MODELLING THE CURE RATE OF AN INFECTIOUS DISEASE (TUBERCULOSIS)

ABSTRACT

In this investigation, the tuberculosis and tuberculosis co-contaminated with HIV information were demonstrated. The information inspected time until a patient is relieved of the infection having a few patients right controlled. With the idea of the information, the proper examination is endurance investigation. This work was roused by the way that irresistible illness and co-disease has become defense in our general public. The investigation analyzed the variables impacting the fix pace of the illnesses among age and sexual orientation. This assists with making mindfulness and advance approach development and guideline on the fix of these illness. The examination target fitting suitable models to the TB and TB/HIV co-disease information.

Different models have been utilized in fitting the fix pace of the irresistible sickness, yet this examination expanded the models utilized for the information to 3 and 4 boundaries parametric models instead of outstanding and 2 types of weibull utilized by Balogun in his exploration proposal for these information. The discoveries of the examinations shows that none of the covariates was critical for the tuberculosis information (model 1) in both the semi-parametric and parametric models. Summed up gamma supposedly was the best fit among the parametric models fitted. In the tuberculosis co-tainted with HIV information (model 2), both age and sex are factually huge in both semi-parametric and parametric models. Be that as it may, gamma model fits the information best among the parametric models fitted. The peril rate for females was about 62% higher than that of guys.

INTRODUCTION

An irresistible sickness is an infection that is brought about by the attack of a host by specialists whose exercises hurt the host's tissue and can be communicated to others. There are six significant kinds of contamination, they are: microbes, infections, organisms, protozoa, helminths and prions. Tuberculosis is a possibly genuine irresistible illness that mostly influences the lungs. The microorganisms that cause tuberculosis are spread starting with one individual then onto the next through small beads delivered to the air by means of hacks and sniffles. Tuberculosis(TB) is brought about by microorganisms MY-COBACTERIUM TUBERCULOSIS. TB is reparable and preventable. As per world wellbeing association, around one-fourth of the total populace has inert TB, which means individuals have been tainted by TB microbes however are not sick with the illness and can't trans-mit the disease(WHO,2018). Individuals contaminated with TB microscopic organisms have a 5-15 % lifetime danger of becoming sick with TB. In any case, people with bargained safe frameworks, for example, individuals living with HIV, ailing health or diabetes or individuals who use tobacco, have a lot higher danger of becoming sick when an individual creates dynamic TB disease(WHO,2018).

The symptoms(such as hack, fever, weight reduction or night sweat) might be gentle for a long time. This can prompts delay in looking for care and result in transmission of the microscopic organisms to other people. Individuals with dynamic TB can influence 10-15 others through close contact throughout a year. Without legitimate treatment,45 percent of HIV-pessimistic individuals with TB by and large and practically all HIV-positive individuals with TB will die(WHO,2018). HIV pos-itive kids with pneumonic TB are grouped into WHO clinical stage 3 while those with extra-aspiratory TB are put in stage 4.

Who is most at risk?

TB generally influence grown-ups in their most profitable year. Be that as it may, all age bunches are in

danger. Individuals who are contaminated with HIV are 20 to multiple times bound to create dynamic TB. The danger of dynamic TB is additionally more noteworthy in people experiencing different conditions that impede the insusceptible framework. 1,000,000 youngsters (0-14 years) became sick with TB and 230000 kids (in-cluding kids at HIV related TB) passed on from the sickness in 2017. Tobacco use enormously increment the danger of TB illness and demise. 7.9% of TB cases overall are ascribed to smoking.

Symptoms

Normal side effects of dynamic lung TB are hack with sputum and blood now and again chest agony or torment with breathing and hacking , shortcoming, accidental weight reduction, fever and night sweat, exhaustion, chills, loss of craving. Tuberculosis can likewise influence other piece of your body including your kidneys, spines, or cerebrum. At the point when TB happens outside your lungs, signs and indications shift as per the organs in question. In spite of the fact that your body may hold the microscopic organisms that cause tuberculosis, your insusceptible framework ordinarily can keep you from getting debilitated. Specialists make a differentiation between LATENT TB AND ACTIVE TB.

LATENT TB : in this condition, you have a TB infection but the bacteria remains in your body in an inactive state and cause no symptoms. Latent TB also called inactive TB or 5

TB infection isn’t contagious. It can turn into active TB so treatment is important for the person with latent TB and to help control the spread of TB. An estimated 2 billion people have latent TB.

ACTIVE TB: this condition makes you sick and can spread to others. It can occur in the first few weeks after infection with the TB bacteria or might occur years later (WHO,2018).

Diagnosis

Diagnosing multi-drug safe and broadly drug safe TB just as HIV-associ-ated TB can be mind boggling and costly. In 2016, 4 new diagnostics tests were suggested by WHO-a quick atomic test to distinguish TB at fringe wellbeing focuses where Xperts MTB/RIF(Mycobacterium Tuberculosis with Rifampicin) can't be utilized, and three tests to identify protection from first and second line TB medications. Prepared research facility professionals see sputum tests under a magnifying instrument to check whether TB microbes are available (WHO,2018). Microscopy identifies just a large portion of the quantity of TB cases and can't identify drug-obstruction.

The most ordinarily utilized analysis are:

1. Chest X-ray
2. Culturing bacteria to test for TB
3. Fluorescent microscopy

4. Serological tests
5. Sputum smear microscopy
6. TB drug susceptibility tests
7. TB skin test

(WHO,2018)

Causes

TB is brought about by microbes that spread from individual to individual through tiny beads delivered into the air. This can happen when somebody with the untreated, dynamic type of tuberculosis hacks, talks, sniffle, spits, giggles or sings. In spite of the fact that TB is infectious, it is difficult to get. You are substantially more prone to get TB from somebody you live with or work with than from outsiders. A great many people with dynamic TB who've had suitable medication treatment for atleast fourteen days are not, at this point infectious.

Complications

Without treatment TB can be fatal.Untreated dynamic sickness normally influences your lungs yet it can spread to different pieces of your body through your circulation systems e.g spinal rope, joint harm, expanding of the films that cover your brain(meningitis), liver or kidney issues, heart problems.

Treatment

TB is a treatable and curable disease. Active drug susceptible TB disease is treated with a standard 6 month course of 4 anti-microbial drugs that are provided with information, suspension and support to the patient by a health worker or trained volunteer. Between 2000 and 2017, an estimated 54 million lives were saved through TB diagnosis andtreat- ment(WHO,2018).

Prevention

Protect your family and friends from active TB if you have because it takes a few weeks of treatment with TB medications before you're not contagious anymore. Follow the tips below:

* Stay home
* Ventilate the room
* Cover your mouth
* Wear a mask
* Finish your course of medication
* Vaccinations

Infants are often vaccinated with Bacillus Calmotte-Guerin(BCG). Dozens of new TB vac- cines are in various stages of growth and testing (www.mayoclinic.com)

Multi-Drug Resistance Tuberculosis

Medication obstruction arises when hostile to TB meds are utilized improperly through mistaken solution by medical services suppliers, low quality medications and patients halting therapy rashly. Multidrug-obstruction TB (MDR-TB) is a type of TB brought about by microbes that don't react to ioniazid and rifampicin, the two most impressive first line against TB drugs. MDR-TB is treatable and reparable utilizing second-line drugs. In any case, second line treatment alternatives are restricted and require broad chemotherapy(up to 2 years) with prescriptions that are costly and poisonous at times, more serious medication obstruction can create. Broadly drug-obstruction TB (XDR-TB) is a more genuine type of MDR-TB brought about by microbes that don't react to the best second line hostile to TB drugs, frequently leaving patients with no further treatment choices (WHO,2018).

Why survival analysis

Survival model will be appropriate in this study since the time until a patient is cured was recorded. In everyday life, we want to know the time it will take for a person to recover from a particular disease, or time until death of an individual infected with a disease. This type

of situation is called survival analysis.

Generally, survival analysis is a collection of statistical techniques for data analysis for which the outcome variable of interest is time duration until an event occurs. The time could be in years, months, weeks, or days from the starting point of follow-up of a person until an event hap- pens; on the other hand, time can imply the age of a person when an event occurs. The event means death, disease incidence, relapse from remission, recovery (e.g., return to work or cure) or any assigned experience of intrigue that may happen to a person. Albeit more than one event might be considered in a similar analysis, we will expect that just a single event is of assigned interest. At the point when more than one event is considered (e.g., demise from any of a few causes), the statistical issue can be described as either a repetitive event or a competing risk problem. In a survival analysis, we majorly consider the time variable as survival time, which gives the time period that a particular individual has survived. If the event of interest more often than not is death, disease incidence, or some other negative individual experience then the event is considered as a failure. In any case, survival time might be ”time to cure after a treatment procedure,” in which case failure is a positive event (as in case of this study). A study that focus on how long patients survival after receiving a heart transplant can be allude to an example of survival analysis issue. This example considers the event of ”death” with the out- come being ”time until death (number of months after getting a transplant).”

Survival analyses consider a key analytical issue called censoring. Generally, censoring hap-

pens when we have some data about individual survival time, however we don't know the survival time precisely. If for a given patient, the study ends while the patient doesn't get the event, at that point that patient's survival time is viewed as censored. For this individual, the survival time is in any event is the period as long as that the individual has been followed, however in the event that the individual encounters the event after the study ends, we can have some idea about the total survival time. There are commonly three reasons why blue censoring may happen:

1. an individual not experiencing the event before the study ends;
2. an individual is lost to follow-up amid the study period;
3. an individual pulls back from the study as a result of death (suppose death is not the event of interest) or some other reason such as negative medication response or another conflicting risk.

Right-censored data: can happen when an individual's actual survival time is greater than the individual's observed survival time. That is, the individual's actual survival time at the right side of the follow-up period becomes incomplete, happening when the study ends or when the individual is lost to follow-up or is pulled back. For these data, the unknown total survival time interval has been cut off (censored) from the right side of the observed survival time interval. Despite the fact that data can likewise be left-censored, most survival data is right-censored.

Left-censored data: can happen when an individual's actual survival time is less than or

equal to that individual's observed survival time. For an instance, the event that following people until they become HIV positive, when a subject first tests positive for the infection can be taken as a failure. Nonetheless, we may not aware about the specific time of initial exposure to the infection, and via this manner we are not aware precisely when the failure happened. Hence, the survival time is changed on the left side as the genuine survival time (ends at exposure), is less than the follow-up time (ends when the subject's test is positive). As such, if an individual is left-censored at time t, we realize that an event has occurred between time 0 and t, yet we are un-evident about the specific time of event occurrence.

Data of survival analysis can be interval censored, when if a subject's unobserved actual survival time falls within a certain known specified time interval. For instance, again thinking about HIV surveillance, when two HIV tests performed on a subject, where the subject was HIV negative at the first time (t1) of the first test and HIV positive at the second time (t2) of the second test. In such a situation, the subject's actual survival time happened after time t1 and before time t2, (between t1 and t2) i.e., the subject is said to be interval censored between the time interval (t1, t2). The idea of Interval censoring really incorporates with right-censoring as well as left-blue censoring as special cases. When whatever the point the estimation of t1 is 0 and t2 is a known upper bound on the true survival time left-censored data occurs. Interestingly, when whatever the point the estimation of t2 is infinity and t1 is a known lower bound on the true survival time, right-censored data occurs. If an individual is right-censored due to the occurrence of a competing event such as death from another cause, at that point in this unique circumstance, the investigating thoughts needed to be carried about the

true survival time if the competing event had not happened. At the end of the day, when we express that the value of the upper bound for the true survival time is infinity for right-censored data, we are considering what might have happened without a competing risk.

When it comes to censoring, three assumptions needed to be considered:

1. Independent(vs Non independent) censoring
2. Random (vs Non-random) censoring
3. Non-informative (vs Informative) censoring

Within any subgroup of interest, the subjects who are censored at time t are representative of all the subjects in that subgroup that remained at risk at time t with respect to their survival experience is assumed in independent censoring.

Random censoring essentially means that, suppose at a time t, the subjects who are censored assumed to be representative of the entire study subject who remained at risk at time t with respect to their survival experience.

Non-informative censoring: whether censoring is informative or not depends on two distributions-

1. The distribution of the time-to-event random variable and
2. The distribution of time-to censorship random variable.

Non- informative censoring occurs if the distribution of survival time (T) provides no information about the distribution of censorship time (C) and vice versa. Note , however that

the data must still identify which subjects are or are not censored.

The assumption of independent censoring is the most useful of the three types for drawing correct inference that compare the survival experience of two or more groups.

(Kleinbaum D.G. and Klein M., 2012)

Some Terminologies and Notations

We denote T as the random variable for a person's survival time. T can be greater than or equal to zero.

t is any specific value of interest for a random variable T.

d(0,1) is a random variable indicating either failure or censorship.

♣S(t) is the survival function and it gives the probability that a person survives longer than some specified time t.i.e $s(t) = P(T > t)$.

All S(t) are

* non increasing at

$$t = 0,\ s(t) = S(0) = 1 \tag{1.1}$$

$$t = (\infty),\ S(\infty) = 0 \tag{1.2}$$

♣h(t): is the hazard function which gives the instantaneous potential per unit time for the event to occur, given that the individual has survived upto time t. Mathematically,

$$h(t) = \lim_{\Delta t \to 0} \frac{(t \leq T < t+\Delta t | T \geq t)}{Ot} \tag{1.3}$$

The hazard function is sometimes called a condition failure rate. in particular, for a specified value of t, the hazard function $h(t)$ has the following characteristic:

* It is alinary non-negative, i.e $t > 0$, $h(t)>=0$
* It has no upper bound.

Relationship between the hazard and survival function

$$h(t) = \frac{f(t)}{S(t)} \tag{1.4}$$

$$S(t) = exp[-\int_0^t h(u)du \tag{1.5}$$

$$h(t) = -[\frac{dS(t)/dt}{S(t)}] \tag{1.6}$$

When either of survival or hazard function is known, you can get the other.

Aims and objectives

The aim of this study is to appropriately model the tuberculosis and tuberculosis/HIV co-infection data. However, the objectives are to

- Fit various models; Non-parametric, semi-parametric and parametric models to the datasets.
- Assess the best model using appropriate criteria.
- Identify covariates that influence the cure rate of tuberculosis and tuberculosis co- infected with HIV patients.

Significance of the study

Nowadays, statistical models are needed in work related to an illness. This research helps to provide information about the threat TB and TB co-infected with HIV poses to health and gives insight on how to prevent and fight the manifestation of these disease. It also helps to identify risk factors and inform government and health sectors on planning and policy

formation and implementation against the disease. Educational sector will find the study useful in creating space for innovative methods and discovering ways of fighting the spread of TB and TB/HIV co-infection. The significance of the study is to know the cure rate of TB after being infected with the virus using survival analysis estimation process which include:

1. Parametric survival models
2. Semi-parametric survival models
3. Non-parametric survival models

DATA ANALYSIS

This chapter contains the data presentation and empirical analysis of the infectious disease. All analysis was done using the statistical package R.(R Core Team, 2018)

Data Collection

The data used is a secondary data gotten from the Medical Record Unit of University Of Ilorin Teaching Hospital (UITH), Ilorin, Kwara State over the period of 2000-2015 on patients with TB (Example 1) and TB co-infected with HIV(Example 2). The data was previously used by Balogun (2018) in his research thesis. He used exponential and 2 forms of Weibull models in his study, while this study included the 3 and 4 parameters parametric models.

Descriptive Statistics of the TB data

Example 1

This data is from the UITH and is of patient treated of TB disease in the year 2000-2015. There are 518 observations. The age(in years), time(in months) ,censoring(event=1,censored=0) and gender(male=1, female=0) of each patient were recorded.

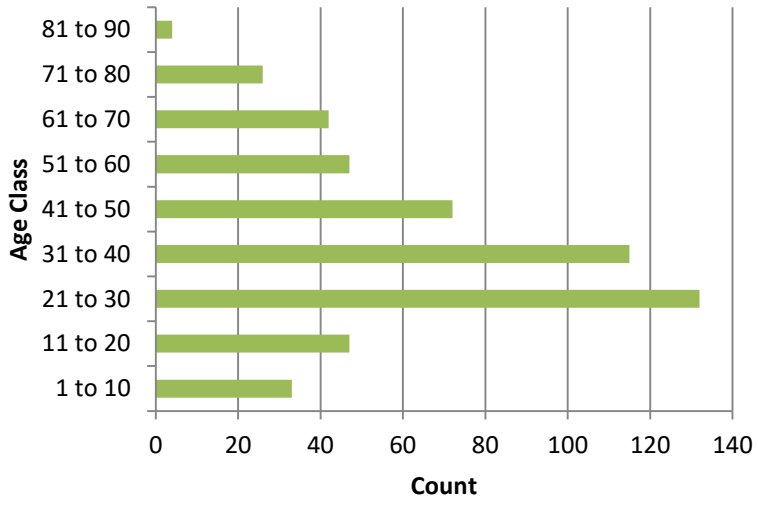

Figure 1: Age separation of TB patients.

Table 1a: Sample of the TB patients' data

S.No	Ag	Tim	Cens	Gend
1	33	15	1	0
2	27	40	1	0
3	48	5	1	0
4	54	32	1	0
5	50	14	1	1
.	.	.	.	.
514	30	21	1	0
515	25	40	1	1
516	60	31	1	0
517	54	30	1	0
518	70	4	1	1

Source: Medical Record Department U.I.T.H., Ilorin

Table 1b: Description of the TB data

Covariates	Description
S/N	Serial number of patients
Gend	male=1, female=0
Ag	age in years of patients
Tim	Time in months until a patient survives
Cens	Censoring status censored=0, uncensored=1

Table1c:Preliminary Data Analysis

Covariates		
Age	minimum age	1
	maximum age	88
	average age	37.57239
Gender	number of males	298
	number of females	220
Censoring	event	448
	censored	70
Time	minimum time	1
	maximum time	129
	average time	20.14093

Table 1d: Test for the PH assumption

	rho	chisq	p
age	-0.0404	0.767	0.381
Gender	-0.0159	0.114	0.735
Global	NA	0.906	0.636

The p-value of the covariates were not significant($p-value$0.05) from the test above(individual test of age and gender with p=0.381 and 0.735 respectively, and global test with p-value=0.636) and that the PH assumption was satisfied.

Kaplan Meier estimate

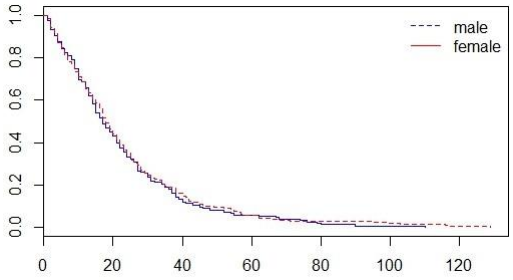

Figure 1a:Kaplan meir estimate of TB data

Result of the parametric estimate

The parametric models is being fitted to the TB data with distributions; exponential, weibull, log-normal, log-logistic, gompertz,gamma, gengamma and genf.

Table 1e: Parametric estimates of TB data

	exp	weibul	lnorm	llogis	gomp	gamma	geng	genF
Age	0.00197	-0.00235	-0.00305	-0.00312	0.00244	0.00241	-0.00252	-0.00251
Gender	-0.05032	0.05583	0.04236	0.03148	-0.06604	-0.05079	0.04613	0.0446
LL	-1857.868	-1844.631	-1856.735	-1855.2	-1854.321	-1841.817	-1840.687	-1840.679
Df	3	4	4	4	4	4	5	6
AIC	3721.735	3697.262	3721.47	3718.401	3716.643	3691.633	3691.375	3693.359
P-value(age)	0.4486	0.274	0.214	0.184	0.349	0.2667	0.256	0.2625
P-value(gender)	0.5987	0.480	0.646	0.724	0.491	0.5256	0.578	0.5941

Result of the Semi-parametric estimate

Table 1f: Semi-parametric estimate

Covariate	coef	exp(coef)	s.e(coef)	z	p-value
Age	0.002455	1.002458	0.002619	0.937	0.349
Gender	-0.062120	0.939770	0.096028	-0.647	0.518

Likelihood ratio test=1.24 on 2df, p=0.5

wald test=1.25 on 2df, p=0.5 score(logrank)

test=1.25 0n 2df, p=0.5 AIC= 4652.09 on

2df

Model:

$$\hat{h}(t) = \hat{h}_0(t)e^{0.002455age - 0.062120gender} \tag{3.1}$$

Result

From the analysis above, of the 518 TB patients, 298 were males and 220 were females. the best fitting model of the parametric model is the gengamma having the least AIC (3691.375). None of the covaraites was significant ($Pvalue > 0.05$). Also, the cox model gives a likelihood ratio test of 1.24 and an AIC of 4652.09. None of the covariates also was significant ($Pvalue > 0.05$). The p-value of the covariates age and gender are 0.394 and 0.518 respectively.

Example 2

The data is of TB patients co-infected with HIV. There are 133 observations with covariates age and gender. The survival time and censoring was also recorded.

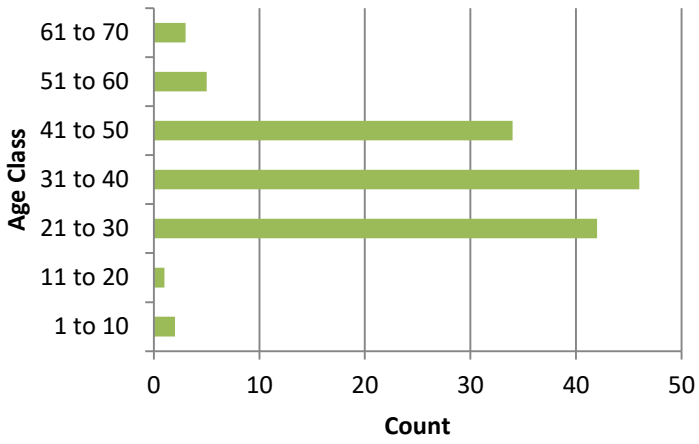

Figure 2: Age separation of TB patients co-infected with HIV.

Table 2a:Sample of the TB/HIV co-infected patient data

S.NO	age	time	censoring	gender
1	25	13	1	0
2	40	23	1	1
3	27	19	1	0
4	27	34	1	1
5	35	37	1	0
.	.	.	.	.
128	39	22	1	1
129	30	35	1	1
130	47	24	1	1
131	40	15	1	0
132	45	35	1	0
133	70	17	1	0

Source: Medical Record Department U.I.T.H., Ilorin

Preliminary data Analysis

Table 2b:Preliminary data Analysis of TB/HIV co-infected patient data

Covariates		
Age	minimum age	2
	maximum age	70
	average age	36.69925
Gender	number of males	53
	number of females	80
Censoring	event	111
	censored	22
Time	minimum time	1
	maximum time	88
	average time	19.05263

Table 2c: Test for the assumption of PH

	rho	chisq	p
Age	0.1475	2.4287	0.119
Gender	0.0101	0.0115	0.914
Global	NA	2.5299	0.282

The result shows that all the covariates are not significant(*P value* 0.05) from the test

above(individual age and gender with Pvalue=0.119 and 0.914 respectively and a global test with p-value=0.282), hence, the PH assumption holds.

Kaplan Meier estimate Of TB/HIV Co-infection Data

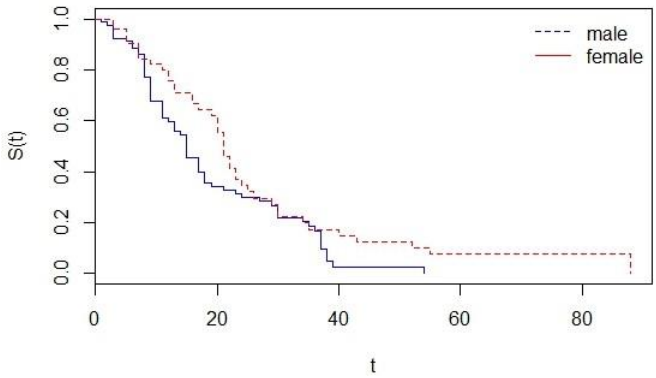

Figure 2a:Kaplan-Meir estimate 2

Result of the Parametric estimate of TB/HIV Co-infection Data

The parametric model is being fitted to the TB data with distributions; exponential, weibull, log-normal, log-logistic, gompertz,gamma, gengamma, genf.

Table 2d: Parametric estimates of TB/HIV co-infected patient

	exponen	weibull	lnorm	llogis	gomp	gamma	geng	genF
Age	0.01808	-0.01925	-0.01508	-0.013009	0.02868	0.01766	-0.01768	-0.0177
Gender	-0.34264	0.34669	0.35054	0.339186	-0.56454	-0.34140	0.34145	0.342
LL	-454.7954	-440.2774	-443.3006	-443.0412	-447.4042	-439.7842	-439.7839	-439.7844
Df	3	4	4	4	4	4	5	6
AIC	915.5907	888.5549	894.6012	894.0824	902.8083	887.5683	889.5678	891.5688
P-value(age)	0.0324	0.000	0.025	0.045	0.001	0.0022	0.003	0.0028
P-value(gender)	0.0796	0.006	0.021	0.022	0.007	0.0108	0.011	0.0107

Result of the Semi-parametric estimate of TB/HIV Co-infection Data

Table 2e: Semi-parametric estimate of TB/HIV data

Covariate	coef	exp(coef)	s.e(coef)	z	p-value
Age	0.02713	1.02750	0.00895	3.031	0.00244
Gender	-0.47920	0.61928	0.20604	-2.326	0.02003

Likelihood ratio test=12.24 on 2df, p=0.002, n=133
wald test = 12.9 on 2df, p=0.002 Score (logrank) test=12.9 on 2df, p=0.002 AIC=855.3267 on 2df

Model:

$$\hat{h}(t) = h_o(t)e^{0.02713age - 0.47920gender} \tag{3.2}$$

Result Of the 133 TB/HIV patients, 53 were males and 80 were females. The best fitting model of the parametric model is the gamma with AIC of 887.5683. The covariates has p-

values less than 0.05 which are significant, hence affect the cure rate of TB/HIV co-infection. Also,the cox model gives a likelihood ratio test of 12.24 and an AIC of 855.3267 with a significant age and gender having p-values of 0.0024 and 0.02003 respectively (p-values less than 0.05).

CONCLUSION

From the analysis above, the generalized gamma fit TB data better with an AIC of 3691.375 on 5 from the parametric models being fitted. Gamma fits better in example 2 with aic of 887.5683 of the parametric models being fitted. The estimated hazard ratio for males relative to females is 0.62 i.e there is 62% lower incidence for males than for females. Lastly, in TB data; None of the covariate influence the cure rate of TB in the parametric and semi- parametric model.

In TB/HIV data, both age and gender influence the cure rate of TB co-infected with HIV in parametric and semi-parametric model.

Bibliography

[1] Anderson R.M and May R.M, (1992). *Infectious Disease of Humans*: Dynamics and control, New York: Oxford University press.

[2] Abera, B., Kate, F., Zelalam, H. and Andrew, F. (2009). The association of TB with HIV infection in Oromia Regional National State, Ethiopia in 2006/7. *J*ournal of Ethiopia Health Dev., 23(1): 63-67.

[3] Nwabuko C. O., Ejele, O. A., Chuku, A., Nnoli, M. A. and Chukwuonye, I. I. (2012). Prevalence of Tuberculosis-HIV coinfection and relationship between Tuberculosis and CD4/ESR in HIV patients in Niger Delta region of Nigeria.*I* OSR Journal of Dental and Medical Sciences (JDMS), 2(4): 1-4.

[4] Kleinbaum, D.G. and Klein M. (2012):*S* urvival Analysis ; A self-learning Text , Third Edition, Statistics for Biology and Health, Springer Science + Business Media; ISBN 9781441966452 .

[5] Balogun (2018), Modelling Cure Rate of Infectious Disease with Co-infection (Unpub- lished Thesis).

[6] Brust J.C.M., Lygizos M., Chaiyachati K., Scott M., Van der Merwe T.L., et al. (2011) Culture Conversion Among HIV Co-Infected Multidrug- Resistant Tuberculosis Patients in Tugela Ferry, South Africa.*P*LoS ONE 6: e15841 doi: 10.1371/journal.pone.0015841.

[7] Kalbfleisch J.D. and Prentice R.L. (2002). The Statistical Analysis of Failure Time Data-Second Edition.*J* ohn Wiley and Sons, New York.

[8] Peterson E. et.al., (2019). "It's TIME" — International Journal of Infectious Diseases Tuberculosis Theme Series *I*nternational Journal Of Infectious Diseases, 80:S1-S5.

[9] Mulisa G, Workneh T, Hordofa N, Suaudi M, Abebe G and Jarso G.,(2015). Multi-drug Resistant Mycobacterium Tuberculosis and associated risk factors in the Oromia region of Ethiopia, *I* nternational Journal of Infectious Diseases, 39:57-61.

[10] Terefe A.N., Gebrewold L.A.,(2018). Modelling Time to recovery of Adult TB patients in Mizan-Tepi University Teaching Hospital, South-West Ethopia.*M* ycobact Dis 8: 258. DOI:10.4172/2161-1068.1000258.

[11] Corbett E.L, Watt C.J, Walker N., Maher D., Williams B.G., Raviglione M.C.and Dye C., (2003). The Growing Burden Of Tuberculosis: Global Trend and Interactions with the HIV epidemic. Arch Intern Med.;*1* 63(9):1009-21.

[12] Kamenju P. and Aboud S.,(2011). Tuberculosis-HIV Co-infection Among Patients in Muhimbili National Hospital in Dare es Salam, Tanzania *T* azan J Health Res,Jan; 13(1): 21-6.

[13] Kaplan E.L and Meir P., (1958). Non-parametric Estimation From Incomplete Observation.*J* Ameri Statist Asso 53: 457-481.

[14] Cox D.R.,(1972). Regression Models and Lifetables.*J* Royal Stati Society 34: 187-220.

[15] Weibull W., (1951). "A Statistical Distribution Function Of Wide Applicability". *J* ournal Of Applied Mechanics, 18:293-297.

[16] Aitchison J. and Brown J.A.C., (1957). "The Log-normal Distribution", *C* ambridge University Press.

[17] Akaike H., (1974). A New Look At The Statistical Model Identification, *I* EEE Transactions On Automatic Control, 19(6):716-723, DOI: 10.1109/TAC:1974.1100705.

[18] Gompertz B.,(1825). The Nature Of The Function Expression Of The Law Of Hu- man Mortality And On A New Mode Of Determining The Value Of Life contigen- cies,*P* hilosophical Transactions Of The Royal Society Of London, Vol. 115:513-583.

[19] Gloria A.A, John D.W., Kamran S. and xaolin W., (2012). Assessing The Impact Of TB/HIV Services Integration On TB Treatment Outcomes And Their relevance In TB/HIV Monitoring In Ghana *I* nfectious Disease Of Poverty 1:3.

[20] Asik S., Setyaningsih B., Nasutwn M.S., Parwati G.C., Yuzwar E.Y., Osberg M., Hanson C.L., Hymoff A., Mingkwan P., Makayova J., Gebhard A., and Waworuntu W., (2017). Quality Tuberculosis Care in Indonesia: Using Patient Pathway Analysis to Optimize Public-Private Collaboration, *T*he journal Of Infectious Diseases, 216(7):S724-S732

[21] Nagu T.J., Aboud S., Mwiri R., Matee M., Rao M.,Fawzi W.W., Zumla A., Macurer M.J., and Mugusi F., (2017). Tuberculosis Associated Mortality In A Prospective Cohort In Sub-Saharan Africa: Association With HIV and Anti-Retroviral Therapy, *I*nternational Journal Of Infectious Diseases, 56:39-44. doi: 10.1016/j.ijid.2017.01.023

[22] Goldin A.G.and Baker W.T., (1985). Tuberculosis of the Female Genital Tract. *J*. K.Y Med Assoc 83: 75

[23] Varma T.,(2008). Tuberculosis of The female Genital Tract, *G*lob.libr.Women's Med., (ISIN : 1756-2228). Doi: 10.3843/GLOWM.10034

[24] Mekonnen D., Awoke D., Desalegn E., (2015). "Tuberculosis Co-infected With HIV Among Associated Factors Among Patients On Directly Observed Treatment Short Course In Northeastern Ethopia" *B* MC Research Notes, 8(1). Doi: 10.1186/s13104- 015-1664-0

[25] World Health Organization WHO Report 2011: Global Tuberculosis Control www.WHO.int

[26] Therneau T.M. and Grambsch P.M. (2000).*M*odeling Survival Data: Extending the Cox Model, Springer.

[27] Rao M., Ippolito G., Mfinanga S., Ntoumi F., Yeboah-Manu D., Vilaplana C., Zumla A., Maeurer M., (2019). Latent TB Infection (LTBI)— Mycobacterium Tuberculosis Pathogenesis And The Dynamics Of The Granuloma Battleground. *I* nternational Journal Of Infectious Diseases, 80:S58–61.

[28] Lienhard J.H and Meyer P.L., (1967). A Physical Basis For The Generalized gamma*Q* uarterly Of Applied Mathematics; Vol. XXV:3.

[29] Cox D.R. and Oakes D., (1984). Analysis of Survival Data. *C* hapman and Hall, New York.

[30] R Core Team (2018). R: A language and environment for statistical computing. R Foun- dation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Appendix

R Program

tbdat = read.csv(”*tbdata.csv*”)

attach(*tbdat*)

tbhivdat = read.csv(”tb_h*ivdata.csv*”)

attach(*tbhivdata*)

PLOT HAZARD FUNCTION

To compute a locally optimal estimate

library(”*survival*”)

library(”*muhaz*”)

library(*flexsurv*)

library(*survminer*) *for ggsurvplot function*

$fit1 < -muhaz(tim, cens)$

plot(*fit*1, *col* = ”*blue*”, *lty* = 1, *lwd* = 3, *main* = ”*HAZARDFUNCTION*”)

summary(*fit*1)

To compute a globally optimal estimate $fit2 <$

$-muhaz(tim,$ *cens, bw.method* = ”*g*”)

plot(*fit*2)

library(*RTCGA.clinical*)

fit < −*survfit*(*Surv*(*tim, cens*) *gend, data = tbdat*)

Visualize with survminer

ggsurvplot(*fit, data = tbdat, risk.table = TRUE*)

fitaKaplan − Meierandplotit

fit < −survfit(*Surv*(*tim, cens*) *gend, data = tbdat*)

plot(*fit, lty* = 2 : 3)

legend(100, .8, *c*("1 = *Male*", "0 = *Female*"), *lty* = 2 : 3)

COX's Regression

fitcox < −coxph(*Surv*(*tim, cens*) *age + gend, data = tbdat*)

summary(*fitcox*)

confint(*fitcox*)

exp(*cbind*(*coef* (*fitcox*), *ci*))

A Check of the PROPORTIONAL HAZARD ASSUMPTION

asmp < −cox.zph(*fitcox, transform* = "*km*", *global = TRUE*) *asmp*

ggcoxzph(*asmp*)

temp < −cox.zph(*fitcox*)

print(*temp*) *display the results*

par(*mfrow* = *c*(1, 2))

plot(*temp*) *plot curves*

summary(*coxph*(*Surv*(*tim, cens*) *age + gend, data = tbdat*))

summary(*coxph*(*Surv*(*tim, cens*) *age* + *factor*(*gend*), *data* = *tbdat*))

Compare generalized gamma fit with Weibull

fitg < −*flexsurvreg*(*formula* = *Surv*(*tim, cens*) 1, *data* = *tbdat, dist* = ”*gengamma*”)

fitg

summary(*fitg*)

fitgg < −*flexsurvreg*(*formula* = *Surv*(*tim, cens*) *age*+*gend, data* = *tbdat, dist* = ”*gengamma*”)

fitgg

summary(*fitgg*)

fitggint < −*flexsurvreg*(*formula* = *Surv*(*tim, cens*) *age* ∗ *gend, data* = *tbdat, dist* = ”*gengamma*”)

fitggint

fitw < −*flexsurvreg*(*formula* = *Surv*(*tim, cens*) *age*+*gend, data* = *tbdat, dist* = ”*weibull*”)

fitw

summary(*fitw*)

fitex < −*flexsurvreg*(*formula* = *Surv*(*tim, cens*) *age* + *gend, data* = *tbdat, dist* = ”*exp*”)

fitex

summary(*fitex*)

fitln < −*flexsurvreg*(*formula* = *Surv*(*tim, cens*) *age*+*gend, data* = *tbdat, dist* = ”*lnorm*”)

fitln

summary(*fitln*)

```
fitlnint < −flexsurvreg(formula = Surv(tim, cens) age ∗ gend, data = tbdat, dist =
"lnorm")
fitlnintfitgom < −flexsurvreg(formula = Surv(tim, cens) age+gend, data = tbdat, dist =
"gompertz")fitgom
summary(fitgom)
fitgf < −flexsurvreg(formula = Surv(tim, cens) age+gend, data = tbdat, dist = "genf ")
fitgf
summary(fitgf )
fitg < −flexsurvreg(formula = Surv(tim, cens) age+gend, data = tbdat, dist = "gamma")
fitg
summary(fitg)
AIC(fitgg, fitw, fitex, fitln, fitgom, fitgf, fitg)
par(mfrow = c(3, 3))
plot(fitgg)
plot(fitw, col = "blue", lwd.ci = 1, lty.ci = 1);
plot(fitex, col = "green", lwd.ci = 1, lty.ci = 1);
plot(fitln, col = "purple", lwd.ci = 1, lty.ci = 1);
plot(fitgom, col = "pink", lwd.ci = 1, lty.ci = 1);
plot(fitgf, col = "red", lwd.ci = 1, lty.ci = 1);
plot(fitg, col = "yellow", lwd.ci = 1, lty.ci = 1)
```

```
weib = survreg(Surv(tim, cens) age + gend, data = tbdat, dist ="weibull", scale = 1)
expo = survreg(Surv(tim, cens) age + gend, data = tbdat, dist = "exponential") logn
= survreg(Surv(tim, cens) age + gend, data = tbdat, dist = "lognormal") logn2 =
survreg(Surv(tim, cens) age *gend, data = tbdat, dist = "lognormal")
logl = survreg(Surv(tim, cens) age + gend, data = tbdat, dist = "loglogistic")
logis = survreg(Surv(tim, cens) age + gend, data = tbdat, dist = "logistic")
AIC(weib, expo, logn, logn2, logl, logis)
```