



# An empirical comparison of time-to-event models to analyse a composite outcome in the presence of death as a competing risk

Ndamonaoghenda Haushona<sup>a, \*</sup>, Tonya M. Esterhuizen<sup>a</sup>, Lehana Thabane<sup>b, c, d</sup>, Rhoderick Machekano<sup>a, e</sup>

<sup>a</sup> Division of Epidemiology and Biostatistics, Stellenbosch University, South Africa

<sup>b</sup> Department of Health Research Methods, Evidence, and Impact, McMaster University, Hamilton, ON, Canada

<sup>c</sup> Biostatistics Unit, Research Institute at St Joseph's Healthcare—Hamilton, Hamilton, ON, Canada

<sup>d</sup> Departments of Paediatrics and Anaesthesia, Schools of Rehabilitation Sciences and Nursing, McMaster University, Hamilton, ON, Canada

<sup>e</sup> Biostatistics Unit, Research Department, Elizabeth Glaser Pediatric AIDS Foundation, Washington DC, USA

## ARTICLE INFO

### Keywords:

Competing risks  
Composite endpoint  
Cause-specific hazard  
Sub-distribution hazard  
Accelerated failure time

## ABSTRACT

**Introduction:** Competing risks arise when subjects are exposed to multiple mutually exclusive failure events, and the occurrence of one failure hinders the occurrence of other failure events. In the presence of competing risks, it is important to use methods accounting for competing events because failure to account for these events might result in misleading inferences.

**Methods and Objective:** Using data from a multisite retrospective observational longitudinal study done in Ethiopia, we performed sensitivity analyses using Fine-Gray model, Cause-specific Cox (Cox-CSH) model, Cause-specific Accelerated Failure Time (CS-AFT) model, accounting for death as a competing risk to determine baseline covariates that are associated with a composite of unfavourable retention in care outcomes in people living with Human Immune Virus who were on both Isoniazid preventive therapy (IPT) and antiretroviral therapy (ART). Non-cause specific (non-CSH) model that does not account for competing risk was also performed. The composite outcome comprises of loss to follow-up, stopped treatment and death. Age, World Health Organisation (WHO) stage, gender, and CD4 count were the considered baseline covariates.

**Results:** We included 3578 patients in our analysis. WHO stage III-or-IV was significantly associated with the composite of unfavourable outcomes, Sub-hazard ratio (SHR) = 1.31, 95% confidence interval (CI):1.04–1.65 for the sub-distribution hazard model, hazard ratio [HR] = 1.31, 95% CI:1.05–1.65, for the Cox-CSH model, and HR = 0.81, 95% CI:0.69–0.96, for the CS-AFT model. Gender and WHO stage were found to be significantly associated with the composite of unfavourable outcomes, HR = 1.56, 95% CI:1.27–1.90, HR = 1.28, 95% CI: 1.06–1.55 for males and WHO stage III-or-IV, respectively for the non-CSH model.

**Conclusions:** Results show that WHO stage III-or-IV is significantly associated with unfavourable outcomes. The results from competing risk models were consistent. However, results obtained from the non-CSH model were inconsistent with those obtained from competing risk analysis models.

## 1. Introduction

Competing risks arise when subjects are exposed to multiple mutually exclusive causes of failure and failure due to one cause precludes the occurrence of failure from other causes. A competing risk is defined as an event that prevents the observation of an event of interest or that in principle modifies the probability of occurrence of the out-

come of interest in the study [12]. Death is one of the common competing risks in health studies, as the occurrence of an event of interest will not be observed once the participant is dead. For instance, in our study, death precludes the occurrence of other unfavourable retention in care outcomes such as loss to follow-up or stopped treatment among people living with HIV (PLHIV). The death event competes with observing other events under investigation, and it prevents us

\* Corresponding author. Faculty of Medicine and Health Sciences, Department of Epidemiology and Biostatistics, Stellenbosch University, P.O. Box 241, Cape Town, 8000, South Africa.

E-mail address: [ndamonahaushona@gmail.com](mailto:ndamonahaushona@gmail.com) (N. Haushona).

<https://doi.org/10.1016/j.conctc.2020.100639>

Received 26 April 2020; Received in revised form 21 July 2020; Accepted 9 August 2020

Available online 14 August 2020

2451-8654/© 2020 The Authors.

Published by Elsevier Inc.

This is an open access article under the CC BY-NC-ND license

(<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

from knowing when an individual would have experienced another unfavourable retention in care outcome had they been alive.

Typically, time to event data are usually analysed using standard survival analysis techniques such as the Kaplan Meier curve and the Cox proportional hazard model to estimate the effect of treatment and risk factors on the hazard of the outcome of interest. These methods assume non-informative censoring and failures from competing risks are treated as censored [2]. However, in the competing risks paradigm, the occurrence of competing events can result in informative censoring [2–6]. Failure to account for competing events may reduce statistical power, overestimate probability of outcomes of interest and result in biased inferences [7,8,9].

Accordingly, extended survival techniques to account for competing risks have been developed, such as the sub-distribution hazards model by Fine and Gray [10] and cause-specific models suggested by Prentice et al [11]. Cause-specific models are appropriate if the interest is in understanding etiological questions. This model is modelled via an appropriate time-to-event model to estimate the effect of covariates on the cause-specific hazard function keeping other events failing as censored. For example, the application through a Cox-proportional hazard model, or estimating the effect of covariates on the cumulative incidence function (CIF) via the accelerated failure time framework [1]. In the presence of competing risks, an unbiased estimator of the cause-specific event probabilities can be obtained by using its CIF [12,13].

On the other hand, most studies also report composite outcomes with the aim to account for competing risks [7,14]. Lunn-McNeill introduced an approach of combining all failure types in one outcome called a composite outcome in the presences of competing risks [15]. The composite outcome is a combination of multiple individual outcomes into a single endpoint. Combining individual outcomes into composite outcomes can increase the overall event rate and the statistical power for a study because of higher event rates [16,17]. Ideally, a composite outcome comprises of outcomes that in principle are assumed to have similar importance, similar relative frequencies of occurrence, similar underlying etiology, similar precision of measurement, and similar magnitude and direction of the treatment effects [14,18]. Composite outcomes are deemed appropriate when the treatment effect across individual outcomes within a composite outcome is homogeneous [7,9]. However, if it is designed to quantify risk benefits or capture competing risks, the assumption of homogeneous treatment effect across individual components can be relaxed [14].

While there exist, a wide-ranging literature comparing traditional survival models and competing risks models, literature comparing these methods with application to empirical data are scarce. In addition, there is a lack of articles focusing on composite outcomes in the application of competing risk models. Majority of articles focused on the application of clinical trials data [8,18–22]. As a result, this study aims to compare time-to-event models by analysing a composite outcome in the presence of death as a competing risk. Since death as a competing risk was not accounted for in the main study, we re-analysed data from the TB breakthrough observational longitudinal study using competing risk analyses to determine baseline covariates associated with the composite of unfavourable retention in care outcomes among people living with HIV who were on IPT plus ART, while explicitly accounting for death and allowing for within-hospital clustering.

We employed the Fine and Gray model in the primary analysis. In addition, we performed sensitivity analyses using cause-specific models through the Cox proportional hazard model, accelerated failure time model, and a Cox-proportional (non-cause-specific hazard) model to assess robustness of the findings. Performing sensitivity analyses is an essential step in analysis of health studies to assess the robustness or consistency of the results under different models, or assumptions to establish credibility of study findings [23,24].

## 2. Methods

The dataset used in this study is publicly available through the PLOS ONE policy, and comprehensive details on the selection of study participants that provided data analysed in this study are described elsewhere [5]. The following is a brief description of the study design, population, setting, exposure variables and outcomes, data collection procedures and ethical consideration.

### 2.1. Study design, population and setting

This study is a secondary analysis of a multisite retrospective observational longitudinal study that was done in Ethiopia involving  $n = 4484$  participants with HIV, for a period of nine years (from the year 2005–2014) who were initiated on IPT [5]. The main study aimed to assess the magnitude and factors associated with tuberculosis breakthrough among people living with HIV who were initiated on IPT.

The study includes People Living with HIV (PLHIV) from three regions in Ethiopia: Addis Ababa, Gambella, and Southern Nations Nationalities and Peoples (SNNP) region. These regions were selected because they used a similar data management system. The population consists of 35 hospitals in these regions from which 11 hospitals were randomly selected for the sample. PLHIV visited the 11 randomly selected hospitals between September 2005 and October 2013 and were included in the sample. The main study was approved by the National Research Ethics Review Committee (NRERC) of Ethiopia and no informed consent was required because patient's information was extracted from non-identifiable and non-linked databases.

### 2.2. Sample size

As per the rule of thumb, a minimum of 10 events per explanatory variable is efficient to avoid model overfitting, however, when performing sensitivity analysis this rule may be relaxed [25]. In our study, there were at least 10 unfavourable retention in care outcomes per predictor variable, and we performed sensitivity analyses. Therefore, by the rule of thumb, a sample size of 3578 patients (Fig. 1) who were on both IPT and ART is adequate to fit our models.

### 2.3. Inclusion criteria

Only individuals who did not have tuberculosis (TB) were initiated on IPT. Last status was only recorded for patients who were on both IPT plus ART in order to understand retention in care, therefore, only patients who were on both IPT and ART were included in our analysis.

### 2.4. Study measurements

#### 2.4.1. Explanatory and exposure variables

The following baseline characteristics and clinic data were extracted from the databases: patient age, gender, baseline WHO Stage, and baseline CD4 count. The time from starting IPT to the last observation was also recorded. All covariates collected were dichotomised in the primary study.

#### 2.4.2. Composite outcome variable

The primary outcome of our study was defined as a composite of unfavourable retention in care outcomes, which comprises of loss to follow-up or stopped treatment or death as per the primary study [5]. These outcomes were measured during the last visit to understand retention in care. No multiple unfavourable outcomes were recorded per subject. In addition, no missingness of the outcome nor covariates were observed in our data.

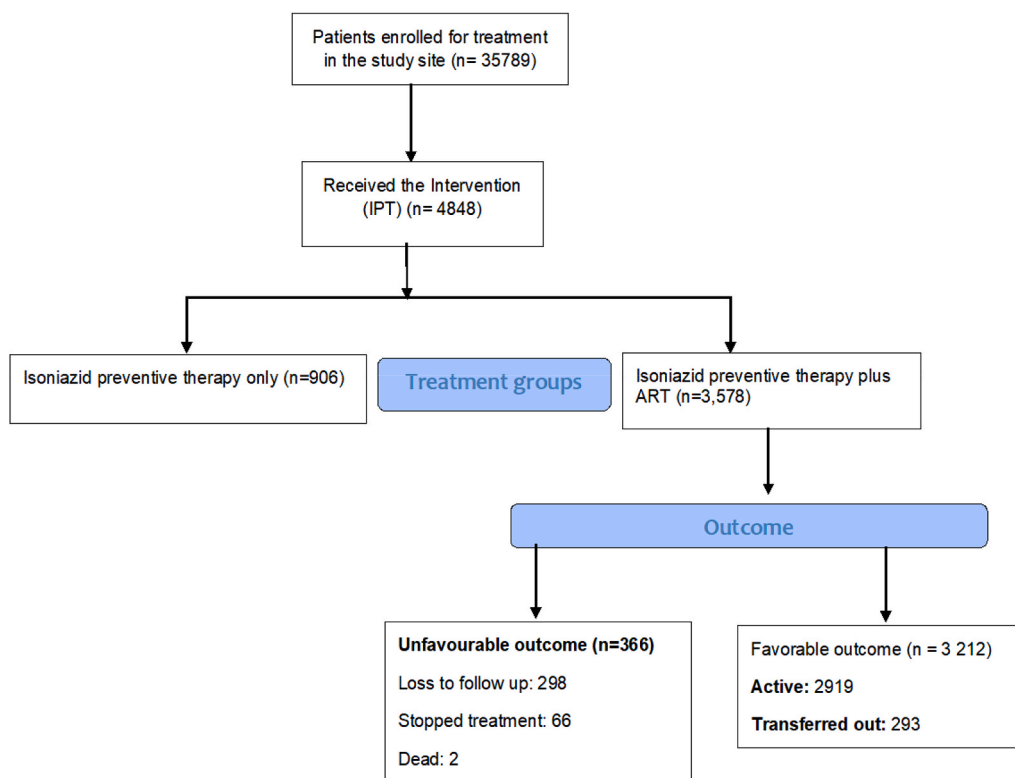


Fig. 1. Study participant's flow diagram.

2.5. Ethics

The data used in this study was obtained from non-identifiable and non-linked databases and no informed consent was required. In addition, it is publicly available for use. The study was approved for exemption of ethics review by the Stellenbosch Human Research Ethics Committee (HREC) under Project ID 10953 and Ethics Reference Number X19/08/029.

2.6. Statistical methods

2.6.1. Fine-Gray sub-distribution hazard model

Fine and Gray [10] introduced a semi-parametric regression analysis used to estimate the effects of treatment and covariates in the presence of competing risks called sub-distribution hazard model. This model is used when the interest is in answering prognostic questions. That is, it models the probability of an event happening. Fine-Gray sub-distribution hazard model estimates the effects of covariates on the sub-distribution hazard function. The sub-distribution hazard function shows the instantaneous rate of occurrence of the event of interest at time t for subjects who are either event free or have experienced the competing event before time t. The crude incidence of the occurrence of the jth event while accounting for competing risks is estimated using the cumulative incidence function F(t) also referred to as sub distribution function.  $F_j(t)$  is the cumulative probability of failure from a cause j prior to time t in the presence of competing risks and it describes the incidence of the occurrence of an event while accounting for competing risks. The cumulative incidence function is given by  $F_j(t) = P [T \leq t, J = j]$ . The sub-distribution hazard function is mathematically defined as  $\lambda_j^{sd}(t) = \lim_{\Delta t \rightarrow 0} \frac{P[t \leq T < t + \Delta t, J = j | (T > t) \cup (T < t \cap J \neq j)]}{\Delta t}$ ,  $J = 1, \dots, j$ , where T is time to failure, and J is the type of event for that subject.

2.6.2. Cox cause-specific (Cox-CSH) regression model

The cause specific model is a semi-parametric model introduced by Kalbfleish and Prentice 2002 used to analyse time to event data in the presence of competing risks [26] in order to address etiological or biological questions. The CSH model estimates the effect of covariates on the cause-specific hazard function. The estimation of the effect of covariates on the cause-specific hazard functions is modelled via an appropriate standard time to event model. For example, the application of a cox-proportional hazard model, and an accelerated failure time model were used in this study. In the presence of competing risks, an unbiased estimator of the cause-specific event probabilities can be obtained by using its cumulative incidence functions (CIFs) [12,13]. The CIF of the cause-specific hazards is defined as the probability of failing from the event of interest by time t while still at risk of failing from other competing events. Mathematically, it is defined as  $F_j(t) = P [T \leq t, J = j]$ . The cause-specific hazard function for the jth failure is given by  $\lambda_j^{cs}(t) = \lim_{\Delta t \rightarrow 0} \frac{P[T \leq t + \Delta t, J = j | T \geq t]}{\Delta t}$ ,  $J = 1, \dots, j$ . Failures from competing events are treated as censored. Hazard ratios obtained from this model are interpreted as the association with the rate of the event, in this case the association between the covariate and the rate of the unfavourable retention in care outcomes.

2.6.3. Cause-specific AFT (CS-AFT) model

Cause-specific AFT is a linear competing risks regression model used to analyse time to event data in the presence of competing risks [26]. The AFT model is used to assess the effect of covariates on mean survival time [1]. Typically, when using cause-specific hazard models in the presence of competing risks, there is no direct effect of covariates on the failure time. Furthermore, failure due to all other causes except for the events of interest are treated as censored. This model is mathematically written as  $\log(T|Z) = \alpha_j Z' + \epsilon_j$ , where T denote the latent failure time due to the unfavourable outcome j,  $\alpha_j$  denote a vector of parameters estimates measuring the crude covariate effects on the failure time of the unfavourable outcome j only, Z denote a vector

of covariates, and  $\epsilon_j$  denote an error term with an unspecified distribution function.

### 2.6.4. Non-cause-specific hazard (non-CSH) model

The Cox proportional hazard model is a semi-parametric model used to assess the effect of treatment and covariates on a single event of interest [27]. It is represented by a hazard function, baseline hazard function and the covariates, and it is mathematically represented as  $h(t) = h_0(t) \exp(\beta' X)$ , where  $h(t)$  is the hazard function,  $h_0(t)$  is the baseline hazard function which is left unspecified,  $\beta$  is a column vector of the regression coefficients, and  $X$  is a column vector of the covariates. The Cox regression model assumes that there is the proportionality of the hazard rate. Proportional hazards mean that the hazard ratio is constant over time such that the effect of a covariate is the same at all points in time. Typically, when multiple outcomes are combined in a composite outcome, it is deemed appropriate to perform a non-cause-specific hazard model to estimate the effect of covariates on the unfavourable outcomes by combining all outcomes without censoring patients who experienced the competing event.

### 2.7. Statistical analyses

Descriptive statistics for participants' baseline characteristics were presented as frequencies and percentages. In all the analyses, participants' age categorized < 15 years and  $\geq 15$  years; gender, WHO Stage categorized I or II, and III or IV, and CD4 cell count categorized < 100, 100–349, and > 349 were considered as covariates. In addition, a graphical examination using cumulative incidence function curves were used to estimate the incidence of unfavourable retention in care outcomes at any time point between baseline and time  $t$ .

For the primary analysis, a multivariable Fine and Gray sub-distribution hazard model was performed to assess the baseline covariates associated with the composite of unfavourable retention in care outcomes, while accounting for death as a competing risk and allowing for within hospital clustering. Also, sensitivity analyses were performed using Cox-CSH, AFT model and a non-cause-specific hazard model (where we combined all the three unfavourable retention in care outcomes). Hazard ratios (HR), corresponding 95% confidence intervals (CIs) and associated p-values were reported for all the models.

Graphical examinations based on Schoenfeld residuals were used to check for the assumption of proportional hazards. In addition, a confirmatory test was performed by including an interaction between time and the covariates in the model using a tvc (time varying covariates) option in STATA. The tvc option is used together with texp to create interactions of the predictors and a function of survival time in the model. Covariates violating the proportional hazard assumption i.e. covariates with a p-value of less than 0.05 under the tvc function were included in the models as time-varying covariates. The criteria for statistical significance was set at 0.05 for all tests. All analyses were performed in STATA Version 15.1 (Stata Corp., College Station, TX) and cumulative incidence functions graphical display for the composite endpoint, and the individual components of the composite were done in R-version 3.6.1.

## 3. Results

### 3.1. Baseline characteristics

Descriptive statistics for baseline covariates of participants who were on IPT plus ART against their last status are shown in Table 1. Briefly, 3212 (89.8%) participants had favourable retention in care, and 366 (10.2%) patients had one of the unfavourable retention in care outcomes (loss to follow up or stop treatment or death) at last visit or observation. The unfavourable retention in care composite

**Table 1**

Baseline characteristics of patients on IPT plus ART (n = 3578).

Baseline covariates	Subcategories	Favourable outcome (n = 3212)	Unfavourable outcome (n = 366)
Age (years) n (%)	< 15	292 (94.81)	16 (5.19)
	$\geq 15$	2920 (89.30)	350 (10.70)
Gender n (%)	Male	1165 (87.14)	172 (12.86)
	Female	2047 (91.34)	194 (8.66)
WHO Stage n (%)	I or II	1495 (91.33)	142 (8.67)
	III or IV	1717 (88.46)	224 (11.54)
CD4 Count n (%)	< 100	701 (88.62)	90 (11.38)
	100–349	2278 (89.83)	258 (10.17)
	$\geq 350$	233 (92.83)	18 (7.17)

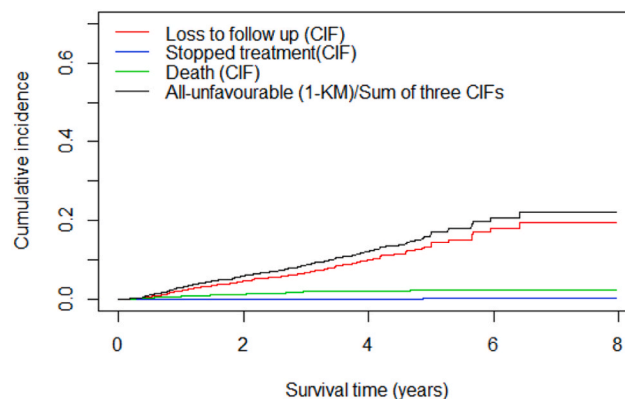
WHO- World Health Organisation, CD4<sup>-</sup>cluster of differentiation 4.

outcome comprises of death (66), stopped treatment (2) and loss to follow-up (298). Unfavourable outcomes were frequently recorded among adults (age  $\geq 15$  years) (10.7%), and only 5.19% was recorded among children (age < 15 years). There was a minor discrepancy in the occurrence of unfavourable outcomes by gender. Furthermore, 11.54% of patients with WHO stage III or IV had unfavourable retention in care outcomes, and 8.67% of patients with WHO stage I or II had unfavourable outcomes. 11.38% patients who had less than 100 CD4 cells per cubic millimeter of blood had unfavourable retention in care outcomes. The median follow-up time to the last status was 3.16 years.

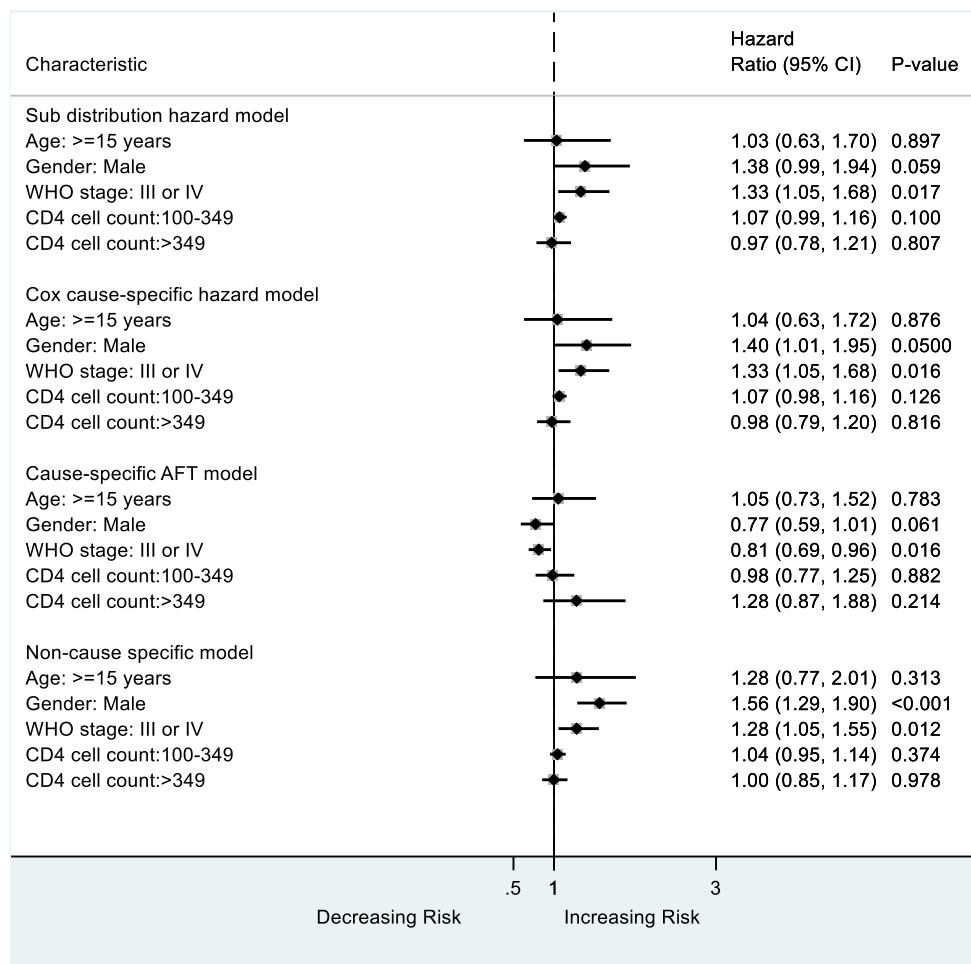
Fig. 2 shows the CIF curves of loss to follow up, stopped treatment, and death, along with the CIF curve for the composite outcome of all the unfavourable retention in care outcomes. The cumulative incidence of the composite outcome is equal to the sum of all the cumulative incidences of all the cause-specific unfavourable retention in care outcomes.

### 3.2. Primary analysis: Fine and Gray sub-distribution hazard model

Results from the sub distribution hazard model are presented in Fig. 3. CD4 cell count violated the assumption of proportional hazards i.e. under the tvc option CD4 count had a p-value = 0.001 which is less than 0.05, indicating the violation of proportional hazards. As a result, we included an interaction between CD4 cell count and survival time in order to account for the violation of the proportional hazards assumption. A significant difference in the risk of unfavourable retention in care outcomes between WHO stages was observed. The estimated adjusted SHR and associated 95% confidence



**Fig. 2.** Cumulative incidence functions, and KM Kaplan-Meier. WHO- World Health Organisation, CD4<sup>-</sup>cluster of differentiation 4, CI- Confidence Interval.



WHO- World Health Organisation, CD4- cluster of differentiation 4, CI- Confidence Interval

Fig. 3. Forest plot for baseline covariates based on different models.

interval was 1.33(1.05–1.68) for WHO stage III or IV compared to WHO stage I or II after controlling for sex, age and CD4 cell count in the model. WHO stage III or IV was significantly associated with a 33% increase in the incidence of unfavourable retention in care outcomes in subject who had favourable outcomes or experienced the competing event (death). There was an insignificant increase in the incidence of unfavourable retention in care outcomes among adults ( $\geq 15$  years) than in children among individuals who had favourable outcomes or who died. The interaction between CD4 cell count and time was insignificant. There was no evidence that the effect of an increase in CD4 cell count levels decreased the incidence of unfavourable outcomes over time in both models. Lastly, there was no significance difference in unfavourable retention in care outcomes between males and females.

3.3. Sensitivity analyses

Fig. 3 shows results obtained from three multivariable sensitivity analyses models. CD4 cell count violated the assumption of proportional hazards i.e. under the tvcoption CD4 count had a p-value = 0.001 which is less than 0.05, indicating the violation of proportional hazards. As a result, we included an interaction between CD4 cell count and survival time in all models.

WHO stage III or IV was significantly associated with the composite of unfavourable retention in care outcomes in both CS models, adjusted HR = 1.33, 95% CI: 1.06–1.68, p-value = 0.016, HR = 0.81,

95% CI: 0.69–0.96, p-value = 0.016 for the Cox-CSH and CS-AFT model, respectively. From the Cox-CSH model, WHO stage III or IV is significantly associated with a 33.2% increase in the incidence of unfavourable retention in care outcomes among patients who were having favourable retention in care outcomes. WHO stage III or IV was significantly associated with 18% decrease in the mean survival to unfavourable retention in care outcomes in those who were currently having favourable retention in care outcomes for the CS-AFT model.

Age, sex, and CD4 cell count were not significantly associated with the unfavourable retention outcomes as they were observed with p-values equal to or greater than 0.05 (Fig. 3).

Similarly, age group, sex and CD4 cell count were not significantly associated with the unfavourable outcomes for the CS-AFT model adjusted HR = 1.05, 95% CI: 0.73–1.52, p-value = 0.783 for adults (age  $\geq 15$  years); adjusted HR = 0.77, 95% CI: 0.59–1.01, p-value = 0.061 for males compared to females; adjusted HR = 0.98, 95% CI: 0.77; 1.25, p-value = 0.88 and adjusted HR = 0.78, 95% CI: 0.87–1.88, p-value = 0.21 for CD4 cell count 100–349, and  $\geq 350$ , respectively.

Additionally, the non-CSH model shows being Male and WHO stage III or IV to be significantly associated with the composite of the unfavourable retention in care outcomes after controlling age and CD4 cell count in the model. Males were more likely to have unfavourable retention in care outcomes (loss to follow up, stopped treatment or death) as compared to females (adjusted HR = 1.56, 95% CI: 1.29–1.90, p-value<0.001). Patients with WHO stage IIV or



IV were significantly more likely to have unfavourable retention in care outcomes as compared to those with WHO stage I or II.

#### 4. Discussion

We performed competing risk analyses to determine baseline covariates that are associated with a composite of unfavourable retention in care outcomes (stopped treatment or loss to follow up or death) among people living with HIV who were on both IPT and ART, accounting for death as a competing risk. Competing analyses (sub-distribution hazard model and CS models) showed that WHO stage III or IV is significantly associated with an increase in the incidence of unfavourable retention in care outcomes. The AFT model shows that WHO stage III or IV is significantly associated with a reduction in survival mean time of unfavourable retention in care outcomes among individuals who are currently having favourable retention in care outcomes. Additionally, gender was significantly associated with unfavourable retention in care outcomes from the non-CSH model.

Findings from three competing risk analyses were consistent. However, results obtained from the CS-AFT model has effect estimates going in the opposite direction of those obtained in the sub-distribution hazard model, and Cox-CSH model. This is due to the fact that AFT models measures the effect of covariates on the survival mean time of the outcome instead of hazard ratio contrary to the sub-distribution hazard model and Cox-CSH hazards models [1]. Effect of covariates obtained from the AFT models are interpreted as time ratios and the ratio denotes the acceleration factor. A time ratio less than one implies that the event is more likely to happen, similarly to when the hazard ratio is greater than. These models reached the same conclusion though a negative correlation between the AFT, Cox-CSH and the sub-distribution hazard model is observed. Results from sub-distribution hazard model, Cox-CSH, and non-CSH model were roughly similar. CSH and sub-distribution hazard yields similar results when there are few events or when a covariate only affects one of the cause-specific hazards [13].

Moreover, findings from competing risk models were not in agreement with results from the non-CSH model, which showed both gender and WHO stage to be significantly associated with unfavourable retention in care outcomes. A composite outcome of all failure events is deemed appropriate to capture competing risks without censoring individuals who experience competing event.

However, results from the non-CSH model were not consistent with those obtained from the Fine and Gray sub-distribution hazard model and CSH models because of the difference in baseline characteristics of participants that were significantly associated with unfavourable retention in care outcome. Therefore, estimates from the non-CSH model might not be the most accurate in this case hence the results from the non-CSH model need to be interpreted with thoughtfulness. Our findings are consistent with those found in Dignam JJ et al. [22], that when the rate of competing events is low, the two competing risk models tends to have similar results. Additionally, this was also confirmed that when the incidence of the competing events is small, there will be minimal bias after performing traditional survival methods that ignores competing events, however, bias escalates as the incidence of the competing risks increases [28].

There were some limitations to our study. The major limitation of our study is bias resulting from residuals of loss to follow-up outcomes that might have been classified as outcome of interest since we could not confirm whether individuals who are lost to follow up are dead or not because we utilised secondary data. The non-discreteness of death from lost to follow up might have overestimated the probability of the outcome of interest and underestimate the probability of those who died. Another significant limitation is that our study used empirical data for analysis, thus any conclusion made from this study may not mirror findings in other settings hence results from this study

cannot be generalised to other population. Additionally, the use of secondary data could not spare us on limitations that arises from the use of existing records such as residual confounding when crucial variables that needed to be accounted for in the analysis are omitted during data collection. Also, the aim of the study was to explore the performance of models via these data not to make clinical conclusions. Our study will contribute to literature to help researchers identify effective modelling approaches in competing risk settings to enable unbiased inferences.

Literature around the paradigm of competing risks in comparison to traditional time to event models found that ignoring competing risks may lead to an overestimation of cumulative incidence and hence leading to misleading results [2,4,6,21,22,26,29-32]. Considering the fact that competing risks arises in most studies be it observational studies where routine data is collected at health facilities or in clinical trials where patients are subjected to multiple outcomes, comparison of time to event models catering for different scenarios such as competing events are infrequently applied to empirical data. Also, rarely composite outcomes are considered. This paper offers an empirical comparison of such methods with an application to empirical data focusing on composite outcomes and it will help researchers around this area with analysis as well as provide a grasp of how to interpret results from various competing risks models. Our study will contribute to literature to help researchers identify effective modelling approaches in competing risk settings, particularly when analysing survival data from observational studies as this literature is limited as opposed to literature focusing on data from randomised clinical trials.

#### 5. Conclusions

In both competing risk analyses, we found evidence that WHO stage was found to be significantly associated with the composite of unfavourable retention in care outcomes. However, we did not find any significant effect of gender, age and CD4 cell count on the unfavourable retention in care outcomes was found in these analyses. Non-competing risk analysis showed evidence of gender being associated with unfavourable outcomes, in addition to the WHO stage. Results from competing risk analyses were consistent, however, not in agreement with those obtained from the non-CSH model. Therefore, results from the non-CSH model need to be interpreted with caution. We have provided an empirical comparison of the results from different models. These models may result into different results as they address different research questions, cause specific addresses etiological questions, while sub-distribution addresses prognostic questions; hence it is advisable to perform both models in the presence of competing risks to ensure valid inferences. In conclusion, the model choice must be guided by the type of research question to be addressed.

#### Authorship contribution

Ndamonaonghenda Haushona led data acquisition, data analysis and interpretation of results. Ndamonaonghenda Haushona, Tonya Esterhuizen and Lehana Thabane were responsible for the concept and design of the study. Ndamonaonghenda Haushona drafted the first version of the article. Rhoderick Machezano is the co-Principal Investigator of the Fogarty Grant and internally reviewed the paper. Ndamonaonghenda Haushona, Lehana Thabane, and Tonya Esterhuizen revised the paper critically for important intellectual content. All authors reviewed the manuscript and contributed to the final submission.

#### Declaration of competing interest

No potential conflict of interest was reported by the authors.

## Acknowledgments

This analysis was carried out as a part of the MSc studies for N. Haushona who received a scholarship from the African Centre for Biostatistical Excellence supported by the Fogarty International Centre of the National Institutes of Health for funding her postgraduate studies under Award Number D43 TW010547. *The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.* Additionally, we would like to thank Ms. Kesetebirhan Delele who provided us with the data.

## References

- [1] S. Choi, H. Cho, Accelerated failure time models for the analysis of competing risks, *J. Korean Surg. Soc.* 48 (2018) 315–326, <https://doi.org/10.1016/j.jkss.2018.10.003>.
- [2] Z. Zhang, *Survival Analysis in the Presence of Competing Risks*, vol. 5, 2009 3, <https://doi.org/10.21037/atm.2016.08.62>.
- [3] H.T. Kim, Cumulative incidence in competing risks data and competing risks regression analysis, *Clin. Canc. Res.* 13 (2 I) (2007) 559–565, <https://doi.org/10.1158/1078-0432.CCR-06-1210>.
- [4] B. Haller, G. Schmidt, K. Ulm, Applying competing risks regression models: an overview, *Lifetime Data Anal.* 19 (1) (2012) 33–58, <https://doi.org/10.1007/s10985-012-9230-8>.
- [5] K.D. Yirdaw, A.M. Teklu, A.T. Mamuye, S. Zewdu, Breakthrough tuberculosis disease among people with HIV - should we be worried? A retrospective longitudinal study, *PLoS One* 14 (2) (2019) 1–11, <https://doi.org/10.1371/journal.pone.0211688>.
- [6] P.C. Austin, J.P. Fine, Practical recommendations for reporting Fine-Gray model analyses for competing risk data, *Stat. Med.* 36 (27) (2017) 4391–4400, <https://doi.org/10.1002/sim.7501>.
- [7] L.K. Mell, J.H. Jeong, Pitfalls of using composite primary end points in the presence of competing risks, *J. Clin. Oncol.* 28 (28) (2010) 4297–4299, <https://doi.org/10.1200/JCO.2010.30.2802>.
- [8] S. Parpia, J.A. Julian, L. Thabane, A.Y.Y. Lee, F.R. Rickles, M.N. Levine, Competing events in patients with malignant disease who are at risk for recurrent venous thromboembolism, *Contemp. Clin. Trials* 32 (6) (2011) 829–833, <https://doi.org/10.1016/j.cct.2011.07.005>.
- [9] G. Li, D.J. Cook, M.A.H. Levine, et al., Competing Risk Analysis for Evaluation of Dalteparin versus Unfractionated Heparin for Venous Thromboembolism in Medical-Surgical Critically Ill Patients, vol. 94, 2015, pp. 1–8 36, <https://doi.org/10.1097/MD.0000000000001479>.
- [10] J.P. Fine, R.J. Gray, A proportional hazards model for the subdistribution of a competing risk, *J. Am. Stat. Assoc.* 94 (446) (1999) 496–509, <https://doi.org/10.1080/01621459.1999.10474144>.
- [11] R.L. Prentice, J.D. Kalbfleisch, A.V. Peterson, N. Flournoy, V.T. Farewell, N.E. Breslow, The analysis of failure times in the presence of competing risks, *Biometrics* 34 (4) (2006) 541, <https://doi.org/10.2307/2530374>.
- [12] G. Rauch, M. Kieser, S. Ulrich, et al., Competing Time-To-Event Endpoints in Cardiology Trials : A Simulation Study to Illustrate the Importance of an Adequate Statistical Analysis, 2016, <https://doi.org/10.1177/2047487312460518>.
- [13] S. Van Der Pas, R. Nelissen, M. Fiocco, Different competing risks models for different questions may give similar results in arthroplasty registers in the presence of few events: illustrated with 138,234 hip (124,560 patients) and 139,070 knee (125,213 patients) replacements from the Dutch Arth, *Acta Orthop.* 89 (2) (2018) 145–151, <https://doi.org/10.1080/17453674.2018.1427314>.
- [14] J. Pogue, P.J. Devereaux, L. Thabane, S. Yusuf, Designing and analyzing clinical trials with composite outcomes: consideration of possible treatment differences between the individual outcomes, *PLoS One* 7 (4) (2012), <https://doi.org/10.1371/journal.pone.0034785>.
- [15] D.G. Kleinbaum, Survival analysis, a self-learning text, *Biom. J.* 40 (1) (1998) 107–108, [https://doi.org/10.1002/\(SICI\)1521-4036\(199804\)40:1<107::AID-BIMJ107>3.0.CO;2-9](https://doi.org/10.1002/(SICI)1521-4036(199804)40:1<107::AID-BIMJ107>3.0.CO;2-9).
- [16] G. Rauch, K. Kunzmann, M. Kieser, K. Wegscheider, J. König, C. Eulenburger, A weighted combined effect measure for the analysis of a composite time-to-first-event endpoint with components of different clinical relevance, *Stat. Med.* 37 (5) (2018) 749–767, <https://doi.org/10.1002/sim.7531>.
- [17] E.J. Mascha, D.I. Sessler, Design and analysis of studies with binary-event composite endpoints: guidelines for anesthesia research, *Anesth. Analg.* 112 (6) (2011) 1461–1471, <https://doi.org/10.1213/ANE.0b013e31821796d3>.
- [18] G. Li, D.J. Cook, M.A.H. Levine, et al., Competing risk analysis for evaluation of dalteparin versus unfractionated heparin for venous thromboembolism in medical-surgical critically ill patients, *Med (United States)* 94 (36) (2015) 1–8, <https://doi.org/10.1097/MD.0000000000001479>.
- [19] H.J. Lim, X. Zhang, R. Dyck, N. Osgood, Methods of competing risks analysis of end-stage renal disease and mortality among people with diabetes, *BMC Med. Res. Methodol.* 10 (2010), <https://doi.org/10.1186/1471-2288-10-97>.
- [20] R. Sapir-Pichhadze, M. Pintilie, K.J. Tinckam, et al., Survival analysis in the presence of competing risks: the example of waitlisted kidney transplant candidates, *Am. J. Transplant.* 16 (7) (2016) 1958–1966, <https://doi.org/10.1111/ajt.13717>.
- [21] I.D. Ha, N.J. Christian, J.H. Jeong, J. Park, Y. Lee, Analysis of clustered competing risks data using subdistribution hazard models with multivariate frailties, *Stat. Methods Med. Res.* 25 (6) (2016) 2488–2505, <https://doi.org/10.1177/0962280214526193>.
- [22] J.J. Dignam, Q. Zhang, M. Kocherginsky, The use and interpretation of competing risks regression models, *Clin. Canc. Res.* 18 (8) (2012) 2301–2308, <https://doi.org/10.1158/1078-0432.CCR-11-2097>.
- [23] L. Thabane, L. Mbuagbaw, S. Zhang, et al., A tutorial on sensitivity analyses in clinical trials: the what, why, when and how, *BMC Med. Res. Methodol.* 13 (1) (2013), <https://doi.org/10.1186/1471-2288-13-92>.
- [24] R.J. De Souza, R.B. Eisen, S. Perera, et al., Best (but oft-forgotten) practices: sensitivity analyses in randomized controlled trials, *Am. J. Clin. Nutr.* 103 (1) (2016) 5–17, <https://doi.org/10.3945/ajcn.115.121848>.
- [25] E. Vittinghoff, C.E. McCulloch, Relaxing the rule of ten events per variable in logistic and cox regression, *Am. J. Epidemiol.* 165 (6) (2007) 710–718, <https://doi.org/10.1093/aje/kwk052>.
- [26] A.R.L. Prentice, J.D. Kalbfleisch, A.V. Peterson, N. Flournoy, V.T. Farewell, N.E. Breslow, The analysis of failure times in the presence of competing risks published by : international biometric society stable, REFERENCES Linked references are available on JSTOR for this article : You may need to log in. *Biometrics*. 1978;34 (4):541-554 <https://www.jstor.org/stable/2530374>.
- [27] R.S. Society, Regression models and life-tables author ( s ) : D . R . Cox Source, J. Roy. Stat. Soc. B 34 (2) (2019) Published by : Wiley for the Royal Statistical Society Stable URL [https://www.jstor.org/stable/34\(2\):187-220](https://www.jstor.org/stable/34(2):187-220).
- [28] H. Abdel-Qadir, J. Fang, D.S. Lee, et al., Importance of considering competing risks in time-to-event analyses: application to stroke risk in a retrospective cohort study of elderly patients with atrial fibrillation, *Circ Cardiovasc Qual Outcomes* 11 (7) (2018) 1–11, <https://doi.org/10.1161/CIRCOUTCOMES.118.004580>.
- [29] T.A. Gooley, W. Leisenring, J. Crowley, B.E. Storer, Estimation of failure probabilities in the presence of competing risks: new representations of old estimators, *Stat. Med.* 18 (6) (1999) 695–706, [https://doi.org/10.1002/\(SICI\)1097-0258\(19990330\)18:6<695::AID-SIM60>3.0.CO;2-O](https://doi.org/10.1002/(SICI)1097-0258(19990330)18:6<695::AID-SIM60>3.0.CO;2-O).
- [30] W. Yang, C. Jepson, D. Xie, et al., Statistical methods for recurrent event analysis in cohort studies of CKD, *Clin. J. Am. Soc. Nephrol.* 12 (12) (2017) 2066–2073, <https://doi.org/10.2215/CJN.12841216>.
- [31] B. Freidlin, E.L. Korn, Testing treatment effects in the presence of competing risks, *Stat. Med.* 24 (11) (2005) 1703–1712, <https://doi.org/10.1002/sim.2054>.
- [32] P.C. Lambert, Europe PMC Funders Group The Estimation and Modelling of Cause-specific Cumulative Incidence Functions Using Time-dependent Weights, vol. 17, 2018, pp. 181–207 1.