Home Assignment 3

Liang-Jen Huang
Hadeel Elhassan
Department of Statistics, Uppsala University

January 2026

1. Provide a reference to the chosen article.

The impact of different censoring methods for analyzing survival using real-world data with linked mortality information: a simulation study by Wei-Chun Hsu, Aaron Crowley and Craig S. Parzynski. `https://doi.org/10.1186/s12874-024-02313-3`

2. What is the purpose of the study?

The purpose of this study is to examine how different censoring strategies affect survival analysis results when mortality information is incomplete, which is common in real-world data. Specifically, the authors compare censoring at the last observed activity versus censoring at the data cutoff date and evaluate their impact on estimates of median overall survival and hazard ratios. Using simulation, the study aims to quantify the bias and inferential performance of these approaches under varying degrees of missing mortality information, with the goal of informing appropriate censoring choices in real-world survival studies

3. Describe the analysis method used/investigated:

The authors conducted a Monte Carlo simulation study to evaluate how different censoring strategies affect survival analysis when mortality information is incomplete, a common situation in real-world data (RWD).

They simulated a non-randomized comparative effectiveness study with two treatment groups and overall survival as the outcome. For each simulated patient, three time variables were generated: time to death, time to last database activity, and time to data cutoff. Time to death represented the true event time and was generated from an exponential distribution, with the hazard rate modified by a predefined hazard ratio (HR) to represent different treatment effects. Time to last activity reflected loss to follow-up typical of electronic health record data, while time to data cutoff represented administrative censoring at the end of data availability.

Deaths occurring before the last activity date were always observed within the database. Deaths occurring after the last activity but before the data cutoff were considered observable only through linked external mortality data. To reflect imperfect mortality capture in real-world settings, a proportion of these deaths was randomly set to missing, resulting in incomplete death ascertainment. The missingness mechanism was assumed to be missing completely at random (MCAR).

Two commonly used censoring schemes were then applied. In the first scheme, patients without observed death were censored on their last activity date, reflecting a conservative approach that limits follow-up to observed healthcare encounters. In the second scheme, patients without observed death were censored on the data cutoff date, corresponding to administrative censoring and assuming patients remain at risk until the end of data collection.

For each simulated dataset, Kaplan-Meier estimators were used to estimate median overall survival, and Cox proportional hazards models were fitted to estimate log hazard ratios between treatment groups. The Cox models included only treatment as a covariate to isolate the impact of censoring mechanisms.

The simulations were repeated under multiple scenarios varying sample size, length of follow-up, treatment effect size, and degree of missing mortality information. The number of simulation replications was chosen to ensure adequate precision, as measured by the Monte Carlo standard error.

Performance of the censoring strategies was evaluated using bias, variance, coverage probability, and rejection rate (power), allowing the authors to assess both estimation accuracy and inferential validity under each censoring approach.

4. Why is the investigated methodological aspect of interest?

The methodological aspect investigated in this study the choice of censoring strategy in survival analysis is of interest because censoring decisions directly affect the validity of survival estimates, particularly when mortality information is incomplete, as is often the case in real-world data. In such settings, deaths may be missed or recorded with delay, and standard censoring approaches can either exclude patients too early or incorrectly extend their time at risk. These issues can introduce substantial bias, especially in estimates of median overall survival derived from Kaplan–Meier estimators. Understanding how different censoring strategies perform under varying levels of incomplete mortality capture is therefore crucial for producing reliable survival estimates and for making valid comparisons between treatment groups in real-world evidence studies.

5. How censoring is taken into consideration?

Censoring is explicitly incorporated into the study design through the simulation of realistic follow-up processes and the comparison of two commonly used right-censoring schemes. For each simulated patient, the authors generate a time to death, a time to last database activity, and a time to data cutoff. Depending on the relative ordering of these times and whether death is observed, patients may experience an event or be right-censored.

The study evaluates two censoring strategies. In the first, patients without an observed death are censored at their last recorded activity, reflecting loss to follow-up in real-world databases. In the second, patients without an observed death are censored at the data cutoff date, representing administrative censoring. By deliberately removing a proportion of deaths occurring after the last activity date, the authors create scenarios where censoring may occur either before or after the true death time, thereby inducing bias. This framework allows the authors to assess how different censoring rules interact with incomplete mortality information and affect survival estimates obtained from Kaplan–Meier and Cox regression methods.

6. If a simulation study is described, how do they motivate the choice of sample size and number of replications?

In the simulation study, the choice of sample size and number of replications is motivated by realism and statistical precision. The sample sizes were selected to reflect typical real-world oncology studies, with scenarios representing both small and large study populations as well as short and extended follow-up. The treatment effect sizes were chosen based on values commonly observed in clinical oncology research.

The number of simulation replications varied across scenarios and was determined to achieve an acceptable Monte Carlo standard error (MCSE) for the primary performance measure, namely the bias of the estimated median overall survival. By selecting the number of repetitions to control MCSE rather than using an arbitrary fixed number, the authors ensured that the simulation results were sufficiently precise and that comparisons between censoring strategies were statistically reliable

7. How is the aspect of covariates discussed?

The role of covariates in the analysis is intentionally minimized in this study. The authors include treatment group as the only covariate in the Cox proportional hazards model, while no additional baseline covariates are considered. This design choice is made to isolate the effect of censoring strategies and incomplete mortality information on survival estimates, without introducing additional complexity from confounding or model misspecification. By excluding other covariates, the study focuses on how censoring alone influences the performance of Kaplan-Meier and Cox regression estimators, allowing for a clearer interpretation of bias attributable to censoring mechanisms rather than covariate adjustment.

8. Are the assumptions of the method evaluated in the article? If not, is it reasonable that they hold? If the assumptions don't hold, what impact does

that have on the conclusion?

Not all of the Kaplan-Meier assumptions are clearly assumed or discussed in this study. Independent censoring was assumed explicitly. Since this is a simulation study and the analyzed data is simulated, random sampling assumption holds given the nature of the study design. Secondly, independent samples' assumption also holds. Again, given that the samples are simulated. Thirdly, right censoring assumption is satisfied. All censored observations correspond to patients who their follow-up ended without an observed event (death) either at last activity date or at data cut-off date. Right censoring is satisfied at the data structure level. However, the incomplete mortality capture may result in censoring times happening after the true death time which might lead to biased estimation. Finally, and most importantly in this study is the noninformative censoring assumption, as the authors highlight earlier there is a gap or (ghost time) between the actual death time and observed death time. Missing deaths lead patients to be censored after death. This may violate the non-informative censoring assumption by making censoring informative, even though the missingness is considered to be completely at random.

Furthermore, the Cox proportional hazards model's assumptions were not clearly stated. The assumptions that the Cox model shares with the Kaplan-Meier estimator discussed above apply to the Cox model as well from the study. However, the assumption of proportional hazards is satisfied by design. Moreover, the large sample assumption for the Cox regression was satisfied as well since the authors deliberately vary the sample size based on different scenarios.

9. Do you miss anything in the article? Think in terms of assumptions, regression diagnostics, ethics considerations

The simulation study lacks the explicit discussion of some of the model assumptions for the Kaplan-Meier and the Cox proportional hazards models. While the study uses standard survival analysis methods such as Kaplan-Meier and Cox proportional hazards model, how the study satisfies or violates the undelying assumptions of these models is missing. leaving it to the reader to infer assumption validity from the study design and results.

10. Summarize the results/conclusions of the article.

The study shows that the choice of censoring strategy has a substantial impact on estimates of median overall survival when mortality information is incomplete. Censoring at the data cutoff date produces unbiased median survival estimates when mortality is fully captured but leads to overestimation when deaths are missing due to the inclusion of ghost time. In contrast, censoring at the last activity date tends to underestimate median survival, although the bias decreases as mortality capture becomes poorer.

In comparison, estimates of hazard ratios obtained from Cox proportional hazards models are relatively robust to incomplete mortality capture, with only minimal bias observed across censoring strategies and simulation scenarios. However, missing mortality information still affects statistical power and coverage, indicating that inference may be impacted even when point estimates appear stable.

In conclusion, the authors conclude that inappropriate censoring choices can lead to substantial bias in survival analyses using real-world data. They emphasize that understanding the completeness and quality of mortality data is critical when selecting a censoring strategy and advocating for validation studies and greater transparency from data providers to support sound methodological decisions.