Applied statistics in vascular surgery Part IV: Introduction to survival analysis

Constantine N. Antonopoulos[1], Efthymios D. Avgerinos[2], John, Kakisis[1]

[1]Department of Vascular Surgery, Athens University Medical School, "Attikon" Hospital, Athens, Greece
[2]Department of Vascular Surgery, University of Pittsburgh, UPMC Presbyterian Hospital, Pittsburgh, PA

Abstract:
Survival or 'time to event' analysis is used to estimate the lifespan of a particular population under study. The goal is to estimate the time for an individual or a group of individuals to experience an event of interest. In this paper, we present the basic concepts of survival analysis.

INTRODUCTION

When conducting a clinical study, the type of the experimental design and the type of the variables which should be analyzed are important factors highlighting the type of statistical analysis needed. It is evident that mortality after open repair of abdominal aortic aneurysm (AAA) is a binary outcome variable (equal to 1 if the patient died and 0 if the patient survived). In that case, the researcher can estimate the odds of experiencing the outcome of death within a specific time frame (eg, 30-day mortality after surgery), when one or more covariates are present by conducting logistic regression analysis.[1] However, if we want to estimate the lifespan of this particular population, the time to death is the observed outcome and this type of analysis is not appropriate.

In the last scenario, the research question involves the length of time until death occurs, and one might think that length of time is a continuous outcome variable and as a result, linear regression analysis might be appropriate. However, in this case, death will usually have occurred only in some, but not in all patients, by the time the study ends. Furthermore, for those who are lost to follow-up before the end of the observation period or those who haven't yet reached the reported time endpoint, full survival times are unknown. It seems that when length of time until death is the outcome of interest, a combination of whether death has occurred (binary outcome) and when death has occurred (continuous outcome) should be taken into consideration[1] and the appropriate analysis is called "survival analysis". However, it should be highlighted that despite the name ''survival analysis,'' this method can be used in any time-to-event outcome analysis, such as the time until a patient experiences a stroke, a myocardial infarction, a bypass or stent failure etc.[2]

Author for correspondence:
Constantine N. Antonopoulos
Department of Vascular Surgery, Athens University Medical School, "Attikon" Hospital, Athens, Greece
E-mail: kostas.antonopoulos@gmail.com
ISSN 1106-7237/ 2019 Hellenic Society of Vascular and Endovascular Surgery Published by Rotonda Publications

WHICH ARE THE BASIC COMPONENTS OF SURVIVAL ANALYSIS?

Firstly, a clear definition of the outcome of interest defined as "event" (eg. death after open repair of AAA) and the exact starting point (eg. time of open AAA repair) are crucial. The length of follow-up should be enough to ensure that a sufficient number of events are observed, but at the same time not too long to allow changes to the variable under investigation during the study period, due to uncontrolled factors (eg change in a surgical technique over time, changes in endograft design which cause improvements in endograft performance etc).[1,3]

SURVIVAL TIME AND CENSORING

After defining the event and the starting point, the time until it occurs, (eg. the time interval between the open AAA repair and death) should be captured. This is also a synonym for "survival time" or "failure time" or "time-to-event" or "freedom from the event" depending on the definition of the event. For some patients, the event (eg. death) occurred, and we were able to measure when it occurred (eg. three months after AAA repair). However, for some patients, the event was not observed during follow-up and the researcher does not know the exact time of the event (if happened), or the patients were lost to follow-up or not have reached the reported follow up. This is called "censoring". In this case, we have incomplete observation of survival times and the subjects with the incomplete observation are referred to as "censored". There are many types of censoring. In one of the most frequent types patients do not complete the follow-up, because they are lost to follow-up, or they are uncooperative and refuse to remain in the study. This is called "random censoring" and it occurs when follow-up ends for reasons that are not controlled by the investigator. Another type is when the patient is known to have experienced the event (eg. death) before the start of the observation period (eg. the patient died, while he was waiting for open AAA repair). This is called "left censoring". In another situation, we know that the event occurred (or will occur) sometime after the date of last follow-up, but the observation of the patient is terminated before the event occurs, which means that the actual time-to-event, if it were to occur, is longer than the observation time. This is called "right censoring". In other words, for some patients the event did not occur during the time we observed the individual, and we only know the total number of days in which it did not occur (eg. we do not kwon the exact time of death for a patient with AAA repair; we only know that the patient was followed up for 6 months, during which he was alive). There are two types of right censoring; type I and II. In type I censoring, the study stops at a predetermined time, at which point any subjects remaining are right-censored, while in type II censoring, the study stops when a predetermined number of patients are observed to have failed; the remaining subjects are then right-censored.[3]

ASSUMPTIONS OF SURVIVAL ANALYSIS

Before conducting a survival analysis, the researcher should check if specific assumptions are fulfilled.[4] Firstly, the outcome for a patient can either be censored or the event has occurred; it cannot be both. For example, the patient is either dead or censored at the end of follow-up and at least one of these two states have occurred. Secondly, the "survival time" should be measured with precision (eg. the researcher should know the exact time of death or censorship after open AAA repair for each study participant). Thirdly, left censorship should be as low as possible. Moreover, independence of censoring and event is required. This means that a patient is "censored" not because he/she is at greater risk for the event (eg. death). This is based on the assumption that the reasons for someone to quit follow-up are unrelated to whether the event occurs or not. For example, some patients become too ill during the study and then withdraw and then die. These patients will skew survival analysis, as they will be counted as non-events without ever contributing an event when in fact they should have counted as an event. Furthermore, the observation period should not be too long to introduce bias by allowing confounders to affect the likelihood of the event; this is called "secular trend". Lastly, in case of comparison of survival between two or more groups, the possibility and pattern of censorship should be similar among the different groups.[5,6]

CONCLUSION

Survival or "time-to-event" analysis is a set of methods to measure the time until a specific event occurs in a study sample. Due to the fact that not everyone in the sample will experience the event, either because the study ended before they had the event or because they were lost to follow-up, there is incomplete observation of time-to-event and the subjects with the incomplete observations are referred to as "censored". Unlike regression models, survival methods correctly incorporate information from both censored and uncensored observations in estimating important model parameters.

REFERENCES

1 Schober P, Vetter TR. Survival Analysis and Interpretation of Time-to-Event Data: The Tortoise and the Hare. Anesth Analg. 2018;127(3):792-8.

2 George B, Seals S, Aban I. Survival analysis and regression models. J Nucl Cardiol. 2014;21(4):686-94.

3 Rich JT, Neely JG, Paniello RC, Voelker CC, Nussenbaum B, Wang EW. A practical guide to understanding Kaplan-Meier curves. Otolaryngol Head Neck Surg. 2010;143(3):331-6.

4 Thuijs D, Hickey GL, Osnabrugge RLJ. Statistical primer: basics of survival analysis for the cardiothoracic surgeon. Interact Cardiovasc Thorac Surg. 2018;27(1):1-4.

5 Kleinbaum DG, Klein M. Survival analysis: A self-learning text 3rd ed. New York: Springer; 2012.

6 Kaplan EL, Meier P. Nonparametric Estimation from Incomplete Observations. Journal of the American Statistical Association. 1958;53(282):457-81.