Journal of Chemical Health Risks

www.jchr.org

JCHR (2024) 14(2), 878-884 | ISSN:2251-6727

Survival Analysis of Time to Cure on Tuberculosis Patients in Dhemaji District, Assam, India

Achinta Saikia[1], Bipin Gogoi[2]

Department of Statistics, Dibrugarh University, Dibrugarh-786004, Assam, India

(Received: 07 January 2024 Revised: 12 February 2024 Accepted: 06 March 2024)

KEYWORDS

Censoring, Cox-regression, Kaplan-Meier, Sample size, Survival, Tuberculosis

ABSTRACT:

Introduction: We conducted the records of 1500 confirmed cases of tuberculosis patients were reviewed in Dhemaji district of Assam, 2019.

Objectives: The aim of this study was to study the recovery time of tuberculosis patients using survival analysis technique and to assess the association and impact of covariates (TB risk factors) to event status and Survival time.

Methods: Kaplan-Meier analysis and the log-rank test were used to assess the differences in survival among the patients, while Cox-regression model was used for multivariate analysis. Multivariate analysis was performed using binary logistic regression analysis. The significance levels for all the tests were set at 0.05.

Results: In this study, the 1500 TB patients in Dhemaji were assessed out of 1038 male patients, 512 (49.33%) were censored and 526 (50.67%) were cured of TB. Again out of 462 female, 226 (48.92%) were censored and 236 (51.08%) were cured of TB. Overall, the median recovery time of TB patients in Dhemaji was 180 days (approx.), which means that the recovery time of patients is within the recommended treatment interval of 160 to 220 days or longer given close monitoring of patients while taking drugs. The results of binary logistic regression analysis show that disease type (OR = 2.171, 95% CI 1.771–2.662), diabetes status (OR = 1.966, 95% CI 1.632–2.364) and residence (OR = 1.241, 95% CI 1.048– 1.469), were risk factors for tuberculosis. These results can provide insights on local tuberculosis early increase public health awareness, intervention and strengthen the control of factors that may affect the survival and tuberculosis patients.

Conclusions: In conclusion, overcrowding, smoking, tuberculosis (TB) were important risk factors and negatively affected the survival rates of TB patients in Dhemaji. At the same time, Disease Type, Diabetes status, Type of Patients and Residence, were significant risk factors and negatively affected the survival rates of TB patients in Dhemaji. The results suggested that the mycobacterium tuberculosis drug sensitivity test should be strengthened.

1. Introduction

According to WHO tuberculosis (TB) is caused by bacteria that are resistant to themost effective anti-tuberculosis drugs (Isoniazid and Rifampicin) [WHO 2010; Faustini et. al., 2006]. TB results either from primary infection or develop in the course of treatment of a patient due to human error, poor supply management, poor quality anti-TB drugs and/or improper treatment [WHO 2010; Kundu et. al 2018, Singh et. al., 2007]. In addition, poor infection control practice has also been identified as a major factor for the spread of TB and -TB has different recovery time for different patients [Koul et. al., 2011]. TB is being an increasing global problem, and in 2016, 153,119 cases were notified from which 129,689 enrolled for treatment, of which only 22% started treatment [WHO. Global tuberculosis report 2017]. Assefa et al also noted that 3.7% new and 20% previously treated TB cases were identified [Assefa et al 2017].

The burden and incidence of TB is increasing and varying significantly from country to country. The countries with the largest number of TB cases (47% of the global total) were China, India and the Russian Federation [WHO. Global tuberculosis report 2017]. The highest (28%) rates of new TB cases are from the Soviet Union including regions that share borders with the European Union [Eldholm & Balloux, 2016].

Journal of Chemical Health Risks

www.jchr.org

JCHR (2024) 14(2), 878-884 | ISSN:2251-6727

In Africa, an estimated 69,000 cases emerged of which about 1.2% were new. 12% of re-treatment cases were from Ethiopia of which 1.6 and 12% TB patients were new and previously treated TB cases respectively [Getachew, Bayray and Weldearegay B., 2013]. In addition, Ethiopia is one among the 20 countries with the highest absolute estimated number of incidents of TB and TB [WHO. Global tuberculosis report 2017]. In comparison to drug-susceptible TB, that takes about 6 to9 months to treat, recommended treatment for TB lasts 18 to 24 months or longer [Falzon, 2017], and requires the second line medicines that are not effective as first-line medicines commonly prescribed to treat TB [Falzon, 2017]. Previous studies indicated that drug-resistant strains of Mycobacterium tuberculosis are of great concern as they are more toxic and more expensive than the first-line regimen [Espinal, 2001]. Hence, monitoring closely patients while they take these drugs is critical, as the medications may lead to other serious health problems such as damage to the kidneys, liver, or heart; loss of vision or hearing; and changes in behaviour or mood including depression or psychosis [Tyrrell et. al., 2013]. As India is one of the 20 high burdens TB countries and TB has been a major health problem of the society in the Dhemaji district of Assam, a strategy to provide culture and drug susceptibility testing services has been designed [WHO, 2010, Assefa, Seyoum and Oljira, (2017)]. Even though various studies on the prevention and control of the cross-transmission of healthcare-acquired infections between hospitalized patients have been carried out, the prevalence is still increasing [Tacconelli, et al., 2014; Alrabiah et al., 2016]. Importantly, the appearance and transmission of TB is increasing in hospitals worldwide [Tarai, Das and Kumar, 2013]. TB poses therapeutic difficulties in the twenty-first century, with only a few antibiotics continuing effective [Fair & Tor, 2014]. Consequently, controlling and pre- venting the emergence and overflow of TB organisms is of vital importance. Thus, the aim of this study is to investigate the recovery time of TB patients in Dhemaji, Assam using accelerated failure time and parametricshared frailty models.

This time indicator will basically tell the duration of time that will take from beginning of any follow-up and the occurrence of an event. The time lapse between the starting point and the end point is the outcome variable of interest. In the medical research, the outcome variable or the desired outcome of interest is may be the recurrence of symptoms, death of a patient, relapse from remission, relief from pain, incubation of various diseases like Hepatitis B, AIDS, etc., disease incidence, in clinical traits remission duration of certain disease (Andersen, 1992; Kalbfleisch and Prentice, 1980; Cox and Snell 1968; Cox and Oakes, 1984; Crowley and Hu, 1977; Jenkins, 1997; Miller, 1981; Clayton, 1978). The survival analysis technique can be used in the fields where data have to analyze regarding the duration between the two events. Therefore, survival analysis is also known as life time data analysis, time to event analysis or event history analysis.

2. Objectives

The objective of the study was to the recovery time of tuberculosis patients using survival analysis technique and to assess the association and impact of TB risk factors to event status and survival time.

The relevant information has been collected from the District Tuberculosis unit, Dhemaji (Assam). The records of a total number of 1500 patients suffered from tuberculosis and treated under the DOTS strategy have been considered for the study. The main emphasis is in the category of disease, sex, age and type of tubercular infection as well as event of occurrence of death over a period of 3 years (i.e. from 2018 onwards).

3. Methods

3.1 Data source, sampling design

A retrospective study is carried out in seven hospitals of Dhemaji which have TB treatment center from September 2018 to February 2020. In addition, patients that have no full history about their epidemiological, clinical and laboratory results were excluded from the study using exclusion criteria.

3.2 Determination of sample size (n)

For calculating sample size required for the study we have used formula for computing n.

We have at our hand that:-

Journal of Chemical Health Risks

www.jchr.org

JCHR (2024) 14(2), 878-884 | ISSN:2251-6727

$$n = \frac{\{[Z_\alpha \sqrt{[2\bar{P}\,(1-\bar{P}\,)]} + Z_\beta \sqrt{[P_1\,(1-P_1) + P_2\,(1-P_2)]}\}^2}{(P_1 - P_2)^2}$$

[Lwanga S.K., and Lemeshow S. (1991)]

$$n = \frac{\{1.645\sqrt{[2\times 0.325(1-0.325)]} + 1.28\sqrt{[0.35(1-0.35)+0.30(1-0.30)]}\}^2}{(0.35-0.30)^2}$$

Where, $\bar{P} = \frac{P_1+P_2}{2}$, $\alpha = 0.05$, $\beta = 0.10$

$\bar{P} = 0.325$, $n = 1500$

Accommodating 5% non response, we get n = (105 × 1500)/ 100 = 1575. This non-response occurs due to some incomplete information was found at the time of data collection. I have observed that it will not be more than 5%.

So we have to take 1575 approximately to get our 1500 observations.

Earlier study report says around 35% is male and 30% are female. So, we have taken,

$p_1 = 0.35$, $p_2 = 0.30$

Now, we have used proportional allocation formula-

$$\frac{n_1}{N_1} = \frac{n_2}{N_2} = \frac{n}{N}$$

N = 3356, n = 1500

N_1 = Male patient = 2324, N_2 = Female patient = 1032

$n_1 = 1038$, $n_2 = 462$, $n = n_1 + n_2 = 1038 + 462 = 1500$

3.3 Measurements

The response variable in this study is defined to be the treatment period from the starting time of TB treatment up to the time of its cure. The event of interest was recovery from TB (1 = recovery and 0 = censored). The predictor variables that are included in this study were background characteristics of TB patients and history of epidemiological, clinical and laboratory results.

3.4 Data Analysis

Data were entered into an access database using the statistical package for the social sciences (SPSS, version 16.0, Armonk, NY, USA) and STATA. Kaplan-Meier analysis and the log-rank test were used to assess the differences in survival among patients. Cox regression model was used for multivariate analysis. In the single factor analysis of TB, the qualitative data was tested with chi-square test, multivariate analysis was performed using binary logistic regression analysis, and the significant level for all the tests was set at 0.05.

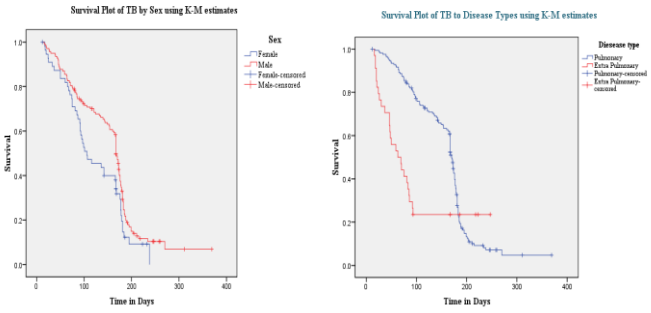
3.5 Statistical analysis

Survival analysis is the analysis of statistical data in which the outcome variable of interest is time until an event occurs. Survival data are almost always incomplete and called censoring that may be a right censor, left censoring and interval censoring. The most common are right censoring that happens when a subject follow-up time to terminate before the outcome of interest ob- served [Fan, 2018]. In any applied set, a survival data can summarize through life tables [Van Der Meulen, 2012], Kaplan-Meier Survival functions [Dabrowska, 1988] and median survival time [Brookmeyer, 2014; Reid, 1981].

To compare the survival pattern between different characteristics of TB patients

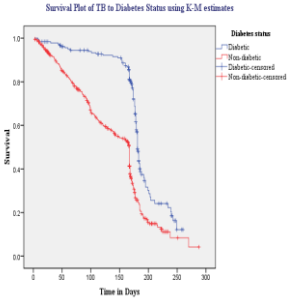
Here, T: Period of difference between the date of diagnosis and date of outcome (recovery) in days.

Using Kaplan-Meier method as given in (2) we get the K-M curve given in Fig1.

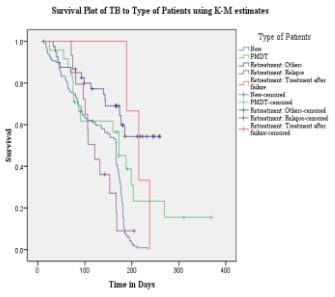


(a)

(b)



(c)

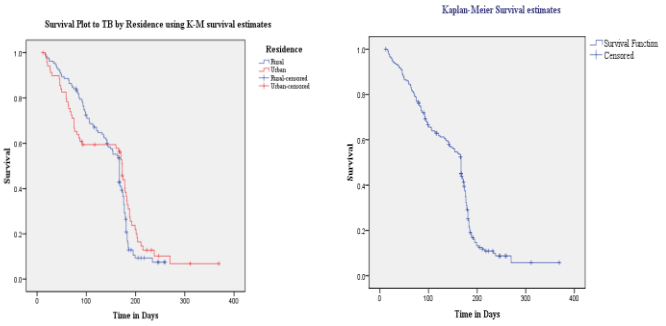


(d)

Journal of Chemical Health Risks

www.jchr.org

JCHR (2024) 14(2), 878-884 | ISSN:2251-6727



(e) (f)

Fig.1:-The K-M Survival function for the recovery time of TB patients in Dhemaji District by (a) Sex (b) Disease type (c) Diabetes Status (d) Type of patients (e) Residence (f) overall

For diabetic patients Fig 1(c) it has been observed that median is higher (185 days) than the non-diabetic (170 days). Also for diabetic patients, the recovery is very slow upto 1st 155 days as the curve remains almost flat and then there is a rapid decline in the KM curve which shows fast recovery of the patients. Some clinical study is necessary here. From the Kaplan-Meier curve Fig1 (e) we found that the median survival time for the combined (male and female) patients is 162 days (approx.). So we can say that 50% of the patient not recovered till 162 days.

But the median survival time for female is much lower than the male patients which are 125 days & 170 days respectively Fig1 (a). The reason for this vast difference can be studied with in depth survey.

Table1: Characteristics of TB patients in Dhemaji

District, Assam

Predictors	Labels	Status of Patients		
		Total	Cured / Event (%)	Censored
Sex	Male	1038	526 (50.67)	512
	Female	462	236 (51.08)	226
Disease Type	Pulmonary	1091	649 (59.49)	442
	Extra Pulmonary	409	113 (27.63)	296
Diabetes status	Diabetic	309	148 (47.90)	161
	Non-diabetic	1191	614 (51.55)	577
Type of Patients	New	1289	651 (50.50)	638
	PMDT	39	24 (61.54)	15
	Retreatment: Others	106	35 (33.02)	71
	Retreatment: Relapse	55	43 (78.18)	12
	Retreatment: Treatment after failure	11	9 (81.82)	2
Residence	Rural	1085	560 (51.61)	525
	Urban	415	202 (48.67)	213

The Kaplan-Meier survival function is an important tool for analyzing censored data [Gutierrez, 2002; Miller, 2011]. The Kaplan-Meier estimator survival curve depicted the overall estimated survivor function and for different groups of predictors.

Table2: The Log-rank and Breslaw test of predictors for the recovery time of TB patients in Dhemaji, Assam

Covariates	categories	Median	Log Rank test			Breslow test		
			Chi-Square	df	p-value	Chi-Square	df	p-value
Sex	Male	174	4.782	1	0.029	3.156	1	0.076
	Female	169						
Disease Type	Pulmonary	168	59.287	1	< 0.001	16.335	1	< 0.001
	Extra-Pulmonary	396						
Diabetes status	Diabetic	181	46.20	1	< 0.001	72.629	1	< 0.001
	Non-diabetic	168						
Type of Patients	New	170	50.343	4	0.001	26.9	4	< 0.001
	PMDT	160						
	Retreatment: Others	184						
	Retreatment: Relapse	222						
	Retreatment: Treatment after failure	216						
Residence	Rural	171	4.057	1	0.044	2.933	1	0.087
	Urban	177						

Journal of Chemical Health Risks
www.jchr.org
JCHR (2024) 14(2), 878-884 | ISSN:2251-6727

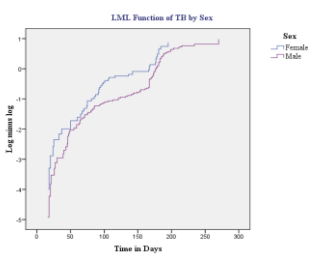
Kaplan-Meier survival function was used to estimate the cumulative probability of survival for different Sex, Disease Type, Diabetes status, Type of Patients and Residence. The results of log-rank test show that the differences of cumulative probability of survival of various factors such as Disease Type, Diabetes status, and Type of Patients are statistically significant ($p < 0.05$), but for sex and residence, the difference is statistically significant.

3.6 Cox proportional hazard model:

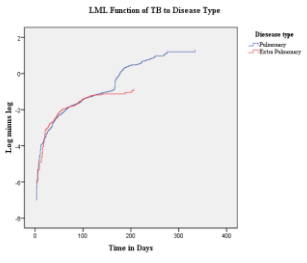
First we have to see whether Cox proportional hazard model is suitable or not. A guide to whether or not the hazard ratio can be regarded as constant is to plot the complementary log transformation, which is log{-log[s(t)]} against log t, as we have illustrate in fig. If the hazard rate does not change with time. Then the resulting plot will be approximately linear. Departures from linearity indicate that the hazard rate is changing with time.

Table 3:- Demographic characteristics of TB patients in Dhemaji District

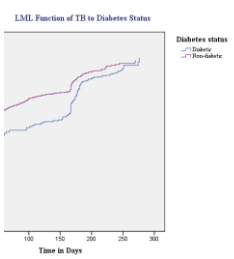
Variables	Categories	Constituent Ratio (%)
Sex	0=Female	462
	1=Male	1038
Diabetes Status	1=Diabetic	309
	0=Non-diabetic	1191
Disease Type	1=Pulmonary	1091
	0=Extra Pulmonary	409
Residence	0=Rural	1085
	1=Urban	415



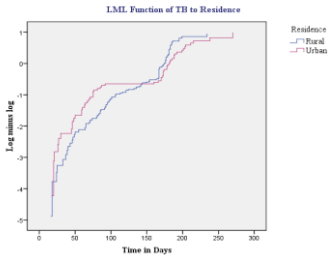
(a)



(b)



(c)



(d)

Fig.2:- log minus log survival plot for (a) Sex (b) Disease type (c) Diabetes Status (d) Residence

Figure shows that PH model is suitable here. We can use Cox's model as given for analyzing our data. Let us define the variables,

T: Survival time, time to recovery for the disease tuberculosis.

The risk factors are defined as follows:

(For Sex) $X_1 = 1$, male, 0, if female

(For Disease type) $X_2 = 1$, if Pulmonary, 0, if Extra pulmonary

(Diabetes Status) $X_3 = 1$, if Diabetic, 0, if Non-diabetic

(Residence) $X_4 = 1$, if urban, 0, if rural

Then Cox model for us becomes-

$$\lambda(t, X) = \lambda_0(t) \exp(\beta' X)$$

Where,

$$\beta = (\beta_1, \beta_2, \beta_3, \beta_4)'$$

$$X = (x_1, x_2, x_3, x_4)'$$

Or, X = (Sex, Disease type, Diabetes Status, Residence)'

Table 4: Multivariable Cox regression of factors influencing survival of TB patients in Dhemaji district

Analysis of Parameter estimates								
Parameter	Parameter estimate (β)	S.E.	Wald	d.f.	p-value	Hazard ratio	95% Hazard ratio confidence Interval	
							Lower	Upper
Sex	0.141	0.079	3.139	1	0.076	1.151	0.985	1.345
Disease Type	0.775	0.104	55.636	1	<0.001	2.171	1.771	2.662
Diabetes Status	0.676	0.095	50.783	1	<0.001	1.966	1.632	2.364

Journal of Chemical Health Risks

www.jchr.org

JCHR (2024) 14(2), 878-884 | ISSN:2251-6727

Residence	0.216	0.086	6.292	1	0.012	1.241	1.048	1.469

Using the table we can compare survival time of patients with different risk factors.

(a) To compare survival time of male with a female

λ (t, X) = λ_0(t, sex), other factors kept at fix.

$= \lambda_0(t)\exp(\beta_1 x_1)$

$= \lambda_0(t)\exp(\beta_1)$, for male

$= \lambda_0(t)$, for female

Now, HR = exp (β_1) = 1.151

The risk of bearing with TB is 15% more for male than female patients but this higher risk is not statistically significant.

This means that the risk of TB is less for patients with female than male.

(a) To compare survival time of pulmonary TB with a extra pulmonary TB

λ (t, X) = λ_0(t, Disease Type), other factors kept at fix.

$= \lambda_0(t)\exp(\beta_2 x_2)$

$= \lambda_0(t)\exp(\beta_2)$, for pulmonary

$= \lambda_0(t)$, for extra pulmonary

Now, HR = exp (β_2) = 2.17059

While comparing recovery time of pulmonary with extra pulmonary TB patient, it is found that there is a significant difference (table 6) and the risk is 2.17 times more for pulmonary than extra pulmonary. It means the treatment time is much longer while affected with pulmonary TB.

(a) To compare survival time of TB with diabetic and a non-diabetic

λ (t, X) = λ_0(t, Diabetes status), other factors kept at fix.

$= \lambda_0(t)\exp(\beta_3 x_3)$

$= \lambda_0(t)\exp(\beta_3)$, for diabetic

$= \lambda_0(t)$, for non-diabetic

Now, HR = exp (β_3) = 1.966 $\cong$ 2

Comparing diabetic with non-diabetic also we have found that risk of suffering with TB is 2 times as longer for TB patient with diabetes and this higher risk is statistically significant also. Thus TB patient with diabetes takes longer treatment time to be cured from the disease.

(a) To compare survival time of urban TB patients with a rural TB patients

λ (t, X) = λ_0(t, residence), other factors kept at fix.

$= \lambda_0(t)\exp(\beta_4 x_4)$

$= \lambda_0(t)\exp(\beta_4)$, for urban

$= \lambda_0(t)$, for rural

Now, HR = exp (β_4) = 1.242

This shows that the urban area TB patients are 24% more than the rural TB patients and this risk is statistically significant.

4. Conclusion & Discussion

In conclusion, overcrowding, smoking, tuberculosis (TB) were important risk factors and negatively affected the survival rates of TB patients in Dhemaji. At the same time, Disease Type, Diabetes status, Type of Patients and Residence, were significant risk factors and negatively affected the survival rates of TB patients in Dhemaji. The results suggested that the mycobacterium tuberculosis drug sensitivity test should be strengthened. In this study, the 1500 TB patients in Dhemaji were assessed out of 1038 male patients, 512 (49.33%) were censored and 526 (50.67%) were cured of TB. Again out of 462 female, 226 (48.92%) were censored and 236 (51.08%) were cured of TB. Overall, the median recovery time of TB patients in Dhemaji was 180 days (approx.), which means that the recovery time of patients is within the recommended treatment interval of 160 to 220 days or longer given close monitoring of patients while taking drugs.

References

1. Agerie N W. (2013): Determinants of smallholder rural farm households' participation in small scale irrigation and its effect on income in North Gondar zone: a cross-sectional approach (evidence from Dembia Woreda). Ethiopia: Mekelle University.
2. Alrabiah K, Al Alola S, Al Banyan E, Al Shaalan M, Al Johani S. (2016): Characteristics and risk factors of hospital acquired–methicillin-resistant Staphylococcus aureus (HA-

Journal of Chemical Health Risks

www.jchr.org

JCHR (2024) 14(2), 878-884 | ISSN:2251-6727

MRSA) infection of pediatric patients in a tertiary care hospital in Riyadh, Saudi Arabia. Int J Pediatr Adolesc Medi.; 3(2):71–7.

3. Assefa D, Seyoum B and Oljira L. (2017): Determinants of Multidrug-resistant tuberculosis in Addis Ababa, Ethiopia, Infect Drug Resist, 10, 209.
4. Brook Meyer R (2014): Median survival time, Wiley Stats Ref. Statistics ReferenceOnline.
5. Brueffer C, Vallon-Christersson J, Grabau D, Ehinger A, Häkkinen J, Hegardt C, Malina J, Chen Y, Bendahl P.O & Manjer J. (2018): Clinical value of RNA sequencing–based classifiers for prediction of the five conventional breast Cancer biomarkers: a report from the population-based multicenter Sweden Cancerome analysis network—breast initiative. JCO Precis Oncol., 2:1–18.
6. Cochran W G. (1977): Sampling Techniques. 3rd Ed. New York: Wiley.
7. Dabrowska D M. (1988): Kaplan-Meier estimate on the plane. Ann Stat. 16 (4):1475–89.
8. Eldholm V, Balloux F (2016): Antimicrobial resistance in mycobacterium tuberculosis; the odd one out, Trends Microbial. 24(8), 637-48.
9. Espinal M.A, Laszlo A., Simonsen L, Boulahbal F, Kim S.J., Reniero A, Hoffner S, Rieder H.L, Binkin N, & Dye C (2001): Global trends in resistance to antituberculosis drugs. N Engl J Med.; 344(17):1294–303.
10. Fair R. J & Tor Y. (2014): Antibiotics and bacterial resistance in the 21st century. Perspect Med Chem.; 6:25–64. https://doi.org/10.4137/PMC.S14459.
11. Falzon D, Schunemann H.J, Harausz E, Gonzalez-Angulo L, Lienhardt C, Jaramillo E and Weyer K. (2017): World Health organization treatment guidelines for drug-resistant tuberculosis, 2016 update, Eur Respir J; 49 (3): 1602308.
12. Fan J. (2018): Local polynomial modelling and its applications: monographs on statistics and applied probability 66. New York: Routledge.
13. Faustini A, Hall A.J., & Perucci C.A. (2006): risk factors for multidrug resistant tuberculosis in Europe, a systematic review. Thorax. 61(2); 158-63.
14. Gutierrez RG. (2002): Parametric frailty and shared frailty survival models. Stata J; 2(1):22–44.
15. Koul A, Arnoult E, Lounis N, Guillemont J, and Andries K (2011): The challenge of new drug discovery for tuberculosis, Nature; 469 (73310, 483.
16. Kundu D, Sharma N, Chadha S, Laokri S, Awungafac G, Jiang L and Asaria M (2018): analysis of multidrug resistant tuberculosis (MDR-TB) financial protection policy: MDR-TB health insurance schemes, in Chhattisgarh state, India. Health econ. Rev. 8(1), 3.
17. Miller Jr, R.G.(2011): Survival Analysis. Vol. 66. Chichaster: Wily; 2011.
18. Nigus D, Lingerew W, Beyene B, Tamiru A, Lemma M and Melaku M. (2014): Prevalence of multi drug resistant tuberculosis among presumptive multi drug resistant tuberculosis cases in Amhara National Regional State, Ethiopia. J Mycobac Dis. 4(152): 2161–1068.1000152.
19. Reid N. (1981): Estimating the median survival time. Biometrika, 68 (3): 601–8.
20. Singh J.A, Upshur R, & Padayatchi N. (2007): XDR-TB in South Africa, no time for denial or complacency, PLoS Med. 4(1), 50.
21. Tacconelli E, Cataldo M, Dancer S, De Angelis G, Falcone M, Frank U, Kahlmeter G, Pan A, Petrosillo N, and Rodríguez-Baño J. (2014): ESCMID guidelines for the management of the infection control measures to reduce transmission of multidrug-resistant gram-negative bacteria in hospitalized patients. Clin Microbiol Infect., 20:1–55.
22. Tarai B, Das P & Kumar D. (2013): Recurrent challenges for clinicians: emergence of methicillin-resistant Staphylococcus aureus, vancomycin resistance, and current treatment options. J Lab Phys.; 5(2):71.
23. Tyrrell F, Stafford C, Yakrus M, Youngblood M, Hill A, & Johnston S. (2017): Trends in testing for mycobacterium tuberculosis complex from US public health laboratories, 2009–2013. Public Health Rep.;132 (1):56–64.
24. Van Der Meulen A. (2012): Life tables and survival analysis. The Hague: Statistics Methods Statistics Netherlands.
25. World health organization Multidrug and extensively drug-resistant TB (M/XDR-TB) (2010): Global report on Surveillance and response.
26. World Health organization Global Tuberculosis report Geneva (2017): World Health organization, License, CC BY-NCSA 3.0 IGO.