Journal of Engineering and Technology for Industrial Applications

ITEGAM-JETIA

Manaus, v.12 n.57, p. 1074-1083. January/February, 2026.
DOI: https://doi.org/10.5935/jetia. v12i57.3235

ISSN ONLINE: 2447-0228

RESEARCH ARTICLE OPEN ACCESS

OF AHERO RICE SURVIVAL ANALYSIS WITH CEMPETING RISKS: INSIGHTS FROM THE FINE AND GRAY MODEL AND COVARIATE ADJUSTMENTS

Sumithra B[1], Anu Swedha Ananthan[2], Zechariah Jebakumar Arulanantham[3]

[1*]Assistant Professor, Department of Mathematics, Faculty of Science & Humanities, SRM Institute of Science and Technology, Ramapuram, Chennai – 600089, Tamil Nadu, India.
[2] Associate Professor, Department of Microbiology, Justice Basheer Ahmed Sayeed College for Women, Teynampet, Chennai – 600018, Tamil Nadu, India.
[3] Lecturer for Biostatistics, Vice Deanship of Postgraduate Studies and scientific research, Prince Sultan Military college of Health science, AD Dammam 343137761, Kingdom of Saudi Arabia.

[1]http://orcid.org/0000-0001-9384-9318, [2]http://orcid.org/0000-0002-4965-3278, [3]http://orcid.org/0000-0002-5978-2436

Email: *sumithrb@srmist.edu.in, anuswedha.a@jbascollege.edu.in, zacbiostat@gmail.com

ARTICLE INFO

Article History
Received: January 4, 2025
Revised: January 20, 2026
Accepted: January 30, 2026
Published: February 28, 2026

Keywords:
Competing Risks
Fine and Gray Model
Survival Analysis
Cumulative incidence Function
Kaplan-Meier Survival

ABSTRACT

Competing risk survival analysis is deemed necessary when manifold events, such as recurrence and death, are present to prevent the incidence of the event of interest. The application of the Fine and Gray model for cumulative incidence function (CIF) estimation and competing risk regression in a simulated dataset is explored in this study. Simulated time-to-event data and covariates are used to estimate the Fine and Gray model both with and without covariates, with comparisons being made against the Kaplan-Meier (KM) method, which does not account for competing risks. It is demonstrated that the Fine and Gray method provides a more accurate representation of event-specific incidence in the presence of competing risks. Competing risks being ignored, as in the Kaplan-Meier approach, can lead to an underestimation of the event of interest. Additionally, the inclusion of covariates like age and tumor size into the Fine and Gray model is shown to significantly impact the incidence of recurrence and competing events. This comparative analysis is provided.

Copyright ©2026 by authors and Galileo Institute of Technology and Education of the Amazon (ITEGAM). This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

I. INTRODUCTION

I.1 COMPETING RISKS

In survival analysis, the primary focus is typically placed on the time awaiting an event to occur, such as death, failure of a system, or relapse in disease. However, competing risks are encountered when multiple types of events can be practiced by individuals in a study, and the occurrence of one event may avert the occurrence of others. For example, in a cancer study, a patient may die from cancer (the primary event of interest) or from another cause like a heart attack (a competing event). In the presence of competing risks, traditional survival analysis methods must be adjusted to account for the fact that competing events can censor or modify the probability of the primary event being observed.

I.2 KM ESTIMATOR

The KM estimator is a non-parametric statistic that is used to estimate the survival function from life span data. An estimate of the probability that an individual will survive further than a certain time point is offered by it. However, in the context of competing risks, the occurrence of competing events is not adequately accounted for by the KM method, as all competing events are treated as censored. As a result, alternative methods, such as the CIF, are frequently used to estimate the probability of precise events in competing risks settings.

I.3 EXPONENTIAL DISTRIBUTION

The exponential distribution is a common method for survival information where the event rate (hazard) is constant over time. Time to event data is often modeled using this distribution when events occur at a constant hazard rate, without any competing risks. The exponential distribution is memoryless, denoting that the probability of an event occurring in the next time interval is independent of how much time has already passed. While being simple and mathematically tractable, this model can be considered overly simplistic for real-world survival data, which often has hazards that vary over time or are influenced by covariates.

I.4 FINE AND GRAY METHOD

The Fine and Gray sub-distribution hazard method is a regression model designed to handle competing risks data. Unlike traditional Cox proportional hazards models, which are designed for cause-specific hazards, the sub-distribution hazard is focused on by the Fine and Gray model, which quantifies the risk of the event of interest occurring, while accounting for the presence of competing events. More accurate estimation of the cumulative incidence is allowed by this method and it is broadly used in medical research where competing risks are prevalent. In summary, when dealing with time-to-event data, the appropriate statistical model must be considered, particularly when competing risks are present. The Kaplan-Meier estimator is found to be useful in simpler cases, while the Fine and Gray method is deemed more appropriate in the presence of competing events. A basic understanding of survival time is offered by the exponential distribution, but more sophisticated approaches are often required by real-world data.

II. THEORETICAL REFERENCE

The semi-parametric proportional sub-distribution hazards model was introduced to directly model cumulative incidence in competing-risks settings [1]. Estimation via weighted partial likelihood is presented in the paper, and it is shown how direct interpretation of covariate effects on the marginal probability (CIF) rather than cause-specific hazards is allowed by the model. This is regarded as the foundational reference for sub-distribution hazard modeling. Best practices for reporting Fine–Gray analyses [2] (interpretation of sub-distribution hazard ratios, model checks, and presentation of absolute risks) are discussed. Transparent reporting of censoring assumptions, proportionality checks, and presentation of CIFs alongside SHR is emphasized. Common misinterpretations are to be avoided by applied researchers. Cause-specific hazards and sub-distribution approaches are reviewed, interpretations are contrasted, and available software is surveyed [3]. Tradeoffs are highlighted: cause-specific models for etiologic questions vs. Fine–Gray for absolute risk prediction, and comments on covariate adjustment strategies are made. According to [4] Diagnostic checks for Fine–Gray models (proportionality, influential observations, mis-specification) are presented. Residuals and time-varying coefficient checks adapted to sub-distribution hazards are recommended important because standard Cox diagnostics do not directly translate. Fine–Gray is extended to clustered/correlated observations (marginal/regression approaches with robust variance). [5] It is shown how covariate effects on CIFs can be estimated while accounting for within-cluster correlation (e.g., multi-center studies). Methods for clustered competing-risks data (cause-specific vs sub-distribution, frailty vs marginal) are systematically compared [6]. It is found that both inference and prediction are affected by the choice of model and covariate adjustment; sensitivity analyses are recommended.

According to [7] A modern compendium is covering classical and ML approaches for competing risks (including Fine–Gray variants, penalized methods, and random survival forests adapted for competing risks). It is noted how covariate adjustment, variable selection, and prediction calibration are being recognized as active research areas. Competing risks in clinical trials [8] It is argued that competing risks are often mishandled in trials; caution is given against mechanical Fine–Gray use without checking assumptions, and it is suggested that sometimes alternative multi-state or cause-specific frameworks for trial inference are used. Pitfalls of covariate adjustment when competing events are informative are highlighted. Simultaneous Fine–Gray models for multiple outcomes [9] Estimation strategies to model multiple CIFs using Fine–Gray are discussed, along with the interpretation limits when multiple sub-distribution hazards are modeled separately. A warning is given about comparing SHRs across causes, and the suggestion is made to report absolute risk differences. Stratified FG and covariate-adjusted censoring weights [10] Stratified proportional sub-distribution hazards with covariate-adjusted censoring weights are proposed to handle covariate-dependent censoring and complex designs (case-cohort), with valid covariate effect estimates on CIF being improved. Imputation & missing covariates for Fine–Gray [11] Multiple imputation strategies for missing predictors when Fine–Gray is used are highlighted, showing that naive complete-case analyses can bias SHR estimates; imputation models respectful of competing-risks structure are recommended. Practical primer [12] An accessible overview of competing risks concepts for clinicians is provided: the difference between cause-specific hazards and sub-distribution hazards is explained, how covariates are interpreted in each is described, and practical examples emphasizing appropriate covariate adjustment for prediction vs causal inference are illustrated.

Pedagogical note / textbook treatment [13] Theoretical properties are reviewed and practical guidance is given on implementing Fine–Gray, with an emphasis placed on the fact that the sub-distribution hazard is a mathematical construct tied to the CIF and requires careful thought about which covariates should be adjusted for depending on the estimand. Recent applied example [14] Fine–Gray is used to study time to loss-to-follow-up with competing events in infectious-disease data; covariate adjustment, model checking, and CIF plotting are demonstrated, illustrating how interpretation changes relative to cause-specific hazard models. A useful concrete application is provided. Cautions on multiple Fine–Gray usage [15] Limitations are pointed out when multiple Fine–Gray models are used to summarize absolute risks of several outcomes simultaneously; careful interpretation and combined reporting strategies (e.g., absolute CIFs) are advised to avoid misleading conclusions from SHRs alone. Considerations in cardiology studies [16] How cardiology researchers should choose between cause-specific and sub-distribution approaches is reviewed, particularly when covariate adjustment is motivated by etiologic vs prognostic aims; recommended checks and sensitivity analyses are described. Debate on default methods & interpretability [17] the tendency to default to cause-specific or Fine–Gray models is critically examined, and clarity on research questions is called for: covariate adjustment must be aligned with whether the question is about marginal risk prediction or conditional/hypothetical causal effects.

Technical report: stratified sub-distribution hazards [18] Technical details for stratified proportional sub-distribution hazards are developed and guidance is given on selecting strata and covariate adjustments when baseline CIF differs across strata. It is considered helpful for complex sampling designs. Fine-Gray under length-biased sampling/extensions [19] Fine–Gray estimation under nonstandard sampling (length bias) is explored, with corrections proposed to maintain valid covariate effect inference on CIF when sampling departs from random. The model's extensibility to practical sampling issues is shown. Imputation combined with sub-distribution weights to handle missing outcomes/covariates in competing-risks analyses is proposed by multiple imputation & sub-distribution-weight imputation [20]; it is demonstrated how SHR estimates and CIF predictions are stabilized through covariate adjustment after appropriate imputation.

III. MATERIALS AND METHODS

The study applied competing risk survival analysis using the **Fine and Gray method** to a simulated dataset. Methodology involved the following steps:

1. **Simulated Dataset**: A dataset of 569 samples was generated, including time-to-event data, event types (recurrence and death), and covariates (age and tumor size). The time-to-event data was drawn from an exponential distribution, while event types (censoring, recurrence, and death) were simulated with predefined probabilities.
2. **Fine and Gray Method**: The Fine and Gray sub-distribution hazard model was employed to estimate cumulative incidence functions (CIFs) for the competing events of recurrence and death. The model was first fit without covariates and then with covariates to analyze their influence on the event of interest and competing risks.
3. **KM Comparison**: Overall survival was approximated using the KM method, with competing risks being ignored. A comparative evaluation between methods was allowed, highlighting the impact of accounting for competing events.
4. **Statistical Analysis**: The R programming language was used to fit Fine and Gray models and plot CIFs with the cmprsk package. Kaplan-Meier estimates were generated using the survival package. The results were compared to show how adjustments for competing risks are made by the Fine and Gray method, in contrast to the approach taken by Kaplan-Meier, which treats competing events as censored.

Dataset

- **Size**: 569 samples.
- Features:
  - **Covariates**: Age and tumor size (simulated from normal distributions).
  - **Time-to-event**: Simulated as of an exponential distribution representing the time in days until an event.
  - **Event types**: Censoring (60%), recurrence (25%), and death (15%).
- **Event of Interest**: Recurrence of disease.
- **Competing Event**: Death before recurrence.

Comparison of Kaplan-Meier and Exponential Distribution

The KM estimator and the Exponential distribution are both widely used in survival analysis to model time-to-event data. However, significant differences are found in their underlying assumptions, use cases, and interpretability.

1. Assumptions

- Kaplan-Meier Estimator:
  - Non-parametric method, meaning it makes no assumption about the fundamental distribution of the survival times.
  - Can handle censoring, where the exact event time is unknown for some subjects.
  - It estimates the survival function straight from the observed data and can accommodate varying hazard rates over time.
- Exponential Distribution:
  - Parametric method, assuming the survival times follow an Exponential distribution**.**
  - Assumes constant hazard rate (i.e., the probability of an event taking place in next time interval is the same regardless of how much time has passed).
  - Simplicity in the model makes it less flexible when hazard rates change over time.

2. Flexibility

- Kaplan-Meier Estimator:
  - Extremely flexible and can be used in a variety of scenarios without making burly assumptions about the form of the survival distribution.
  - It adapts to the data and provides an empirical survival curve, showing how survival probability changes over time.
  - Works welling cases where hazard rates are not constant, such as when risks increase or decrease over time.

- Exponential Distribution:

  - Much more restrictive due to the constant hazard assumption. It works best when the data genuinely follow an exponential survival distribution.
  - If hazard rates are time-dependent (e.g., initially low but increasing with time), the exponential distribution will not provide an accurate fit.

3. Interpretation

- Kaplan-Meier Estimator:

  - Provides a step wise survival curve based on actual observed data, where the survival probability changes at the time of each event.
  - The result is an empirical estimation of survival at any given point in time, giving a clear and direct view of the survival pattern.

- Exponential Distribution:

  - Provides a smooth survival curve based on a mathematical function with a constant hazard rate. The survival function declines exponentially over time.
  - The result is ananalytical expression for survival, which is simple but may not reflect the actual survival patterns of more complex data sets.

4. Use Cases

- Kaplan-Meier Estimator:

  - Often used in clinical trials, medical research, or engineering studies where event times and censoring are observed, but no strong assumptions about hazard rates can be made.
  - Ideal for data sets where the survival probability varies overtime.

- Exponential Distribution:

  - Commonly used in scenarios where the hazard rate is known to be constant, such as in certain types of reliability engineering or failure analysis.
  - Suitable for modeling systems or components with memory less properties, like light bulbs or electronic components with a constant failure rate.

5. Handling Censoring

- Kaplan-Meier Estimator:

  - Handles censored data well, meaning it can accommodate individuals whose event time is unknown or who did not practice the event by the end of the study.

- Exponential Distribution:

  - Also handles censored data but within the constraints of its parametric form. If data do not convene the assumption of constant hazard, censoring can complicate the interpretation of results.

6. Survival Function

- **Kaplan-Meier**: The survival function is computed step wise at each observed event time.

$$\hat{S}(t)= \prod_{t(i)\leq t} \frac{ni-di}{ni}$$

with the convention that $\hat{S}(t) = 1$ if $t < t_{(1)}$.where $d_i$is the number of events at time $t_i$and $n_i$is the number of subjects at risk just prior to time t_i

- **Exponential Distribution**: Survival function has a specific mathematical form.

$$S(t)=e^{-\lambda t} \tag{1}$$

Where $\lambda$is the constant hazard rate, and t is time.

Dataset Simulation

A dataset of 569 samples with two covariates, age and tumor size, will be simulated, and time-to-event data will be generated from an exponential distribution. The event types will include censoring, recurrence, and death, with predefined probabilities.

Explanation of Code:

1. Covariates:

- Age: Simulated from a normal distribution with a mean of 50 years.
- tumor_size:Simulatedfromanormaldistributionbyameanof3cm.

2. Event Times:

- Time-to-event data is drawn from an exponential distribution with a rate of 0.05, representing the number of days that an event is awaited to occur.

3. Event Types:

- 0 for censoring (60%ofthedata),
- 1 for recurrence (25%),
- 2 for death (15%).

4. Time-to-Event (Exponential Distribution):

- Exponential Distribution (rexp ()) is used to generate time-to-eventdata,which represents the time in days until recurrence, death, or censoring happens.
- Lambda (rate = 0.05) controls the hazard rate for event occurrences. In this case, the expected time-to-event is 20 days (1/0.05).

5. Cumulative Incidence Function (CIF):

- The Fine and Gray model is used to estimate the CIF for recurrence (event of interest) and death (competing risk). The cuminc() function in R generates these estimates.
- The plot visualizes CIF for recurrence (blue) and death (red), showing the incidence of each event over time.

6. Kaplan-Meier Method:

- The Kaplan-Meier estimator ignores competing risks and considers only event of interest (recurrence). In this case, competing events like death are treated as censored.
- The Kaplan-Meier curve is plotted in green, showing the estimated survival probability for recurrence.
- The survfit() function estimates overall survival, ignoring competing risks, to compare with the Fine and Gray method.

7. Comparison:

- The Fine and Gray CIF gives the true cumulative incidence of recurrence and death, accounting for competing risks.
- Fine and Gray with Covariates: The crr() function includes covariates (age and tumor size) in the model.
- The Kaplan-Meier method, on the other hand, over estimates the probability of survival because it does not account for the competing risk of death.
- By comparing the two plots, you can see that the Fine and Gray model provides a more realistic estimation when competing events are present.

Table 1: Simulated Dataset.

Time	status	tumor_size
6.067325	0	3.079502
67.558109	0	3.961264
22.648645	0	1.543534
5.144028	1	2.218260
1.103972	0	3.320402
34.969098	2	2.555218

Source: Authors, (2026).

Table 2: Estimated Fine and Gray Model using Competing Risk.

Variables	coef	exp(coef)	se(coef)	z	p-value	2.5%	97.5%
Age	0.0104	1.010	0.00922	1.131	0.26	0.992	1.03
tumor_size	-0.0463	0.955	0.08377	-0.553	0.58	0.810	1.13

Num. cases=569

Pseudo Log-likelihood=-757

Pseudo likelihood ratio test=1.49on2df.

Source: Authors, (2026).

The effect of each variable (age and tumor size) on the hazard of the event of interest (e.g., recurrence) is represented by the coefficient. A positive value (for age: 0.0104) suggests that the hazard of the event is slightly increased by an increase in age. A negative value (for tumor size: -0.0463) suggests that the hazard of the event is slightly decreased by a larger tumor size. This is exponentiated coefficient, representing the hazard ratio (HR).

It tells us the multiplicative result of a one-unit increase in the variable on the hazard:

Age (1.010): 1-year increase in age is connected with a 1% increase in the hazard of recurrence (hazard ratio=1.010).Tumor size (0.955): A 1cm increase in tumor sizeisassociatedwitha4.5% decrease in the hazard of recurrence (hazard ratio = 0.955). Standard Error of the Coefficient represents the uncertainty around the estimated coefficients. Smaller values indicate more precise estimates. Age has a standard error of 0.00922,while tumor size has a standard error of 0.08377, suggesting that the tumor size estimate is more uncertain. Z-statistic is the test statistic used to determine if the coefficient is significantly different from zero. It is calculated as coef / se(coef).For age (1.131) and tumor size (-0.553), the z-values are relatively small, indicating weak facts against the null hypothesis.

Value tests the null hypothesis to the coefficient is zero (i.e., there is no effect). A p-value below 0.05 designates statistical significance. Age (0.26): The p-value is high, indicating that age is not statistically significantly associated with the event of interest. Tumor size (0.58): Similarly, tumor size is not statistically significant. 2.5% and 97.5% Confidence Intervals represent the 95% confidence intervals for the hazard ratios (exp (coef)), providing arrange within which the true hazard ratio is likely to lie. For age(0.992 to 1.03), the confidence interval suggests that the effect of age is very close to 1, further indicating that age has a minima limpact on the hazard. For tumor size (0.810to1.13), the confidence interval crosses 1, also indicating that the effect of tumor size is not significant.

Table 3: Estimated Kaplan-Meier Survival Probability.

Time	No. of risk	No. of event	Survival	Std. error	Lower 95%	Upper 95%
0.111	566	1	0.998	0.00177	0.995	1
0.444	560	1	0.996	0.00251	0.992	1
0.505	557	1	0.995	0.00307	0.989	1
0.608	556	1	0.993	0.00355	0.986	1
0.616	554	1	0.991	0.00397	0.983	0.999
0.783	550	1	0.989	0.00435	0.981	0.998
.	.	.	.	.	.	.
1.68	528	1	0.98	0.00594	0.969	0.992
.	.	.	.		.	.
9.61	341	1	0.886	0.0147	0.858	0.916
9.665	339	1	0.884	0.01489	0.855	0.913
9.731	337	1	0.881	0.01508	0.852	0.911
.	.	.	.	.	.	.
80.67	11	1	0.397	0.05817	0.298	0.529
84.126	10	1	0.357	0.06449	0.251	0.509
94.213	4	1	0.268	0.09122	0.137	0.522

Source: Authors, (2026).

This table presents the survival analysis data, likely from a Kaplan-Meier estimate, with key statistics such as the number of individuals at risk, the number of events (death so recurrences), survival probabilities, and their respective confidence intervals. Here's a step-by-step interpretation:

1. Early Time Points (0.111 to 1.68 months):

- The survival probabilities remain very high, with survival rates between 0.993 and 0.980 over this period.
- Standard errors (Std.error) are small (ranging from 0.00177 to 0.00594), indicating high precision in the survival estimates.
- The 95% confidence intervals are very tight (e.g., 0.995 to 1 or 0.969 to 0.992), suggesting high confidence in the estimates.
- There are very few events (only 1 at every time point), and the number of individuals at risk gradually decreases (from 566 to 528), indicating only a slight drop in the survival probability.

2. Mid Time Points (up to9.665 months):

- By 9.61 months, the survival probability drops to 0.886, indicating that approximately 88.6% of the cohort has survived or remained event-free by this time.
- The standard error increases to 0.0147, reflecting greater uncertainty in the estimate as fewer people remain at risk.
- The confidence intervals widen slightly (e.g., 0.855 to 0.913), showing increasing uncertainty but still relatively precise.

3. Later Time Points (after 80 months):

- By 80.67 months, the survival probability has dropped significantly to 0.397, meaning only around 39.7% of individuals remain alive or event-free.
- The standard error increases significantly to 0.05817, reflecting growing uncertainty due to the small number of individuals remaining at risk (11 on this point).
- Confidence intervals widen dramatically (0.298 to 0.529), indicating high uncertainty in the survival estimate.
- By 94.213 months, survival drops further to 0.268, with very wide confidence intervals (0.137to0.522), suggesting that the survival probability estimate is highly uncertain because only 4 individuals remain at risk.

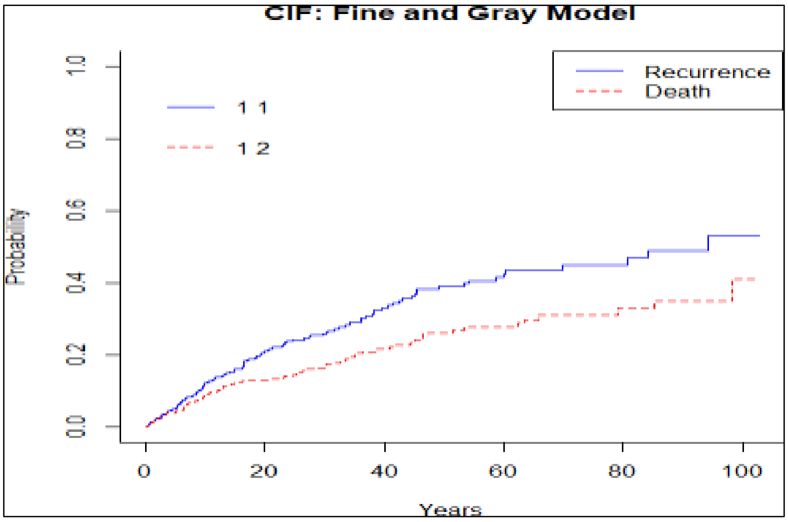

Figure 1: Survival Curve for Fine and Gray Model using Competing risk.
Source: Authors, (2026).

This plot shows the CIF from the Fine and Gray method, used in the context of competing risks survival analysis. The x-axis corresponds to time (in years), and the y-axis demonstrates the probability of events occurring over time. The **Fine and Gray model** adjusts for the presence of competing events, providing a more accurate representation of the incidence of recurrence (event of interest) in the being there of death (competing risk). Results suggest that over time, recurrence is more likely to occur than death, but the risk of death remains significant and should be accounted for in any clinical decision-making.

1. Cumulative Incidence of Recurrence (Blue Line):

- Over time, the probability of experiencing recurrence increases steadily, reaching a probability of approximately 0.6 (or 60%) after 100 years.
- The cumulative incidence of recurrence grows faster during the earlier periods but slows down as time progresses, with fewer new events happening later in the timeline.

2. Cumulative Incidence of Death (Red Dashed Line):

- The competing event of death also increases over time but at a slower rate compared to recurrence.
- By 100 years, the cumulative incidence of death reaches around 0.4 (or40%).
- The gradual increases how that while death is a competing risk, recurrence is more frequent in this dataset.

3. Comparison of Recurrence vs. Death:

- Recurrence (blue line) has a higher cumulative incidence than death (red dashed line), indicating that the event of interest (recurrence) is more common than the competing event (death).
- This is critical in competing risk analysis because it highlights that not accounting for competing risks could lead to overestimating the probability of recurrence if death is ignored.

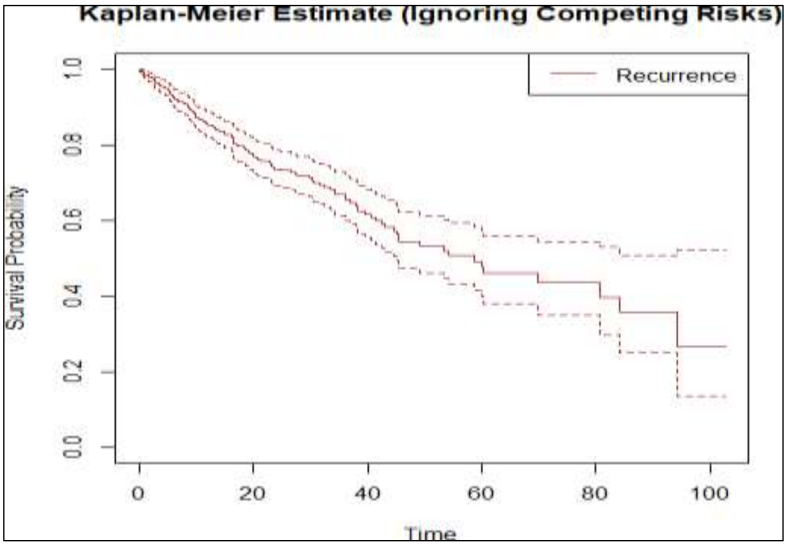

Figure 2: Kaplan-Meier Estimate (Without using Competing Risk).
Source: Authors, (2026).

This plot displays the KM of survival, specifically focusing on the event of recurrence, and it ignores competing risks (such as death). The Kaplan-Meier method assumes that individuals who experience competing events (like death) are censored and does not differentiate between different types of events. Here's the interpretation of the key aspects of the graph:

1. **Y-axis (Survival Probability)**: This axis symbolizes the probability of not practicing the event of interest (recurrence).A survival probability of 1.0 at time zero means that initially, no one has experienced the event.
2. **X-axis (Time)**: The time scale represents the progression in years. Over time, the probability of survival (without recurrence) decreases.
3. **Kaplan-Meier Survival Curve (Solid Line):**

- The solid red line represents the **Kaplan-Meier estimate** of survival, specifically focusing on **recurrence** as the event of interest.As time progresses, the probability of survival decreases, indicating that more patients experience recurrence. After about 100 years, the survival probability drops toaround0.4 (40%), meaning 60%of patients have experienced recurrence by this time.

4. **Confidence Intervals (Dashed Lines):**

- The dashed lines around the survival curve represent the **95% confidence intervals**, providing a range of possible values for the survival probability at each time point.
- The confidence intervals widen as time progresses, indicating increasing uncertainty in the estimates as fewer individuals remain under observation.

Interpretation:

- The **Kaplan-Meier method** provides an estimate of survival (time until recurrence) **without accounting for competing risks** like death. This means that individuals who die before experiencing recurrence are treated as if they were simply censored, which can lead to an **over estimation** of the survival probability if death is a significant competing event.
- As seen from the plot, the survival probability decreases steadily over time, indicating that a significant proportion of individuals eventually experience recurrence. However, this estimate may be **biased upward** since it does not account for individuals who die before recurrence could occur.

Comparison to Competing Risks Model:

In contrast to the **Fine and Gray model**, the Kaplan-Meier method does not adjust for the competing event of death. The **Fine and Gray model** would provide a more accurate estimate of the probability of recurrence by considering death as a competing risk, potentially showing a **lower incidence** of recurrence compared to the Kaplan-Meier method.

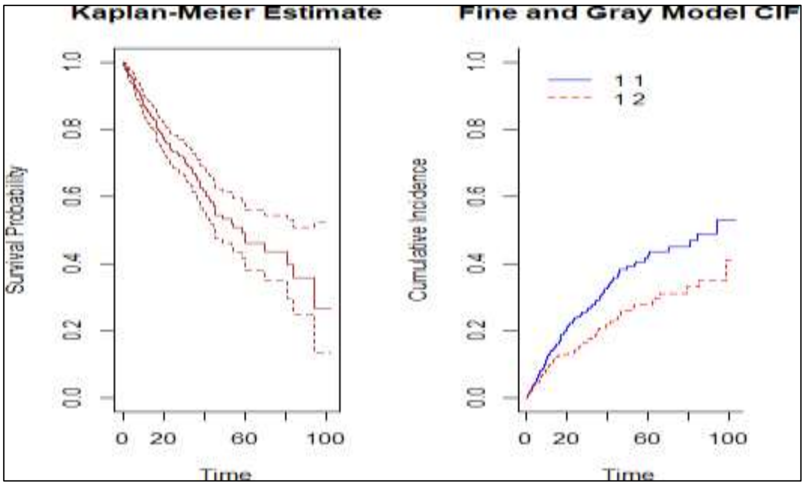

Figure 3: Comparison between Fine and Gray Model vs. Kaplan-Meier Method.
Source: Authors, (2026).

Table 4: Differences between Fine and Gray Model vs. Kaplan-Meier Method.

Feature	Fine and Gray Model	Kaplan-Meier Method
Purpose	Estimates cumulative incidence functions (CIF), accounting for competing risks.	Estimates overall survival, treating competing events as censored.
Event of Interest	Focuses on the event of interest (e.g., recurrence),while accounting for the Competing event (e.g., death).	Focuses on the event of interest only, treating competing events as censored.
Handling of Competing Events	Incorporates competing risks into the model, providing accurate event- specific incidence estimates.	Ignores competing risks, potentially leading to overestimation of survival probabilities.
Hazard Type	Estimates sub-distribution hazard, which allows for	Estimates the cause-specific hazard, which does not

	the cumulative probability of the event of interest in the presence of competing risks.	account for competing risks.
Model Flexibility	Can include covariates (e.g., age, tumor size) to analyze their effect on the event of interest and competing risks.	Covariates can be included but do not adjust for competing risks.
Risk Estimation	Provides the true cumulative incidence of each event, adjusting for competing risks.	Provides a biased survival estimate when competing risks are present, leading to overestimated survival for the event of interest.
Use Case	Appropriate when competing risks like death can prevent the event of interest (e.g., recurrence) from Happening.	Suitable when there are no competing risks or when their impact can be ignored.
Visual Output	Cumulative incidence curves for each competing event, showing the proportion of patients experiencing each event over time.	Survival curve showing the probability of the event of interest over time, assuming no competing events.
Impact of Ignoring Competing Risks	Not applicable—competing risks are built into the model, resulting in realistic estimates.	Ignoring competing risks leads to an underestimation of the event of interest, such as recurrence, due to censoring competing events like death.

Source: Authors, (2026).

- Fine and Gray model provides a more accurate and realistic representation of event-specific outcomes in the presence of competing risks.
- Kaplan-Meier method can lead to misleading results when competing risks a represent, as it does not properly account for events like death that can prevent the primary event of interest (e.g., recurrence).
- Covariates like age and tumor size can significantly impact the results in the Fine and Gray model but are less impactful in the Kaplan-Meier approach when competing risks are ignored.

Table 5: Differences between Kaplan-Meier vs. Exponential Distribution.

Feature	Kaplan-Meier Estimator	Exponential Distribution
Model Type	Non-parametric – makes no assumptions about the underlying survival distribution.	Parametric–assumes survival times follow an exponential distribution.
Assumption on Hazard Rate	Makes no assumptions about the hazard rate; it can vary over time.	Assumes a constant hazard rate over time (memory less property).
Flexibility	Highly flexible, adapts to different survival patterns, including time- varying hazard rates.	Limited flexibility, as it assumes a fixed rate of event occurrence.
Handling of Censoring	Handles censored data efficiently, updating survival probability as events occur.	Can handle censored data but less flexible in adjusting to varying hazards.
Survival Curve	Produces a stepwise survival curve, with changes occurring at observed event times.	Produces a smooth exponential curve, declining steadily over time.
Use Case	Suitable for a wide range of survival analysis applications, especially when hazard rates change over time or are unknown.	Best suited for scenarios where the hazard rate is truly constant (e.g., reliability analysis of mechanical systems).
Interpretability	Empirically estimates survival probabilities at different time points based on observed data.	Provides an analytical formula for survival, with simple interpretation but limited to constant hazard scenarios.
Time-Varying Hazards	Capable of handling time-varying hazards since no constant hazard rate is assumed.	Inappropriate for data with time- varying hazard rates due to its constant hazard assumption.
Computational Simplicity	Requires more data to construct the survival function but provides a detailed and accurate estimate.	Simple to compute and interpret, but the model's fit can be poor when assumptions of constant hazard are not met.

Source: Authors, (2026).

- Kaplan-Meier is preferred for flexibility in estimating survival without assuming a constant hazard, making it more widely applicable to real-world scenarios with varying risks over time.
- Exponential distribution is useful for simple, constant hazard rate scenarios but may not provide accurate results if hazard rates fluctuate or are unknown.
- The Fine and Gray model correctly adjusts for competing risks (such as death), providing an accurate cumulative incidence function (CIF) for recurrence.
- Kaplan-Meier overestimates survival probabilities because it treats competing events as censored, ignoring the risk of events like death.

This gives a more comprehensive comparison between the Kaplan-Meier survival estimates and the Fine and Gray cumulative incidence functions, highlighting the importance of using appropriate methods when competing risks are present.

IV. CONCLUSIONS

Neither age nor tumor size appears to be significant predictors of recurrence in this analysis. Although the hazard ratio for age (HR = 1.010) suggests a slight increase in the hazard of recurrence with each additional year, the effect is not statistically significant (p = 0.26), and the confidence interval (0.992 to 1.03) is too narrow to draw meaningful conclusions. Similarly, the hazard ratio for tumor size (HR = 0.955) suggests a small protective effect, but this too is not statistically significant (p = 0.58), and the wide confidence interval (0.810 to 1.13) indicates considerable uncertainty. Therefore, age and tumor size may not be reliable predictors of recurrence in this context. The Kaplan-Meier estimate shown here assumes that recurrence is the only event of interest and treats deaths as censored observations. While it gives a simple estimate of survival probabilities, it may over estimate the risk of recurrence in the presence of competing risks liked earth, making it less suitable when competing risks are significant.Table 3 shows the conclusion of early period (first few months) the high survival probabilities with minimal events, low uncertainty, and tight confidence intervals. Inter mediate period(~9months), survival begins to decrease moderately, with more events occurring, and confidence intervals widen as the population at risk decreases. Late period (after 80 months), Survival drops sharply, and estimates become increasingly uncertain due to the small number of individuals at risk and wide confidence intervals, indicating greater variability in the remaining population. This data shows a gradual decline in survival, with significant decreases occurring in later months.

V. AUTHOR'S CONTRIBUTION

Conceptualization: Sumithra B, Anu Swedha Ananthan, Zechariah Jebakumar Arulanantham.
Methodology: Sumithra B, Anu Swedha Ananthan.
Investigation: Sumithra B, Anu Swedha Ananthan.
Discussion of results: Sumithra B, Anu Swedha Ananthan, Zechariah Jebakumar Arulanantham.
Writing – Original Draft: Sumithra B.
Writing – Review and Editing: Sumithra B, Anu Swedha Ananthan.
Resources: Anu Swedha Ananthan.
Supervision: Sumithra B, Anu Swedha Ananthan
Approval of the final text: Sumithra B, Anu Swedha Ananthan, Zechariah Jebakumar Arulanantham.

VI. REFERENCES

[1] Fine, J. P. & Gray, R. J. (1999). A Proportional Hazards Model for the Subdistribution of a Competing Risk. Journal of the American Statistical Association. JSTOR

[2] Austin, P. C. et al. (2017). Practical recommendations for reporting Fine-Gray model analyses. (Practical reporting guidance; PMC). PMC

[3] Geskus, R. B. (Annual Review). Competing risks: concepts, methods, and software. Annual Review of Statistics and Its Application. Annual Reviews

[4] Li, J. et al. (2014). Checking Fine and Gray Subdistribution Hazards Model: Diagnostics and Goodness-of-Fit. (PMC article). PMC

[5] Zhou, B., Fine, J., Latouche, A., & Labopin, M. (2012). Competing risks regression for clustered data. Biostatistics. OUP Academic

[6] Schmitt, S. et al. (2023). Systematic comparison of approaches to analyze clustered competing risks data. BMC Medical Research Methodology. SpringerLink

[7] Monterrubio-Gómez, K. et al. (2024). A review on statistical and machine learning competing risks methods. Biometrical Journal. Wiley Online Library

[8] Gregson, J. et al. (2024). Competing risks in clinical trials: Do they matter and how to handle them? (2024 article). ScienceDirect

[9] Austin, P. C. (2021). Fine-Gray subdistribution hazard models to simultaneously estimate absolute risks. Statistics in Medicine. Wiley Online Library

[10] Xu, Y. et al. (2022). Stratified proportional subdistribution hazards model with covariate-adjusted censoring weights. (PMC article). PMC

[11] Bonneville, E. F. et al. (2025). Multiple imputation of missing covariates when using the Fine-Gray model. (PMC/2025). PMC

[12] Basak, R. (2021). Understanding Competing Risks (Clinical primer). redjournal.org

[13] Putter, H. (text/note). The Fine–Gray model (pedagogical resource). scholarlypublications.universiteitleiden.nl

[14] Demelash, R. et al. (2025). Competing risk analysis of predictors of time to lost to follow-up (application using Fine–Gray). Nature

[15] Austin, P. C. (2022). Issues With the Use of Multiple Fine-Gray Subdistribution Models. Circulation: Cardiovascular Quality and Outcomes (or commentary). AHA Journals

[16] Li, Y. et al. (2022). Considerations of competing risks analysis in cardiology (review). PMC

[17] Nakamura, T. et al. (2025). On concerns with cause-specific incidence and interpretability in competing risks (methodological critique). SpringerLink

[18] Kim, S. (technical report). Stratified proportional subdistribution hazards model with covariate-adjusted censoring weights (MCW technical report). mcw.edu

[19] Zhang, F., Peng, H., & Zhou, Y. (ResearchGate). Fine–Gray model for competing risks under length-biased sampling / extensions. ResearchGate

[20] Gorgi Zadeh, S. et al. (2022). An imputation approach using subdistribution weights for missing data in competing-risks analysis. Scientific Reports. Nature