Competing Risks Model to Evaluate Dropout Dynamics Among the Type 1 Diabetes Patients Registered with the Changing Diabetes in Children (CDiC) Program

Noora Al-Shanfari, Ronald Wesonga, Amadou Sarr and M. Mazharul Islam*

Department of Statistics, College of Science, Sultan Qaboos University, Oman

Abstract: Understanding the survival dynamics of registered patients on a disease control program is a vital issue for the success of program objectives. Dropout of registered patients from such a program is a critical issue, hindering the effectiveness of the program. This study aimed to identify the risk factors of dropout of patients who were registered on the Changing Diabetes in Children (CDiC) program, taking a case of Uganda. Survival analysis was done by integrating competing risk of factors associated with attrition from the CDiC program. The data for the study was obtained from patients with type 1 diabetes mellitus (T1DM) registered during 2009-2018 at health units with specialized pediatric diabetes clinics from various regions in Uganda. The study considered follow-up data of 1132 children with T1DM. Our analysis revealed that the Body Mass Index (BMI) significantly influences dropout time, with patients classified as underweight showing higher hazards than those with normal BMI. Moreover, when considering competing risks, dropout hazards increased. Comparing the Cox model with the Fine and Gray model shows the latter exhibiting a smaller AIC value, which indicates its superiority in the time-to-dropout analysis. Thus, utilizing methods that integrate competing risks for CDiC dropout analysis is preferable and recommended for related studies. These findings provide actionable insights for enhancing CDiC program efficacy.

Keywords: Competing risks, Time to dropout, CDiC program, Subdistribution hazard, Fine-Gray model, Cox model.

1. INTRODUCTION

A great amount of investment has been made in the health systems to prolong the lives of patients with non-communicable chronic diseases. The non-communicable chronic diseases (NCDs), including diabetes, are the leading causes of death worldwide and represent an emerging global health burden [1]. According to the International Diabetic Federation (IDF), 537 million people had diabetes globally in 2021, and the number is expected to grow to 643 million by 2030, and 783 million by 2045 [2]. Unfortunately, the majority of these diabetes patients are living in low-and middle-income countries, and 1.5 million deaths are directly attributed to diabetes each year [3]. There is an increasing trend in the prevalence and incidence of both type 1 and type 2 diabetes globally. Type 1 diabetes mellitus (T1DM), once known as juvenile diabetes or insulin-dependent diabetes, is a chronic condition in which the pancreas produces little or no insulin by itself. Of the total population with T1DM in 2022, 1.52 million (17.0%) were younger than 20 years [4]. There is no prevention or cure for type 1 diabetes, and as such, it is a lifelong condition that requires daily management and care, as well as a high level of patient knowledge.

The increased prevalence and incidences of T1DM, especially among children and adolescents, has led to specialized T1DM Clinics in some developing countries. A number of programs have been undertaken to follow up with patients to control or help cure the disease in different countries. The Changing Diabetes in Children (CDiC) program is one example of such program. The CDiC is a global initiative led by Novo Nordisk to address childhood diabetes worldwide [5]. Novo Nordisk is a Danish multinational pharmaceutical company headquartered in Bagsværd with production facilities in nine countries and affiliates or offices in five countries. Currently, the CDiC program is implemented in 14 African and Southeast Asian countries. In this program, all children with diabetes below 18 years of age are encouraged to enroll in the program by attending the nearest established Type1 Diabetes Mellitus (T1DM) clinic. In Uganda, for example, such a program started in the year 2009. For optimal program benefits, patients must remain engaged for the expected duration. Unfortunately, some patients either drop out or become inactive, failing to attend scheduled appointments with CDiC program [6]. Therefore, understanding the dropout dynamics of T1DM patients and assessing the survival time from treatment initiation to dropout, as well as the risk factors for dropout, are essential for designing time-relevant intervention strategies and thus improving the effectiveness of the program.

In survival analysis, various methods are available to examine data sets defined in terms of the time from a well-defined time origin to the occurrence of a particular event, depending on different risks and the survival time [7]. The risk measure of interest could be the survival time, the time lapse between the registration on the program, time to dropout, time to lost to follow up and the event such as death,

*Address correspondence to this author at the Department of Statistics, College of Science, Sultan Qaboos University, P.O. Box 36, PC 123, Muscat, Oman; Tel: +968 98264437; E-mail: mislam@squ.edu.om

E-ISSN: 1929-6029/24

recurrence, or treatment failure. For example, an event of specific interest to this study is the dropping out of a T1DM patient from the CDiC program. In survival analysis, subjects who do not experience the event of interest during the follow-up period are considered censored data. For example, to estimate the survival time for patients with T1DM, patients are followed from a baseline date (such as the date of diagnosis or registration) until the date of death or study closing date, a patient who dies during the follow-up period or lost to follow-up would be considered censored data. In this domain, the Kaplan-Meier method and Cox proportional hazards models are two widely utilized statistical techniques, each offering distinct advantages in analyzing survival data. The Kaplan-Meier is a nonparametric estimator used to estimate the survival function, whereas the Cox proportional hazards model is a semi-parametric regression model used to assess the association between covariates and the hazard rate. Moreover, the log-rank test is often conducted before employing the Cox proportional hazards model. Renowned for its simplicity and capability to discern differences in survival trends among groups, this test is instrumental in uncovering factors influencing survival duration. When applying the Kaplan-Meier method or Cox model, it is assumed that censoring occurs independently of other factors. This means that a patient's survival time is presumed to be unaffected by the reasons for censoring [8]. However, in reality, patients may encounter various outcomes that prevent the occurrence of the event of interest. Consequently, the assumption of independent censoring is violated. For instance, when studying mortality associated with T1DM, patients who die from a different cause will not die due to T1DM. Such events that prevent the event of interest from occurring are known as competing risk events. When competing risks are present, the Kaplan-Meier, log-rank test, and Cox proportional hazard model aren't suitable. Ignoring these competing risks can result in an overestimation of cumulative incidence and predicted risk [9-11].

The determinants of time to dropout from CDiC program can be examined by considering death event as a censored observation in survival analysis. However, T1DM patients admitted to the CDiC program are followed to achieve the desired result and death appears as a competing risk event when dropout is the primary event. In this situation, standard survival methods such as KM method or Cox model are inadequate to analyze survival data in the presence of competing risk [12,13]. Death from any cause in T1DM patients precludes observation of dropout from treatment and thus a competing of dropout. A major limitation of modeling time-to-event data in the presence of competing events is that when estimating regression parameters under a specific cause, individuals failing from causes other than the cause of interest are considered censored observations. In the presence of competing risks, the cumulative incidence function (CIF) is the most widely used technique, which estimates the marginal probability for each competing event [14]. We utilize the CIF to calculate the total occurrence of the event of interest. The CIF for a specific event indicates the probability of experiencing that event before a certain time and before the occurrence of another type of event. In survival analysis, when there are no competing risks, the hazard function is a fundamental concept. It indicates the instantaneous rate at which the event of interest occurs in individuals who remain at risk of the event. While the regression coefficient of the Cox model explains the proportional impact of covariates on the event's hazard, this relationship also holds in the absence of competing risks [13]:

$$S(t) = S_0(t)\exp(X^T\beta)$$

where S(t) represents the survival function for an individual with a given set of covariates, X, and $S_0(t)$ signifies the baseline survival function (i.e., the survival function for individuals with covariates equal to zero). Consequently, the impact of a covariate on the event's hazard is equivalent to its impact on the natural logarithm of the survival function [13,15]. In scenarios where competing risks are present, two distinct types of hazard functions emerge: the cause-specific hazard function and the subdistribution hazard function. The cause-specific hazard function reflects the instantaneous rate at which the event of interest occurs among individuals who have not encountered any other type of event. On the other hand, the subdistribution hazard function indicates the instantaneous risk of experiencing the event of interest among individuals who have not yet encountered such an event. Notably, this risk set encompasses individuals who are currently free from any events as well as those who have previously experienced a competing event [15]. By modeling both hazard functions, competing risks are appropriately accounted for. However, modeling the subdistribution hazard function explicitly establishes the relationship between the subdistribution hazard and its impact on the incidence of an event. Consequently, this approach enables the estimation of the effect of covariates on the Cumulative Incidence Function (CIF) for the event of interest [15]. Therefore, when addressing prognostic inquiries in the presence of competing risks, the subdistribution hazard model, also known as the Fine and Gray model, is employed.

In recent studies, researchers have shown interest in pinpointing factors linked to patient attrition or loss to follow-up. Some have employed conventional analyses

[16-18], while others have explored competing risk analyses [19,20]. To investigate factors affecting participation duration in the CDiC program in Uganda, Wesonga *et al*. (2023) utilized the Kaplan-Meier method to estimate survival duration within the program. Additionally, they employed the Cox model to identify contributing factors. However, it's important to note that patients who pass away during the follow-up period are considered dropouts, thus violating the assumption of independent censoring. It's essential to recognize that patients who drop out and those who pass away do not experience the same likelihood of censoring [21]. Therefore, incorporating competing risk methods becomes crucial, as it accounts for varying probabilities of censoring and ensures accurate estimation by considering the weights associated with different outcomes.

In this study, firstly, we sought to investigate the factors influencing the duration of participation in the CDiC program, considering both traditional survival analysis methods and competing risk analysis. Secondly, we examined how various factors, including demographic characteristics and disease severity, affect the time until dropout from the CDiC program, and how this relationship changes when accounting for competing risks such as mortality. Lastly, we determined a better statistical approach between the traditional survival analysis methods and the competing risk analysis that could offer the most accurate and interpretable analysis of time to dropout from the CDiC program. Through these inquiries, we generated insights that can enhance decision-making processes and ultimately improve the effectiveness of the CDiC program in Uganda, and wherever it may be implemented. The specific objectives of the study were to:

1. Analyze the time-to-dropout from the CDiC program among patients with T1DM, considering both methods; with and without competing risks, such as death.

2. Identify and assess the factors influencing the dropout time of the CDiC program.

3. Estimate the impact of the identified factors on the dropout rates, accounting for both the absence and presence of competing risks.

4. Evaluate and determine the most appropriate approach for analyzing dropout time of the CDiC program, considering factors such as statistical robustness and interpretability.

In brief, this research holds importance as it addresses a critical problem in the analysis of CDiC program data and offers insights that could lead to more accurate, informed decision-making and ultimately improve the effectiveness of the CDiC program.

2. MATERIALS AND METHODS

2.1. Study Design

The data used in this study were derived from patients with T1DM registered at health units with specialized pediatric diabetes clinics from various regions in Uganda, which are supported by the CDiC project [6]. Two main hospitals, Nsambya and Mulago, were used as key focal patient tracking health centers. All patients registered during the period beginning from January 2009 to the end of June 2018 were included in the study. Patients who had been seen once in the clinic and had no single HbA1C recorded were excluded. Additionally, patients with a program survival time exceeding nine years were considered outliers and were therefore removed from the dataset. The Total dataset after cleaning was 1132 patients. Each patient was categorized as "Active," "Inactive," "Lost to Follow-up," "Drop Out," or "Dead." An "Inactive" status indicates that a patient missed two consecutive appointments but was still alive. A patient labeled as "Lost to Follow-up" meant that they could not be located, and their vital status was uncertain. For our analysis, we treated patients with "Inactive" or "Lost to Follow-up" status as dropouts from the program, considering it an event of interest. At the same time, death was regarded as a competing event. To evaluate the influence of covariates on the time to dropout from the program, variables such as place of residence, gender, body mass index, and hypertension level were included in the analysis.

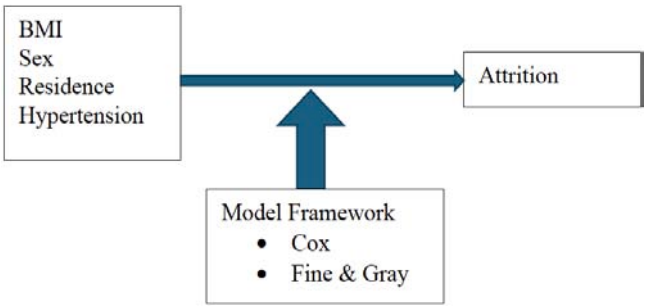

Figure 1: Conceptual framework for the study.

2.2. Statistical Analysis

In this study, we initially determined the median dropout time. When competing risks are not considered, individuals who dropped out, became inactive, passed away, or were lost to follow-up were all classified as dropouts, and the median for the dropout time was calculated accordingly. However, when competing

risks were considered, patients who dropped out, became inactive, or were lost to follow-up were considered as dropouts, and their median time was calculated. To identify factors affecting dropout time, we utilized univariable methods. In the absence of competing risks, we employed the log-rank test to identify the significant factors, whereas in the presence of competing risks, we used the Gray test. Subsequently, significant factors were integrated into multivariable regression models to estimate their impact on dropout time. While the Cox proportional hazard model examines the effect on the hazard rate in the absence of competing risks, the subdistribution hazard model assesses the effect of factors on dropout incidence when competing risks are present. We compared the two models using the Akaike Information Criterion (AIC). The model with the smaller value indicates a more suitable model. Finally, we assessed the assumptions of both models. For the Cox proportional model, we verified the assumption of proportional hazard (PH) by contrasting the PH model for each predictor with an alternative model, allowing the predictor's regression coefficient to vary smoothly over time [21]. Conversely, we examined the assumption of proportionality of the hazard of the Cumulative Incidence Function (CIF) by plotting log(-log(1-CIF)) against log(time) [12]. This assumption is deemed valid when the two graphs do not intersect. All statistical analyses were conducted using R [22].

The log-rank test provides a nonparametric hypothesis test statistic, which was utilized to compare survival distributions between two groups. Essentially, the log rank test compares the actual number of events observed in each group with the expected number under the assumption of identical survival curves.

H_0: The two survival curves are identical, or $(H_0: S1_t = S2_t)$

H_1: The two survival curves are not identical at least for one time point, or $(H_1: S1_t \neq S2_t)$

At a significance level of α=0.05, the test statistic for the log-rank test is approximately distributed as a chi-square test statistic and is calculated as the summation of the differences between observed and expected event counts squared over all event times, divided by the sum of the expected event counts squared. This test statistic has degrees of freedom equal to the number of comparison groups minus one (k-1), where k represents the number of comparison groups. The log-rank test statistic is approximated as follows [23]:

$$\sum_i \frac{(O_i - E_i)^2}{E_i} \sim \chi^2_{k-1} \quad (1)$$

Where the $\mathrm{O_i}$ is the total number of observed events in groups I, respectively, and $\mathrm{E_i}$ is the total number of expected events.

One of the most popular regression techniques for survival outcomes is the Cox proportional hazards regression model. It was used to investigate the association between the survival time of individuals and predictor variables or covariates. The Cox proportional hazards regression model can be written as follows:

$$h(t) = h_0(t) \exp(X^T \beta) \quad (2)$$

where h(t) is the expected hazard at time t, $h_0(t)$ is the baseline hazard and represents the hazard when there are no predictors or all of the predictors (or independent variables) X_1, X_2 , X_p are equal to zero, β is a vector of unknown regression parameters. For each t_k, the risk set R_k immediately before time t_k is the set of patients who are still active (been right-censored), which represents the candidate set of patients for whom a dropout was observed at time t_k. Conditionally, on that some patient in this risk set R_k dropped out at time t_k, the probability that it is a particular patient $i_k \in R_k$ is :

$$\frac{h_i(t_k)}{\sum_{j \in R_k} h_j(t_k)} \quad (3)$$

This is also referred to as the instantaneous rate of dropout for patient i_k to the sum of rates for all candidate patients [24]. Under the proportional hazard model, that is :

$$\frac{h_0(t_k) \exp(\beta^T X_j)}{\sum_{j \in R_k} h_0(t_k) \exp(\beta^T X_j)} \quad (4)$$

Notice that this quantity does not depend on the baseline $h_0(t_k)$ since they cancel. By taking the product of expression (4) over all observed dropout times, we obtain the partial likelihood function as defined by Cox [25]:

$$L(\beta) = \prod_{i=1}^{n} \frac{\exp(\beta^T X_i)}{\sum_{j \in R_k} \exp(\beta^T X_j)} \quad (5)$$

For making inference about the coefficients β, Cox recommended treating equation (5) as an ordinary likelihood. The likelihood function in equation (5) can be expressed as:

$$L(\beta) = \prod_{i=1}^{n} \left\{ \frac{\exp(\beta^T X_i)}{\sum_{j \in R_i} \exp(\beta^T X_j)} \right\}^{\delta_i} \quad (6)$$

where $\delta_i = 1$ if the event (dropout) occurred and 0 if censored. The values of β can be estimated using the maximum likelihood method, by taking the log of the equation (6), and then differentiating with respect to each β, after setting the equation equal to zero.

2.3. Competing Risks Model for Patients in the CDiC Program

Let T represent failure time as the time from registration in the CDiC program to the time of exiting the program. We considered two causes for exiting the program (K={1,2}), drop out from the program and death. For competing risks data, the observed follow-up time of patient i is defined by $T_{ik} = \min(T_{ik}, C_{ik})$ which represents failure time and censoring time for patient i with k cause of failure. When competing risks are not present, one compares the survival functions of groups using the log-rank test to identify the significant factors. However, when competing risks are present, the cumulative incidence functions are compared using Gray's test statistic [26]. This test is a modified chi-square test, an alternative to the log-rank test. It compares the weighted averages of the hazard of the subdistribution functions for the event of interest. The general form of the score group i is:

$$Z_i = \int_0^\tau W_i(t)\{\gamma_i(t) - \gamma_0(t)\}\, dt \quad (7)$$

where τ is the maximum time observed in both groups, W_i is a weight function, $\gamma_i(t)$ is the hazard of the subdistribution for group i and γ_0(t) is the hazard of the subdistribution for all groups together. To estimate the effect of covariates on the absolute risk of dropping out, the Fine and Gray model was used (15), which can be presented as:

$$\lambda_k(t\,; Z) = \lambda_{k0}(t) \exp(\beta_k^T Z) \quad (8)$$

where $\lambda_k(t\,; Z)$ is the sub-distribution hazard function, $\lambda_{k0}(t)$ is the baseline sub-distribution hazard for a patient with all covariates equal to zero, β_k is the vector of regression coefficients of the variables Z. For event 1, the sub-distribution hazard function (SDH) is the probability of exiting from the program from cause 1 (dropping out) in a small-time interval Δt, given that no other event than 1 occurred for a subject before time t [13].

$$\lambda_1(t) = \lim_{\Delta t \to 0} \frac{P(t \le T < t+\Delta t, K=1 \mid T > t \cup (T \le t \cap K \ne 1))}{\Delta t} \quad (9)$$

Note that patients who did not drop out and died before time t are in the risk set for all future failure times. The SDH rates for the included covariates are assumed to be proportional. The most important feature of this function is that it can be directly linked to CIF as follows:

$$F_k(t; Z) = 1 - \exp\big(-\Lambda_k(t; Z)\big) = 1 - \exp\Big(-\int_0^t \lambda_k(s; Z)\, ds\Big) \quad (10)$$

To estimate the parameters in the Fine and Gray model, the partial likelihood approach is used [12]

$$L(\beta) = \prod_{i=1}^{n} \frac{\exp(\beta^T Z_i)}{\sum_{l \in R(t_l)} w_{il} \exp(\beta^T Z_i)} \quad (11)$$

where w_{il} is the weight of the patient i in the event of interest and defined as

$$W_{il} = \frac{\hat{G}(t_i)}{\hat{G}(\min(t_i, t_l))} \quad (12)$$

where $\hat{G}(t_i)$ is the estimate of survival function from survival time i patient, $\hat{G}(\min(t_i, t_l))$ is the estimation of the survival function from the minimum value between the survival time of the patient i and the patient on the event 1 or event of interest (18). We can see that equation (5) and equation (11) are similar except that equation (11) includes weight.

3. RESULTS

Out of 1132 patients considered in the study, almost half (49.6%) of them were male, and two-third (65%) were from urban areas. Every 2 out of 5 (41.3%) children with T1DM were underweight, and another 39.5% had high blood pressure. Among the 1132 registered T1DM patients, 20 (1.8%) of them dropped from the program, 43(3.8%) died, and 1069 (94.4%) were active. According to the findings presented in Table **1**, individuals residing in urban areas demonstrated a slightly higher rate of active participation in the program in comparison to those in rural areas (95% versus 93.4%). Conversely, individuals residing in rural areas exhibited a slightly lower dropout rate (1.3%) in contrast to their urban counterparts (2%). Additionally, mortality rates were observed to be higher among individuals in rural settings (5.3%) as opposed to those in urban settings (3%). The results also indicated similar levels of participation in the program among both males and females. However, males had a slightly higher dropout rate (2.3%) compared to females (1.2%), while females showed a slightly higher mortality rate (4.8%) compared to males (2.9%). According to the BMI variable, individuals classified as overweight demonstrated the highest level of program engagement (96.9%), closely trailed by those with a normal BMI (96.1%). Conversely, underweight patients exhibited the highest mortality rates (5.4%) and dropout rates (3%). Furthermore, it was observed that individuals with high blood pressure showed the highest level of program engagement (97%). In comparison, those with low blood pressure exhibited the lowest program participation rate (92.5%). Although mortality rates were highest among patients with low blood pressure (6.7%), dropout rates were highest among those with high blood pressure (2%), followed by individuals with normal blood pressure (1.8%).

In the absence of competing risks, the average duration until dropout from the program was 471 days,

Table 1: Distribution of Variables Based on the Patients CDiC Program Status

Factor	Category	Active, n (%)	Dropout, n (%)	Died, n (%)	Total n=1132
Residence	Rural	370 (93.4)	5 (1.3)	21 (5.3)	396
	Urban	699 (95.0)	15 (2.0)	22 (3.0)	736
Gender	Female	537 (94.0)	7 (1.2.0)	27 (4.8)	571
	Male	532 (94.8)	13 (2.3)	16 (2.9)	561
BMI	Normal	394 (96.1)	5 (1.2)	11 (2.7)	410
	Overweight	247 (96.9)	1 (0.4)	7 (2.7)	255
	Underweight	428 (91.6)	14 (3.0)	25 (5.4)	467
Hypertension	Normal	525 (92.7)	10 (1.8)	31 (5.5)	566
	High BP	434 (97.0)	9 (2.0)	4 (1.0)	447
	Low BP	110 (92.5)	1 (0.8)	8 (6.7)	119

whereas when accounting for competing risks, the average duration until dropout extended to 522.5 days. Estimating the CIFs for dropout from the program and death events, results in Figure **2** showed that after 9 years from follow-up, the probability of dropping out is about 12.5%, and the probability of death is about 20%. To explore the factors influencing dropout time, we initially employed univariable methods. In scenarios without competing risks, we utilized the log-rank test, while in situations involving competing risks, we used the Gray test.

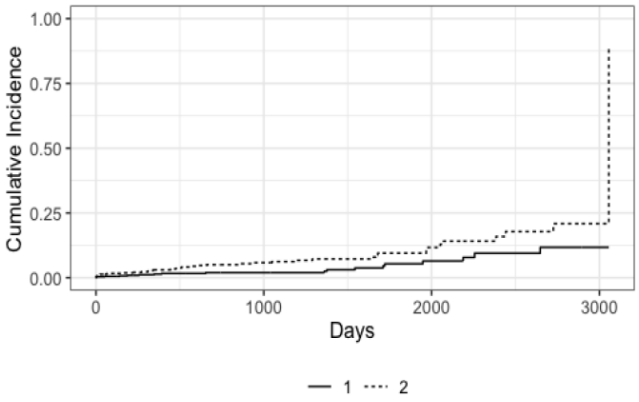

Figure 2: Cumulative incidence function for event 1 (dropout) and event 2 (death).

Table **2** presents the statistical test results for examining the association between various factors and dropouts from the program. Testing the influence of the residence variable on dropout time, both tests generated non-significant p-values (0.6 for the Log-rank test and 0.085 for the Gray test), indicating no notable association between residence and dropout time when competing risk is ignored, while residence has marginally significant ($p<0.10$) effect on dropout when competing risk is accounted for. Similarly, the gender of the patients showed no significant association with dropout, with p-values of 0.8 for the Log-rank test and 0.12 for the Gray test. Conversely, the BMI variable exhibited highly significant results in both tests, with p-values of 0.0002 for the Log-rank test and 0.009 for the Gray test. This underscores a strong association between BMI and dropout from the program.

Figure **3** shows that patients with underweight had a higher estimated probability of dropping out than patients with category normal and overweight BMI. Regarding the Hypertension variable, while the Log-rank test yielded a p-value of 0.2, indicating non-significance, the Gray test showed a p-value of 0.22, suggesting no significant association between hypertension and dropout. The results highlight strong correlation between BMI and dropout from the program. It is worth noting that even when there are small number (<5) of events in some categories, the log-rank test and the Gray test can still be applied [27].

To assess the impact of the variable on dropout duration, we utilized the Cox regression model for scenarios without competing risks, while for situations involving competing risks, the Fine and Gray model was employed. Results presented in Table **3** revealed that in the absence of competing risks, the hazard of dropout for underweight patients was 3.088 times higher compared to normal-weight patients. However, when accounting for competing risks, it was observed that the hazard of dropout for underweight patients increased to 3.43 times that of patients with normal weight. Our findings are in agreement with [32] whose systematic review determined that whereas underweight group was shown to have a 3.4 times greater risk of mortality than the normal-weight group, there was no significant difference in mortality risk between normal weight group and overweight group. Although, it is not so clear the reason for this difference, some studies, such as [33, 34] have associated this difference with increased impairments in insulin secretion, undernutrition, and epigenetic alterations to

Table 2: Nonparametric Tests for Significance of Variables with and without Competing Risks

Variable	Log-rank test	P-value	Gray test	P-value
Residence	0.5	0.6	2.97	0.085
Gender	0.1	0.8	2.38	0.12
BMI	16.7	0.0002	9.34	0.009
Hypertension	3	0.2	10.2	0.22

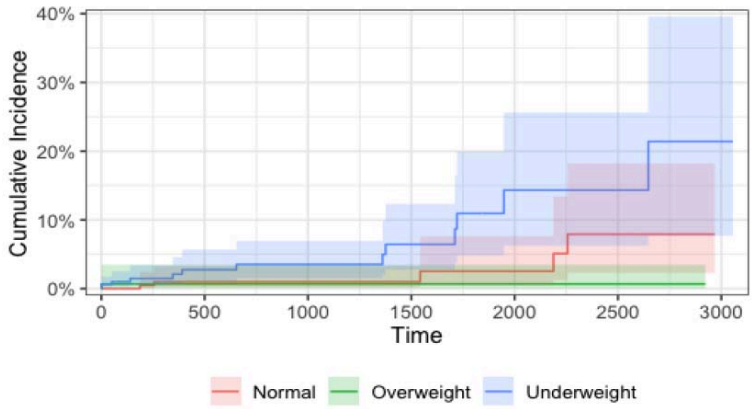

Figure 3: Cumulative incidence function for BMI.

the genome in the underweight as compaired to the overweight BMI among patients.

Despite the smaller standard error observed for the estimated coefficients in the Cox model, there was a significant decrease in the AIC value from 689 to 218 upon utilizing the Fine and Gray model, indicating better fitting of data by Fine and Gray model than Cox model.

To assess the proportional hazard assumption, which is a fundamental assumption of the Cox regression model, the “cox.zph” function from the survival package is used. Based on the p-value results in Table **4**, there was no strong evidence to suggest a violation of the proportional hazard assumption for any of the variables. The global p-value of 0.71 also indicates no strong evidence of violating the assumption for any of the variables in the Cox model. Unfortunately, R lacks a direct function for assessing the proportionality assumption regarding the Fine and Gray model. To investigate the proportionality assumption for the FG model, log(-log(1-F)) can be plotted against log(time) [12], where F is the CIF for the event dropping out. If the curves do not cross, we can say that the assumption is not violated. Figure **3** shows that the assumption was violated for hypertension and residence variables. For the BMI variable, we can see that the curve for the overweight BMI was horizontal, which crosses the other curves. This may be because there is only one event in this group. Ignoring this group, we could say that the variable BMI does not violate the assumption. Regarding the gender variable, the curves show no evidence of violating the assumption.

4. DISCUSSION

Exploring factors contributing to dropouts from the CDiC program for T1DM patients is crucial for enhancing the program's efficacy in addressing specific challenges. This study employs both non-competing risk methods and competing risk methods to assess

Table 3: Regression Models for the Effect of BMI on Patient Time to Drop Out

		Cox Model			FG Model		
		β (SE)	exp (β)	P-value	β (SE)	exp (β)	P-value
BMI	Normal	-	-	-	-	-	-
	Overweight	0.35(0.43)	1.417	0.421	-0.62(1.1)	0.54	0.57
	Underweight	1.12(0.30)	3.088	0.000	1.23(0.51)	3.43	0.015
Log-L		-342.5			-107		
AIC		689			218		

Table 4: Checking the Assumption of Proportionality for the Cox Model

	Chisq	df	P-value
BMI	1.993	2	0.37
Gender	0.383	1	0.54
Residence	0.463	1	0.50
Hypertension	1.040	2	0.59
GLOBAL	3.728	0	0.71

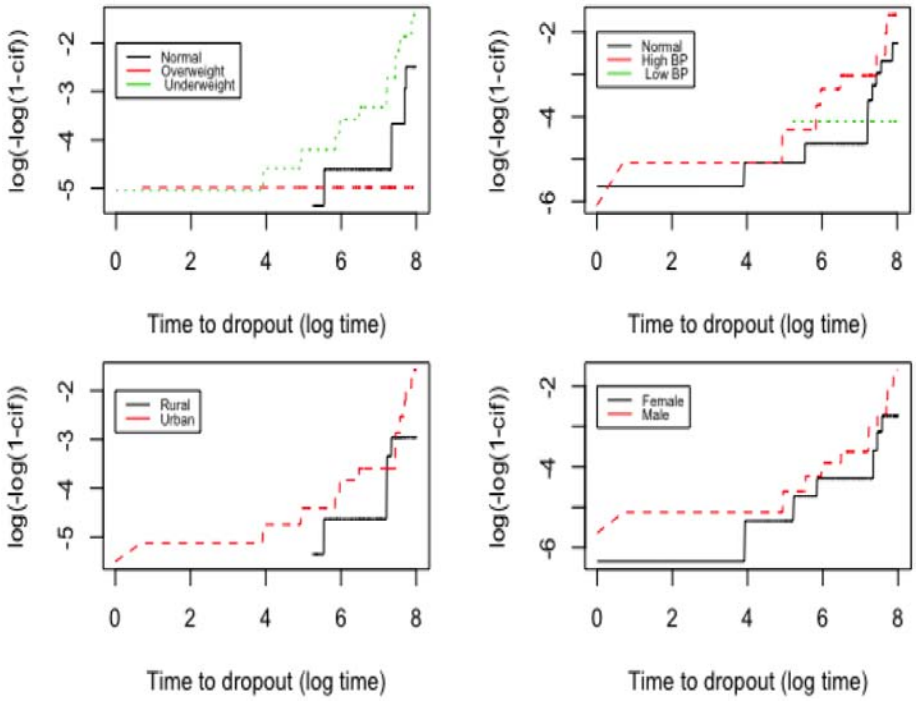

Figure 4: Checking the assumption of proportionality for the FG model.

the average time to dropout, identify factors influencing dropout time, and estimate the effect of significant variables. Additionally, our aim is to compare the two methods to determine the preferred approach for analyzing the time to dropout.

From the results, we found that the average time for dropping out of the program was 471 days. When accounting for competing risks, the average dropout time increased by 51.5 days, indicating that considering death as a dropout event led to an underestimation of dropout time. It's worth noting that while our focus aligns with [16] regarding analyzing CDiC program data in Uganda to investigate the reasons behind dropping out of the program, our methodology differs. Wesonga *et al*. [16] treated "Active" as the event of interest, and therefore, patients who dropped out were considered censored data. While we considered dropout as the event of interest and death as a competing event. Consequently, our study exhibited a higher percentage of censoring compared to theirs (94%). Nonetheless, our study shares similarities with the work of Fufa *et al*. (2023), where their percentage of censoring was 81%.

When pinpointing factors influencing the dropout time, both non-competing methods and competing methods indicated that the BMI variable had a significant effect on dropout time, while residence showed a marginal significant ($p<0.10$) effect on dropout. Gender , and hypertension variables showed no influence on dropout time. In contrast, [16] reported that gender and residence variables were significant factors. However, it is noteworthy that their analysis focused on survival time in the program, whereas ours focused on dropout time. Therefore, it is not necessarily expected that a factor influencing survival time would also affect dropout time. Conversely, BMI was found to be a significant factor influencing both survival time and dropout time.

To estimate the effect of factors affecting dropout time, multivariable models were used. In the absence of competing risks, we employed the Cox proportional hazard model, and the BMI variable was included. It

was found that the hazard of dropping out for patients with underweight BMI is three times that for patients with normal BMI. However, when considering competing risks, the results of the Fine and Gray model indicated that the hazard of dropping out for patients with underweight is 3.43 times that for those with normal BMI. This enhancement can be attributed to the inclusion of weight in the Fine and Gray model formula, allowing for a more comprehensive assessment of the factor's impact on the outcome. Furthermore, it was determined that the AIC value for the Fine and Gray model (218) was lower than that of the Cox model (689), indicating that the Fine and Gray model is more suitable for analyzing time to dropout. These differences suggest that when competing risks are present, it is better to apply competing risk models, which is consistent with the study conducted by Nolan EK and Chen HY [28].

In most studies that evaluate disease outcomes, statistical approaches such as the Kaplan-Meier (KM) survival analysis and Cox proportional hazards regression are normally used to account for unequal follow-up time, for instance when patients die or drop out before study completion. However, these survival analysis methods were originally developed to describe all-cause mortality in the presence of loss to follow-up independent of the study outcome [24,29]. Therefore, when KM and Cox proportional hazards models are used to describe outcomes other than all-cause mortality in the presence of a significant and related competing risk, such as death, it is likely that biased results may result [30,31].

5. CONCLUSION

In this study, we sought to determine the importance of utilizing competing risks methods in dropout analysis within the CDiC program for T1DM patients. By employing both non-competing and competing risk approaches, we comprehensively assessed dropout patterns, identified influential factors, and estimated their effects on dropout time. Our findings underscore the necessity of considering competing risks, such as death, to avoid underestimating dropout time.

While our methodology differed from previous studies, particularly in handling dropout events and censoring, we consistently found BMI to be a significant determinant of dropout time. Multivariable modeling revealed the substantial impact of underweight BMI, with the Fine and Gray model providing a more refined estimation compared to the Cox model. This improvement, attributed to the inclusion of weight, enhances our understanding of dropout determinants. Moreover, our comparison of model performance using AIC values highlights the superiority of the Fine and Gray model in analyzing time to dropout. These insights contribute to optimizing interventions within the CDiC program and similar initiatives, ultimately enhancing patient care and program effectiveness in addressing dropout challenges among T1DM patients. Therefore, our study pinpoints two major implications that is, a patient dropout prevention strategy should adopt statistical modeling that would track all patients registered on the CDiC program to identify at-risk patients and provide them with additional support. And secondly, that a competing risks hazard model for predicting the probability of dropping out should become part of a powerful tool to identify patients at risk of dropping out.

FUNDING

The study received no funding from any source.

CONFLICT OF INTEREST

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

ACKNOWLEDGEMENTS

The authors would like to thank Dr. Silver Bahendeka, a Senior Consultant Physician, Diabetes and Endocrinology, who is the lead investigator for the CDiC program in Uganda. The authors are indebted to the participants of the CDiC program in Uganda for providing valuable information. The views expressed herein are solely those of the authors and do not necessarily reflect the views of any institution or organization. All procedures followed in this study were under the ethical standards of the responsible committee on human experimentation (institutional and national) and with the Helsinki Declaration of 1975, as revised in 2000.

REFERENCES

[1] Richards NC, Gouda HN, Durham J, Rampatige R, Rodney A, Whittaker M. Disability, noncommunicable disease, and health information. Bull World Health Organ 2016; 94(3): 230-232. https://doi.org/10.2471/BLT.15.156869

[2] International Diabetes Federation. A summary of the figures and key findings. The IDF Diabetes Atlas. 10th edition. Brussels: International Diabetes Federation; 2021. https://diabetesatlas.org/idfawp/resource-files/2021/07/IDF_Atlas_10th_Edition_2021.pdf

[3] Global Burden of Disease Collaborative Network. Global Burden of Disease Study 2019. Results. Institute for Health Metrics and Evaluation, https://vizhub.healthdata.org/gbd-results/ (2020).

[4] International Diabetes Federation (IDF). Type 1 diabetes estimates in children and adults. International Diabetes Federation (IDF) Atlas Report 2022. Sponsored by Novo Nordisk, MSD, and Sanofi 2022. Available from: https://www.diabetesatlas.org.

[5] Novo Nordisk, Roche, International Society for Pediatric and Adolescent Diabetes (ISPAD), World Diabetes Foundation

(WDF). 10 years of Changing Diabetes in Children 2009 – 2020. 2023. Available at: https://www.novonordisk.com/content/dam/nncorp/global/en/sustainable-business/pdfs/changing-diabetes-in-children/cdic-10-year-report.pdf.

[6] Bahendeka S, Mutungi G, Tugumisirize F, Kamugisha A, Nyangabyaki C, Wesonga R, Sseguya W, Mubangizi D, Nalunkuma C, Were TP. Healthcare delivery for paediatric and adolescent diabetes in low resource settings: Type 1 diabetes clinics in Uganda. Global Public Health 2019; 14(12): 1869-1883. https://doi.org/10.1080/17441692.2019.1611897

[7] Clayton D, Cuzick J. Multivariate generalizations of the proportional hazards model. J R Stat Soc Ser A Gen 1985; 148(2): 82-117. https://doi.org/10.2307/2981943

[8] Abdel-Qadir H, Fang J, Lee DS, Tu JV, Amir E, Austin PC, Anderson GM. 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 2018; 11(7): e004580. https://doi.org/10.1161/CIRCOUTCOMES.118.004580

[9] Schuster NA, Hoogendijk EO, Kok AAL, Twisk JWR, Heymans MW. Ignoring competing events in the analysis of survival data may lead to biased results: A nonmathematical illustration of competing risk analysis. J Clin Epidemiol 2020; 122: 42-48. https://doi.org/10.1016/j.jclinepi.2020.03.004

[10] Verduijn M, Grootendorst DC, Dekker FW, Jager KJ, le Cessie S. The analysis of competing events like cause-specific mortality—beware of the Kaplan-Meier method. Nephrol Dial Transplant 2011; 26(1): 56-61. https://doi.org/10.1093/ndt/gfq661

[11] Lacny S, Wilson T, Clement F, Roberts DJ, Faris PD, Ghali WA, Marshall DA. Kaplan-Meier survival analysis overestimates the risk of revision arthroplasty: A meta-analysis. Clin Orthop Relat Res 2015; 473(11): 3431-3442. https://doi.org/10.1007/s11999-015-4235-8

[12] Pintilie M. Competing Risks: A Practical Perspective. John Wiley & Sons 2006. https://doi.org/10.1002/9780470870709

[13] Austin PC, Lee DS, Fine JP. Introduction to the analysis of survival data in the presence of competing risks. Circulation 2016; 133(6): 601-609. https://doi.org/10.1161/CIRCULATIONAHA.115.017719

[14] Fine JP, Gray RJ. A proportional hazards model for the subdistribution of a competing risk. J Am Stat Assoc 1999; 94(446): 496-509. https://doi.org/10.1080/01621459.1999.10474144

[15] Austin PC, Fine JP. Practical recommendations for reporting Fine-Gray model analyses for competing risk data. Stat Med 2017; 36(27): 4391-4400. https://doi.org/10.1002/sim.7501

[16] Wesonga R, Sarr A, Bahendeka S. Survival of Type 1 Diabetes Mellitus Patients on Disease Control Programme. [Preprint]. Available at SSRN: https://ssrn.com/abstract=4642031

[17] Akessa GM, Tadesse M, Abebe G. Survival analysis of loss to follow-up treatment among tuberculosis patients at Jimma University Specialized Hospital, Jimma, Southwest Ethiopia. Int J Stat Mech 2015; 2015: 923025. https://doi.org/10.1155/2015/923025

[18] Woldeyohannes D, *et al*. Predictors of mortality and loss to follow-up among drug resistant tuberculosis patients in Oromia Hospitals, Ethiopia: a retrospective follow-up study. PLoS One 2021; 16(5): e0250804. https://doi.org/10.1371/journal.pone.0250804

[19] Fufa DB, Diriba TA, Dame KT, *et al*. Competing risk models to evaluate the factors for time to loss to follow-up among tuberculosis patients at Ambo General Hospital. Arch Public Health 2023; 81: 117. https://doi.org/10.1186/s13690-023-01130-2

[20] Kim HW, Min J, Ko Y, Oh JY, Jeong YJ, Lee EH, Yang B, Koo HK, Lee SS, Park JS, Park KJ, Chang JH, Joh J, Lee MK, Kim JS. Risk of loss to follow-up among tuberculosis patients in South Korea: whom should we focus on? Front Public Health 2023; 11: 1247772. https://doi.org/10.3389/fpubh.2023.1247772

[21] Grambsch PM, Therneau TM. Proportional hazards tests and diagnostics based on weighted residuals. Biometrika 1994; 81(3): 515-526. https://doi.org/10.1093/biomet/81.3.515

[22] R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria; 2023. Available from: https://www.R-project.org/.

[23] Kleinbaum DG, Klein M. Survival Analysis: A Self-Learning Text, Third Edition (Statistics for Biology and Health). Springer; 2011. https://doi.org/10.1007/978-1-4419-6646-9

[24] McGregor DE, Palarea-Albaladejo J, Dall PM, Hron K, Chastin SFM. Cox regression survival analysis with compositional covariates: Application to modeling mortality risk from 24-h physical activity patterns. Stat Methods Med Res 2020; 29(5): 1447-1465. https://doi.org/10.1177/0962280219864125

[25] Cox DR. Regression models and life-tables. J R Stat Soc Ser B Methodol 1972; 34(2): 187-220. https://doi.org/10.1111/j.2517-6161.1972.tb00899.x

[26] Gray RJ. A class of K-sample tests for comparing the cumulative incidence of a competing risk. Ann Stat 1988; 16(3): 1141-1154. https://doi.org/10.1214/aos/1176350951

[27] Collett D. Modelling Survival Data in Medical Research. 3rd ed. Chapman and Hall/CRC; 2014. https://doi.org/10.1201/b18041

[28] Nolan EK, Chen HY. A comparison of the Cox model to the Fine-Gray model for survival analyses of re-fracture rates. Arch Osteoporos 2020; 15(1): 86. https://doi.org/10.1007/s11657-020-00748-x

[29] Gu J, Fan Y, Yin G. Reconstructing the Kaplan–Meier estimator as an M-estimator. Am Stat 2022; 76(1): 37-43. https://doi.org/10.1080/00031305.2021.1947376

[30] Waller M, Mishra GD, Dobson AJ. A comparison of cause-specific and competing risk models to assess risk factors for dementia. Epidemiol Methods 2020; 9(1): 20190036. https://doi.org/10.1515/em-2019-0036

[31] Angriman F, Ferreyro BL, Harhay MO, Wunsch H, Rosella LC, Scales DC. Accounting for competing events when evaluating long-term outcomes in survivors of critical illness. Am J Respir Crit Care Med 2023; 208(11): 1158-1165. https://doi.org/10.1164/rccm.202305-0790CP

[32] Jung HN, Kim S., Jung CH, Cho YK. Association between Body Mass Index and Mortality in Type 1 Diabetes Mellitus: A Systematic Review and Meta-Analysis. Journal of Obesity & Metabolic Syndrome 2023; 32(2): 151. https://doi.org/10.7570/jomes22061

[33] Gujral UP, Weber MB, Staimez LR, Narayan, KV. Diabetes among non-overweight individuals: an emerging public health challenge. Current Diabetes Reports 2018; 18: 1-10. https://doi.org/10.1007/s11892-018-1017-1

[34] Gravina G, Ferrari F, Nebbiai G. The obesity paradox and diabetes. Eating and Weight Disorders-Studies on Anorexia. Bulimia and Obesity 2021; 26: 1057-1068. https://doi.org/10.1007/s40519-020-01015-1

Received on 20-04-2024 Accepted on 10-05-2024 Published on 04-06-2024

https://doi.org/10.6000/1929-6029.2024.13.06