Iraqi Journal of Statistical Sciences

www.stats.mosuljournals.com

Flexible Parametric Survival Model for Analysing Censored Time-To-Event Data Among Tuberculosis Patients

Azeez Adeboye[1]* Osuji Geogeleen[1], Alakija Temitope[2], Odeyemi Akinwumi[1], Mutambayi Ruffin[1],Madu Peter[3]

[1]Department of Statistics, University of Fort Hare, Alice, Eastern Cape, South Africa, [2]Department of Statistics, Yaba College of Technology, Yaba, Lagos State, Nigeria, [3]Department of Mathematical Sciences, Olabisi Onabanjo University, Ago-Iwoye, Ogun State, Nigeria.

Article information

Article history:
Received :August 23, 2024
Accepted : October 27, 2024
Available : December 1, 2024

Keywords:
Flexible parametric survival model, Additive hazard model, Cox model, Time-varying covariate, Tuberculosis.

Correspondence:
Amira Wali Omer
amira.omer@su.edu.krd

Abstract

Tuberculosis (TB) is a globally deadly infectious disease responsible for 10 million new cases and 1.5 million deaths annually. Shorter TB treatment regimens show promise in reducing this problem, but there is an improved treatment success rate in South Africa, while retreatment cases remain a concern. An important feature of time-to-event modelling is its ability to consider transition probabilities of heterogeneous subgroups with different risk profiles. Survival analysis is generally performed to accurately estimate the transition probabilities associated with the risk profiles. This study explored the application of a flexible parametric survival model for analysing censored time-to-event data among TB patients.

The data were obtained from East London Central Clinic-TB unit, Eastern Cape, South Africa. In total, 174 patients were included in the analysis. The goodness of fit of the models was explored using Akaike information criterion (AIC). We estimated the hazard ratios (HR) and baseline cumulative hazards of our model, which are necessary to calculate individual transition probabilities, and compared the model with the Cox model and additive hazard model to determine the survival predictions of TB patients.

The flexible parametric survival model produced hazard ratio and baseline cumulative hazard estimates that were similar to those obtained using the Cox proportional hazards model. The analysis revealed that sex (HR=0.49, 95% CI: 0.38, 0.62), antiretroviral therapy (ART), (HR=0.53, 95% CI: 0.34, 0.78), and diabetes (HR=0.58, 95% CI: 0.41, 0.78) were all statistically significant factors associated with improved treatment survival in tuberculosis patients.

Flexible parametric survival models are a powerful tool for modelling time-to-event data and individual transition probabilities. It is of great importance to fit models by modelling the baseline, which makes it easier to make different types of predictions and allows for non-proportional hazards since it is an interaction.

DOI 10.33899/iqjoss.2024.185256,

1. Introduction

Tuberculosis (TB) is primarily caused by Mycobacterium tuberculosis, typically affecting the lungs [1]. Annually, 10 million people contract TB and 1.5 million people die from the disease, making it one of the deadliest infectious diseases globally. It is the primary cause of death for those with HIV and significantly contributes to antimicrobial resistance [1]. Poor outcomes from TB treatment are driven by the high rate of death and significant loss of follow-up [2], [3]. This issue of loss to follow-up is a major challenge experienced by the South Africa National TB Programme [4]. One of the key advantages of shorter treatment regimens for tuberculosis-resistant treatment is a reduction in the loss to follow-up rate [5].

The treatment success rate in South Africa for new smear-positive and smear-negative/ extrapulmonary TB patients has improved by 79% and 76%, respectively [6]. This was achieved as a result of higher cure rates and a decrease in the treatment default rate. However, the treatment success rate for retreatment cases remains low at 66.3% [6]. Of particular concern is the fact that up to 25% of sputum smear-positive TB cases are lost to follow-up before treatment initiation, which may contribute to ongoing transmission of the disease and an increased risk of death [7]. Furthermore, the mortality rate remains high even after completion of TB treatment, likely due to HIV disease [8]. To address this issue, there is a need to expand access to antiretroviral therapy (ART) for all HIV-infected TB patients to reduce HIV-related mortality among individuals with TB.

Understanding the effect of TB treatment on the time-to-death of TB patients by covariates such as gender, HIV status, age, and many more, may provide valuable insights for health programs in South Africa and globally. To achieve this target, flexible parametric hazards (PH) models and Additive hazard (AH) models were used in survival analysis to study the time-to-event data, which provide a flexible and versatile framework to capture complex hazard functions and assess the impact of covariates on survival probabilities in a more flexible way.

The concept of the flexible PH models is to use restricted cubic splines to approximate the baseline hazard function in the context of the Cox proportional hazards model [9] and the AH model using kernel smoothing techniques [10].

These models are a more adaptable approach to modelling survival data, which accommodates both non-linear and time-dependent effects. The integration of time-dependent covariates is used to examine the changes in risk factors over time. Furthermore, these models offer greater flexibility in capturing a wide range of survival patterns, from monotonically increasing or decreasing hazards to more intricate shapes, by altering the number and placement of spline knots [11]. Parametric models offer distinct advantages, including better suitability for prediction, extrapolation, quantification of risks, modelling time-dependent effects, enhancing understanding, and handling complex large datasets. However, the estimates from flexible parametric survival models are often similar to those obtained from the Cox model.

In Cox regression, the baseline hazard function is not estimated, which can limit its ability to capture the true underlying hazard function. In contrast, flexible parametric models offer an alternative approach by explicitly modelling the baseline hazard using splines, enabling more accurate representations of complex hazard patterns and facilitating better predictions, especially in scenarios with non-proportional hazards. The models provide a parametric estimate of the baseline hazard without usual shape restrictions, making it highly flexible [9]. It can be applied to both standard and relative survival models and is capable of fitting relative survival cure models [12]. It can also be estimated on the log-hazard scale [13].

For this work, we investigated the use of flexible parametric methods to analyse censored time-to-event data in patients with tuberculosis in a small population of South Africa. We described and compared flexible parametric hazards (PH) models and Additive hazard (AH) models using a real-life case study of individual-level censored data from the Tuberculosis Hospital and linked mortality data for the general population and people with HIV status and stratified by ART status.

2. Methods

The Cox Proportional Hazards (Cox PH) model

In survival data analysis, the Cox model is a widely used statistical method to assess the relationship between covariates and the hazard rate over time while making no assumptions about the shape of the hazard function. The model is written as:

$$h(t|X) = h_0(t)e^{\beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_p X_p} \quad (1)$$

$$\text{In}\big(h(t|X)\big) = \text{In}\big(h_0(t)\big) + \beta_1 X_1 + \cdots + \beta_p X_p$$

where $h(t|X)$ is the hazard rate at time t for a given set of covariates X, $h_0(t)$ is the baseline hazard rate, and $\beta_1 + \beta_2 + \cdots + \beta_p$ is the coefficients of the covariates $X_1 + X_2 + \cdots + X_p$ on the hazard rate. The integrated form of the model is:

$$H(t|X) = \left(\int_0^1 h_0(\mu) d\mu \right) e^{\beta_1 X_1 + \cdots + \beta_p X_p} = H_0(t) e^{\beta_1 X_1 + \cdots + \beta_p X_p} \qquad (2)$$

where $H(t|X)$ is the cumulative hazard function.

The model does not assume a specific distribution for survival times but estimates the relative risk of covariates about the shape of the baseline hazard function. But, the underlying shape of the hazard function is often ignored [14].

Additive hazards

The additive hazards models have a general form of $H(t|x:\theta) = \eta(t, x; \theta)$. Recently, the default model specification for survival without specifying the smoothness has cumulative hazard function as:

$$H(t|x;\eta) = \text{B}(t)\eta_\text{B} + t(x'\eta_x)$$

where $\text{B}(t)$ is a natural spline design matrix with parameter η_B, and η_x is the parameter for x. The hazard function is given as

$$h(t|x;\eta) = \text{B}'(t)\eta_\text{B} + (x'\eta_x)$$

Let $\lambda(t)$ be the hazard rate at time t for a specific event of interest, the additive hazards model can be expressed as:

$$\text{Log}\big(\lambda(t)\big) = \beta' X + \alpha(t)$$

where $\lambda(t)$ is the hazard rate for the event, X is the vector of covariates, β' is the vector of regression coefficients, and $\alpha(t)$ is the function of time t that captures the smooth or time-varying effects. The additive hazards model assumes that the log-hazard rates for the event are linearly related to the covariates ($\beta' X$) and include time-varying components $\big(\alpha(t)\big)$. The implementation of Additive hazard models offers flexibility by allowing the modelling of the baseline hazard using splines and accommodating both constant hazards and smooth time-varying effects.

3. Flexible Parametric Survival Models

Let the survival function $S(t|x) = \text{P}(T > t|x)$ for a random variable T at time t with covariate $\boldsymbol{x} = x_i$ written as:

$$S(t|x) = \exp\left(-\exp\left(s(\log(t);\gamma) + \sum_i \beta_i x_i \right) \right) \qquad (3)$$

where γ is the parameter, $\beta_i x_i$ is the coefficient of covariate indexed i, and $s(\mu;\gamma)$ is the parametric smooth function. Within this framework, a smooth function is used to model the baseline log cumulative hazard function and a linear predictor to model the covariates. However, the hazard function and cumulative hazard function are modelled as:

$$H(t|x) = -\log\big(S(t|x)\big) = \exp\left(s(\log(t);\gamma) + \sum_i \beta_i\, x_i\right) \tag{4}$$

$$h(t|x) = \frac{d}{dt}H(t|x) = \exp\left(s(\log(t);\gamma) + \sum_i \beta_i\, x_i\right) \times \frac{d\big(s(\log(t);\gamma)\big)}{dt} \tag{5}$$

Now, considering two sets of covariates, $x = x_{1i}$ and $x = x_{2i}$, the hazard ratio can be expressed as:

$$\frac{h(t|x_2)}{h(t|x_1)} = \frac{\exp(s(\log(t);\gamma) + \sum_i \beta_i\, x_{2i}) \times \dfrac{d\big(s(\log(t);\gamma)\big)}{dt}}{\exp(s(\log(t);\gamma) + \sum_i \beta_i\, x_{1i}) \times \dfrac{d\big(s(\log(t);\gamma)\big)}{dt}}$$

$$= \exp\left(\sum_i \beta_i\,(x_{2i} - x_{1i})\right) \tag{6}$$

If $x_{2i} = x_{1i} + 1$ and $x_{2i}{}' = x_{1i}{}'$ for $i' \neq i$, then the hazard ratio is equal to $\exp(\sum_i \beta_i)$ for all t and other covariate values. The model can be incorporated with time-dependent effects of covariates x on the log-hazard scale given as:

$$\mathrm{In}\big(h(t|x)\big) = s(\log(t);\gamma) + x\beta + \sum_{p=1}^{P} s\big(\log(t);\gamma_p\big)X_p \tag{7}$$

where $s(\log(t);\gamma_0)$ is the restricted cubic spline function, $s\big(\log(t);\gamma_p\big)$ denotes the spline function for the pth time-dependent effect, and P is the number of time-dependent effects.

The restricted cubic splines within the models are employed to model the log cumulative hazard or the log cumulative odds [9], [11]. These splines are piecewise cubic functions connected at specific positions referred to as knots. To ensure smoothness, the first and second derivatives of the overall function are enforced to be continuous at the knots, and the function is linearly constrained before the first knot and after the last knot. The complexity of these spline functions is dictated by user-defined degrees of freedom, which equate to the number of knots minus one. Knot positions can either be defined by the user or set to be evenly spaced percentiles of the observed event-time distribution [13].
However, $s(\log(t);\gamma_0)$ with knots $k_1 + k_2 + \cdots + k_\kappa$ is expressed as:

$$s(\log(t);\gamma_0) = \gamma_{00} + \sum_{l=1}^{\kappa-1} \gamma_{0l}{}^{\pi_l(a)}$$

where $\pi_l(a)$ is the basis function for the lth time defined as:

$$\pi_l(a) = \begin{cases} a & \text{for } l = 1 \\ (a-k_1)_+^3 - \lambda_1(a-k_1)_+^3 - (1-\lambda_1)(a-k_\kappa)_+^3, & \text{for } l = 2,\dots,K=1 \end{cases}$$

where k_1 and k_κ are the knots boundaries and $\lambda_1 = {(k_\kappa - k_l)}/{(k_\kappa - k_1)}$.

Using restricted cubic splines in flexible parametric survival models would help to capture both simple and complex hazard functions in situations where standard parametric models may have challenges [15].

Study setting and design

This was a hospital-based retrospective individual-level censored data in TB patients reported for the treatment in East London Central Clinic-TB unit, Eastern Cape, South Africa. This clinic is a specialized facility funded by the provincial government, dedicated to the diagnosis, treatment, and prevention of TB, especially in patients with HIV co-infections. The clinic offers a range of services, including antiretroviral treatments (ARTs), TB services, and diagnostic tools with standard TB treatment protocols. Patient medical histories, from their initial consultation to discharge, are recorded in the TB treatment registry, which is the official record-keeping system for TB treatment under the Department of Health in South Africa.

Data collection

In this study, data were gathered from medical records of hospitalized TB patients with HIV coinfection. Information was extracted from TB record files and patients' medical files using a standardized TB card format recommended by South Africa's Department of Health. The information includes gender, age, location, TB category, HIV status, diabetes, weight, and antiretroviral therapy (ART).

Statistical analyses

We initiate our analysis by applying basic proportional hazard models to the TB dataset to identify the contributing covariate factors. Initially, we employed a Cox regression model was used to analyze each single covariate to determine whether the covariate is associated with improved survival of the TB patient. A flexible parametric survival model was fitted with an additional argument (df =4) to specify four different degrees of freedom for the baseline smoother. We compare the survival estimate of the flexible parametric survival model with predictions from non-parametric Kaplan-Meier and Additive hazard model curves. All analyses were done in R using rstpm2 and flexsurv packages.

Ethics

Ethical clearance for the study was obtained from the Ethics Committee, University of Fort Hare, and Department of Health, Eastern Cape chapter, South Africa.

Results

More than half of the patients were male (63.2%). The mean age of TB patients on treatment was 39.4 ± 17.3 years (range: 14-80 years). Of all the TB patients, 79 (45.4%) were HIV-positive TB patients, 78.2% had pulmonary TB, 77.6% had drug-resistance TB, 21.8% were placed on antiretroviral therapy and almost two-thirds (60.9%) of the patients were treated without diabetes (Table 1).

Table 1: The Demographic Characteristics of the TB patients

Variables	levels	Number (%)
Sex	Female	64 (36.8%)
	Male	110 (63.2%)
Age	Mean ± SD	39.4 ± 17.3
Weight	Mean ± SD	86.5 ± 17.4
HIV status	Positive	79 (45.4%)
	Negative	95 (54.6%)
Disease class	ExtraPTB	38 (21.8%)
	PTB	136 (78.2%)
TB type	DR-TB	135 (77.6%)

	MDR-TB	39 (22.4%)
ART	Yes	38 (21.8%)
	No	136 (78.2%)
Diabetes	No	106 (60.9%)
	Yes	68 (39.1%)
Alcohol	Yes	97 (55.7%)
	No	77 (44.3%)
Smoking	No	111 (63.8%)
	Yes	63 (36.2%)
Substance use	No	134 (77.0%)
	Yes	40 (%)

As we did not assume proportional hazards for TB treatment risk and aimed to evaluate the risk ratio instead of the attributable risk, we opted for a flexible parametric survival model. Unfortunately, the results from the additive hazards regression models did not provide satisfactory outcomes. We considered three different models for fitting the data, and based on the Akaike Information Criterion (AIC=1421.628) and the log-likelihood estimate (-2 log L=1389.628), we determined that the flexible proportional hazards (PH) model was the most suitable model. The results are summarized in Table 2, with the flexible PH model being the preferred choice.

Table 2: Selection criteria for best model fit

Model	**AIC**	**-2 log L**
Cox-PH model	1913.6087	1891.6087
Additive hazards	1419.842	1391.842
Flexible PH model	**1421.628**	**1389.628**

The analysis results revealed that certain variables, including sex, ART, and diabetes, were identified as treatment risk factors affecting the survival of TB patients, while the other clinic characteristics were not statistically significant (Table 3). Notably, the result of the Cox proportional hazards model was similar to the flexible parametric survival model in detecting the TB treatment risk factors, whereas the results from the additive hazards model were notably different. From the model output. the hazard ratios in the flexible model for estimating the TB treatment risk factors were lower and had more narrow confidence intervals compared to the Cox regression model. Specifically, the analysis showed that sex was significantly associated with improved treatment survival for TB patients (HR=0.49, 95% CI: 0.38, 0.62). Furthermore, ART was found to be statistically significantly associated with improved treatment survival for TB patients (HR=0.53, 95% CI: 0.34, 0.78), and diabetes exhibited a similar statistically significant association with improved treatment survival for TB patients (HR=0.58, 95% CI: 0.41, 0.78).

Table 3: Hazard estimates of TB treatment factors based on Cox model and Flexible model

	Cox PH		**Flexible PH**	
Variables	**HR (95% C.I)**	**Pr(>|z|)**	**HR (95% C.I)**	**Pr(>|z|)**
Sex				
Female	1		1	
Male	0.510 (0.326, 0.798)	0.003	0.490 (0.382, 0.617)	0.002
Age	0.995 (0.983, 1.006)	0.354	0.994 (0.989, 0.998)	0.278
Weight	0.994 (0.981, 1.007)	0.373	0.993 (0.991, 0.995)	0.299
HIV-status				
-tive	1		1	
+tive	1.041 (0.694, 1.562)	0.846	1.058 (0.804, 1.359)	0.784
Class				
EPTB	1		1	
PTB	0.772 (0.464, 1.285)	0.320	0.789 (0.635, 0.966)	0.360

TB type				
MDR-TB	1		1	
DR-TB	1.321 (0.789, 2.212)	0.230	1.346 (0.919, 1.886)	0.256
ART				
No	1		1	
Yes	0.538 (0.313, 0.922)	0.024	0.526 (0.336, 0.777)	0.019
Diabetes				
No	1		1	
Yes	0.607 (0.391, 0.941)	0.026	0.579 (0.414, 0.784)	0.015
Alcohol				
No	1		1	
Yes	0.884 (0.576, 1.358)	0.574	0.892 (0.697, 1.121)	0.601
Smoking				
No	1		1	
Yes	0.715 (0.464, 1.102)	0.128	0.708 (0.493, 0.978)	0.117
Substance				
No	1		1	
Yes	0.835 (0.527, 1.322)	0.441	0.861 (0.571, 1.235)	0.521

Flexible parametric survival models are capable of estimating a wide range of parameters. However, the prediction estimates from the flexible model were compared with predictions from the Additive model and non-parametric Kaplan-Meier curves (Figure 1). The plot shows that the Flexible parametric survival model has a better-predicted survival probability at each time point compared to other models. The shape of the flexible model indicates a higher and improved survival rate among TB patients, followed by the additive model and KM model suggests a lower survival rate among TB patients. The overall pattern of the curves is steadily decreasing indicating consistent association with improved treatment survival for TB patients, which contributes to the survival estimates up to the last observed time.

In addition to the mortality rates, we obtained smooth predicted survival curves to facilitate comparisons among different covariate groups within various sub-groups and time intervals. Figure 1 illustrates the predicted survival probabilities on the time scale for gender and the use of antiretroviral therapy (ART), which were found to significantly contribute to improved survival outcomes among TB patients. The left panel shows the smooth predicted survival curves for the three models over the time since the initiation of TB treatment, categorized by patient gender. The right Panel shows the corresponding smooth predicted survival curves for patients on ART.

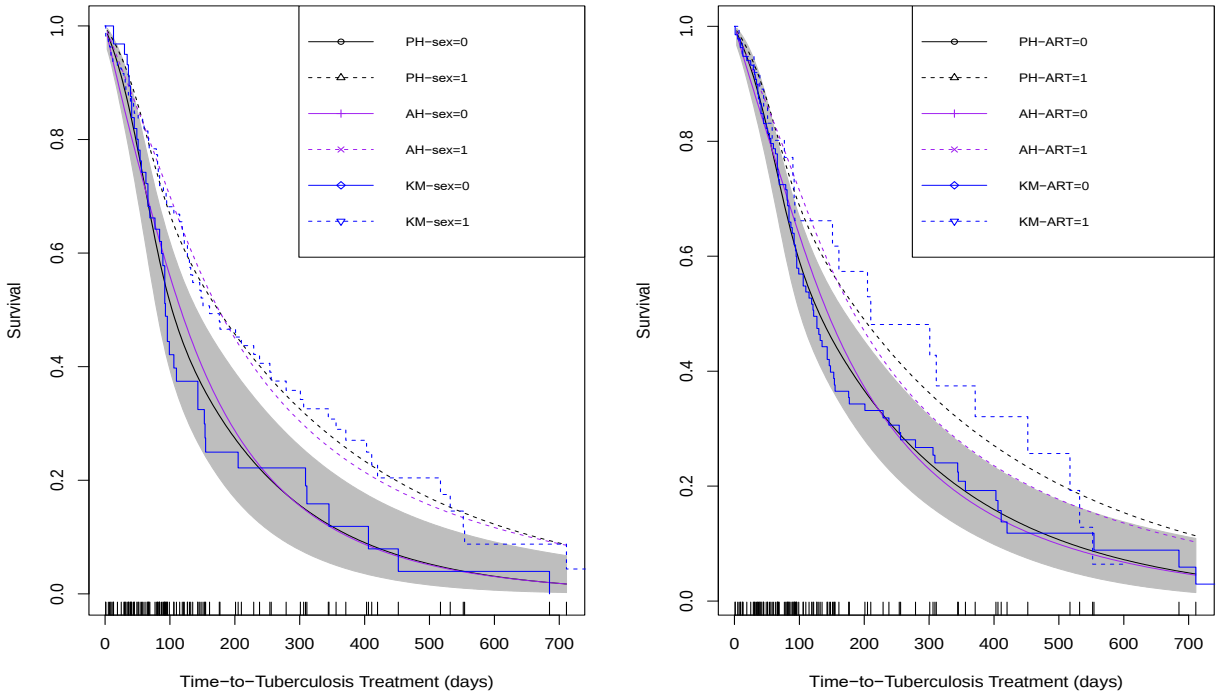

Figure 1: Predicted survival rate for covariate sex and ART among patients on TB treatment

Figure 2 displays the predicted survival probabilities on the time scale for diabetes and HIV status, both of which were identified as significant contributors to enhanced survival rates among TB patients. The left Panel exhibits the smooth predicted survival curves for the three models over the time since the TB treatment initiation for diabetic patients. Meanwhile, the right Panel displays the equivalent curves for patients with HIV status. It can be seen from both figures that the survival proportions are higher for TB patients in the flexible model compared to other models.

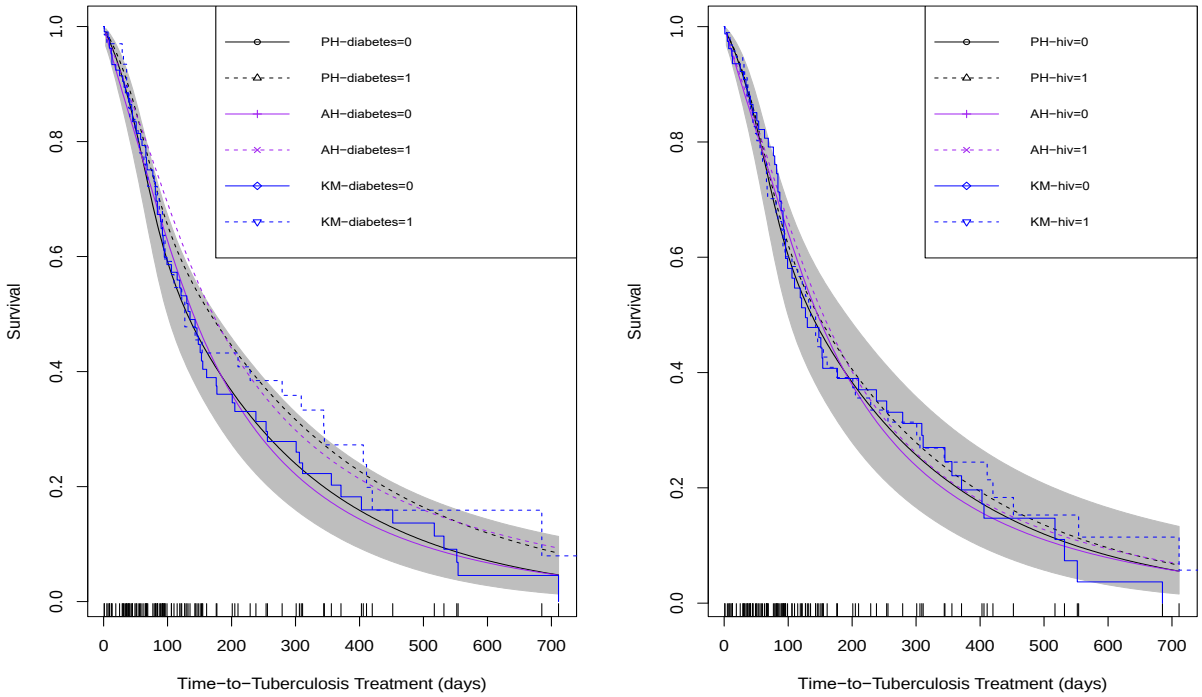

Figure 2: Predicted survival rate for covariate diabetes and HIV status among patients on TB treatment

The spline coefficients are not interpretable on their own but they are used to predict the shape of the hazard surface at different covariate values. Figure 3 displays four panels for viewing the estimated mortality rates among TB patients from the flexible parametric model. The upper left panel shows the estimated mortality rates of female patients on TB treatment and male patients on TB treatment. The upper right panel shows the estimated mortality rates of patients without ART on TB treatment and patients with ART on TB treatment. The lower left panel shows the estimated mortality rates of patients without diabetes on TB treatment and diabetic patients on TB treatment. The lower right panel shows the estimated mortality rates of HIV-negative patients on TB treatment and HIV-positive patients on TB treatment.

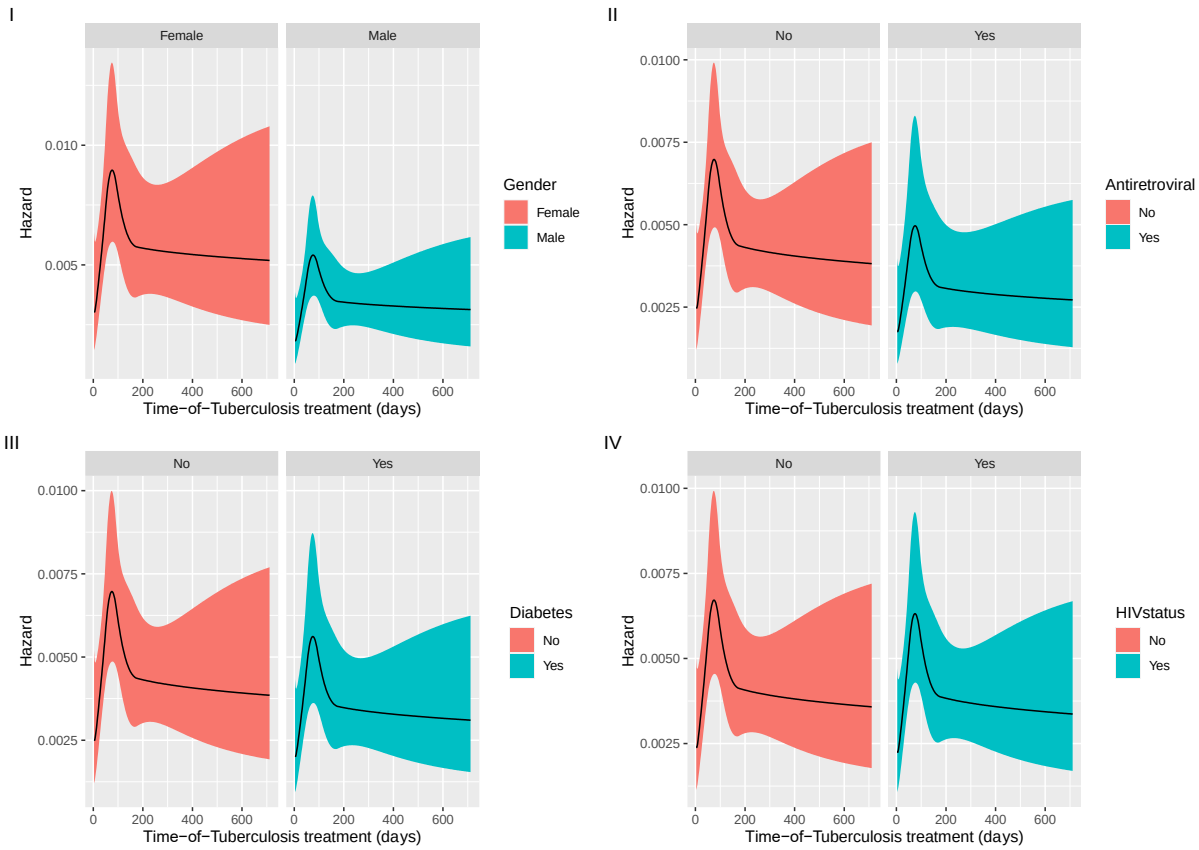

Figure 3: Predicted estimate of Hazard ratios with 95% C.I for sex, ART, diabetes, and HIV status groups among TB patients

The covariate sex panel shows that the mortality rates for both females and males are decreasing with male patients having significantly lower mortality rates (improved survival rates) compared to female TB patients. There was a complete overlap between the sex group's estimated mortality rates and survival proportions. Patients treated with ART therapy also have lower mortality rates than patients without ART therapy. There was a small overlap between the ART group's estimated mortality rates. Diabetic patients were observed to have a lower mortality rate compared to patients without diabetes. The high mortality among patients without diabetes can be due to a more severe disease for this group of patients. The overlap between the diabetes group's estimated mortality rates is very small. There was little or significant overlap between the HIV group's estimated mortality rates and survival proportions (Figure 4). Moreover, the peak of the surface, marked with green and red bands, reflects group hazard ratios of 95% CI. As time progresses on the time scale, the mortality rate surface widens, showing predicted clinical treatment observations [16].

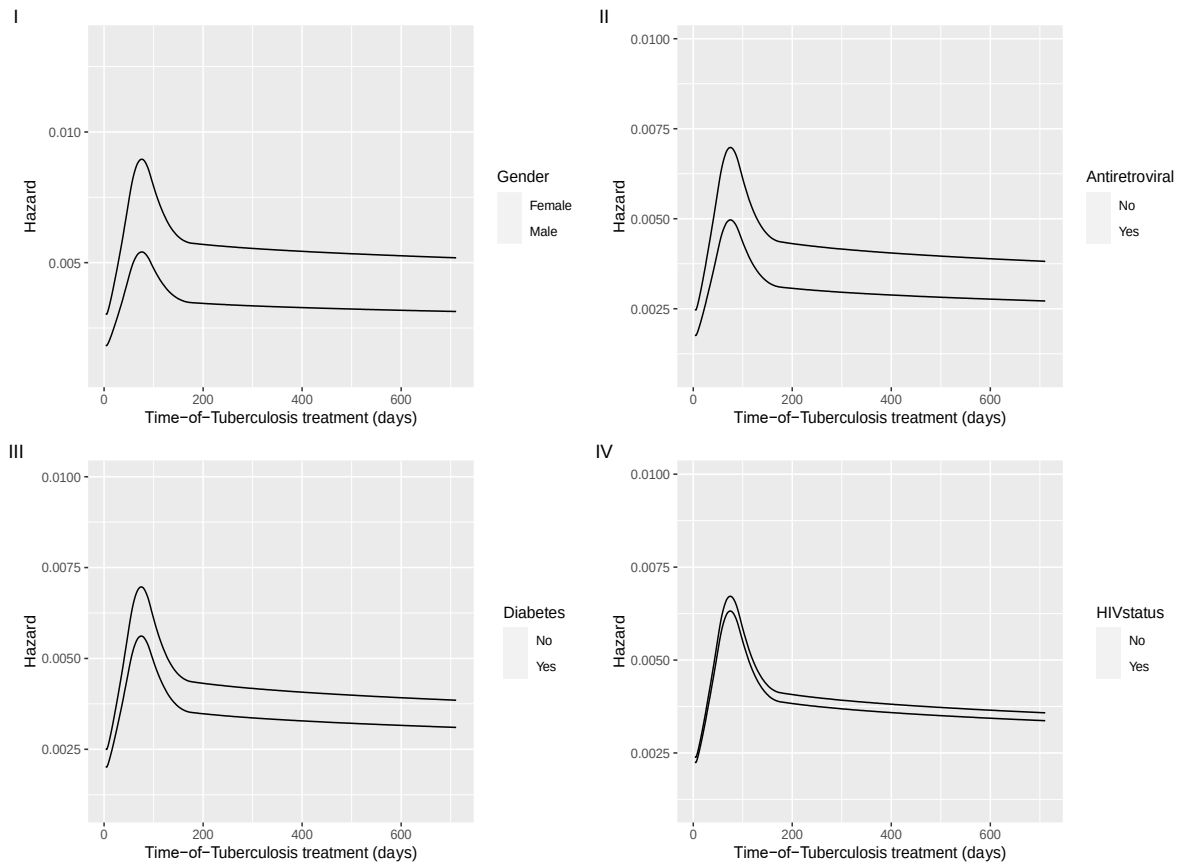

Figure 4: Overlap Predicted estimate of Hazard ratios with 95% C.I for sex, ART, diabetes, and HIV status groups among TB patients

We assessed the time-varying covariate effects on TB treatment by estimating survival differences and hazard differences to compare hazard and mortality rate ratios along with their 95% confidence intervals over time for each group (Figure 5). We initially defined the survival differences based on the covariate pattern and subsequently transformed them into an 'exposed' covariate pattern using the exposed function. From Figure 5, we observed that the relative risk effect (hazard differences) of sex, ART, and diabetes slightly increases and the mortality effect (survival differences) decreases rapidly with time after 100 days of TB treatment initiation. Meanwhile, the relative risk effect (hazard differences) and mortality effect (survival differences) of HIV status among TB patients are the same (no increase or decrease) after 100 days of TB treatment initiation.

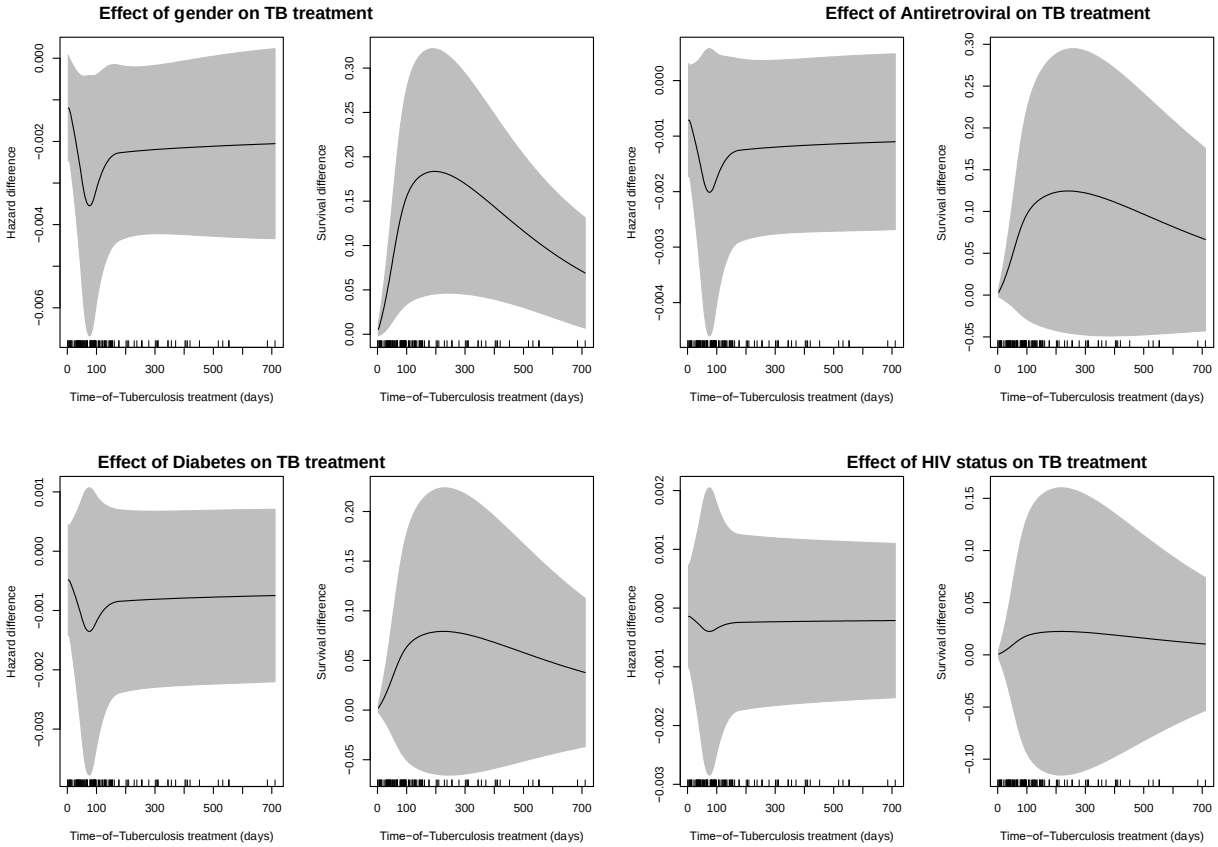

Figure 5: The time-varying covariate effects on TB treatment using survival and hazard differences

Discussion

The flexible parametric survival model has been widely used in various fields, including applications in relative survival and clinical decision-making [9], [11], [17]–[21]. However, its utilization in the context of TB treatment has been relatively limited. In this article, we have employed the flexible parametric survival model to estimate survival probabilities for patients undergoing TB treatment and juxtapose the results with additive hazard models and the Cox proportional hazards model.

The flexible parametric model is an alternative approach for estimating survival probabilities. Unlike conventional methods that rely on a priori transition probabilities, this approach uses individual patient data directly to model survival. The integration of individual patient data with additional covariate information could be worthwhile and lead to improved accuracy in predicting survival probabilities [22]. In our analysis, the flexible parametric survival model provided reliable and smooth estimates of the baseline cumulative hazards [18].

According to the analyses of the flexible parametric survival models, the results obtained indicate that sex is significantly associated with improved survival rates of patients on TB treatment. The result indicates that male is significantly associated with a 51% lower risk of mortality rate [0.490 (95% CI: 0.382, 0.617), p-value=0.002] compared to female. This result is consistent with other earlier studies [23]–[25] in which standard survival analyses were applied. Hence, it can be concluded that male patients had a higher improved survival rate of TB treatment. There may be biological differences between males and females that can impact how they respond to TB treatment, such as hormonal differences, which play a role in immune response and may explain this pattern, which has also been observed in TB research studied in other areas [26], [27].

Our findings indicated that individuals who began the ART regimen experienced a 47.4% improved survival rate following the initiation of TB treatment [0.526 (95% CI: 0.336, 0.777), p-value=0.0019]. This result supports the recommendations to commence ART at earlier stages for all symptomatic individuals, irrespective of their CD4

cell counts. Previous studies have consistently demonstrated that the initiation of ART serves as a significant predictor of mortality among TB patients concurrently receiving ART [28]–[33].

The findings from our study also revealed that diabetes was significantly associated with improved survival rates among TB patients. The result showed that TB patients have a 42% lower risk of mortality rate [0.579 (95% CI: 0.414, 0.784), p-value = 0.0015]. This is consistent with some studies with lower risk [34]–[36] and dissimilar to other studies with higher diabetes risk among TB patients in different countries [37]–[39]. The observation of improved survival rate among the patients may be attributed to several factors. TB and diabetes are known to have complex biological interactions. TB can lead to a temporary state of insulin resistance, which can affect glucose metabolism. This can result in lower blood sugar levels, reducing the likelihood of a diabetes diagnosis during TB infection. TB and diabetes share some common symptoms. These overlapping symptoms might make it more challenging to diagnose diabetes in TB patients, potentially leading to underdiagnosis. TB can be a severe disease, and individuals with TB may not survive long enough to develop diabetes. This selective survival effect could contribute to the observed improved survival rate among TB patients.

Conclusion

Using the flexible parametric survival model hazard can avoid splitting the survival data into intervals and avoid making an assumption of constant hazard rates within the time intervals in estimating the hazard rates and probabilities. The study highlights the importance of fitting models by modelling the baseline, which is easier to make different types of predictions and allows for non-proportional hazards since it is an interaction. The results from applying the three models revealed that sex, ART, and diabetes covariates were found to statistically significantly improve the survival rate of TB treatment prognostic. The flexible parametric survival model hazard ratios were compared with the additive hazard model and Cox proportional hazard model, and the effect estimates obtained from the hazard models were different. The study recommends hospital authorities pay attention to female TB patients who are HIV positive, on ART, and diabetic. The rate of mortality during TB treatment was high for these groups. TB patients should be advised to adopt healthier lifestyles during treatment, which can include better nutrition and increased physical activity. These lifestyle changes can reduce the risk of developing diabetes.

Acknowledgments

Not applicable.

Authors' contributions

AA, OG, AT, and OA designed the study. AA, OA, and MR performed the data analysis and drafted the manuscript. OA, MP, AT and RM supported the analysis and provided revisions and improvements to the manuscript. All authors have read and approved the final manuscript.

Funding

No funding

Availability of data and materials

Data will be available on request. All data analysis codes used can be found on my GitHub page.

Ethics approval and consent to participate

The study was approved by the Ethic Committee of the University of Fort Hare (REC-270710-028-RA) with reference number: QIN11SAZE01 and also approved by the Eastern Cape Department of Health (EC_202008_007). Informed consent was not required as we had no contact with the study individuals nor did our analyses contain any personal identifiers. All methods were performed according to relevant guidelines and regulations.

Consent for publication

Not applicable.

Competing interests

We declare that they have no conflict of interest.

Author Details

Reference

1. World Health Organization (WHO), "Tuberculosis-Health Topics." Accessed: Sep. 18, 2023. [Online]. Available: https://www.who.int/health-topics/tuberculosis
2. Schnippel, K., Ndjeka, N., Maartens, G., Meintjes, G., Master, I., Ismail, N., Hughes, J., Ferreira, H., Padanilam, X., Romero, R. and Te Riele, J., (2018). Effect of bedaquiline on mortality in South African patients with drug-resistant tuberculosis: a retrospective cohort study. *The lancet respiratory medicine*, *6*(9), pp.699-706.
3. Walker, I.F., Shi, O., Hicks, J.P., Elsey, H., Wei, X., Menzies, D., Lan, Z., Falzon, D., Migliori, G.B., Pérez-Guzmán, C. and Vargas, M.H., (2019). Analysis of loss to follow-up in 4099 multidrug-resistant pulmonary tuberculosis patients. *European Respiratory Journal*, *54*(1).
4. Ndjeka, N., Campbell, J.R., Meintjes, G., Maartens, G., Schaaf, H.S., Hughes, J., Padanilam, X., Reuter, A., Romero, R., Ismail, F. and Enwerem, M., (2022). Treatment outcomes 24 months after initiating short, all-oral bedaquiline-containing or injectable-containing rifampicin-resistant tuberculosis treatment regimens in South Africa: a retrospective cohort study. The Lancet Infectious Diseases, 22(7), pp.1042-1051.
5. Abidi, S., Achar, J., Neino, M.M.A., Bang, D., Benedetti, A., Brode, S., Campbell, J.R., Casas, E.C., Conradie, F., Dravniece, G. and du Cros, P., (2020). Standardised shorter regimens versus individualised longer regimens for rifampin-or multidrug-resistant tuberculosis. European Respiratory Journal, 55(3).
6. Churchyard, G.J., Mametja, L.D., Mvusi, L., Ndjeka, N., Pillay, Y., Hesseling, A.C., Reid, A. and Babatunde, S., (2014). Tuberculosis control in South Africa: successes, challenges and recommendations: tuberculosis control-Progress towards the Millennium Development Goals. South African Medical Journal, 104(3), pp.244-248.
7. Claassens, M.M., Du Toit, E., Dunbar, R., Lombard, C., Enarson, D.A., Beyers, N. and Borgdorff, M.W., (2013). Tuberculosis patients in primary care do not start treatment. What role do health system delays play?. The International Journal of Tuberculosis and Lung Disease, 17(5), pp.603-607.
8. Abdool Karim, S.S., Naidoo, K., Grobler, A., Padayatchi, N., Baxter, C., Gray, A., Gengiah, T., Nair, G., Bamber, S., Singh, A. and Khan, M., (2010). Timing of initiation of antiretroviral drugs during tuberculosis therapy. New England Journal of Medicine, 362(8), pp.697-706.
9. Royston, P. and Parmar, M.K., (2002). Flexible parametric proportional-hazards and proportional-odds models for censored survival data, with application to prognostic modelling and estimation of treatment effects. Statistics in medicine, 21(15), pp.2175-2197.
10. Li, J., Fine, J. and Brookhart, A., (2015). Instrumental variable additive hazards models. *Biometrics*, 71(1), pp.122-130.
11. Royston, P., & Lambert, P. C. (2011). Flexible parametric survival analysis using Stata: beyond the Cox model (Vol. 347). College Station, TX: Stata press.
12. Andersson, T. M., Dickman, P. W., Eloranta, S., & Lambert, P. C. (2011). Estimating and modelling cure in population-based cancer studies within the framework of flexible parametric survival models. BMC medical research methodology, 11, 1-11.
13. Bower, H., Crowther, M.J. and Lambert, P.C., (2016). strcs: A command for fitting flexible parametric survival models on the log-hazard scale. The Stata Journal, 16(4), pp.989-1012.
14. Cox, D. R. (1972). Regression models and life-tables. Journal of the Royal Statistical Society: Series B (Methodological), 34(2), 187-202.
15. Rutherford, M. J., Crowther, M. J., & Lambert, P. C. (2015). The use of restricted cubic splines to approximate complex hazard functions in the analysis of time-to-event data: a simulation study. Journal of Statistical Computation and Simulation, 85(4), 777-793.
16. Dent, R., Valentini, A., Hanna, W., Rawlinson, E., Rakovitch, E., Sun, P., & Narod, S. A. (2014). Factors associated with breast cancer mortality after local recurrence. Current oncology, 21(3), 418-425.
17. Burke, K., Jones, M. C., & Noufaily, A. (2020). A flexible parametric modelling framework for survival analysis. Journal of the Royal Statistical Society Series C: Applied Statistics, 69(2), 429-457.

18. Hinchliffe, S. R., Scott, D. A., & Lambert, P. C. (2013). Flexible parametric illness-death models. The Stata Journal, 13(4), 759-775.
19. Batyrbekova, N., Bower, H., Dickman, P.W., Ravn Landtblom, A., Hultcrantz, M., Szulkin, R., Lambert, P.C. and Andersson, T.M., (2022). Modelling multiple time-scales with flexible parametric survival models. BMC medical research methodology, 22(1), p.290.
20. Michael, V. A., & Bolarinwa, I. A. (2020). Parametric survival modeling of Tuberculosis data-A case study of Federal Medical Centre, Bida, Nigeria. Modern Applied Science, 14(7), 1-37.
21. Du, X., Li, M., Zhu, P., Wang, J., Hou, L., Li, J., Meng, H., Zhou, M. and Zhu, C., (2018). Comparison of the flexible parametric survival model and Cox model in estimating Markov transition probabilities using real-world data. PLoS One, 13(8), p.e0200807.
22. Holford, T. R. (1983). The estimation of age, period and cohort effects for vital rates. Biometrics, 311-324.
23. Fufa, D. B., Diriba, T. A., Dame, K. T., & Debusho, L. K. (2023). Competing risk models to evaluate the factors for time to loss to follow-up among tuberculosis patients at Ambo General Hospital. Archives of Public Health, 81(1), 117.
24. Masini, E.O., Mansour, O., Speer, C.E., Addona, V., Hanson, C.L., Sitienei, J.K., Kipruto, H.K., Githiomi, M.M. and Mungai, B.N., (2016). Using survival analysis to identify risk factors for treatment interruption among new and retreatment tuberculosis patients in Kenya. PloS one, 11(10), p.e0164172.
25. Dangisso, M. H., Datiko, D. G., & Lindtjørn, B. (2014). Trends of tuberculosis case notification and treatment outcomes in the Sidama Zone, southern Ethiopia: ten-year retrospective trend analysis in urban-rural settings. PloS one, 9(12), e114225.
26. Nhamoyebonde, S., & Leslie, A. (2014). Biological differences between the sexes and susceptibility to tuberculosis. The Journal of infectious diseases, S100-S106.
27. Gupta, M., Srikrishna, G., Klein, S. L., & Bishai, W. R. (2022). Genetic and hormonal mechanisms underlying sex-specific immune responses in tuberculosis. Trends in immunology, 43(8), 640-656.
28. Lumu, I., Musaazi, J., Semeere, A., Handel, I., & Castelnuovo, B. (2023). Survival and predictors of mortality after completion of TB treatment among people living with HIV: a 5-year analytical cohort. BMC infectious diseases, 23(1), 238.
29. Bisson, G., Bastos, M., Campbell, J.R., Bang, D., Brust, J.C., Isaakadis, P., Lange, C., Menzies, D., Migliori, G.B., Pape, J.W. and Palmero, D., (2020). Mortality in Adults with MDR-TB and HIV by ART and TB Drug Use: Individual Patient Data Meta-Analysis.
30. Gatechompol, S., Kawkitinarong, K., Suwanpimolkul, G., Kateruttanakul, P., Manosuthi, W., Sophonphan, J., siwimol Ubolyam, S., Kerr, S.J., Avihingsanon, A. and Ruxrungtham, K., (2019). Treatment outcomes and factors associated with mortality among individuals with both TB and HIV in the antiretroviral era in Thailand. Journal of Virus Eradication, 5(4), pp.225-230.
31. Kaplan, R., Hermans, S., Caldwell, J., Jennings, K., Bekker, L. G., & Wood, R. (2018). HIV and TB co-infection in the ART era: CD4 count distributions and TB case fatality in Cape Town. BMC infectious diseases, 18, 1-9.
32. Yan, I., Bendavid, E., & Korenromp, E. L. (2016). Antiretroviral treatment scale-up and tuberculosis mortality in high TB/HIV burden countries: an econometric analysis. PLoS One, 11(8), e0160481.
33. Harries, A. D., Kumar, A. M., Kyaw, N. T. T., Hoa, N. B., Takarinda, K. C., & Zachariah, R. (2016). The role of antiretroviral therapy in reducing TB incidence and mortality in high HIV-TB burden countries. Asian Pacific journal of tropical disease, 6(3), 243-247.
34. Huber, F.G., Kristensen, K.L., Holden, I.K., Andersen, P.H., Bakir, B., Jørgensen, A., Lorentsson, H.J.N., Bjorn-Mortensen, K., Johansen, I.S. and Ravn, P., (2022). The prevalence of diabetes among tuberculosis patients in Denmark. BMC Infectious Diseases, 22(1), p.64.
35. Araia, Z.Z., Mesfin, A.B., Mebrahtu, A.H., Tewelde, A.G., Osman, R. and Tuumzghi, H.A., (2021). Diabetes mellitus and its associated factors in tuberculosis patients in maekel region, eritrea: analytical cross-sectional study. Diabetes, Metabolic Syndrome and Obesity, pp.515-523.
36. Fwoloshi, S., Hachaambwa, L. M., Chiyeñu, K. O., Chirwa, L., Hoffman, T. W., Ngalamika, O., & Bailey, S. L. (2018). Screening for diabetes mellitus among tuberculosis patients: findings from a study at a Tertiary Hospital in Lusaka, Zambia. Canadian Journal of Infectious Diseases and Medical Microbiology, 2018(1), 3524926.
37. Khalil, N. H., & Ramadan, R. A. (2016). Study of risk factors for pulmonary tuberculosis among diabetes mellitus patients. Egyptian Journal of Chest Diseases and Tuberculosis, 65(4), 817-823.
38. Yoo, J. E., Kim, D., Han, K., Rhee, S. Y., Shin, D. W., & Lee, H. (2021). Diabetes status and association with risk of tuberculosis among Korean adults. JAMA network open, 4(9), e2126099-e2126099.

39. Tenaye, L., Mengiste, B., Baraki, N., & Mulu, E. (2019). Diabetes mellitus among adult tuberculosis patients attending tuberculosis clinics in Eastern Ethiopia. BioMed Research International, 2019(1), 7640836.

نموذج بقاء حدودي مرن لتحليل بيانات الوقت إلى الحدث الخاضعة للرقابة بين مرضى السل

عزيز أديبوي [*1]، أوسوجي جيوجيلين [1] ، ألاكيجا تيميتوبي [2] ، أوديمي أكينومي [1] ، موتامباي روفين 1، مادو بيتر [3]

[1] قسم الإحصاء ، جامعة فورت هير ، أليس ، كيب الشرقية ، جنوب إفريقيا ، [2]قسم الإحصاء ، كلية يابا للتكنولوجيا ، يابا ، ولاية لاغوس ، نيجيريا، [3]قسم العلوم الرياضية ، جامعة أولابيسي أونابانجو ، آغو–إيوي ، ولاية أوغون ، نيجيريا.

الخلاصة: السل مرض معد مميت على مستوى العالم مسؤول عن 10 ملايين حالة جديدة و 1.5 مليون حالة وفاة سنويا . تظهر أنظمة علاج السل الأقصر وعدا في الحد من هذه المشكلة ، ولكن هناك معدل نجاح محسن للعلاج في جنوب إفريقيا ، بينما لا تزال حالات إعادة العلاج مصدر قلق .من السمات المهمة للنمذجة من وقت إلى حدث قدرتها على النظر في احتمالات الانتقال لمجموعات فرعية غير متجانسة ذات ملفات تعريف مخاطر مختلفة .يتم إجراء تحليل البقاء على قيد الحياة بشكل عام لتقدير احتمالات الانتقال المرتبطة بملفات تعريف المخاطر بدقة .استكشفت هذه الدراسة تطبيق نموذج بقاء حدودي مرن لتحليل بيانات الوقت إلى الحدث الخاضعة للرقابة بين مرضى السل .تم الحصول على البيانات من عيادة شرق لندن المركزية–وحدة السل ، كيب الشرقية ، جنوب أفريقيا .في المجموع ، تم تضمين 174 مريضا في التحليل .تم استكشاف الخير من تناسب النماذج باستخدام معيار المعلومات أكايك)إيك .(وقد قدرنا نسب الخطر والمخاطر التراكمية الأساسية لنموذجنا ، وهي ضرورية لحساب احتمالات الانتقال الفردية ، وقارننا النموذج بنموذج كوكس ونموذج الخطر الإضافي لتحديد تنبؤات بقاء مرضى السل .أنتج نموذج البقاء البارامتري المرن نسبة الخطر وتقديرات المخاطر التراكمية الأساسية التي كانت مماثلة لتلك التي تم الحصول عليها باستخدام نموذج المخاطر النسبية كوكس .وكشف التحليل أن الجنس)الموارد البشرية=0.49 ، 95 ٪ سي 0.38 :، 0.62 (، العلاج المضاد للفيروسات القهقرية (الفن) ، (الموارد البشرية=0.53 ، 95 ٪ سي: 0.34 ، 0.78) ، ومرض السكري (الموارد البشرية=0.58 ، 95 ٪ سي: 0.41 ، 0.78) كلها عوامل ذات دلالة إحصائية مرتبطة بتحسين بقاء العلاج في مرضى السل .تعد نماذج البقاء البارامترية المرنة أداة قوية لنمذجة بيانات الوقت إلى الحدث واحتمالات الانتقال الفردية .من الأهمية بمكان ملاءمة النماذج من خلال نمذجة خط الأساس ، مما يسهل عمل أنواع مختلفة من التنبؤات ويسمح بمخاطر غير متناسبة لأنها تفاعل.

الكلمات المفتاحية: نموذج بقاء حدودي مرن ، نموذج خطر مضاف ، نموذج كوكس، متغير متغير زمنيا ، مرض السل.