0.551, 36.353)

Table 2: Outcome of Bivariate Binary Model: HIV and TB under FGM copula

Parameter	OR	SE	95% CI
HIV			
Age	1.372	0.135	(1.091, 2.114)
Gender	2.873	0.253	(1.138, 4.223)
ROT	5.611	1.574	(2.893, 26.432)
Occupation	4.627	1.282	(1.482, 7.324)
Residence	1.825	0.831	(1.121, 3.473)
TB			
Age	1.245	0.143	(0.913, 1.865
Smoking	3.122	0.986	(1.628, 7.421)
Residence	1.386	0.668	(0.739, 3.116)
Occupation	3.192	1.154	(1.433, 6.310)
Gender	2.521	0.431	(0.957, 3.264)
Correlation (α)	2.084	0.382	(0.891, 4.884)

The risk factors of HIV and tuberculosis were identified using the preliminary analysis. The variables that were found to be significantly associated with HIV and TB in the preliminary analysis were further considered for multivariate analysis. Age, gender, route of transmission (ROT), occupation were associated with both infections. We did a multivariate analysis using these variables as covariates in a joint response model. The results of our multivariate analysis are reported in Table 2. Odds ratios together with their corresponding standard errors and 95% confidence intervals are presented. Despite the relatively small sample, we found a significant association between HIV and TB. The results indicate that higher prevalence of HIV was associated among younger age patients. Male had significantly higher prevalence of HIV and TB than female. The odds of having HIV in younger people were 1.372 times higher than that of older people (OR = 1.372; 95% CI: 1.091–2.114). Men were associated with higher rates of HIV (OR = 2.873; 95% CI: 1.138–4.223) and TB (OR = 2.521; 95% CI: 0.957–3.264) infections. Heterosexual route of transmission was highly associated with higher prevalence of HIV (OR = 5.611; 95% CI: 2.893–26.432). The odds of having HIV and TB in urban patients were 1.825 times (OR = 1.825, 95% CI: 1.121-3.473) and 1.186 times (OR = 1.186, 95% CI: 0.439-3.116) higher. HIV was significantly associated with TB (OR = 2.084; CI: 0.891 – 4.884).

4. Conclusion

In this paper, we developed a joint response model for HIV and TB infections. Our study aimed to assess the association of HIV and TB infections with their determinants among peoples in Assam, north-eastern part of India. Socio-demographic characteristics such as age, gender and occupation as well as biological risk factors such as route of transmission of HIV and TB were considered. One hundred and eighty-four patients coming with suspected infection with HIV or TB were included in the study at Department of TB and Chest Disease/RNTCP, Assam Medical College and Hospital.

Assam has HIV prevalence of about 0.06%, compared to India's prevalence of 0.22% and TB prevalence of about 0.001%. The prevalence rates of HIV and TB among the study participants were 65.8% and 22.8%, respectively. The rate of TB in HIV infected patients was 2.6 times higher than rate of TB in non-HIV infected patients. Moreover, the rate of HIV in TB positive was 1.380 times higher the rate of HIV among TB negative. These indicate that the chance of infected from a disease increases for patients who have already infected from other disease. Thus, a joint response model is considered to accommodate the interdependence between the two infections of HIV and TB with the potential risk factors.
We found high prevalence of HIV among younger patients. The odds of having HIV in younger patients were 1.372 times higher than in older patients (OR = 1.372; 95% CI: 1.091–2.114). This finding may be result of demographical and social structure of India where most of the people are from young generation and they are more exposed to the western culture than their older counterpart. Men were associated with higher rates of HIV and tuberculosis infections. These findings are also associated with the social structure of India where women do not get much freedom and have restricted social life. Heterosexual route of transmission was highly associated with higher prevalence of HIV. Smoking played a pivotal role in developing TB. Study showed people from urban area were more prone to develop the co-infection of HIV and TB. The study also found that HIV was significantly associated with TB. Previous studies of HIV and TB co-infection modelled each infection separately ignoring the potential biological association between the two infections.

Acknowledgements

The authors express their gratefulness to the reviewers and the editor for their invaluable comments and suggestions.

References

Albalak, R., O'Brein, R. J., Kamemerer, S., *et al*. (2007). Trends in tuberculosis/Human Immunodeficiency Virus comorbidity, United States, 1993-2004. *Archives of Internal Medicine*, **167**(**22**), 2443-2452.

Bhattacharya, J., Goldman, D., and McCaffrey, D. (2006). Estimating probit models with self-selected treatments. *Statistics in Medicine*, **25**(**3**), 389–413.

Cabrera-GaytaÂn, D. A., Niebla-Fuentes, Md. R., Padilla-VelaÂzquez, R., Valle-Alvarado, G., Arriaga-Nieto, L., Rojas-Mendoza, T., *et al*. (2016). Association of pulmonary tuberculosis and HIV in the Mexican Institute of Social Security, 2006-2014. *PLoS One*, **11**(**12**):e0168559. doi:10.1371/journal.pone.0168559.

Carvalho, A. C. C., DeRiemer, K., Nunes, Z. B., *et al*. (2001). Transmission of mycobacterium tuberculosis to contacts of HIV-infected tuberculosis patients. *American Journal of Respiratory and Critical Care Medicine*, **164**(**12**), 2166-2171.

Corbett, E. L., Watt, C. J., Walker, N., Maher, D., Williams, B. G., Raviglione, M. C., *et al*. (2003). The growing burden of tuberculosis: Global trends and interactions with the HIV epidemic. *Archives of Internal Medecine*, **163**(**9**), 1009-1021.

Ghebremichael, M. (2015). Joint modeling of correlated binary outcomes: HIV-1 and HSV-2 co-infection. *Journal of Applied Statistics*, **42**(**10**), 2180-2191.

Kolev, N., and Paiva, D. (2009). Copula-based regression models: A survey. *Journal of Statistical Planning and Inference*, **139**(**11**), 3847-3856.

Lawn, S., and Churchyard, G. (2009). Epidemiology of HIV associated tuberculosis. *Current Opinion in HIV and AIDS*, **4**(**4**), 325-333.

Mendelson, M. (2007). Diagnosing tuberculosis in HIV-infected patients: challenges and future prospects. *British Medical Bulletin*, **81-82**, 149-165.

Nelson, R. B. (2006). *An Introduction to Copulas*. Springer, Berlin (ISBN 978-0-387-28678-5).

Sklar, A. (1959). Fonctions de répartition à n dimensions et leurs marges. *Publications de l'Institut Statistique de l'Université de Paris*, **8**, 229–231.

Sklar, A. (1973). Random variables, joint distributions, and copulas. *Kybernetika*, **9**, 449-460.

Song, P. X. K., Li, M., and Yuan, Y. (2009). Joint regression analysis of correlated data using gaussian copula. *Biometrics*, **65**, 60-68.

Winkelmann, R. (2012). Copula bivariate probit models: with an application to medical expenditures. *Health Economics*, **21**(**12**), 1444-1455.

Young, D. B., Perkins, M. D., Duncan, K., *et al*. (2008). Confronting the scientific obstacles to global control of tuberculosis. *The Journal of Clinical Investigation*, **118**(**4**), 1255-1265.