Journal of Statistics and Computer Science
Vol. 4, Issue 2, 2025, pp. 135-153
https://doi.org/10.47509/JSCS.2025.v04i02.05

Performance of a Rayleigh Regression Model with Interval-Censoring (IC) and Informative Dropout: A Simulation Study Under the Estimand Framework

[1]**Suman Kapoor and** [2]**Arindam Gupta**

Department of Statistics
The University of Burdwan
West Bengal, INDIA
[1]skapoor@scholar.buruniv.ac.in, [2]agupta@stat.buruniv.ac.in

ABSTRACT

This article addresses the analytical challenges posed by interval-censored (IC) survival data when intercurrent events (ICEs), such as study discontinuation due to lack of efficacy or Adverse Events (AE), introduce informative censoring. We assess the performance of the Rayleigh regression model with fixed co-variates within the Estimand framework to handle this issue. A simulation study was conducted to compare parameter estimates under three scenarios: a baseline analysis without imputation, a composite strategy using non-responder imputation (NRI), and a hypothetical strategy using logistic regression multiple imputation (LRMI). Performance was evaluated using Bias, Root Mean Square Error (RMSE), and coverage probability across various sample sizes and censoring proportions. Results show that while RMSE improves with larger sample sizes across all methods, the hypothetical strategy with LRMI is demonstrably superior. It yields less biased estimates and maintains a higher, more consistent coverage probability compared to the conservative NRI approach, which systematically underestimates treatment effects. We conclude that properly handling ICEs through a principled imputation method like LRMI is crucial for obtaining accurate and reliable results when modeling interval-censored data with the Rayleigh distribution.

Key Words: *Rayleigh, Interval-censoring, Maximum likelihood estimation, Estimand, Hypothetical strategy, Composite strategy, Survival analysis, R.*

MSC Classification: 62N01, 62N02, 62P10.

1 Introduction

Exploring the relation between time-to-event data with different co-variates is an emerging topic in survival analysis especially in oncology trials. It's evident that the information gathered about the condition of the participant at the time of commencement of the study also have impact on the duration of event occurrence according to Jammalamadaka and Leong (2015). In survival analysis, we mainly explore the duration until the event of interest (i.e. Progression of the tumor, death etc.) occurs. However, it's not possible to follow individual participants until the event of interest is experienced as described in Schober and Vetter (2018). As a result, the

Article History
Received : 15 November 2025; Revised : 10 December 2025; Accepted : 15 December 2025; Published : 30 December 2025

To cite this paper

Suman Kapoor and Arindam Gupta (2025). Performance of a Rayleigh Regression Model with Interval-Censoring (IC) and Informative Dropout: A Simulation Study under the Estimated Framework. *Journal of Statistics and Computer Science.* 4(2), 135-153.

accurate prediction of the survival time is undetermined. This phenomenon is called censoring, and this situation must be used while drawing any inference about the lifetime data. Survival analysis is the only area in statistics where censored data is used to draw inference about the data. IC is a very common and the most complicated type of censoring in Survival analysis. When T, which is the survival time, is not observed appropriately but can be identified to materialise within a specified interval $[L_i, R_i]$, where i denotes number of participants up to n. In this scenario, n denotes the number of total evaluations or experiments that occurred. In case the event never occurred during a stipulated time frame of a clinical trial, that means most participants are censored from treatment arms, as a result, we can not draw a reliable conclusion about the treatment due to a lack of information. In the IC context, those participants are right censored (i.e. no event of interest was seen until the last disease assessment) as mentioned in Ozenne et al. (2020). Most common reasons of right censoring are, if the particular treatment is working and participant did not experience any event during the time frame of a clinical trial, participant might be experiencing different side effects and participant decided to leave the study ("Discontinued due to Adverse event"), participant might think that the treatment is not working well so decided to drop out from the study ("Discontinued due to Lack of efficacy") or participant might discontinue due to any other reason which is not related to the study. This also called informative censoring, which means the probability of being censored is not independent of the participant's risk of the event as defined in Shih (2002). According to various literature, incomplete information should be addressed in such a way that we can utilize maximum information out of it. To minimize the number of censor, a missing data imputation technique can be deployed as described in Leung et al. (1997). There is a high possibility that a participant who discontinued the clinical trial due to "adverse event related to study medication" or "lack of study medication efficacy" might experience the event of interest in reality if the study were continued. We can think of another scenario where a participant who received any medication which is prohibited in that clinical trial and did not experienced any event by end of the trial, in reality if participant had not taken that prohibited medication or rescue medication and continued might experience an event of interest at some point of time. In the clinical trial universe, these are called inter-current events . The event which occurred during the course of the trial is called inter-current events specifically. If the above scenarios are not handled carefully, it would lead to a biased conclusion for the comparison of certain clinical endpoints like PFS (Progression-free-survival), TTP (Time-to-progression), recurrence of the disease as described in Rufibach (2019). Participants whose data are censored due to intercurrent events—such as dropping out of a study or experiencing an unrelated medical condition—require special handling in survival analysis. To better understand the true likelihood of the event of interest (e.g., disease progression) occurring in a real-world scenario, it is important to account for these censored cases. This is typically done through imputation techniques, which estimate the potential event times that might have occurred had the intercurrent events not taken place. This study aims to address following research questions:

1. To quantify, within the estimand framework, the impact of (a) a hypothetical strategy implemented via multiple imputation (MI) with logistic regression and (b) a composite strategy based on non-responder imputation (NRI) on the finite-sample performance of

parameter estimates from a Rayleigh regression model for interval-censored outcomes, evaluated by bias, root mean squared error (RMSE), and coverage probability.

2. To examine how these performance metrics vary as a function of sample size and the proportion of censoring.

Once imputation is completed, the resulting interval-censored data can be analyzed using appropriate statistical models. Traditionally, parametric distributions such as the exponential and Weibull have been employed to model time-to-event data. However, recent observations suggest that in certain contexts, the failure times may follow a Rayleigh distribution, which can offer a better fit based on the data's characteristics and the underlying survival function. This shift highlights the importance of selecting a distribution that accurately reflects the characteristics of the data, as it can significantly influence the conclusions drawn from the analysis. Extensive work has been done using the Rayleigh distribution in survival analysis for participants with type 2 tuberculosis in Elviana and Purwadi (2020). When the rate of failure linearly increases over time, then Rayleigh distribution is recommended, as mentioned in Anis et al. (2024).

2 Methodology

2.1 Estimand Framework and Imputation Method

In case of time-to-event data , analyzing treatment outcomes despite intercurrent events which might influence the primary outcome of interest such as progression of the disease or recurrence of the tumor lesion is a burning topic. In recent times, different techniques are developed which are more efficient in measuring the treatment effect than previously used censoring rule as discussed in Rufibach (2019). A particular participant can either be censored or event, if we are not sure if a participant is truly censored or not (discontinued due to Lack of efficacy , Adverse event, or consumed any prohibited/rescue medication), then it could be considered a missing value. The nature of this kind of data is categorical. Hence the missing status of such participants should be imputed using the multiple imputation method for categorical variables as discussed by Lee and Simpson (2014). Non-responder imputation is one of the easiest method to impute the missing data. Participants who experienced any intercurrent events then that participant will be considered as an event, irrespective of the occurrence of actual targeted event (e.g. Progression, death in case of progression free survival) and the next scheduled visit will be used to calculate the duration of the event. Since this is a single imputation method and also very conservative approach, it attracts lots of criticism as mentioned in Psychogyios et al. (2023). The main reason is, the method is very simple but it does not account for uncertainty in the prediction. Non-responder imputation generally applied under composite strategy as mentioned in Keene (2019). Moreover, single imputation methods often fail to incorporate individual-level characteristics—such as age, gender, smoking status, and other clinically relevant factors—that can significantly influence disease progression. In the context of cancer research, these variables are particularly important, as they are known to affect the aggressiveness and trajectory of the disease. For instance, older individuals or those with a history of smoking may exhibit a higher susceptibility to developing the disease compared to younger,

non-smoking counterparts. By ignoring these co-variates, single imputation can lead to biased or oversimplified estimates, ultimately reducing the accuracy and reliability of the analysis. Therefore, more sophisticated imputation techniques that account for participant-specific attributes are essential for producing meaningful and personalized predictions. A hypothetical strategy is a widely accepted approach, particularly by the European Medicines Agency (EMA) according to Heinrich et al. (2025), for handling data affected by intercurrent events. These events are identified as those that took place after treatment commencement but before the outcome is measured. This strategy is frequently employed in clinical trials to address such occurrences. For example, in a hypothetical scenario, one might estimate the treatment effect as if these intercurrent events (like treatment discontinuation or use of rescue medication) did not occur. When imputing a missing binary outcome using logistic regression under the MAR (i.e. missing at random) presumption, here the goal is to predict the missing values based on observed co-variates (predictors) that are related to the missingness. MAR assumes that the missingness mechanism can be explained by the observed predictors X. Logistic regression leverages these predictors to estimate $P(Y = 1|X)$, effectively "reconstructing" the unavailable or missing Y values founded based on the patterns of the observed data. - By modelling the relation in between X and Y Logistic model guarantees that predicted or imputed results are consistent with the observed data, satisfying the MAR assumption. The advantage of a hypothetical Estimand framework is , it enables us to understand the treatment effect under ideal conditions (if participant had adhered). The complete simulation programming is prepared using R 4.5.0 software.

2.2 Logistic Regression Multiple Imputation (MI) in General context

Once the status of the participant is set as missing, then it would be imputed applying the logistic regression MI method. The variable of interest with missing information would served as the response variable, and the other factors as the predictors. Let us consider the variable of binary response that takes only two values 0 or 1, with corresponding probabilities p and $1 - p$. Hence, the Y which is the dependent variable would follow a Bernoulli distribution with the parameter $E(Y) = p$. The form of the model would be $Y_i = E(Y_i) + \epsilon_i$. Here ϵ_i is the error term, and the error term distribution would completely depend on the distribution of the response variable which is Y_i. If we consider the independent variables X_1, X_2..X_m are m predictors, then the logistic regression model with multiple independent variables can be written as

$$E(Y_i) = \pi_i = \frac{\exp(\beta_0 + \beta_1 X_1 + \beta_2 X_2 + ...\beta_k X_m)}{1 + \exp(\beta_0 + \beta_1 X_1 + \beta_2 X_2 + ...\beta_k X_m)} \tag{2.1}$$

Using the maximum likelihood estimation method, logistic regression can be fitted. Log-likelihood is the most commonly used likelihood function to fit the parameters. A complete data set will be formed, which means for each and every co-variate, there would be no missing dependent variable. To form a complete dataset, a random sample will be simulated from the posterior distribution. For each case of missing information Bernoulli draw will be made, based on the fitted Logistic regression equation. According to Rubin (1996), using the fitted logistic model for binary categorization, a completely new logistic model will be simulated from

the predictive posterior distribution for the parameters under investigation and will be utilized to generate the missing status of the dependent variable. Finally, imputed values will be created either 0 or 1 as discussed in detail in Allison (2005). Depending on the bigger probability, either 0 or 1 will be assigned in place of missing values. Simply if the probability of being censored is larger than 0.5, then it would be imputed as censored; otherwise and event. The probability depends on the corresponding co-variate values. Further details are provided in the Appendix A.

2.3 Rayleigh distribution with interval-censored data

It is well established that survival data with IC has been analyzed through various approaches, including nonparametric, parametric, and semiparametric models. In this article, we are going to mainly focus on Rayleigh distribution when interval censored data is present with a baseline co-variate. The fundamental assumption is that the Rayleigh distribution with co-variate dependent scale parameter σ adequately models the survival time, with informative censoring. It was selected for this study due to its characteristic linearly increasing hazard function, $h(t;\sigma) = \frac{t}{\sigma^2}$. This property makes it particularly suitable for modeling scenarios where the failure rate grows steadily over time, a common assumption in processes involving aging. Unlike distributions that assume a constant hazard, the Rayleigh model is specifically recommended when the rate of failure increases linearly. This choice provides a theoretically sound basis for evaluating the proposed estimation methods under this important failure pattern as discussed in Al-Noor and Assi (2020). This distribution also have wide application in medical field like to analyze data received after a magnetic resonance imaging (MRI) scan. Once we impute the participants' status in a probabilistic way then we have complete non-missing interval censored data. The Rayleigh distribution, with the scale parameter $\sigma > 0$, has the CDF and PDF defined as:

$$F(t;\sigma) = 1 - e^{-\frac{t^2}{2\sigma^2}}, \quad t > 0, \tag{2.2}$$

$$f(t;\sigma) = \frac{t}{\sigma^2} e^{-\frac{t^2}{2\sigma^2}}, \quad t > 0. \tag{2.3}$$

The survival function and hazard function are:

$$S(t;\sigma) = 1 - F(t;\sigma) = e^{-\frac{t^2}{2\sigma^2}}, \tag{2.4}$$

$$h(t;\sigma) = \frac{f(t;\sigma)}{S(t;\sigma)} = \frac{t}{\sigma^2}. \tag{2.5}$$

2.4 Maximum Likelihood Estimation with Interval-Censored Data

Under IC data, the time of event of interest t_i for the i-th participant is known to fall in between $[L_i, R_i]$. The interval $[L_i, R_i]$ is not determined arbitrarily but arises from the periodic follow-up structure inherent to longitudinal studies such as clinical trials. The main factor is the predefined schedule of follow-up visits or tests. In clinical trials, participants are typically monitored at regular intervals (e.g., every 3 to 6 weeks, every 3 to 6 months) to check for the event of interest, such as disease progression which is completely depends on the nature of the study,

therapeutic area (mainly aggression of the disease) and availability of the participants. These scheduled appointments create the potential boundaries for the interval. Specifically, for a given participant i, L_i represents the time of the last assessment at which the event of interest had not yet occurred, while R_i denotes the time of the first subsequent assessment at which the event was observed. The true, unobserved event time T_i is thus known only to lie within this interval, such that $L_i < T_i \leq R_i$. The following indicator variables are defined to account for different censoring classification:

$$\delta_{E_i} = \begin{cases} 1, & \text{if uncensored at } t_i, \\ 0, & \text{otherwise,} \end{cases}$$

$$\delta_{I_i} = \begin{cases} 1, & \text{if interval censored in } [L_i, R_i], \\ 0, & \text{otherwise.} \end{cases}$$

If the event time t_i can be exactly identified then it's called uncensored and if the time can not be identified then it's called interval-censored. The likelihood expression for uncensored and Interval censored $i = 1, 2, 3.....n$ samples are given by Sparling et al. (2006).

$$L = \prod_{i=1}^{n} \left[f(t_i; \sigma)^{\delta_{E_i}} \left(F(t_{R_i}; \sigma) - F(t_{L_i}; \sigma) \right)^{\delta_{I_i}} \right]. \tag{2.6}$$

We can derive following equation after replacing the survival function and the distribution of the Rayleigh model into the above function,

$$l = \sum_{i=1}^{n} \left\{ \delta_{E_i} \ln \left(\frac{t_i}{\sigma^2} e^{-\frac{t_i^2}{2\sigma^2}} \right) + \delta_{I_i} \ln \left(e^{-\frac{t_{L_i}^2}{2\sigma^2}} - e^{-\frac{t_{R_i}^2}{2\sigma^2}} \right) \right\}. \tag{2.7}$$

Simplifying:

$$l = \sum_{i=1}^{n} \left\{ \delta_{E_i} \left(\ln t_i - 2 \ln \sigma - \frac{t_i^2}{2\sigma^2} \right) + \delta_{I_i} \ln \left(e^{-\frac{t_{L_i}^2}{2\sigma^2}} - e^{-\frac{t_{R_i}^2}{2\sigma^2}} \right) \right\}. \tag{2.8}$$

The MLE for σ is obtained by maximizing (2.8), typically requiring numerical optimization due to the complexity of the derivatives.

2.5 Rayleigh Distribution with Co-variates

To incorporate co-variates (e.g. gender, smoking status, age group), we parameterize the scale parameter σ as a function of co-variates. In this work we have considered only 3 co-variates, however the number of co-variates could be increased or decreased as per the study requirement. Let $\mathbf{x}_i = (x_{i1}, x_{i2}, x_{i3})$ represent the co-variates for the i-th participant, where:

- x_{i1}: e.g. Gender (1 = male, 0 = female),
- x_{i2}: e.g. Status of smoking (1 = smoker, 0 = non-smoker),
- x_{i3}: e.g. Age group (0 = young, 1 = elderly).

We model the scale parameter as:

$$\sigma_i = e^{\beta_0+\beta_1 x_{i1}+\beta_2 x_{i2}+\beta_3 x_{i3}}, \tag{2.9}$$

where $\beta_0, \beta_1, \beta_2, \beta_3$ are regression coefficients. Substituting σ_i into the Rayleigh distribution functions, the log-likelihood becomes:

$$\begin{aligned} l = \sum_{i=1}^{n} \Bigg\{ & \delta_{E_i} \left(\ln t_i - 2(\beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i3}) - \frac{t_i^2}{2e^{2(\beta_0+\beta_1 x_{i1}+\beta_2 x_{i2}+\beta_3 x_{i3})}} \right) \\ & + \delta_{I_i} \ln \left(e^{-\frac{t_{L_i}^2}{2e^{2(\beta_0+\beta_1 x_{i1}+\beta_2 x_{i2}+\beta_3 x_{i3})}}} - e^{-\frac{t_{R_i}^2}{2e^{2(\beta_0+\beta_1 x_{i1}+\beta_2 x_{i2}+\beta_3 x_{i3})}}} \right) \Bigg\}. \end{aligned} \tag{2.10}$$

The first derivatives with respect to β_j ($j = 0, 1, 2, 3$) are:

$$\begin{aligned} \frac{\partial l}{\partial \beta_j} = \sum_{i=1}^{n} \Bigg\{ & -2\delta_{E_i} x_{ij} + \delta_{E_i} \frac{t_i^2 x_{ij}}{e^{2\eta_i}} \\ & + \delta_{I_i} \frac{\frac{t_{L_i}^2 x_{ij}}{e^{2\eta_i}} e^{-\frac{t_{L_i}^2}{2e^{2\eta_i}}} - \frac{t_{R_i}^2 x_{ij}}{e^{2\eta_i}} e^{-\frac{t_{R_i}^2}{2e^{2\eta_i}}}}{e^{-\frac{t_{L_i}^2}{2e^{2\eta_i}}} - e^{-\frac{t_{R_i}^2}{2e^{2\eta_i}}}} \Bigg\}, \end{aligned} \tag{2.11}$$

where

$$\eta_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i3}$$

where $x_{i0} = 1$ for β_0. Numerical methods are used to solve for the MLEs of $\beta_0, \beta_1, \beta_2, \beta_3$. We can not calculate MLEs from the above equations, since not all partial derivatives calculated above attained a closed form. Hence we can utilize numerical procedures to evaluate the parameter's MLE for this model. Newton-Raphson method can be implemented. In this iterative procedure we mainly set an initial value of the parameter and then start the sequential iterative steps until the equation Coverages. One may obtain values of MLEs in this way; however, one must be careful about the property, and the starting value must be selected at random. The accuracy of the parameter's estimate value of the Rayleigh distribution, a simulation study is conducted assuming IC with co-variates for both with and without the estimand framework.

2.6 Simulation

Different proportion of censoring was considered during the simulation. The sample size was also considered from 20, 40 up to 100 to ensure the unbiased-ness of the results. Different kinds of measures were used to check the unbiased-ness, like Bias, RMSE (Root-mean-square error) and coverage probability using Following formula was used. Details of coverage probability calculation is provided in Appendix B

$$\text{Bias} = E(\hat{\theta}) - \theta$$

$$SE = \sqrt{E([\hat{\theta} - E(\hat{\theta})]^2)}$$

$$RMSE = \sqrt{(SE^2) + (E(\hat{\theta}) - \theta)^2}$$

The calculation was done based on 1000 simulations and for different censoring proportion (cp) criteria, and also for different sample sizes (n).

The selection of the true parameter values, including the baseline coefficient $\beta_0 = 0.5$ and the co-variate effects ($\beta_1 = -0.3$, $\beta_2 = 0.4$, and $\beta_3 = 0.2$), is a necessary step in the design of this simulation study. In a simulation, the primary objective is to evaluate the performance of an estimation method by generating data from a model with known, pre-specified parameters. The specific choice of $\beta_0 = 0.5$ is therefore representative but arbitrary; any other reasonable value would serve the same purpose. These values serve as the benchmark for calculating our performance metrics (Bias, RMSE, and coverage probability) under various sample sizes and censoring conditions. To produce actual data from Rayleigh, we need to generate the data from Uniform distribution with the parameters 0, 1 and then apply following transformation using the inverse transformation method.

$$F(t_i) = 1 - e^{-\left(\frac{{t_i}^2}{2e^{-2(\beta_0+\beta_1 x_{i1}+\beta_2 x_{i2}+\beta_3 x_{i3})}}\right)}, (t_i > 0) \tag{2.12}$$

Equate the above equation with u_i which follows the Uniform distribution $u_i \sim U(0, 1)$

$$u_i = 1 - e^{-\left(\frac{{t_i}^2}{2e^{-2(\beta_0+\beta_1 x_{i1}+\beta_2 x_{i2}+\beta_3 x_{i3})}}\right)}, (t_i > 0) \tag{2.13}$$

Taking log both side we can rewrite the equation

$${t_i}^2 = -2log(1 - u_i)e^{-2(\beta_0+\beta_1 x_{i1}+\beta_2 x_{i2}+\beta_3 x_{i3})} \tag{2.14}$$

Finally, we get the equation for generating the time.

$$t_i = \sqrt{(-2log(1 - u_i)e^{-2(\beta_0+\beta_1 x_{i1}+\beta_2 x_{i2}+\beta_3 x_{i3})}} \tag{2.15}$$

We have done both simulation studies without the Estimand setup (Original) and with the Estimand setup (Imputed). We have also tried to understand the nature of the estimates under different proportion of censoring setups. It is very much possible that we may not have the exact information about reasons of the dropout of participants, in that scenario we may not be able to perform the analysis under the Estimand setup. Without the Estimand setup censoring proportion was generated for different combinations as well to test the model. Combination 1 consists of 20 percent IC, combination 2 consists of 40 percent IC, and combination 3 consists of 60 percent IC.

We have also generated data under Estimand setup, where the dropout reason (Lack of efficacy) for a participant was inserted randomly. Under the Estimand framework we have imputed the status of those participants who dropped out from the study and investigated what would have happened to their censoring status under a hypothetical scenario if they had continued. And further non-responder imputation was also applied for the comparison purpose.

3 Measuring Performance of the Estimates

3.1 Without Estimand Setup - Original Data

In this section we have presented the accuracy of the model under different combinations of censoring. Different sample size also considered to check if the error which are measured using Bias, RMSE and coverage probability.

Table 1: Combined Performance Metrics for Original Data Analysis

		β_0			β_1			β_2			β_3		
n	**cp(%)**	**Bias**	**RMSE**	**CP**	**Bias**	**RMSE**	**CP**	**Bias**	**RMSE**	**CP**	**Bias**	**RMSE**	**CP**
	20	-0.057	0.251	0.946	-0.008	0.273	0.938	0.002	0.284	0.932	0.028	0.258	0.936
20	40	-0.058	0.271	0.902	0.001	0.269	0.946	0.004	0.280	0.936	-0.009	0.285	0.926
	60	-0.072	0.270	0.926	0.017	0.262	0.958	0.002	0.291	0.934	0.009	0.282	0.950
	20	-0.025	0.147	0.948	0.006	0.153	0.946	0.005	0.155	0.942	0.005	0.154	0.940
50	40	-0.018	0.149	0.948	-0.005	0.153	0.938	0.009	0.156	0.952	-0.011	0.149	0.956
	60	-0.007	0.146	0.954	-0.010	0.147	0.966	-0.007	0.162	0.940	-0.004	0.161	0.944
	20	-0.007	0.100	0.936	0.000	0.106	0.938	0.000	0.102	0.964	0.002	0.103	0.952
100	40	-0.010	0.103	0.944	0.009	0.106	0.944	-0.002	0.111	0.948	-0.007	0.105	0.944
	60	-0.013	0.104	0.950	-0.004	0.105	0.952	0.003	0.112	0.938	0.005	0.114	0.930

3.2 Under Estimand Setup - Non-Responder Imputation - Composite strategy

Table 2: Combined Performance Metrics for Non-Responder Imputation

		β_0			β_1			β_2			β_3		
n	**cp(%)**	**Bias**	**RMSE**	**CP**	**Bias**	**RMSE**	**CP**	**Bias**	**RMSE**	**CP**	**Bias**	**RMSE**	**CP**
	20	-0.069	0.256	0.946	-0.011	0.276	0.932	0.006	0.287	0.930	0.033	0.262	0.936
20	40	-0.078	0.280	0.900	-0.008	0.279	0.940	0.010	0.287	0.932	-0.004	0.291	0.926
	60	-0.106	0.293	0.894	0.007	0.275	0.940	0.015	0.306	0.918	0.016	0.298	0.940
	20	-0.036	0.152	0.944	0.003	0.155	0.942	0.009	0.158	0.940	0.007	0.156	0.938
50	40	-0.040	0.157	0.940	-0.012	0.158	0.944	0.017	0.161	0.944	-0.008	0.154	0.944
	60	-0.044	0.159	0.940	-0.020	0.156	0.948	0.007	0.169	0.930	0.003	0.168	0.926
	20	-0.019	0.103	0.936	-0.003	0.107	0.936	0.004	0.103	0.956	0.004	0.105	0.952
100	40	-0.034	0.110	0.914	0.002	0.110	0.934	0.006	0.113	0.938	-0.004	0.109	0.934
	60	-0.052	0.121	0.904	-0.014	0.111	0.934	0.016	0.117	0.936	0.011	0.119	0.922

3.3 Under Estimand setup - Logistic Regression MI- Hypothetical Strategy

Table 3: Combined Performance Metrics for Logistic Regression MI

		β_0			β_1			β_2			β_3		
n	**cp(%)**	**Bias**	**RMSE**	**CP**	**Bias**	**RMSE**	**CP**	**Bias**	**RMSE**	**CP**	**Bias**	**RMSE**	**CP**
20	20	-0.057	0.250	0.948	-0.007	0.272	0.940	0.001	0.284	0.936	0.028	0.257	0.938
	40	-0.058	0.269	0.908	0.001	0.267	0.950	0.003	0.278	0.938	-0.009	0.283	0.934
	60	-0.071	0.268	0.928	0.017	0.260	0.958	0.001	0.289	0.938	0.009	0.279	0.948
50	20	-0.024	0.147	0.946	0.006	0.152	0.946	0.005	0.154	0.946	0.005	0.154	0.940
	40	-0.017	0.148	0.950	-0.004	0.152	0.940	0.009	0.154	0.958	-0.011	0.148	0.960
	60	-0.007	0.143	0.956	-0.010	0.145	0.968	-0.007	0.160	0.942	-0.004	0.159	0.946
100	20	-0.007	0.100	0.940	0.000	0.106	0.940	0.000	0.101	0.962	0.002	0.103	0.952
	40	-0.010	0.102	0.944	0.009	0.105	0.948	-0.002	0.110	0.948	-0.007	0.104	0.944
	60	-0.013	0.102	0.954	-0.003	0.103	0.956	0.003	0.111	0.946	0.004	0.111	0.932

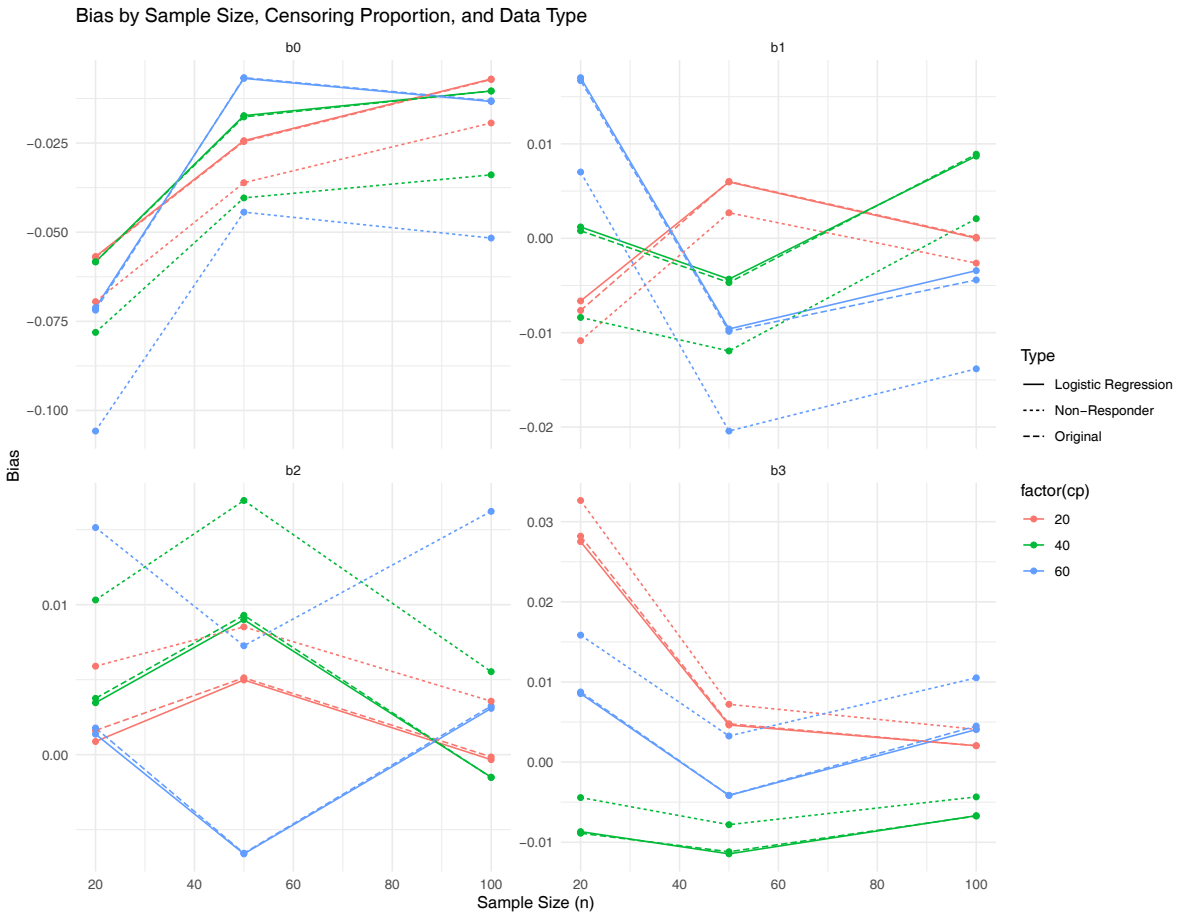

Figure 1: Comparison of Bias for all 3 methods using simulated data

Figure 1 shows the bias of the parameter estimates improves as sample size increases. The Non-Responder imputation method exhibits the most substantial negative bias, particularly for the intercept (β_0) under high censoring. The Logistic Regression MI and Original data methods perform comparably with markedly less bias.

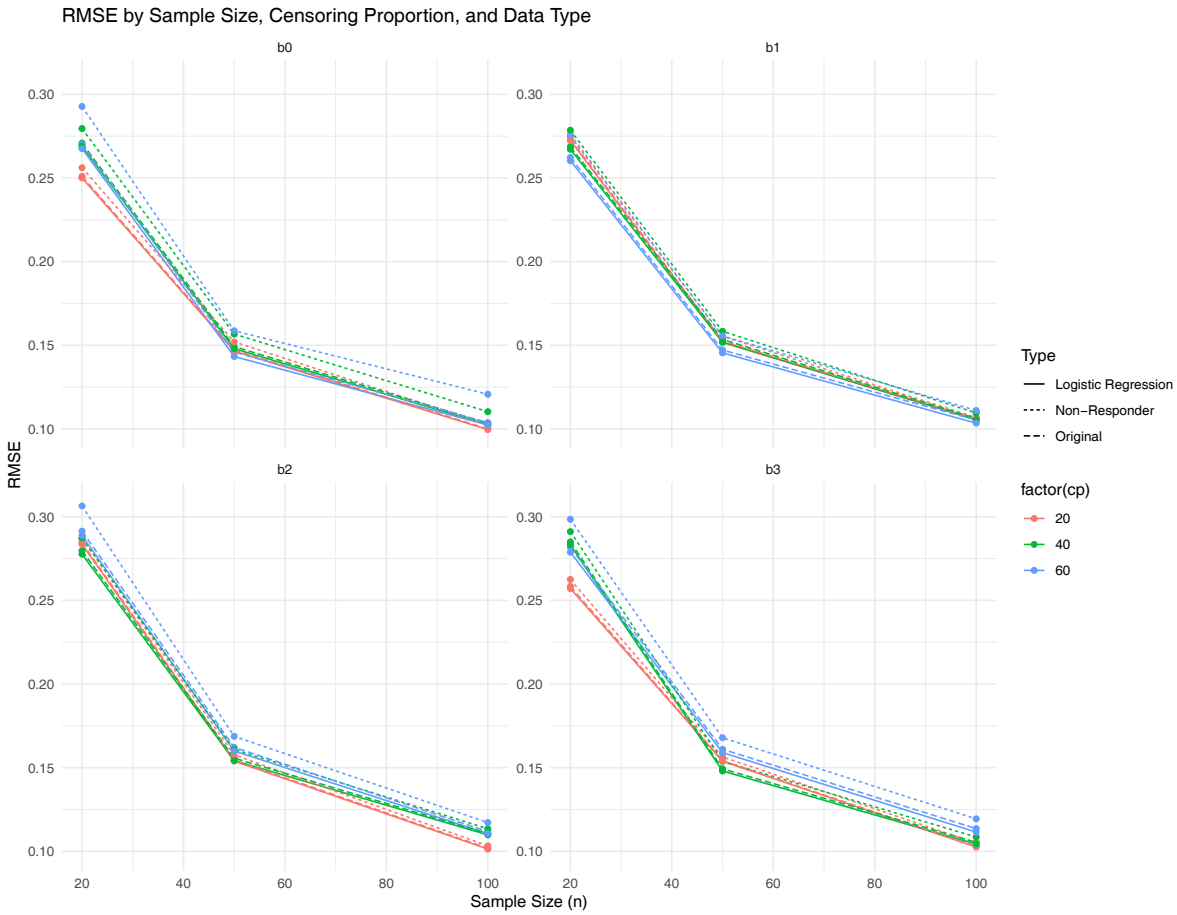

Figure 2: Comparison of RMSE for all 3 methods using simulated data

Figure 2 shows the Root Mean Square Error (RMSE) is predominantly influenced by sample size, decreasing significantly as n increases. All three analytical methods demonstrate comparable performance in terms of overall error, as indicated by their closely clustered RMSE values across all tested conditions.

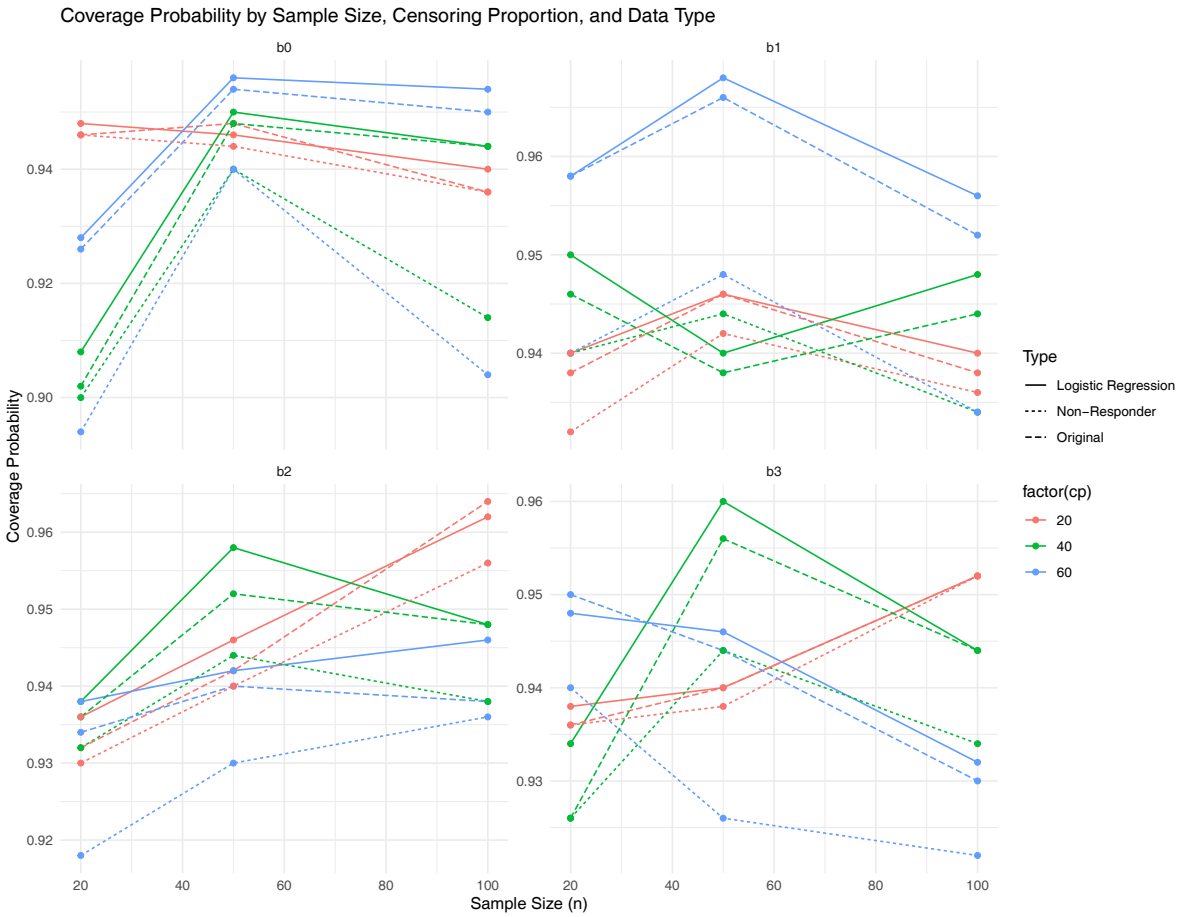

Figure 3: Comparison of Coverage probability for all 3 methods using simulated data

Figure 3 shows the coverage probability for both the Logistic Regression MI and Original data methods remains close to the nominal 0.95 level. In contrast, the Non-Responder imputation frequently results in significant under-coverage, a direct consequence of its larger bias, which causes the corresponding confidence intervals to fail to capture the true parameter value more often.

4 Discussion

This simulation study aimed to test how well a co-variate-adjusted Rayleigh regression model works for interval-censored data when informative censoring is present, as defined by the Estimand framework. We compared three analysis strategies:

- Original (Naive) approach
- Composite strategy using Non-Responder Imputation (NRI)
- Hypothetical strategy using Logistic Regression Multiple Imputation (LRMI)

Our results, shown in Tables 1–3 and Figures 1–3, highlight the statistical impact of each choice. The results confirm that the Maximum Likelihood Estimates for the Rayleigh model are asymptotically consistent. As sample size (n) increased from 20 to 100, the Root Mean Square Error (RMSE) — a measure of precision—dropped steadily across all three methods (Figure 2). This pattern, supported by Tables 1–3, shows that the model becomes more precise with larger samples, regardless of the censoring strategy. Interestingly, RMSE values were similar across methods, meaning the real differences lie in accuracy and bias, not precision.

Furthermore, examination of the Original Data Analysis (Table 1) confirms the robustness of the standard Rayleigh MLE, showing that when censoring is non-informative, bias remains minimal (e.g., $|\text{Bias}| < 0.013$ for $n = 100$) and coverage probability is maintained near the nominal 0.95 level across all tested conditions. Critically, Figure 2 highlights that all the methods tested, regardless of the imputation strategy used, exhibit similarly low Root Mean Square Error (RMSE) values, indicating that the sample size is the dominant factor controlling the precision of the estimation.

Bias tells the bigger story (Figure 1). The NRI method, a conservative composite approach, introduced a strong negative bias, especially for the intercept (β_0). For example, Table 2 shows a bias of -0.106 for β_0 at $n = 20$ with 60% censoring — much higher than other methods. This isn't a coincidence; it's built into NRI's design. By treating all participants who drop out as "failures," NRI overestimates failure rates, skewing the baseline parameter.

This bias affects Coverage Probability (CP) too (Figure 3). The 95% confidence intervals from NRI often fell short, with CP values as low as 0.894 (Table 2, $n = 20, cp = 60\%$) and 0.904 ($n = 100, cp = 60\%$). Such under-coverage means the intervals are too narrow, making statistical inference unreliable and increasing Type I error risk.

The hypothetical LRMI approach performed best. Table 3 shows minimal bias across all parameters, even with heavy censoring. The coverage probabilities stayed close to the nominal 0.95 level. By modeling dropout probability and imputing plausible outcomes (rather than assuming failure), LRMI handles informative censoring well and preserves the integrity of estimates.

Our study, therefore, confirms that in the presence of intercurrent events, the choice of an Estimand and the corresponding analytical strategy is not trivial. While the Rayleigh model is a robust tool for this data structure, its application is highly sensitive to the method used for handling informative censoring. The common "conservative" non-responder imputation, while simple to implement, produces biased estimates and invalid confidence intervals.

These findings are subject to limitations. Our analysis was confined to the Rayleigh distribution and a specific set of co-variates. Future work should explore these dynamics using other parametric models, such as the Weibull distribution, to assess generalizability. Furthermore, the challenge of applying these methods in practice remains, as it is based on the collection of high-quality data on the specific reasons for participant discontinuation, a critical component for implementing any Estimand-based analysis correctly.

5 Conclusion

The main goal of this study was to evaluate a robust statistical approach for analyzing interval-censored time-to-event data under a Rayleigh distribution, framed within the modern Estimand perspective, where informative censoring from intercurrent events poses a significant challenge. The contribution of this work lies in the integration of established components into a unified framework: we deliver a thorough assessment of how advanced imputation strategies perform in this increasingly common clinical trial setting. Our findings aim to provide practical guidance for addressing complex censoring patterns that traditional survival analysis methods often fail to capture adequately. Our research confirms that the Rayleigh regression model is a precise and effective tool for handling interval-censored data, especially when paired with a thoughtful imputation strategy.
We believe that participants who stopped the trial treatment because it was not working due to certain personal characteristics. Within the framework of interval-censored data, the Rayleigh regression model augmented with co-variates exhibits a high degree of precision in estimating model parameters. our Findings from the Simulation study are as follows

1. **Model accuracy improves with larger sample sizes.** A clear downward trend in Root Mean Square Error (RMSE) is observed as sample size increases, as illustrated in Figure 2. This indicates that the model demonstrates strong robustness and reliability with more data.

2. **Handling missing data is critical.** Our analysis shows that logistic regression-based multiple imputation consistently outperforms the simpler non-responder imputation approach, yielding estimates with lower bias (Figure 1) and more reliable coverage probabilities (Figure 3).

3. **Simplistic methods can be misleading.** Treating all dropouts due to lack of efficacy as events (the non-responder method) introduces bias because it ignores patient-level heterogeneity that may influence outcomes.

The central contribution of this paper is the clear empirical evidence that the choice of imputation strategy is critical for obtaining unbiased estimates. Among various imputation strategies, logistic regression-based imputation has consistently shown superior performance in terms of coverage probability. In contrast, non-responder imputation where all participants who dropped from the clinical trial due to lack of drug efficacy are treated as having experienced the event can introduce bias. This is because such an approach fails to account for the heterogeneity

in individual characteristics, potentially skewing the results and undermining the validity of the conclusions. Therefore, using an appropriate imputation method such as logistic regression-based multiple imputation followed by a suitable parametric model can significantly improve the accuracy of the study's outcomes. Ultimately, this work demonstrates a validated pathway for researchers to derive more meaningful and reliable conclusions from complex, real-world clinical trial data.

Acknowledgments

We gratefully acknowledge the guidance and sincerely thank the Editor-in-Chief and anonymous reviewers for their valuable feedback that strengthened this work.

Conflict of Interest

No conflicts of interest have been declared by the authors.

References

Allison, P. D. (2005). Imputation of categorical variables with proc mi. *SUGI 30 proceedings*, *113*(30), 1–14.

Al-Noor, N., & Assi, N. (2020). Rayleigh-rayleigh distribution: properties and applications. In *Journal of physics: Conference series* (Vol. 1591). doi: 10.1088/1742-6596/1591/1/012038

Anis, M., Okorie, I., & Ahsanullah, M. (2024). A review of the rayleigh distribution: properties, estimation & application to covid-19 data. *Bulletin of the Malaysian Mathematical Sciences Society*, *47*(1), 6.

Elviana, E., & Purwadi, J. (2020). Parameters estimation of rayleigh distribution in survival analysis on type ii censored data using the bayesian method. In *Journal of physics: Conference series* (Vol. 1503). doi: 10.1088/1742-6596/1503/1/012004

Heinrich, M., Zagorscak, P., Bohn, J., Knaevelsrud, C., & Schulze, L. (2025). Using the ich estimand framework to improve the interpretation of treatment effects in internet interventions. *npj Digital Medicine*, *8*(1). doi: 10.1038/s41746-025-01936-0

Jammalamadaka, S. R., & Leong, E. (2015). Analysis of discrete lifetime data under middle-censoring and in the presence of covariates. *Journal of Applied Statistics*, *42*(4), 905–913.

Keene, O. N. (2019). Strategies for composite estimands in confirmatory clinical trials: examples from trials in nasal polyps and steroid reduction. *Pharmaceutical Statistics*, *18*(1), 78–84.

Lee, K. J., & Simpson, J. A. (2014). Introduction to multiple imputation for dealing with missing data. *Respirology*, *19*(2), 162–167.

Leung, K.-M., Elashoff, R. M., & Afifi, A. A. (1997). Censoring issues in survival analysis. *Annual review of public health*, *18*(1), 83–104.

Ozenne, B. M. H., Scheike, T. H., Stærk, L., & Gerds, T. A. (2020). On the estimation of average treatment effects with right-censored time to event outcome and competing risks. *Biometrical Journal*, *62*(3), 751–763.

Psychogyios, K., Ilias, L., Ntanos, C., & Askounis, D. (2023). Missing value imputation methods for electronic health records. *IEEE Access*, *11*, 21562–21574.

Rubin, D. B. (1996). Multiple imputation after 18+ years. *Journal of the American statistical Association*, *91*(434), 473–489.

Rufibach, K. (2019). Treatment effect quantification for time-to-event endpoints–estimands, analysis strategies, and beyond. *Pharmaceutical statistics*, *18*(2), 145–165.

Schober, P., & Vetter, T. R. (2018). Survival analysis and interpretation of time-to-event data: the tortoise and the hare. *Anesthesia & Analgesia*, *127*(3), 792–798.

Shih, W. J. (2002). Problems in dealing with missing data and informative censoring in clinical trials. *Current controlled trials in cardiovascular medicine*, *3*, 1–7.

Sparling, Y. H., Younes, N., Lachin, J. M., & Bautista, O. M. (2006). Parametric survival models for interval-censored data with time-dependent covariates. *Biostatistics*, *7*(4), 599–614.

A Appendix A: Logistic Regression Multiple Imputation (MI)

A.1 IC in Survival Data

Let T_i denote the true (unobserved) time to the event for participant i. In IC:

- The exact event time is known only to lie within an interval $T_i \in (L_i, R_i)$, where L_i is the time of the last assessment showing no event, and R_i is the first time the event was observed.
- We define a binary variable Y_i for the participant's status, where $Y_i = 0$ indicates an event and $Y_i = 1$ indicates censoring.
- For event participants ($Y_i = 0$), the event time is interval-censored: $T_i \in [L_i, R_i]$.
- For censored participants ($Y_i = 1$), the event time is right-censored: $T_i > R_i$.

In this study, participants who dropped out due to lack of efficacy have an uncertain status Y_i. We address this using logistic regression to model the probability of being censored, followed by multiple imputation.

A.2 Logistic Regression for Censoring Flag Prediction

To predict the censoring flag $Y_i \in \{0, 1\}$, we use a logistic regression model. Let $x_i = (x_{i1}, x_{i2}, x_{i3})$ be the vector of co-variates (e.g., gender, smoking status, age group) for participant i. The logistic model assumes the log-odds of being censored is a linear combination of these co-variates:

$$\text{logit}(P(Y_i = 1|x_i)) = \ln\left(\frac{P(Y_i = 1|x_i)}{P(Y_i = 0|x_i)}\right) = \gamma_0 + \gamma_1 x_{i1} + \gamma_2 x_{i2} + \gamma_3 x_{i3}, \tag{A.1}$$

where γ_0 is the intercept and $\gamma_1, \gamma_2, \gamma_3$ are the regression coefficients for the imputation model. The probability of being censored is then:

$$P(Y_i = 1|x_i) = \frac{\exp(\gamma_0 + \sum_{j=1}^{3} \gamma_j x_{ij})}{1 + \exp(\gamma_0 + \sum_{j=1}^{3} \gamma_j x_{ij})}. \tag{A.2}$$

The probability of having an event is $P(Y_i = 0|x_i) = 1 - P(Y_i = 1|x_i)$. The model parameters $\gamma = (\gamma_0, \gamma_1, \gamma_2, \gamma_3)$ are estimated by maximizing the likelihood function using participants with a known status.

A.3 Multiple Imputation for Uncertain Censoring Flags

We use multiple imputation (MI) to account for the uncertainty in Y_i for participants who dropped out due to lack of efficacy. This process creates M complete datasets.

A.3.1 Imputation Process

1. **Imputation:**
   - Fit the logistic regression model (A.1) using participants with observed Y_i to obtain estimates $\hat{\gamma}$ and their variance-covariance matrix, Cov$(\hat{\gamma})$.
   - For each of the M imputations (indexed by $m = 1, ..., M$) and for each participant i with an unobserved Y_i:
     - Draw a set of coefficients $\gamma^{(m)}$ from a multivariate normal distribution $\mathcal{N}(\hat{\gamma}, \text{Cov}(\hat{\gamma}))$ to reflect parameter uncertainty.
     - Compute the probability of being censored for that participant using the drawn coefficients:
     $$P(Y_i = 1|x_i; \gamma^{(m)}) = \frac{\exp(x_i^T \gamma^{(m)})}{1 + \exp(x_i^T \gamma^{(m)})}.$$
     - Impute the missing status $Y_i^{(m)}$ by drawing from a Bernoulli distribution with the computed probability: $Y_i^{(m)} \sim$ Bernoulli$(P(Y_i = 1|x_i; \gamma^{(m)}))$.
   - This creates M completed datasets.
2. **Analysis with Interval-Censored Survival Models:**
   - For each imputed dataset m, perform the main survival analysis using the Rayleigh model. The likelihood function for the m-th dataset is:
   $$L^{(m)} = \prod_{i:Y_i^{(m)}=0} [S(L_i|\beta) - S(R_i|\beta)] \prod_{i:Y_i^{(m)}=1} S(R_i|\beta), \tag{A.3}$$
   where $S(t|\beta)$ is the survival function of the Rayleigh distribution. As defined in the main text, this is $S(t) = \exp(-t^2/(2\sigma_i^2))$, where the scale parameter σ_i is modeled as a function of the co-variates x_i and the primary model coefficients β:
   $$\sigma_i = \exp(\beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i3}).$$
   - For each dataset, estimate the parameter vector of interest $\hat{\beta}^{(m)}$ and its associated variance-covariance matrix $V^{(m)}$.
3. **Pooling:**

- Combine the results from the M imputed datasets using Rubin's rules. For a single parameter of interest, β_j:
- The pooled point estimate is the average of the individual estimates:

$$\bar{\beta}_j = \frac{1}{M}\sum_{m=1}^{M}\hat{\beta}_j^{(m)}.$$

- The within-imputation variance is the average of the individual variances:

$$\bar{V}_j = \frac{1}{M}\sum_{m=1}^{M}V_{jj}^{(m)}.$$

- The between-imputation variance measures the variation across the imputed datasets:

$$B_j = \frac{1}{M-1}\sum_{m=1}^{M}(\hat{\beta}_j^{(m)} - \bar{\beta}_j)^2.$$

- The total variance combines both sources of uncertainty:

$$T_j = \bar{V}_j + \left(1 + \frac{1}{M}\right)B_j.$$

B Appendix B: Coverage Probability Calculation

The accuracy of the Wald confidence intervals for the Rayleigh model parameters $\beta = (\beta_0, \beta_1, \beta_2, \beta_3)$ is evaluated through the coverage probability.
Let $\hat{\beta}_j$ be the maximum likelihood estimate for a parameter β_j. The standard error, SE($\hat{\beta}_j$), is the square root of the j-th diagonal element of the estimated variance-covariance matrix, $\hat{V}$. The $(1-\alpha)$ Wald confidence interval is:

$$\hat{\beta}_j \pm z_{1-\alpha/2}\sqrt{\hat{V}_{jj}}. \tag{B.1}$$

The coverage probability for β_j, denoted $CP(\beta_j)$, is the probability that this random interval contains the true parameter value:

$$CP(\beta_j) = P(\hat{\beta}_j - z_{1-\alpha/2}\sqrt{\hat{V}_{jj}} \leq \beta_j \leq \hat{\beta}_j + z_{1-\alpha/2}\sqrt{\hat{V}_{jj}}). \tag{B.2}$$

This probability is estimated via simulation by generating N datasets, constructing a confidence interval for each, and calculating the proportion of intervals that contain the true β_j. For each replication $k = 1, ..., N$, we define an indicator variable:

$$I_{jk} = \begin{cases} 1, & \text{if } \beta_j \text{ is in the interval from replication } k, \\ 0, & \text{otherwise.} \end{cases} \tag{B.3}$$

The estimated coverage probability is the mean of these indicators:

$$\widehat{CP}(\beta_j) = \frac{1}{N}\sum_{k=1}^{N}I_{jk}. \tag{B.4}$$

In the simulation, the true parameters $\beta = (0.5, -0.3, 0.4, 0.2)$ are used to generate the data.