CIRUGÍA ESPAÑOLA

www.elsevier.es/cirugia

Methodological letter

Description of survival with numerical and graphic indicators. Basics and mistakes to avoid☆

Descripción de la supervivencia con indicadores numéricos y gráficos. Conceptos básicos y errores que evitar

Guadalupe Gómez Melis, Jordi Cortés Martínez, Erik Cobo Valeri*

Departament d'Estadística i Investigació Operativa, Universitat Politècnica de Catalunya-BarcelonaTECH, Barcelona, Spain

The randomised clinical trial (RCT) by Arezzo et al.[1] studied the overall survival (OS) and time to progression (TTP) in patients with neoplastic colon obstruction after 2 possible interventions, either stent as a bridge to surgery (SBTS) or emergency surgery (ES). Fifty-six participants were assigned to SBTS and 59 to ES. We will use this work to illustrate the above concepts.

Types of variables

Table 1 of the article "photographs" these data: some are binary (gender) or on a nominal scale (the Hartman surgery type), or on an ordinal scale (the ASA, Physical Status Classification System), or with a unit of measurement (the body mass index). Time to events (OS, TTP or DFS, disease free survival) are also described. It is a descriptive snapshot, not intended to infer the population. Let us look at these variables more closely.

The nominal scale classifies patients in such a way that those belonging to the same category are equivalent to each other and different from those in another category. Information on these variables is reported as absolute (n) and relative (%) frequencies. For example, the Hartman-type surgical procedure was used in 11 patients (20.4%) in the SBTS group and in 20 in the ES group (33.9%). A possible graphical representation for this scale is the bar chart.

The ordinal scale allows the calculation of cumulative probabilities. For example, the ASA[2] scale measures the comorbidity status of a patient before an intervention. Ordinal scales have no unit of measurement, so the increase in comorbidity between consecutive categories need not be identical.

Various indicators are available to summarise data with unit of measurement. The mean and standard deviation summarise central tendency and dispersion, respectively. Deviation is most useful with symmetric data, without extreme values or outliers. If either of these conditions is not met, it is better to use "robust" measures such as the median and the interquartile range (IQR), which are not very sensitive to extreme observations. The median is calculated as the central observation of the ordered data and the interquartile range is the interval containing 50% of the central observations. The boxplot and the histogram are the most commonly used graphical representations. The boxplot is based on the robust measures mentioned above. The histogram would allow the detection of bimodal distributions: a large presence of obese and lean people could be missed in a boxplot.

Tables and graphs complement each other. Tables are useful if the precision of the values is relevant or if the variables have different units. Graphs are useful to show trends, patterns or large amounts of data in an efficient way.[3]

☆ Please cite this article as: Gómez Melis G, Cortés Martínez J, Cobo Valeri E. Descripción de la supervivencia con indicadores numéricos y gráficos. Conceptos básicos y errores que evitar. Cir Esp. 2022. https://doi.org/10.1016/j.ciresp.2021.11.017

* *Corresponding author.*
E-mail address: lupe.gomez@upc.edu (G. Gómez Melis).

Survival time

The time that passes to a certain event of interest (e.g., death) is usually an asymmetric measure, with few long and many short times, resulting in an asymmetric, right-tailed distribution.[4] Survival studies require a long period to observe. However, some individuals, termed "censored", will end the follow-up without experiencing the event, indicating that the event-free time is longer than the observed time. Panel A of figure 1 in Arezzo's article represents the survival-to-death (OS) curves. The numbers at the bottom of the figure indicate the individuals "at risk" of the event (those alive) at the beginning of each 12-month interval for each treatment group: at baseline (time 0) all participants are at risk (53 in the SBTS group and 55 in ES); but at month 36, 35 and 40 participants remain in SBTS and ES, respectively. Consequently, 18 (=53−35) and 15 (=55−40) have either died or 'dropped out' of the study, perhaps because they have had less follow-up (e.g., they were included less than 36 months ago). The low number of individuals at risk after 48 months (4 in SBTS and 1 in ES) indicates that from this time onwards the available information comes from few observations and has greater uncertainty.

In short, censoring involves partial information about that individual's time. The most common is right-handed censoring, which occurs when an individual has not yet experienced the event, either because they have missed it during the study or because they have experienced another event that prevents them from observing the event of interest (competing risks).[5]

Survival function

The survival function for a time t is the probability that an individual does not suffer the event of interest before t. The Kaplan-Meier method allows its estimation using those who are still at risk at time t and are therefore likely to suffer the event at t. Panel A of Figure 1 of Arezzo's paper shows the Kaplan-Meier curve, with dips at the time points where deaths are observed and crosses for censorships at the instant they ended their follow-up. At first glance, it can be seen that there are no relevant differences between the 2 groups. At 36 months, survival is almost identical in both groups, with a value around .7, indicating that 70% of patients would survive more than 36 months. To find the median survival time, a horizontal line is drawn at the .5 value on the vertical axis, and thus finds the time for which the survival curve cuts this line: in our example, the medians are 52 and 42 months for the SBTS and ES groups, respectively.

Hazard rate function

The hazard rate function for a time t is the instantaneous rate of suffering the event of interest at time t. This "risk of suffering the event at time t" reports the events per unit of time (rate); it is a more sophisticated concept than the ''probability of surviving at time t'' provided by the survival function and should not be interpreted as a probability: it can be greater than one! It allows us to observe the frequency of the initial events, among a larger number of individuals, and the final ones. Its shape helps to define the statistical analysis, so the clinician must anticipate its expected shape or the general trend.[4] For example, the risk of having to undergo a particular type of surgery (e.g., prostate surgery) in initially healthy individuals in a particular age range (e.g., 45–50 years) may be considered constant. In contrast, the risk of death after highly invasive surgery may be high in the first 24 h and decrease after the second day. An increasing risk can be observed in populations with lethal diseases treated with ineffective treatments.

Competing risks and composite events

Time to disease progression (TTP) competes with time to death (TOD), in the sense that death from another cause precludes observing a time to progression that would have been after death. We are dealing with so-called competing events. Imagine a surgical intervention with high mortality. If one does not take into account that in patients who die it will be impossible to observe recurrence, one could conclude that this intervention decreases the risk of recurrence.

One way to avoid the problem of competing risks is to use composite events, such as the disease-free time variable. This variable captures the time to the first event (death or disease progression). By considering a single time, not only does this avoid dealing with the problem of competing events, but it also eliminates the potential multiplicity problems of analysing multiple responses.[6] Composite response variables also have the advantage of providing a higher probability of detecting a treatment effect if the components are not highly correlated.[7]

Final advice

- Confidence intervals: all relevant measures associated with a study should be reported with their uncertainty.[8]
- Hazard ratio: despite its great popularity, other measures that are based on lifetime gain (e.g., restricted mean survival time, RMST) are more interpretable and can help "informed" decision making.[9]
- Assumptions: if a model (e.g., Cox) is assuming some assumptions, they must be shown to be at least reasonable.
- Censorship: reasons for censorship should be communicated in any study.[10]
- Publication guidelines: review the recommendations of guidelines, e.g., CONSORT[11] in the case of a clinical trial, to increase the transparency and reproducibility of your study.

Financing

This article was funded by the Ministry of Science and Innovation (Spain), PID2019-104830RB-I00/ DOI (AEI): 10.13039/501100011033.

REFERENCES

1. Arezzo A, Forcignanò E, Bonino MA, Balagué C, Targarona E, Borghi F, et al. Long-term oncologic results after stenting as a bridge to surgery versus emergency surgery for malignant left-sided colonic obstruction. Ann Surg. 2020;272:703–8.
2. Hurwitz EE, Simon M, Vinta SR, Zehm CF, Shabot SM, Minhajuddin A, et al. Adding EXAMPLES to the ASA-physical status classification improves correct assignment to patients. Anesthesiology. 2017;126:614–22.
3. González Alastrué JA, Jover L. Los gráficos en la comunicación y el razonamiento científicos: ¿instrumento u ornamento? Med Clin (Barc). 2004;122:3–10.
4. Gómez G, Cobo E. Hablemos de análisis de Supervivencia. Gastroenterol Hepatol. 2004;3:185–91.
5. Clark TG, Bradburn MJ, Love SB, Altman DG. Survival analysis part i: basic concepts and first analyses. Br J Cancer. 2003;89:232–8.
6. Gómez G, Lagakos SW. Statistical considerations when using a composite endpoint for comparing treatment groups. Stat Med. 2013;32:719–38.
7. Bofill M, Cortés J, Gómez G. Decision tool and Sample Size Calculator for Composite Endpoints; 2020. arcXive:2001.03396 [stat.AP].
8. González Alastrué JA. Uso e interpretación de los intervalos de confianza. Med Clin Pract. 2021;4 Suppl 2:100297.
9. Hernán MA. The hazards of hazard ratios. Epidemiology. 2010;21:13–5.
10. Lang TA, Altman DG. Basic statistical reporting for articles published in *Biomedical Journals*: The 'Statistical analyses and methods in the published literature' or the SAMPL guidelines. Int J Nurs Stud. 2015;52.
11. Moher D, Hopewell S, Schulz KF, Montori V, Gøtzsche PC, Devereaux PJ, et al. CONSORT 2010 explanation and elaboration: updated guidelines for reporting parallel group randomised trials. BMJ. 2010;340.