Outbreak, Surveillance, Investigation & Response (OSIR) Journal

Field Epidemiology Training Program, Division of Epidemiology
Department of Disease Control, Ministry of Public Health, Thailand
Tel: +6625903894, Fax: +6625903845, Email: osireditor@osirjournal.net, http://www.osirjournal.net

The Grammar of Science: How do We Count Time?

Jaranit Kaewkungwal

Mahidol University, Thailand

Corresponding author email: jaranitk@biophics.org

Some people say that time cannot be counted as time has no physical properties to measure. What we are really measuring is time intervals, the duration separating two events.[1] In history, people counted time as day and night between sunrise to sunset. Time shown on sundials, pendulum clock, and analog or digital watch tells us loosely about the passage of time. People in different cultures created and use different methods for keeping track of days and larger divisions of time. The Gregorian calendar is the calendar used in most of the world.[2] As introduced in October 1582 by Pope Gregory XIII, the average calendar year is approximately 365.2425 days long according to the Earth's revolution around the Sun.

Cambridge dictionary notes that "Time" is a noun with a number of meanings while it could be countable or uncountable.[3] We may use time as countable to refer to what is measured in seconds, minutes, hours, days, weeks, months and years—*"I was diagnosed and treated for cancer for 6 years since 25 Jul 2016."* On the other hand, we may use time as uncountable—*"He is out of sight for a long time."* However, when we talk about time we usually have a "Reference Time Point" —the point in time that acts as a fixed reference point to an event— *"She has been waiting since 8:00 AM"*. Some people count time from their own loose referential point—*"I have been in this position for only 6 years counting from when I started to work here."* So how do we actually count time in "Time-to-event" analysis?

Time-to-event Analysis

Time-to-event or survival analysis is a statistical procedure that considers amount of time until an event occurs.[4] The event, also called endpoint or outcome, of interest can be good (e.g., cure/recover after treatment) or bad (e.g., death, tumor recurrence). Why do we need to take time into consideration? The answer is that time will give you "rate" (or speed) of the event; it will tell you how fast an event can occur in a certain time period. Figure 1 depicts a scenario of a clinical trial which 10 patients were randomly allocated to either Drug A or Drug B. Without time effect, 3 of 5 patients who received Drug A were cured (incidence proportion=0.6). Similarly, 3 of 5 patients who received Drug B were cured (incidence proportion=0.6). The two groups were not different in terms of disease cure proportion. When considering time each patient was in the study, 3 of 9 months of follow-up among all patients who received Drug A were cured (incidence rate=0.33) while 3 of 20 months of follow-up among all patients who received Drug B were cured (incidence rate=0.15). This informs us that the cure rate per month of Drug A is better than that of Drug B.

In performing time-to-event analysis, we need two pieces of information for every study participant: (1) the time to the event and (2) the event status (whether or not the event occurs).[5] The effect of time to reach the event typically characterizes as "survival function". The function represents the probability of an individual surviving or still not reaching the event beyond time X.[4] In reality, we cannot observe events for all of the study participants as the study may end before the events of some participants occur or the participants may be lost to follow-up, drop out, death from other causes or leave the study. This leads to a concept of censoring; i.e., each participant either has the event (so-called failure case) or have not yet experienced the event (so-called censored case).[4,5] As shown in Figure 2, the time-to-event analysis is applicable to two types of study designs, cohort and experimental studies.

Figure 2 (a) shows time-to-event which could be in a prospective cohort (study starts at present and follow 3 years onward) or retrospective cohort (study starts by reviewing medical records 3 years ago until the closing date of the study). Time counts from date of diagnosis with cancer to date of dead as the endpoint event. For the patient who was not dead, he/she was

censored at date of lost to follow-up (LFU) or date of study closure. Some textbooks call a censored case that his/her time is cut off at the study closure as a "truncated" case. The case with the endpoint as dead due to suicide could be either a failure case or a censored case depending on the definition of the event. If the endpoint is defined as "all causes of death", the case is considered as a failure case; on the contrary, if the endpoint is defined as "death of cancer", the case is a censored case.

Figure 2 (b) shows time-to-event of a clinical trial which study participants were allocated to Drug A or Drug B. The outcome of the study is time from date of treatment initiation to date of cure as the endpoint event. For patient who was not cured, he/she was censored at date of lost to follow-up, date of consent withdrawal or date of study closure. Again, a censored case that his/her time is cut off at the study closure may be called a truncated case.

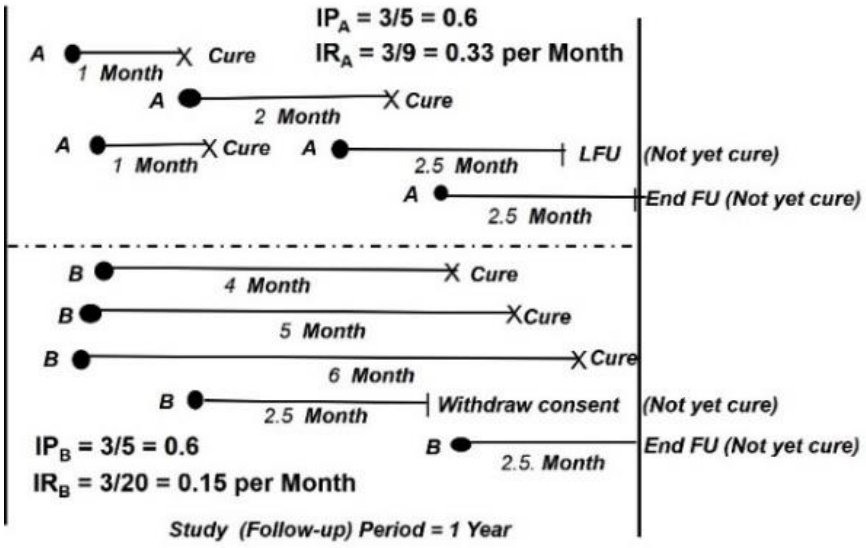
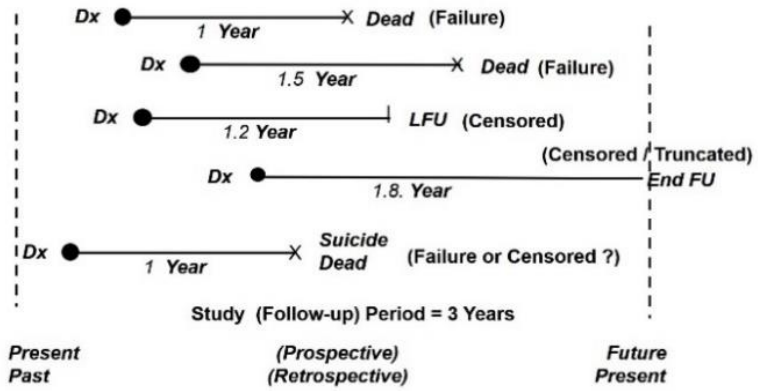
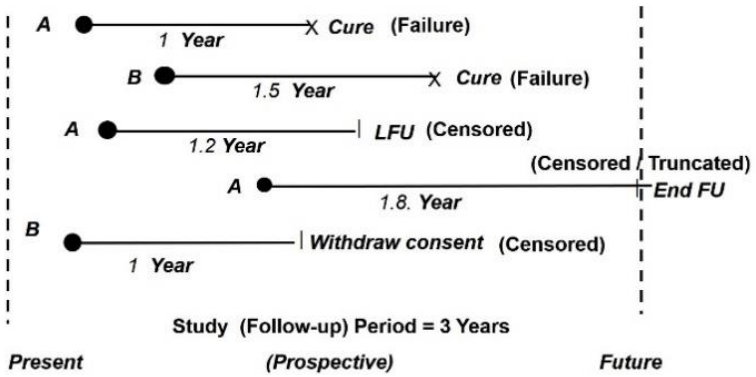


Figure 1. Examples of incidence proportion vs. incidence rate of disease cure



(a) Time to event in cohort study (time from diagnosis to dead)



(b) Time to event in experimental study (time from treatment initiation to cure)

Figure 2. Examples of time-to-event in cohort study and experimental study

Censoring & Truncating Time

When collecting time-to-event data the researchers must consider the study-specific details of recruitment and inclusion criteria.[6] When making predictions with time-to-event data, it is critical to define the risk set appropriately.[7] The study participants in the risk set include those who reach the event (failure) and those who do not have the event (censored) at the particular time point. In general, there are three types of censoring mechanisms: right censoring, left censoring, and interval censoring.[8,9]

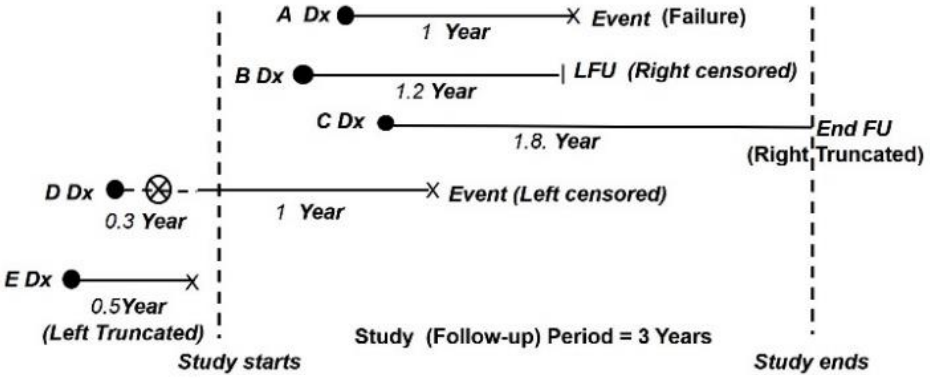
Right-censoring

The most common type of censoring is right-censoring. As previously discussed, right censoring occurs when a study participant drops out or leaves the study before the event occurs, or the study ends before the event has occurred. Right censoring might be imposed due to a competing risk, i.e., the event of interest cannot be observed because of the occurrence of a competing event (e.g., death from other causes).[9] It should be noted that the right-censored case is assumed to follow the same survival distribution after withdrawal as the non-censored cases.[8]

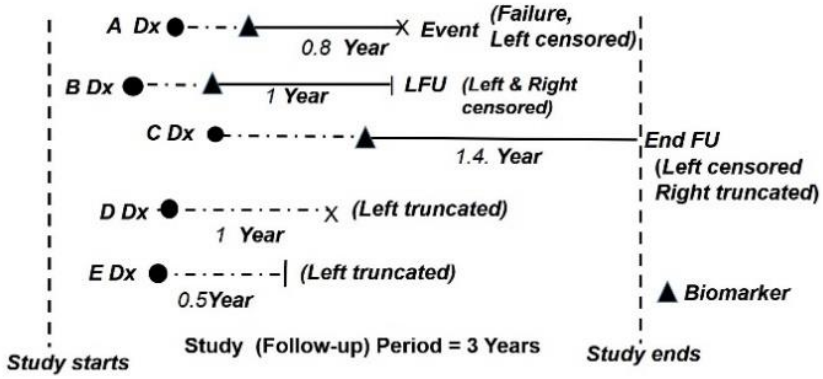
Left Censoring

This is the opposite of right censoring, when the time of a study participant is cut on the left-hand side rather than the right-hand side. There are several situations for a study participant to be considered as a left-censored case. Figure 3 shows different scenarios of left censoring. As shown in Figure 3 (a), a study participant is left censored when his/her event has already occurred prior to enrollment or before the study starts. Such case is sometimes called left truncated case. Patients E reached the event prior to the study starts and thus he is not included in the study. This scenario is very rarely encountered in most study. Patient D was diagnosed prior to the study starts but had been followed until the event occurred within the study time period. In some study, such case may be included as a study participant but the time prior to the study starts is cuff off (censored on the left-hand side).

Another left censoring example is shown in Figure 3 (b) when the time-to-event starts from a certain milestone marker. Study participants who reached the milestone marker (i.e., biomarker in this example) are included in the study (Patients A, B, C) while those who did not are excluded (Patient D, E). Patients A, B, C could be handled in different ways depending on the objective of the study; the time prior to milestone marker can be cut off (as left-hand side censoring) or can be split and treated as a case with 2 time periods (before and after milestone marker) in the analysis model.



(a) Right & left censoring / truncating from initial time point



(b) Right & left censoring / truncating from milestone marker

Figure 3. Examples of right and left censoring cases

Interval Censoring

The censoring occurs when the failure event of interest cannot be observed directly but is known to have occurred during a time interval. Interval censoring is a generalization of left and right censoring.[9] This censoring is common and natural in a clinical trial or longitudinal study in which there is periodic follow-up.[8] Patients have different visit times and durations between visits; the outcome event is measured at each visit. The exact time of event is not observed and is known to fall in an interval between visits.[10] Figure 4 shows 2 classic scenarios of interval censoring cases. Patient A missed a few visits and thus was considered as a LFU case but later on he decided to resume to the study; he reported that he had the event but forgot when the event happened during the missing time period. Patient B had regular visits throughout the study period and he had an event at Visit 6. In a typical time-to-event analysis, it can be simply assumed that he had an event at Visit 6. But in some study, the researchers may decide to model that he had an event some times between Visit 5 and Visit 6, i.e., considering interval censoring between Visit 5 and Visit 6.

In handling interval censoring data by interpolating the event time as the midpoint of the censored interval, it must be cautious that doing so depends strongly on the underlying distributions and the width of the intervals. The survival function based on midpoint event may be biased and the variability of the estimates may be underestimated.[10]

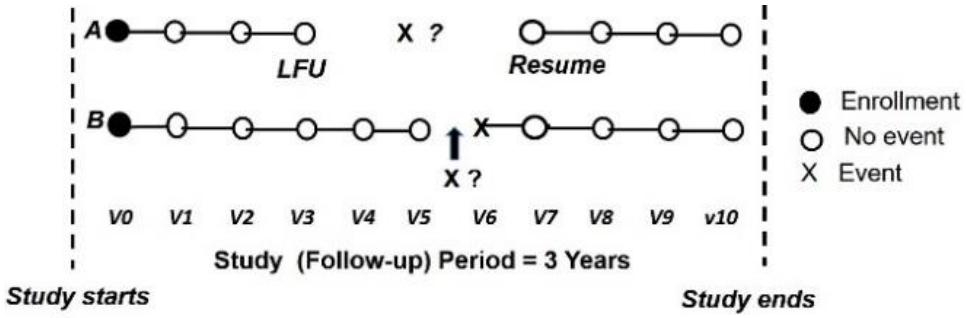


Figure 4. Examples of interval censoring cases

Biases Related to Time-to-Event

There are several biases that should be considered in time-to-event analysis. The researchers should have plan to mitigate such biases that could occur in the study.

Drop-out Bias (Selection Bias)

When a study participant drops out from the study, his/her time is censored at the drop out date. In a typical time-to-event analysis, the distribution of censoring time is assumed to be independent of the distribution of the survival time.[8] In other words, censoring should be random.[4] As an example, in a clinical trial, if there is a certain subgroup (say, younger males) drops out more than the rest of the study participants, the study sample will become biased. Moreover, the reasons for the drop out study participants should not be related to the purpose of the study.[8] Such assumption cannot be met in many studies. For example, in a cancer study, censored cases may be found more among patients who are at a higher risk of progression/death, or among patients who discontinue treatment due to toxicity and have to be shifted to start some other therapy.[8]

If such censoring bias is ignored, there would be selection bias in the data and the survival probability might be overestimated.[4] The researchers should monitor the study whether such bias occurs or not. If so, the researchers should select appropriate methods including, for example, stratification-based techniques, regression adjustment, joint modeling, or censoring weighted estimation.[8]

Length-time Bias

It is also called length-biased sampling or survivorship bias; such bias occurs when time is truncated at a certain cut off point.[4] Analysis at the time cut-off point may affect assessment of survival function among incomplete risk set, not including the number of people who still have not experienced the event. As an example shown in Figure 5, when the researchers want to estimate survival function at 1.5 years (at Month 18) within the study period (3 years), they would assess information from only 3 of 4 patients (Patients A, B, C) while cut off 1 patient (Patient D) who would be diagnosed and experience the event at later time. Incomplete risk set at Month 18 may yield underestimated survival function. Analysis based on complete risk set during the entire 3 years study period, accounting for Patient D, might result in a more precise and correct conclusion.

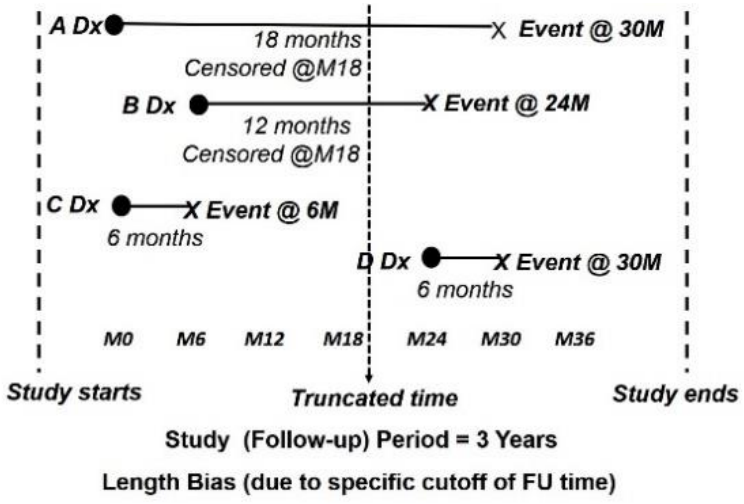


Figure 5. Example for length-time bias

Time-dependent Bias

There are many kinds of time-dependent bias. This bias is also known as immortal time bias or survivor treatment selection bias.[4] Figure 6 shows an example of time-dependent bias in terms of "time-dependent exposure", when an exposure (treatment) varies at different time points among study participants. As shown in Figure 6, treatment was only dispensed when the patient has reached a certain level of biomarker, not at enrollment. There are some patients who had never reached the set level of biomarker and thus they did not get the treatment however, they were followed up/monitored for the endpoint event (Patient A). Patients who reached the set level would receive treatment and followed up for the endpoint events (Patients B, C, D). If the researchers want to compare survival functions between those who received and did not receive the treatment, they must consider time-dependent bias. The researchers cannot simply compare time from diagnosis to the endpoint event between those who received vs. not received treatment (e.g., Patient A vs. Patient B). While those who did not receive treatment had 1 time count (Patient A), those who received treatment did actually have 2 time counts, time before and after treatment (Patient B). To correctly classify the treatment cases, their time should be split into 2 time-to-event periods: time before treatment and no event, and time after treatment and with endpoint event (as shown in Patient C, D). The appropriate time-to-event analytic model must be assessed by taking into consideration of this split time-dependent exposures.

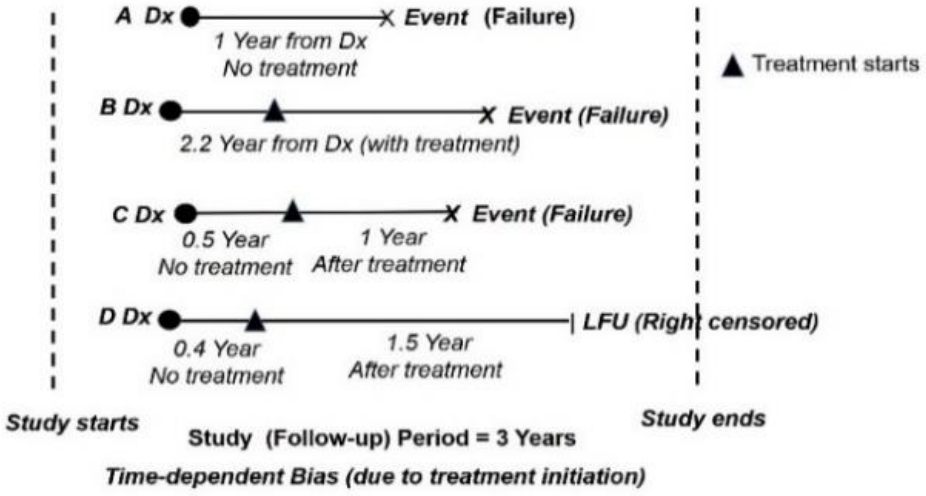


Figure 6. Example of time-dependent bias

Lead Time Bias and Stage Migration

Lead time bias occurs due to the early detection of disease is made before the usual diagnosis based on symptoms, and consequently leads to a fallacious increase in a patient's time to event.[8] As shown in Figure 7, compared to the survival time after usual diagnosis of Patient A, the survival time after early diagnosis of Patient B is longer due to lead time, time gap between early diagnosis and usual diagnosis. In fact, this increase in survival dues only to the lead time and has nothing to do with the survival of the patient.

Particularly in cancer study, another related bias, i.e., stage migration, could occur. Patients at the boundary of cancer stages might be reclassified into the higher stage and thus results in a misleading increase in survival estimation due to earlier detection before the symptoms become evident.[8] As shown in Figure 7, Patient C has much longer survival time due to lead time and gain time as he might get earlier and therefore better treatment outcome than Patient A whose survive time is based on routine practice. When early diagnosis is part of the study procedure, the researchers should acknowledge these potential biases and conclude the estimated survival time by accounting for such lead and/or gain times.

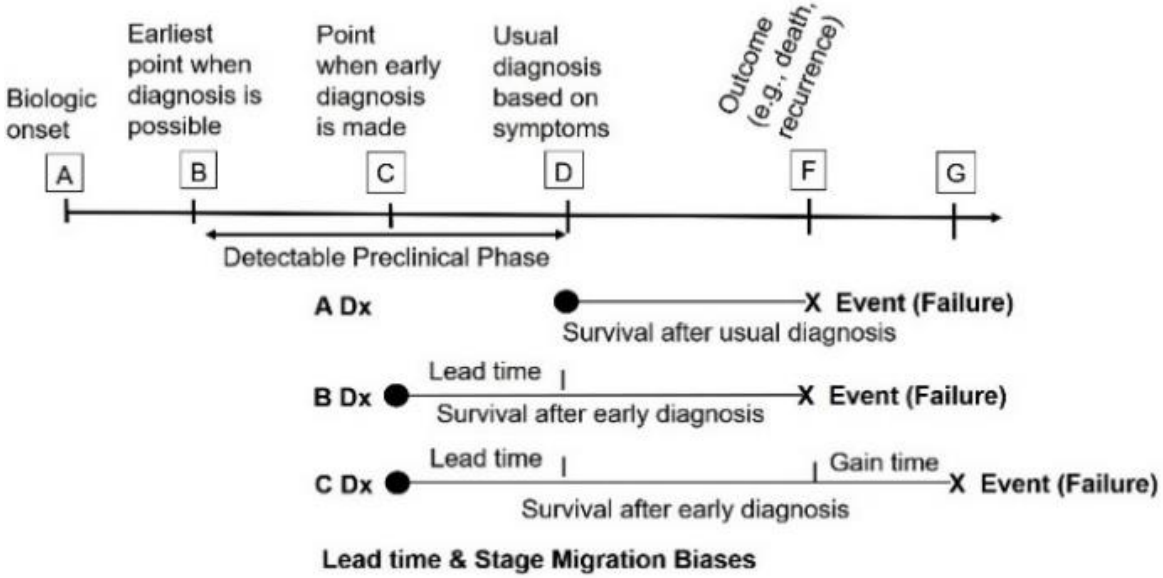


Figure 7. Example of lead time bias

In handling biases, besides procedures within Cox's proportional hazard model, there are several other methods and models that could provide precise survival function including, for examples, interval-censored data models, imputation-based methods, parametric regression models, nonparametric maximum-likelihood estimation, semiparametric regression models, and Bayesian analysis.[10,11]

Conclusion

Time-to-event is not simply counting from the time you start observing the event until the event actually occurs or does not occur. There are situations when time counting is quite complicated due to case censoring and truncating as well as several potential biases related to assessment of time effect. Incomplete information regarding time-to-event of subjects should not be simply discarded as they may reflect certain relevant information for final results of the study. The researchers must understand the concept of time in survival analysis and select the appropriate statistical procedures. Time management for time-to-event analysis need to be predetermined to avoid erroneous conclusion.

Suggested Citation

Kaewkungwal J. The grammar of science: how do we count time? OSIR. 2022 Sep;15(3):99–105.

References

1. National Institute of Standards and Technology, U.S. Department of Commerce. How Do We Measure Time? [Internet]. Gaithersburg (MD): National Institute of Standards and Technology, U.S. Department of Commerce; 2021 Oct 26 [updated 2022 Apr 8, cited 2022 Sep 15]. <https://www.nist.gov/how-do-you-measure-it/how-do-we-measure-time>
2. Wikipedia contributors. History of calendars [Internet]. San Francisco: Wikimedia Foundation; 2022 Aug 11 [cited 2022 Sep 15] <https://en.wikipedia.org/wiki/History_of_calendars>
3. English Grammar Today. Time [Internet]. Cambridge: Cambridge University Press; c2022 [cited 2022 Sep 15]. <https://dictionary.cambridge.org/grammar/british-grammar/time>
4. Liou L. How well do you "really" know survival analysis? [Internet]. Ontario (CA): Towards Data Science; 2021 Feb 5 [cited 2022 Sep 15]. <https://towardsdatascience.com/how-well-do-you-really-know-survival-analysis-221885b89a8e>

5. Quantics Biostatistics. An introduction to survival analysis for clinical trials [Internet]. Edinburgh (GB): Quantics Biostatistics; 2017 Nov 9 [cited 2022 Sep 15]. <https://www.quantics.co.uk/blog/introduction-survival-analysis-clinical-trials/>

6. Cain KC, Harlow SD, Little RJ, Nan B, Yosef M, Taffe JR, et al. Bias due to left truncation and left censoring in longitudinal studies of developmental and disease processes. Am J Epidemiol. 2011 May 1;173(9):1078–84.

7. McGough SF, Incerti D, Lyalina S, Copping R, Narasimhan B, Tibshirani R. Penalized regression for left-truncated and right-censored survival data. Stat Med. 2021 Nov 10;40(25):5487–500.

8. Dey T, Mukherjee A, Chakraborty S. A Practical Overview and Reporting Strategies for Statistical Analysis of Survival Studies. Chest. 2020 Jul;158(1 Suppl):S38–42.

9. Rodrigues AS, Calsavara VF, Silva FIB, Alves FA, Vivas APM. Use of interval-censored survival data as an alternative to Kaplan-Meier survival curves: studies of oral lesion occurrence in liver transplants and cancer recurrence. Applied Cancer Research. 2018 Nov 26;38:16.

10. Yang X. Analyzing interval-censored survival-time data in Stata [PowerPoint presentation]. Proceedings of the 2017 Stata Conference; 2017 Jul 27–28; Baltimore (MD). Texus: StataCorp LLC; 2017 Jul 29 [cited 2022 Sep 15]. 35 p. <https://www.stata.com/meeting/baltimore17/slides/Baltimore17_Yang.pdf>

11. Turkson AJ, Ayiah-Mensah F, Nimoh V. Handling censoring and censored data in survival analysis: a standalone systematic literature review. International Journal of Mathematics and Mathematical Sciences. 2021 Sep 24; 2021: Article ID 9307475 [16 p]. <https://doi.org/10.1155/2021/9307475>