

1 **Competing risks multi-state model for time-to-event data analysis of HIV/AIDS: A** 2 **retrospective cohort national datasets, Ethiopia.**

3 Tsegaye Hailu Kumsa^{1,3} , Andargachew Mulu¹ , Adane Mihret ¹ , Zeytu G. Asfaw²

4
5 1 Armauer Hansen Research Institute (AHRI), Addis Ababa, Ethiopia

6 2 Schools of Public Health, Addis Ababa University, Addis Ababa, Ethiopia

7 3 Department of Statistics Hawassa University, Hawassa, Ethiopia

8 **Abstract**

9 **Introduction:** When HIV/AIDS patients die from opportunistic infections that cause death to
10 happen more quickly, competing dangers are present. The objective of the current study was to
11 estimate the probability of HIV/AIDS patients dying from competing illnesses such as
12 tuberculosis, diarrhea, other opportunistic infections, and unidentified infections.

13 **Methods:** All regional states, including the administrative cities of Addis Abeba and Dire Dawa,
14 a retrospective cohort research was carried out between November 2019 to March 2020. There
15 were 39590 HIV-positive individuals considered. We used competing risk models with a time-
16 to-death horizon of 1212 for the total number of HIV-positive people. A competing event was
17 thought to be death from various causes.

18 **Results:** Out of the total 1212 deaths, 542(44.7%) died competing with other opportunistic
19 infection (TE-Esophageal Candidiasis, TO-oral, CT-CNS Toxoplasmosis, CM-Cryptococcal
20 Meningitis...), 421 (34.7%) died due to tuberculosis and the remaining death were
21 unknown/Not specified infection 222(18.3%) and diarrhea 27(2.2%). Rates of mortality caused
22 by tuberculosis, competing with other opportunistic infection, diarrhea and unknown/Not
23 specified were 3.5, 4.5, 0.2 and 1.8 per 1000 person-months, respectively. Responding to
24 combined Antiretroviral Treatment (cART) 6 months after initiation, receiving Pneumocystis
25 Pneumonia (PCP) prophylaxis, and higher CD4 count at diagnosis reduced the hazard of
26 tuberculosis, other opportunistic infection and unknown and diarrhea causes of death. However,
27 older age, late HIV/AIDS diagnosis, and the last HIV/AIDS WHO clinical stages increased the
28 hazard of tuberculosis and other opportunistic disease mortality. Additionally, male gender, older
29 age and last HIV clinical stages increased the mortality HIV/AIDS patients.

30 **Conclusion:** The findings of this study demonstrated that TB, an opportunistic infection, was the
31 primary cause of death in HIV/AIDS patients, despite the presence of several competing risks,
32 such as diarrhea, other infections, and an undetermined or unclear cause. It's important to use
33 effective techniques to quickly detect those who have HIV or AIDS and provide them with care
34 and treatment to increase their chances of surviving.

35 **Keywords:** *AIDS; Competing risks; Cause-specific hazard; Sub-distribution hazard;*
36 *Cumulative Incidence Function.*

37 **Introduction**

38 According to estimates, HIV has affected 40.1 million people worldwide and is still transmitting
39 in all nations, with 25.6 million of those cases occurring in the African region [1, 2]. HIV
40 infection affected 1.5 million individuals in 2021, and 650000 people passed away from HIV-

41 related causes (1, 3). In 2022, there will be about 10421 fatalities each year in Ethiopia across all
42 age groups (4).

43
44 Antiretroviral therapy (ART) has increased the long-term survival rate of HIV-infected patients
45 and transformed their outlook from a terminal sickness to a protracted and treatable condition.
46 But this depends on a number of risk variables, such as if the HIV patient has TB, diarrhea, other
47 opportunistic illnesses, or an unknown infection. As a result, rather than the virus itself, an
48 HIV/AIDS patient died from TB and opportunistic illness that were growing more common in
49 the population. In order to develop effective intervention strategies, improve patient care, and
50 minimize HIV patient mortality as much as possible, causes of death (CoD) among HIV-infected
51 patients must be evaluated (5). Human immunodeficiency virus (HIV) patients receiving
52 antiretroviral therapy (ART) were the subjects of cohort studies, where participants might go
53 through a variety of circumstances. One must consider the possibility of opportunistic infectious
54 disease while examining the time to death process in HIV positive individuals, for example.
55 Understanding the prognostic variables affecting HIV patients' long-term survival is necessary
56 (6, 7). In addition to the risk of dying from tuberculosis, HIV/AIDS patients also run the risk of
57 dying from other infections or diarrhea. Competing risks arise when participants can witness one
58 or more events, they 'compete' with one another, and the occurrence of one event could make it
59 impossible to witness the other events or change the likelihood of their occurrence. It is
60 customary to achieve this by carrying out separate studies for each end point and the
61 intermediate events, but this is insufficient because it disregards the connections between these
62 events. Multistate models (MSM) are a natural technique to model such complex systems in this
63 context (8).

64
65 A particularly helpful tool for answering a variety of survival analysis concerns that traditional
66 models are unable to address is the MSM framework (9). Figure 1 displays the two simplest but
67 most typical instances. Although the current HIV/AIDS study is primarily focused on deaths
68 caused by tuberculosis, some patients may experience many competing causes of death.
69 Although fatalities from other causes occur in place of tuberculosis deaths, the event of interest is
70 a death from tuberculosis. Comparative risk model assessed cause-specific mortality in patients
71 on highly active antiretroviral treatment (HAART) from tuberculosis, diarrhea, other
72 opportunistic infections, and unknown/unspecified causes (10–12). In patients who died from
73 opportunistic infections while on HAART, there are many international literatures that quantify
74 death from other causes (13–17).

75
76 However, relatively few researches were carried out in Ethiopia, for example, at the Gonder,
77 Dilla, and Pawai hospitals. Their attention was on HIV patients who were lost to follow-up who
78 presented competing risks. In this study, death-competing risks from tuberculosis, diarrhea, other
79 opportunistic infections, and unidentified infections are examined using national statistics. Such
80 research could be a useful contribution to the Ministry of Health's efforts to develop appropriate

81 intervention strategies and enhance patient care. Reduced mortality from AIDS-related
82 opportunistic infections is also very helpful for HIV-positive patients.

83
84 The Kaplan-Meier estimator and Cox proportional hazard model are the traditional classical
85 survival analysis techniques for time-to-event data (18, 19). These techniques work well in
86 investigations where there is just one primary event type of interest. The aforementioned
87 techniques might not fully reveal the relationship when there are several occurrences of interest.
88 Methods that can reveal the underlying connection between the variables, the intermediate
89 outcomes, and the outcomes of interest are required in these circumstances. To expand on tried-
90 and-true techniques, multistate models offer a versatile and expansive framework (20). The
91 competing risks model, which is a special case of a multi-state model and is a cause specific
92 hazard, sub-distribution hazard, and flexible parametric hazard model, was used in this study.
93 Because of this, the main goal of this article was to estimate the risk of death for HIV patients
94 using ART from tuberculosis, diarrhea, other opportunistic infections, and unidentified infections
95 in the lack of covariates and the existence of competing hazards. Finally, the cumulative
96 Incidence Function (CIF) was utilized to predict the likelihood of mortality and the cause
97 specific hazard, sub-distribution hazard, and flexible parametric hazard model were employed to
98 evaluate the covariate impacts on CIF.

99 **Material and Methods**

100 **Source population:** The source population of this study (record review) constituted all medical
101 records of HIV/AIDS patients who initiated and were enrolled on anti-retroviral therapy at all
102 health facilities in Ethiopia.

103 **Study population:** The study population is classified from health facility and individual
104 perspectives as follow:

105 **Inclusion criteria**

106 Health Facilities: Governmental, non-governmental, private or uniformed hospitals and health
107 centers that have been providing ART service for at least one year prior to the study period.

108 Records for review: records of individuals diagnosed with HIV/AIDS and at least one year since
109 they have initiated ART at time of data collection.

110 **Exclusion criteria**

111 Health centers having <100 patients whoever started ART were excluded so as to get adequate
112 records and be cost effective while moving to that site. Catchment Health center which are more
113 than 30 km from the mentoring hospital or Center (Zonal town).

114 **Sample size determination and sampling procedure**

115 **Sampling techniques**

116 Standard probability sampling technique was applied to select nationally representative sample
117 from the study population, and region was the domain of analysis. The primary sampling units
118 are ART patient charts in the hospitals and catchment health centers providing ART.

119 **Sampling**

120 Multi-stage clustered sampling method was used to select study facilities and participants. The
121 steps followed to select the regions, facilities, and allocate ART sample to each facility is
122 depicted in to three strata:

123 **Sample size determination**

124 This study was more or less a replica of the second round ART program effectiveness
125 assessment done back in 2011. Hence, the sample size was derived purposively by using the
126 cohort sampled population size from the previous study of around 40,000 ART patient charts to
127 be reviewed for comparison sake. Finally, at the end of the data collection we ended up
128 retrieving and reviewing a total of 39,590 medical records, which includes all 5,802 pediatric
129 ART patient records in the facilities.

130 **Data Collection Tools and Procedures**

131 The instrument measures age at baseline, gender, baseline WHO stage, baseline functional status,
132 and outcome (alive and on treatment, transferred out, lost to follow-up, known dead). Moreover,
133 achieved ART outcomes explained by quality of life such as CD4 count, weight, functional
134 status, and 7 years survival rates by strata and by different time points were also assessed.
135 Likewise, it also measured self-reported adherence to treatment and care, dropout rate, drug pick
136 up behavior, and retention on 1st line regimen. It detected problems in the flow of the questions,
137 gauge the length of time required for interviews, and identify problems in the understanding of
138 terms and concepts.

139 **Pretest:** The instrument was pretested and piloted during the training of the data collectors, and
140 the modification and amendment were done by all the trainees' feedback and valuable inputs. It
141 At National level, the NAPES team was basically structured into nine routes established based
142 on the main transportation roads in the country. In late- January, three additionally new sub
143 routes were added to assist the existing nine routes. Each route consisted on average of 5 - 7
144 hospitals, and the period of stay at each facility was ranging from 5 -10 days based on the
145 sampled ART patients/case load in that facility. The data collection comprised of gathering GPS
146 coordinates, facility level information by interviewing the facility head, and data extraction of
147 ART patient chart.

148 **Data Quality Assurance**

149 Quality assurance began with the recruitment of data collectors and team leaders with a health
150 background. Data collectors took pre- and post-test to assess their learning and knowledge of the
151 assessment guidelines and standards for data collection. An intensive 8 days training was given
152 to all field team. Each data collector was given a hard copy manual of the assessment guidelines
153 for reference in the field. The central team, performed data quality checks, followed by a data
154 quality control Skype call between the whole team to discuss any data collection or quality
155 issues. Also, based on teams' feedback, frequent modification of the Redcap tool was done to
156 make the tool user friendly and ensure quality.

157 **Data preparation**

158 **Predictors variables**

159 In this survey, a retrospective cohort study design was employed. It was a national study
160 involving nine regional states and two city administrations from November 2019 to March 2020.
161 The study participants were selected randomly from patient registers and medical data those
162 started ART. Age, weight, hemoglobin, sex, marital status, educational level, region, and type of
163 health facility were among the demographic data obtained. The retrospective record review also
164 revealed trends of ART treatment regimen change, adherence and TB status were also
165 determined from the retrospective record review also revealed. The patient has had tests for
166 opportunistic infections, TB status, liver enzyme, Hepatitis B and C, and family planning
167 counseling. Additionally, each visit includes a CD4 count and viral load measurements.

168 **Outcome variable**

169 The primary endpoint for this study was mortality from AIDS related tuberculosis, diarrhea,
170 other opportunistic infection and unknown/not specified causes. The study data collector
171 recorded the causes of death using tablet and in cases where inadequate information is provided
172 to determine the causes of death, the study coordinator contacts personnel at the study site for
173 further information. Patients are classified as having an unknown or not specified cause of death
174 if no further information is obtained.

175 **Multi-state models**

176 A multi-state model is a model for time-to-event data in which all individuals start in one or
177 possibly more starting states (e.g. HIV patient) and eventually may end up in one (or more)
178 absorbing or final state(s) (e.g. TB or Diarrhea or other Opportunistic infection disease or
179 unknown[1]. Some individuals are censored before they reach an absorbing state. Competing
180 risks models are a sub-category of multi-state models: they have one starting state, at least two

181 absorbing states and no intermediate states (see Fig. I) left panel [2]. Their transitions are often
182 indicated as causes of failure. An example of a multi-state model is illustrated in Fig I right panel
183 shown.

184 **Figure I. Title competing risks model with and death as causes of failure (left panel).** Fig I
185 legend Illness-death model with an initial state of illness state and a death state (right panel)

186 **The competing risks multistate model**

187 A special kind of multistate model called competing risk analysis seeks to accurately predict the
188 marginal probability of an event in the presence of competing events[3] When analyzing the
189 marginal probability for cause-specific occurrences, the Kaplan Meier product-limit technique
190 has a tendency to yield estimates that are erroneous since it is not designed to take into account
191 the competing nature of numerous causes to the same event[4]. Cumulative Incidence Function
192 (CIF) has been used to predict the chance of death due to competing risk. In this competing risk
193 analysis, the cause specific hazard model and the sub distribution hazard model were utilized to
194 analyses the covariate impacts on cause specific hazard[5]. A flexible parametric hazard model
195 was also used to determine the importance of covariates[6] and the outcomes of these three
196 strategies are then compared.

197 **3.1 Hazard-based regression models**

198 **Cause-specific hazard regression**

199 The most commonly used regression model for analyzing event-time data is the Cox proportional
200 hazards model[7]. In the presence of competing risks, the standard Cox proportional hazards
201 model is not adequate because the cause-specific Cox model treats competing risks of the event
202 of interest as censored observations [8] In addition, the cause-specific hazard function does not
203 have a direct interpretation in terms of survival probability[9]. An adaptation of Cox regression
204 requiring data augmentation has been proposed. With k competing events, the data for each
205 patient are duplicated k times, one row for each type of failure; then, k-1 indicator variables are
206 created for identifying whether a certain event has occurred. A stratified Cox regression could
207 also be applied to allow non-proportional hazards[10].

208
209 To make complete inference, we performed cause-specific hazard model and cause-specific
210 hazard functions [2, 11-13] are defined as:

$$211 \quad h_k(t) = \log_{\Delta t \rightarrow 0} \frac{p(t \leq T < t + \Delta t, K = k | T \geq t)}{\Delta t}$$

212 (1)

213 Where T be the survival time and K be the cause of death (k=1, 2, 3, 4).

214 The cumulative cause-specific hazard is written as:

215
$$H_q(t) = - \int_0^t h_q(u)du \quad (2)$$

216
$$S_q(t) = \exp (H_q(t)) \quad (3)$$

217 Here we assume that the event of interest is death of AIDS patient due to the opportunistic
 218 infection TB (indexed 1, 2, 3, 4), coded Tuberculosis (as the event of interest), diarrhea, other
 219 opportunistic infections and unknown causes respectively. Fig II tell us the model competing risk
 220 HIV/AIDS patients on ART.

221

222 **Figure II. Title Competing Risks Model for HIV/AIDS Patients on ART**

223 **Sub-distribution hazard regression—the Fine and Gray model**

224 The Fine-Gray[9] sub-distribution hazard model is frequently used in the tuberculosis related
 225 diarrhea, Other opportunistic infection and unknown infections cause death to estimate subject-
 226 specific probabilities of the occurrence of an event of interest over time in the presence of
 227 competing risks. The sub distribution hazard is a key concept in this approach, and it is defined
 228 as the hazard of failing from a given cause in the presence of competing events, given that a
 229 subject has survived or has already failed due to different causes. We can write the sub-
 230 distribution hazard[5] for cause as

231
$$\lambda_r(t) = \lim_{\Delta t \rightarrow 0} \frac{\Pr(t \leq T < t + \Delta t, R = r \mid T \geq t \cup (T \leq t \cap R \neq r))}{\Delta t} \quad (4)$$

232
$$= - \frac{d}{dt} \log (1 - I_r(t))$$

233 Where $I_r(t) = \Pr(T \leq t, R=r)$ is the CIF for cause r ($r=1, 2, 3$).

234 A semi parametric proportional hazards model for the sub-distribution hazard of cause r for a
 235 subject with covariate vector X as follows

236
$$\lambda_r(t/X) = \lambda_{r0}(t) \exp (\beta_r^T X) \quad (5)$$

237 where $\lambda_{r0}(t)$ is the baseline subdistribution hazard of cause r , and β_r is the vector of coefficients
 238 for the covariates. Estimation for this model follows the partial likelihood approach used in a
 239 standard Cox model.

240 **Flexible Parametric Proportional Hazard Model**

241 The flexible parametric model was first developed for use with censored survival data (6). The
 242 log cumulative hazard variant of a Weibull model was picked since it is the most prominent
 243 parametric survival model with both proportional hazards and accelerated failure time model
 244 interpretations. A Weibull distribution's survival function is

245
$$S(t) = \exp (- \lambda t^y) \quad (6)$$

246 Transforming this to the log cumulative hazard scale, we get

$$247 \quad \ln H(t) = \ln[-\ln(S(t))] \quad (7)$$

$$248 \quad \ln H(t) = \ln(\lambda) + \gamma \ln(t) \quad (8)$$

249 This is a linear function of log time, now adding covariates X in this model, we have

$$250 \quad \ln[(t/X)] = \ln(\lambda) + \gamma \ln(t) + X\beta \quad (9)$$

251 Where β are the covariate coefficients, but instead of assuming linearity with $\ln(t)$, adopt a
252 flexible parametric method that employs restricted cubic splines for $\ln(t)$ (14). Under the premise
253 of proportional hazards, covariate effects can be seen as log hazard ratios. Transforming the
254 model parameters yields the survival and hazard functions.

$$255 \quad \ln[H(t/X) = S(\ln(t)/\gamma, n_0) + X\beta] \quad (10)$$

256 One of the primary advantages of the flexible parametric technique is the ease with which time
257 dependent effects can be fitted.

258 **Result**

259 In the present study, we have considered 39590 HIV patients. Out of the total HIV patients, 1212
260 were died during the period of observation from November 2019 to March 2020. The
261 characteristics of these people as at the time of their death are shown in Table 1. Out of the total,
262 34.7% were died due to TB; 44.7% were died of other opportunistic infection; and the remaining
263 18.3% and 2.2 % were died of unknown/Not specified and diarrhea respectively (Table 1 and
264 Fig 1) . At the time of death, the mean age of the HIV patients was 35 ± 14.6 with mean weight of
265 48.7kg and 11.9 ± 6.3 Hemoglobin. The majority of deaths nearly 80% were between WHO stage
266 2 and stage 4 and 639(67.33%) were CD4 count cell and the remaining 19.9% were stage 1. The
267 majority (73.3%) was divorce or widowed and the remaining 18.3% were single. Similarly, the
268 mean weight was 48.7Kg and hemoglobin mean was 11.9 ± 6.31 .

269

270 **Table 1 : Title Characteristics of the respondents at death (N= 1212)**

271

272 **Fig 1: Title Death due to various opportunistic infection**

273 **Survival analysis**

274 Male patients had better survival than female HIV patients for the first 45 months and after 75
275 months than female but female were better than male 46-75 months during the follow-up periods
276 (Fig 2A). In figure 2B, patients had nearly high risk of probability of death due to other
277 Opportunistic Infection and Diarrhea in the first 30 months but probability of death due to
278 unknown were lower as compared to the others causes. After 40 months, Diarrhea had lower risk
279 as compared to the others opportunistic infection and unknown infections.

280

281 **Fig 2. Title Kaplan-Meier curve.** Fig 2 legend Fig.2A:Kaplan-Meier curve by male and
282 female and Fig.2B: Kaplan-Meier curve by Tuberculosis, Diarrhea other opportunistic disease
283 and unknown.

284 **Non-parametric comparison of cumulative incidence** 285 **function (CIFs)**

286 HIV/AIDS patients have higher risk of death from other opportunistic infection and followed by
287 death from TB, Unknown and Diarrhea causes. As the time goes from 6-50 months, Male
288 patients have slightly higher risk of death from Other Opportunistic Infection but from 50-
289 84month Female patients were higher risk of death as compared to male. But in the causes (OI
290 and TB) male had higher risk of death than female. Similarly, there is no significant difference in
291 mortality risk from all causes of death for male versus female patients, since p-values ($p= 0.48$,
292 0.18 , 0.21 , and 0.75 : TB, Diarrhea, other Opportunistic Infection and Unknown respectively)
293 (Fig 3A).

294 **Fig 3.Title Non-parametric estimates of cumulative incidence** functions Fig 3 legend
295 Fig3Adeaths from TB (status==1) Diarrhea (status=2), other Opportunistic Infection (status=3)
296 and unknown causes (status=4).Each outcome is compared between male and female patients.

297 **Fig 3B:** Estimate cumulative incidence in the presence of TB death, Diarrhea, Other OI &
298 Unknown

299

300 Fig 3B above showed that the death from tuberculosis, Diarrhea, other opportunistic infection
301 and unknown causes. From 6 month to 60 months, TB had higher risk of mortality than with the
302 other related death followed by other opportunistic infections and unknown infections but lower
303 death from diarrhea as compared to other infections. After 65 months, unknown infections had
304 higher risk death as compared to the other infections. This result indicated that the unknown
305 infection should be differentiated, listed out and given due attention by health professionals.
306 There is no significant difference in mortality risk from these causes of infections, since p-value
307 ($p= 0.025 < 0.05$).

308 **Cause-specific hazard regression**

309 Estimation of the risk of a male patient, age category 31-40 years and whose CD4 count is
310 between $201-350 \text{ cells/mm}^3$, the output shows that the cumulative incidences of death of this
311 patient from other OIs at time points of 30, 50, 80 months were 0.31,0.37 and 0.40, respectively.
312 The cumulative incidences of death of these patients from TB at the same time point were 0.28,
313 0.33 and 0.35. Unknown infection at time points of 30, 50, 80 months were 0.15, 0.19 and 0.20

314 respectively and diarrhea were 0.017, 0.019 and 0.020 respectively. Based on these numerical
315 figures we can conclude that death due to opportunistic infection is more likely happened than
316 the others causes.

317 **Sub-distribution hazards (SHs) model**

318 Sub-distribution hazards (SHs) model is also known as Fine-Gray model. It is a Cox proportional
319 regression model but the cumulative incidence is associated with SHs. The motivation for Fine-
320 Gray is that the effect of a covariate on cause-specific hazard function may be quite different
321 from that on Cumulative incidence function (CIF). In other words, a covariate may have strong
322 influence on cause-specific hazard function, but have no effect on CIF[2]. The difference
323 between cause-specific hazard and sub-distribution is that the competing risk events are treated
324 differently. The former considers competing risk events as non-informative censoring, whereas
325 the latter takes into account the informative censoring nature of the competing risk events[14].

326

327 **Prediction**

328 Here we have considered four hypothetical individuals and using certain covariate combinations
329 using the fitted Fine-Gray model selected based on the significance of the model. Age, sex, and
330 starting CD4 levels are defined for individuals as covariates. The first individual, female, 29
331 years old and CD4 count is ≥ 501 . Referring with Fig 4, she is belonging to younger age but
332 she is more likely to be died. This mean even if she had higher CD4 count, being a female was
333 high risk of death early the start of ART.

334

335 **Fig 4. Title Parametric estimates of cumulative incidence functions for three patients with**
336 **given covariate**

337

338 The second individual is a man who is 60 years old and his CD4 count is < 200 . This 60 year old
339 HIV patient has lower risk of death. This means, male patient has at lower risk than female
340 during the first 3 years of the follow-up time.

341

342 Fig5 graphical output gives CIF for patients with characteristics specified in extracted data had
343 three variables: Age, CD4 count, WHO clinical stage and Patient Residence. The link argument
344 controls the link function to be used, “prop” for the Fine-Gray regression model with values
345 age=52, CD4 count, WHO clinical stage and Residence compared in the four-graph panel with
346 four of causes: tuberculosis, other opportunistic infection and diarrhea; and unknown causes.
347 Then other opportunistic infection, tuberculosis, unknown and diarrhea causes were higher risk
348 of death followed respectively. when we compared in each cause, those patients had higher
349 WHO clinical stage higher risk as compare to other combination in the patient with the variable

350 combination due to tuberculosis causes and other opportunistic but in contrast with tuberculosis
351 cause the patient who had lower CD4 count high risk of death than the other combination. Blow
352 Fig 5, as the baseline cd4 count increases, then decreases significantly the risk of death in all
353 causes of death.

354 **Fig 5. Title Estimated CIF Fig 5 legend Figure 5A;TB, Figure 5B:Diarrhea, Figure5C:Other**
355 **Infection and Figure 5D;Unknow Causes of Deaths of HIV/AIDS Patients by CD4 cell count**
356 **using cause-specific hazard models.**

357

358 **Factors associated with opportunistic infections related mortality**

359 In this study, we estimated the cumulative incidence of death from tuberculosis, diarrhea, other
360 opportunistic diseases, and unknown infections using a cause-specific, sub-distribution, and
361 flexible parametric hazard model. In univariate analysis, prognostic factors with p-values less
362 than 0.2 were chosen, and the results are reported in Tables 2 & 3 below, and age group, patient
363 residence, WHO stage, and baseline CD4 count met the criteria for multivariate analysis. Age
364 group, WHO stage, residence and baseline CD4 count were statistically significant (P-
365 value<0.05), Participant with age group 21-30 years had 1.78, 1.43 and 2.02 times higher risk
366 of hazard of death due to tuberculosis causes as compared to age group <=20 years in the cause
367 specific, sub-distribution and flexible parametric hazard model [HR= 1.78 , 95% CI (1.24 -
368 2.541), SHR=1.43, 95% CI(.002 -1.98); HR=2.02, 95% CI(1.41-2.88)] respectively. But Death
369 due to diarrhea cause the patient age group 40-50 years less like as compared to age group <20
370 years with HR =0.24 and 0.23 which slightly similar in the two models cause specific and
371 flexible parametric hazard model [HR= 0.24 , 95% CI (0.06-0.93), HR=0.23, 95% CI(0.06-
372 0.92)] respectively but not significant in sub-distribution model. Since age is a confounding
373 variable and majority of the patients 56% were died at early of the follow up period (6 months).

374 The patient live in the rural area were higher risk death due to tuberculosis in the three models
375 and nearly similar hazard ratio HR=1.37, 1.37 & 1.29 as compared to patients live in the Urban
376 area in the cause specific, sub-distribution and flexible parametric hazard model [HR= 1.37, 95%
377 CI (1.11-1.68), HR=1.37, 95% CI(1.13-1.65) and HR=1.29, 95% CI(1.04-1.58)] respectively.
378 But patient live in rural area had 0.77, 0.76 and 0.78 less risk of death due to other opportunistic
379 infection in almost similar in three models as compared to patients live in the urban area in the
380 cause specific, sub-distribution and flexible parametric hazard respectively.

381 The cause specific, sub-distribution and flexible parametric hazard model an increase in baseline
382 CD4 count cells/mm³ is found to be strongly associated with the survival of HIV/AIDS patients
383 and it is found to be a significant factor for the probability of death due to tuberculosis cause (P-
384 value <0.05) .CD4 count category 351-500 count cells/mm³ less risk death due to tuberculosis
385 than CD4 count category <200 count cells/mm³ with HR=0.58 and HR=0.54 in cause specific
386 and flexible parametric hazard model (HR= 0.58, 95% CI (0.34, 0.98) and HR=0.54, 95%
387 CI(0.32-0.92)] respectively. But sub-distribution was not significant death due to tuberculosis,
388 diarrhea, other opportunistic infection and other unknown infection causes. WHO stage is found

389 to be a significant factor for the probability of death due to any cause in the cause specific model
390 (P-value <0.05).The result indicated that higher WHO stage nearly similar result in three models
391 in the cause specific, sub-distribution and flexible parametric hazard ratio more than 2 times
392 higher risk of death due to tuberculosis as compared to lower WHO stage. The reason WHO
393 clinical stage increases the risk of death in all causes, the patient immunology decreases and they
394 are exposed by opportunistic infectious

395

396 **Table 2: Result of Cause specific Hazard Model and Sub-distribution Hazard**
397 **Model**

398 **Table 3. Result of Flexible Parametric Proportional Hazard Model**

399 **Discussion**

400 In this study, we used a national dataset that included all regional states and two administrative
401 cities, and so it can give evidence to national health policy on the causes of mortality of
402 HIV/AIDS patients under the ART category. It could be considered as an extension from the
403 previous few studies already done in one region and health center from November 2019 to march
404 2020 patients under ART. We have discussed the three common methods for handling competing
405 risks and their applications to regression setting.

406 The cause specific hazard model, sub-distribution hazard model and flexible parametric hazard
407 model shows in the analysis output Table 2 & 3 produced nearly similar results with regard to
408 the effect of covariates. Tuberculosis is the most common opportunistic infection cause of death
409 among HIV/AIDS patients supported by previous studies [15-18]. During the period of
410 observation, 34.7% were died due to tuberculosis, 44.7% were died due to other opportunistic
411 infection but the others died due to diarrhea and unknown infection causes were 20.6%.This
412 study agree with the study[15] and higher the study conducted in France[19]. According to all
413 the methodologies, age group, CD4 count and WHO stage were significant indicators for
414 HIV/AIDS patients.

415

416 Male patients had better survival than female HIV patients the start of ART treatment than
417 female but female were better than male after 3 years of treatment periods [15]. Patient had
418 nearly high risk of probability of death due to other opportunistic infection and Diarrhea in the
419 first 2 years but probability of death due to unknown were lower as compared to the others
420 causes. After 3 years, diarrhea had lower risk as compared to the others opportunistic infection
421 and unknown infections and in agreement with the study[20] but in contrast with the study[21].

422

423 According to all the sub-distribution hazard models, flexible parametric and cause specific
424 hazard models approaches, CD4 cells count is a key indicator for HIV/AIDS patients. Cause

425 specific, sub-distribution and flexible parametric hazard ratios for tuberculosis deaths were found
426 that the higher CD4 cell count cells/mm³ patients were less risk of death due to tuberculosis
427 compared to the patients with lower CD4 cell count cells/mm³. This study was lower result than
428 with the studies [15, 17].

429 The patient with higher age group were higher risk of death as compared to patient with lower
430 age group and in agreement with the study [22]but death due to diarrhea were less as compared
431 to lower age group this in agreement. This could be the cofounding effects of age.

432 The results of cause specific, sub-distribution and flexible parametric hazard model are found to
433 be more than 2 times higher risk death in tuberculosis, diarrhea and opportunistic infection
434 patient in ART treatment in WHO stage IV as compared to lower WHO stage I lower with the
435 studies done previously[21]. Because of this, it may be wise on the side of treatment and care
436 providers to focus on early detection of opportunistic infection and fast action should be taken to
437 reduce morbidity and improve quality of life for persons with HIV/AIDS. Early detection and
438 treatment of tuberculosis in HIV-infected people need to be prioritized. The prediction of patient
439 risk with fewer combinations of variables weight, age and sex, CD4 cell count had lower risk as
440 the patient had increase cd4 count.

441
442 There are several important benefits to this study. To our knowledge, this study is the first in
443 Ethiopia compared death risks due to tuberculosis and other opportunistic infection in Ethiopian
444 HIV/AIDS positive individuals. The overestimation of cumulative incidence and bias in the
445 covariate effects might both be avoided. Second, our study had a sizeable sample of 39590
446 participants from across the country and a protracted follow-up period of (84 months). Last but
447 not least, we looked into the connection between the time between starting ART and mortality in
448 HIV/AIDS patients with high and low CD4+ cell counts.

449 According to the current study findings, late HIV/AIDS diagnosis at an older age, in stages 3 or
450 4, or with a lower CD4 count may raise the risk of death considerably. These findings were
451 consistent with those of other investigations[15, 21].

452 This study has some limitations, the dynamic fluctuations in CD4+ cell counts over time and
453 causes of death record not appropriate this could affect to see the survival of HIV/AIDS patients.
454 Due to the study's record review methodology, some variables including height and hemoglobin
455 were missing. Due to the incompleteness of the data, BMI and anemia status were excluded from
456 the current analysis.

457 **Conclusion**

458 The results of this study clearly shown that tuberculosis, diarrhea, opportunistic infection and
459 unknown cause are substantially accelerating death for HIV/AIDS patients in Ethiopia. Applying
460 effective strategies is needed to achieve on time diagnosis of individuals with HIV and provide
461 them with HIV/AIDS care and treatment services to enhance the survival of the patients.
462 Moreover, female younger age living with HIV/AIDS requires more attention when receiving

463 HIV/AIDS care and treatment as they experience a higher risk of death the first 3 years after
464 ART initiation. The findings of the current research revealed that early HIV/AIDS care and
465 treatment could substantially reduce opportunistic infection death among HIV/AIDS patients.
466 Therefore, to reduce mortality rate HIV/AIDS patient, the current strategies should be revised to
467 improve the timing of treatment initiation and also to optimize the adherence to the treatment.

468

469 **Abbreviations**

470	ART	Antiretroviral therapy;
471	HAART	Highly active antiretroviral therapy;
472	AHRI	Armauer Hansen Research Institute
473	AIDS	Acquired Immunodeficiency Syndrome
474	CD4	Cluster of Differentiation 4 or Classification Determinant
475	FHAPCO	Federal HIV/AIDS Prevention & Control Office
476	HIV	Human Immunodeficiency Virus
477	NAPES	National ART Program Effectiveness Study
478	IRB	Institutional Review Board

479 **Ethics approval and consent to participate**

480 Ethical clearance was initially obtained back in June/July 2019 from ALERT/Armauer Hansen's
481 Research Institute's (AHRI's) IRBs [P048/18]. The need for consent was waived by the above
482 mentioned ethics review committee as the study is based on a secondary data(record review).
483 Consecutively, before dispatching the team to the different regions, a formal official approval
484 and support letter describing the purposes of the study was written from Federal MoH to the
485 respective Regional health bureaus and in turn to each zonal offices and hospitals and then from
486 Zonal Health Bureaus to the Woreda Health Offices before reaching the health facilities.

487 **Consent for participation (informed consent)**

488 Since the study design was retrospective and the data obtained from record review and informed
489 consent is not applicable (N/A).

490 **Availability of data and materials**

491 The data sets analyzed in this study available from the corresponding author on reasonable
492 request. The R code used to analyze the data provided as a supplement of the article.

493

494 **Competing interests**

495 The authors declare that they have no competing interests.

496

497 **Funding**

498 N/A

499

500 **Authors' contributions**

501 Tsegaye H Kumsa and Zeytu G Asfaw generated the idea, the corresponding author THK did the
502 data analysis and interpretation and ZGA played an advisory role. THK and Andargachew M ulu
503 (AM) have generated the data and Andargachew Mulu were the project investigator. Adane
504 Mihret reviewed the manuscript. All authors read and approved the final manuscript.

505 Acknowledgements

506
507 We would like to also express our gratitude to MoH, Regional, Zonal and Distract Health
508 Authorities for the administrative effort made for the data collectors. We are also thankful for the
509 collaboration effort both from Armauer Hansen Research Institute, Federal HAPCO and Ministry
510 of Health and Regional and City Administration Health Bureaus. We thank also the data
511 management team, finance and procurement, general service and the entire AHRI top and middle
512 level management team for the integrated service and leadership that they provided for the study
513 team during the data collection

514 Authors' information (optional)

515 Reference

- 516 1. Andersen, P.K., S.Z. Abildstrom, and S. Rosthøj, *Competing risks as a multi-state model*.
517 Statistical methods in medical research, 2002. **11**(2): p. 203-215.
- 518 2. Putter, H., M. Fiocco, and R.B. Geskus, *Tutorial in biostatistics: competing risks and multi-state*
519 *models*. Statistics in medicine, 2007. **26**(11): p. 2389-2430.
- 520 3. Andersen, P., *Competing risks as a multi-state model*. journals.sagepub.com, 2002.
- 521 4. Ma, S., et al., *Predicting the HIV/AIDS epidemic and measuring the effect of AIDS Conquering*
522 *Project in Guangxi Zhuang Autonomous Region*. Plos one, 2022. **17**(7): p. e0270525.
- 523 5. Gray, R.J., *A class of K-sample tests for comparing the cumulative incidence of a competing risk*.
524 The Annals of statistics, 1988: p. 1141-1154.
- 525 6. P, R., *Flexible parametric proportional hazards and proportional odds models for censored*
526 *survival data, with application to prognostic modeling and estimation of treatment effects*. Stats
527 Med, 2002. **21**; **2175-2197**.
- 528 7. Cox, D.R., *Regression models and life-tables*. Journal of the Royal Statistical Society: Series B
529 (Methodological), 1972. **34**(2): p. 187-202.
- 530 8. Kalbfleisch, J.D. and R.L. Prentice, *The statistical analysis of failure time data*. 2011: John Wiley &
531 Sons.
- 532 9. Fine, J.P. and R.J. Gray, *A proportional hazards model for the subdistribution of a competing risk*.
533 Journal of the American statistical association, 1999. **94**(446): p. 496-509.
- 534 10. Lunn, M. and D. McNeil, *Applying Cox regression to competing risks*. Biometrics, 1995: p. 524-
535 532.
- 536 11. Kalbfleisch, J. and R. Prentice, *The Statistical Analysis of Failure Time Data, Hoboken*. Wiley.
537 Semiparametric estimation of random effects using the Cox model based on the EM algorithm.
538 Biometrics, 2002. **48**: p. 795-806.

- 539 12. Tsiatis, A., *Competing risks. Encyclopedia of biostatistics*. 1999, New York: Wiley.
- 540 13. Prentice, R.L., et al., *The analysis of failure times in the presence of competing risks*. Biometrics,
541 1978: p. 541-554.
- 542 14. Scrucca L, *Regression modeling of*
543 *competing risk using R: an in depth guide for clinicians*.
544 *Bone Marrow Transplant*. 2010. **45:1388-95**.
- 545 15. Grover, G., P.K. Swain, and V. Ravi, *A competing risk approach with censoring to estimate the*
546 *probability of death of HIV/AIDS patients on antiretroviral therapy in the presence of covariates*.
547 *Statistics Research Letters Journal (SRLJ)*, 2014. **3(1)**: p. 7-16.
- 548 16. Takalkar, A.A., et al., *Study of opportunistic infections in HIV seropositive patients admitted to*
549 *Community Care Centre (CCC), KIMS Narketpally*. Biomed Res, 2012. **23(1)**: p. 139-142.
- 550 17. Ghate, M., et al., *Incidence of common opportunistic infections in HIV-infected individuals in*
551 *Pune, India: analysis by stages of immunosuppression represented by CD4 counts*. International
552 *Journal of Infectious Diseases*, 2009. **13(1)**: p. e1-e8.
- 553 18. Srirangaraj, S. and D. Venkatesha, *Opportunistic infections in relation to antiretroviral status*
554 *among AIDS patients from south India*. Indian Journal of medical microbiology, 2011. **29(4)**: p.
555 395-400.
- 556 19. Bonnet, F., et al., *Opportunistic infections as causes of death in HIV-infected patients in the*
557 *HAART era in France*. Scandinavian journal of infectious diseases, 2005. **37(6-7)**: p. 482-487.
- 558 20. Jung, I.Y., et al., *Trends in mortality among ART-treated HIV-infected adults in the Asia-Pacific*
559 *region between 1999 and 2017: results from the TREAT Asia HIV Observational Database*
560 *(TAHOD) and Australian HIV Observational Database (AHOD) of IeDEA Asia-Pacific*. Journal of
561 the International AIDS Society, 2019. **22(1)**: p. e25219.
- 562 21. Gheibi, Z., et al., *Determinants of AIDS and non-AIDS related mortality among people living with*
563 *HIV in Shiraz, southern Iran: a 20-year retrospective follow-up study*. BMC Infectious Diseases,
564 2019. **19**: p. 1-11.
- 565 22. Teklu, A.M., et al., *Factors associated with mortality of TB/HIV co-infected patients in Ethiopia*.
566 *Ethiopian journal of health sciences*, 2017. **27(1)**: p. 29-38.
- 567

medRxiv preprint doi: <https://doi.org/10.1101/2023.10.25.23297525>; this version posted October 25, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a [CC-BY 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

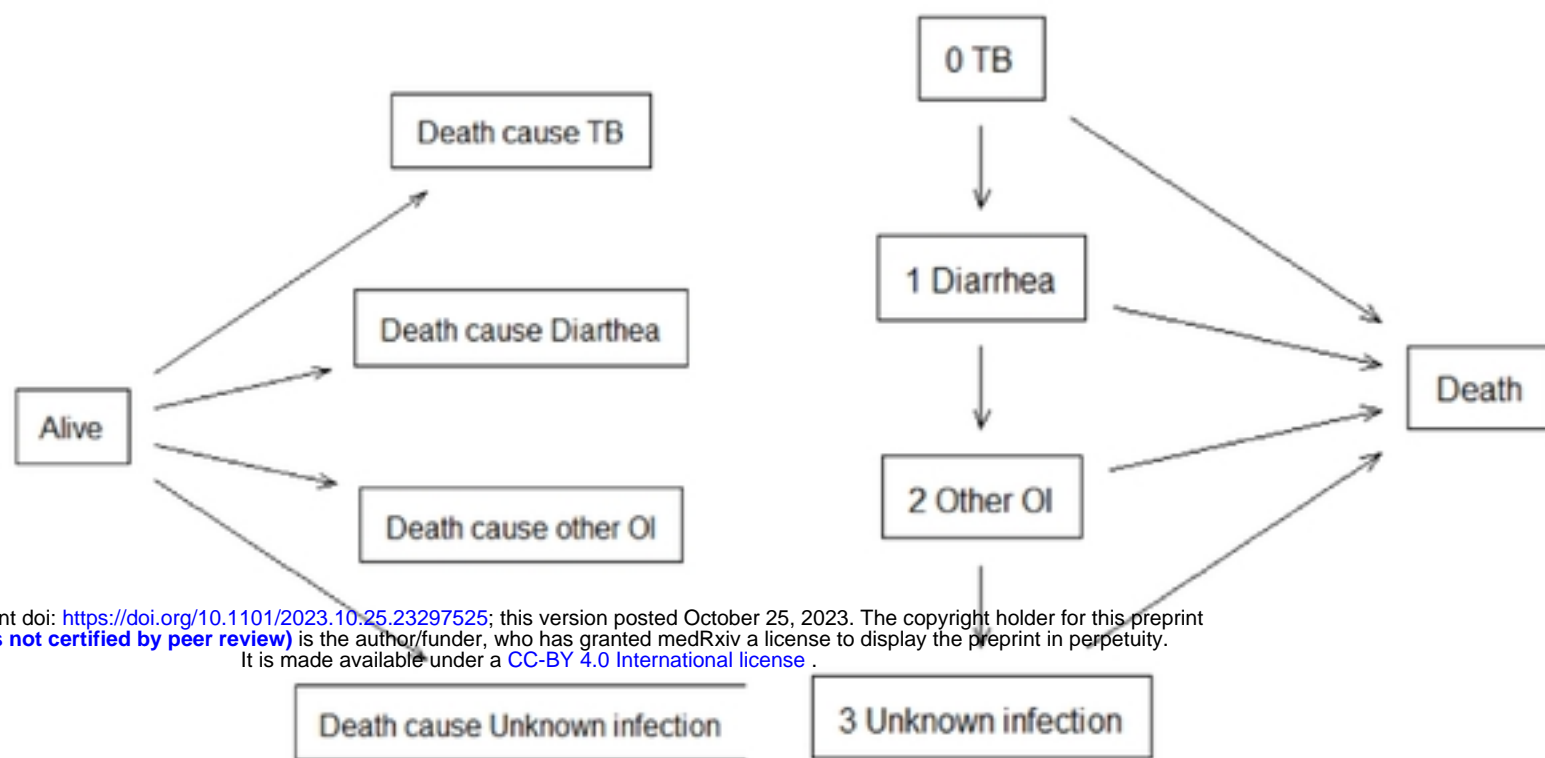


Figure I. Title Competing risks model with and death as causes of failure (left panel). FigI legend Illness-death model with an initial state of illness state and a death state (right panel)

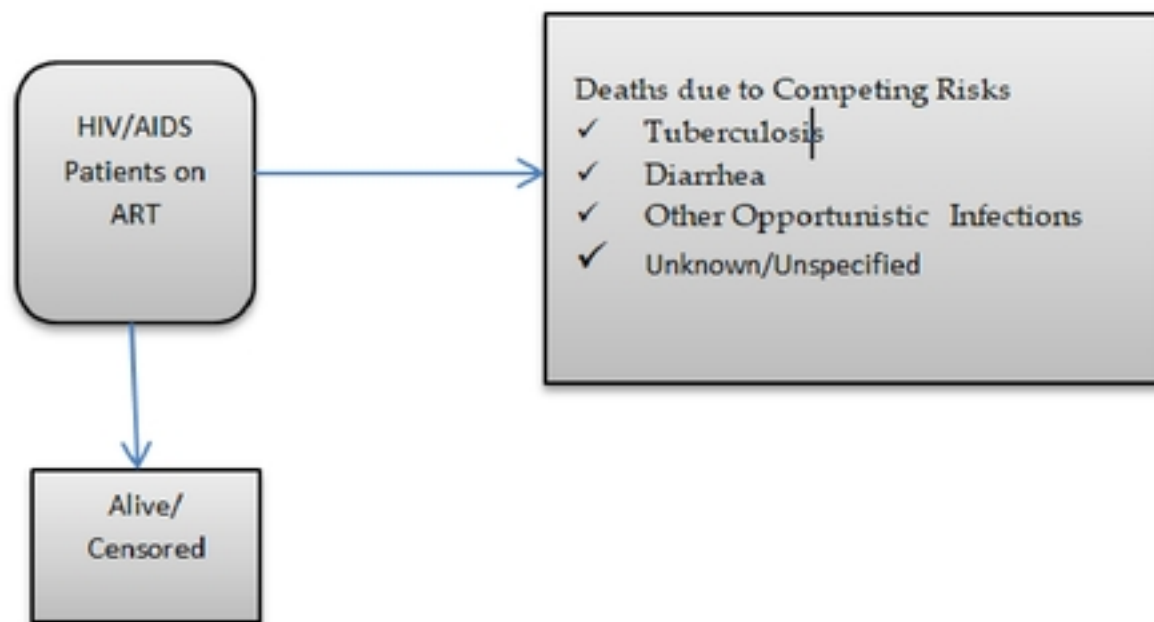
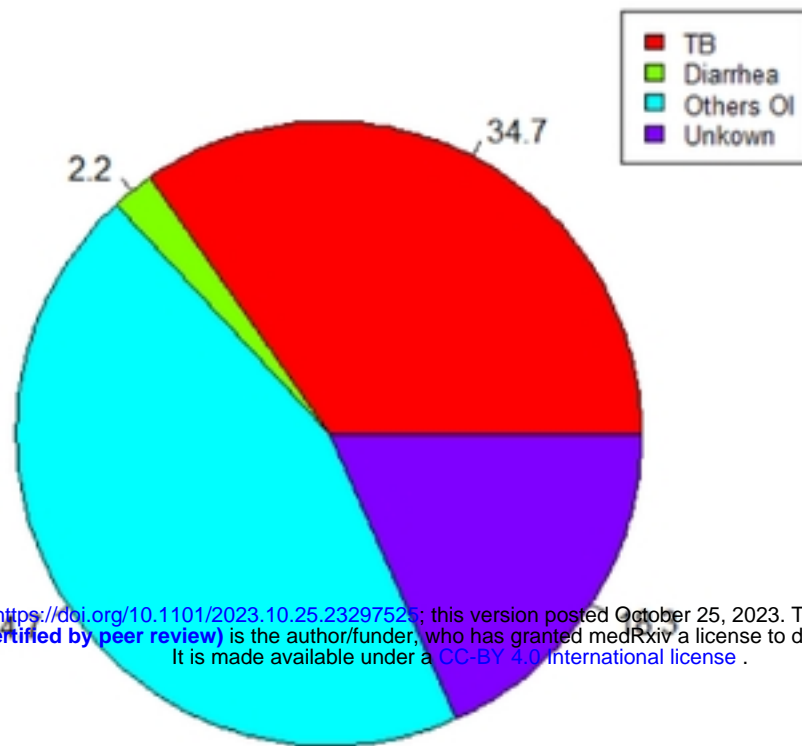


Fig II. Competing Risks Model for HIV/AIDS Patients on ART

Death due to Various Opportunistic Infection



medRxiv preprint doi: <https://doi.org/10.1101/2023.10.25.23297524>; this version posted October 25, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a [CC-BY 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

Fig1: Death due to various opportunistic infection

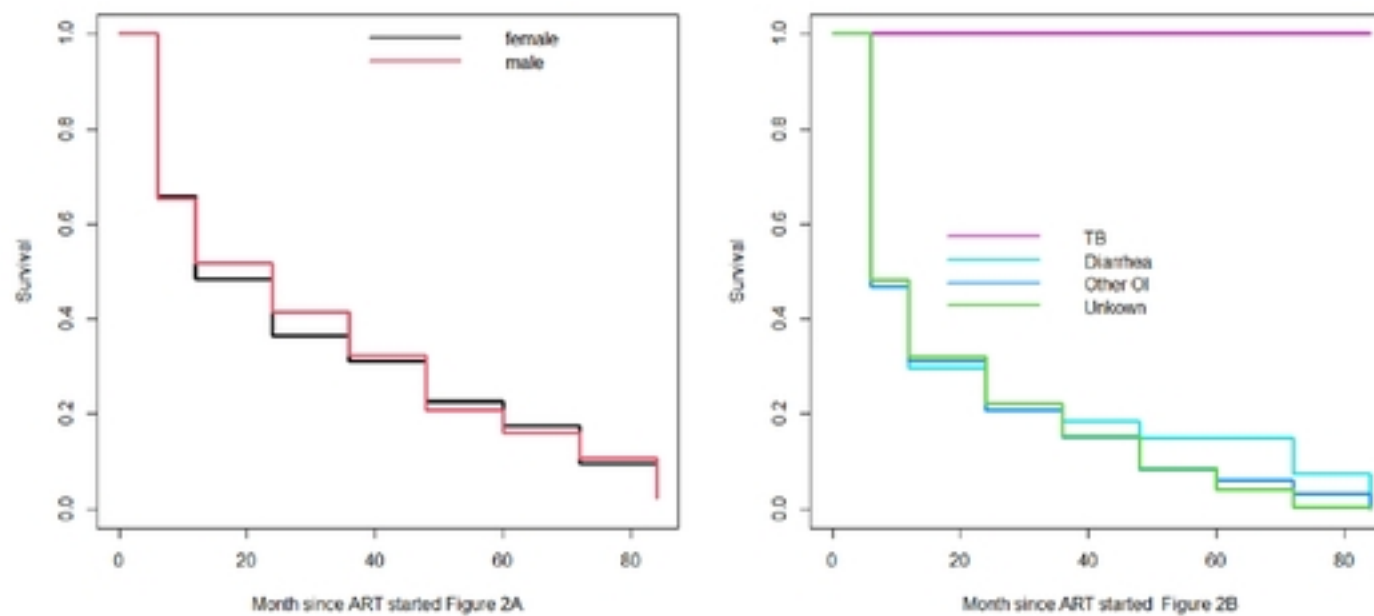
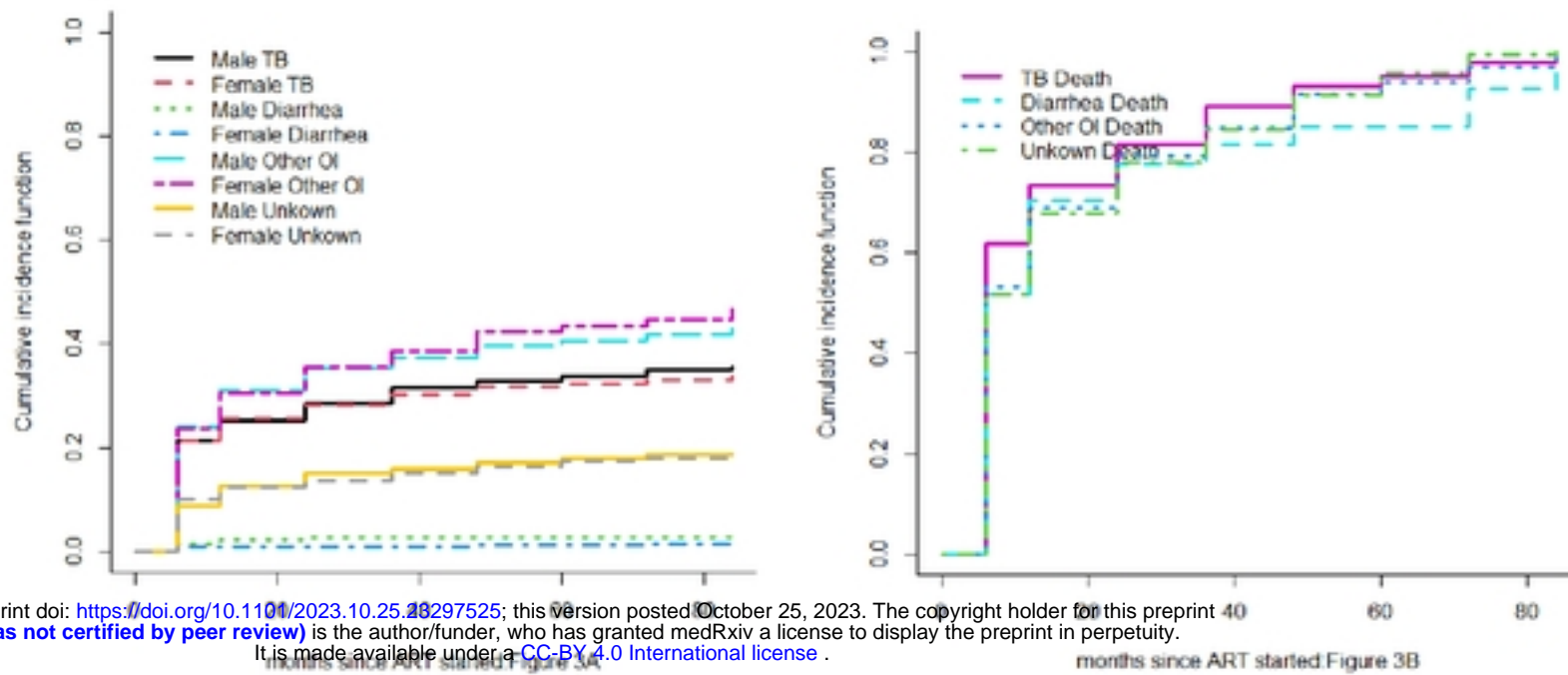


Fig2. Title Kaplan-Meier curve Fig 2 Legend Fig 2A:Kaplan-Meier curve by male and female and Fig 2B: Kaplan-Meier curve by Tuberculosis, Diarrhea other opportunistic disease and unknown.



medRxiv preprint doi: <https://doi.org/10.1101/2023.10.25.20297525>; this version posted October 25, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a [CC-BY 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

Fig 3 Title Non-parametric estimates of cumulative incidence functions **Fig3A:** Non-parametric estimates of cumulative incidence functions for deaths from TB (status==1) Diarrhea (status=2), other Opportunistic Infection (status=3) and unknown causes (status=4).Each outcome is compared between male and female patients. **Fig 3B:** Estimate cumulative incidence in the presence of TB death, Diarrhea, Other OI & Unkown

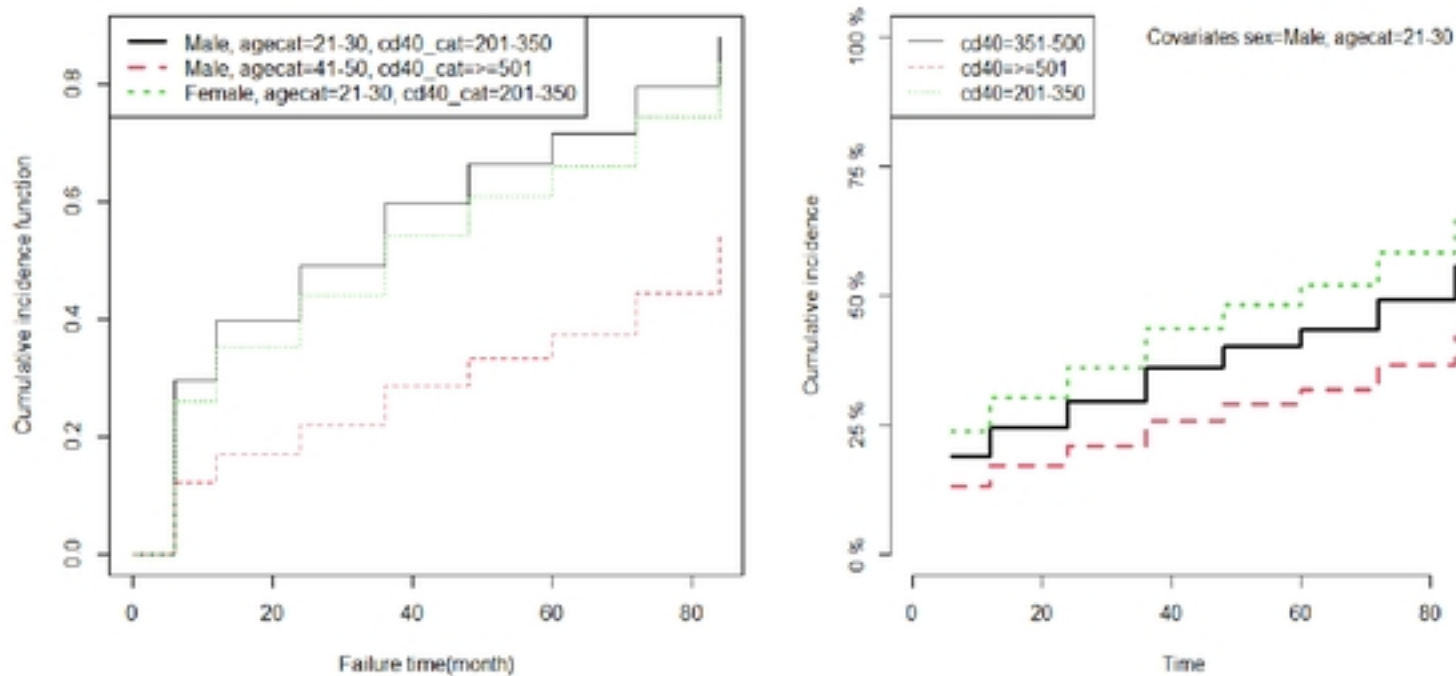
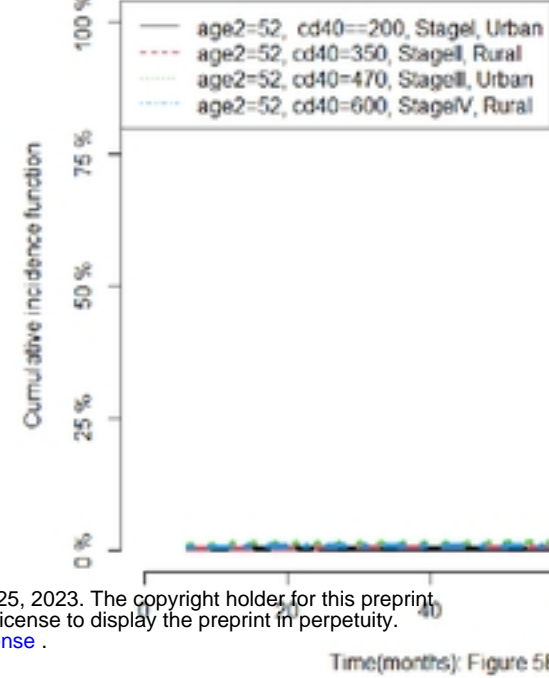
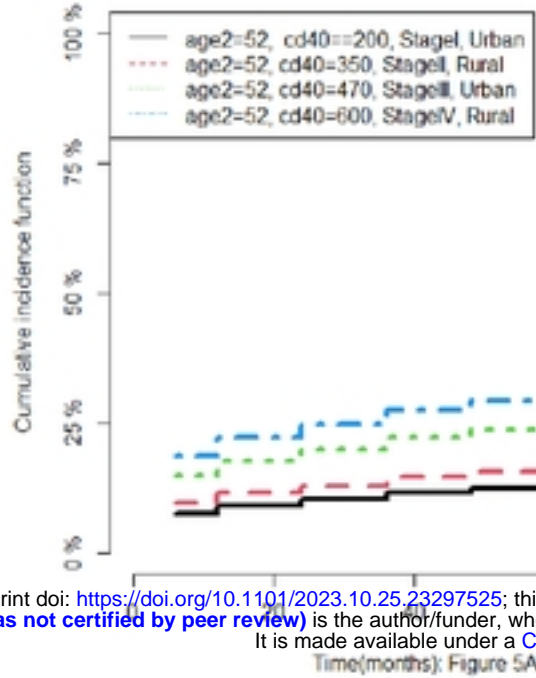


Fig4. Title Parametric estimates of cumulative incidence functions for three patients with given covariate



medRxiv preprint doi: <https://doi.org/10.1101/2023.10.25.23297525>; this version posted October 25, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a [CC-BY 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

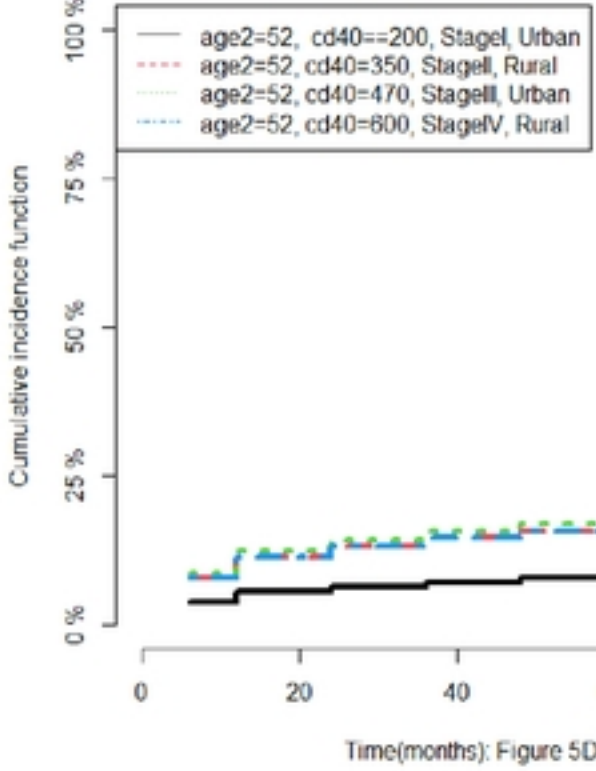
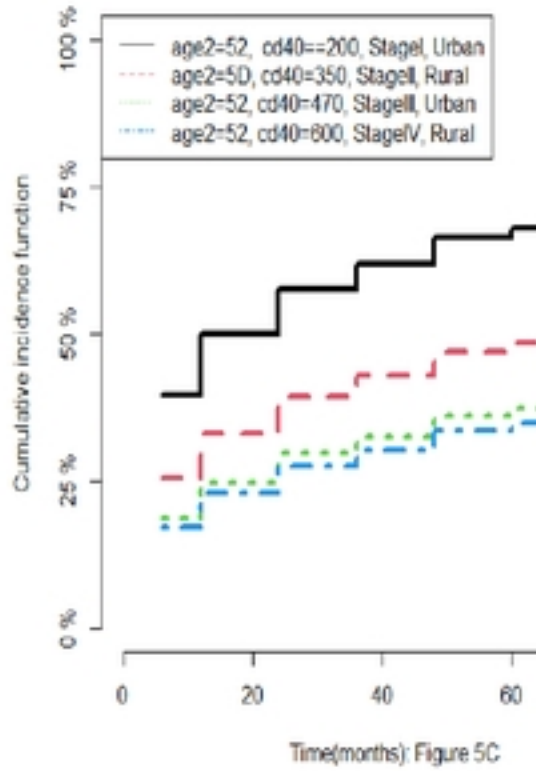


Fig 5. Title Estimated CIF. Fig 5 legend Fig.5A;TB, Fig.5B:Diarrhea, Fig.5C:Other Infection and Fig.5D;Unknow Causes of Deaths of HIV/AIDS Patients by CD4 cell count using cause-specific hazard models.