This is an postprint (Accepted Manuscript) of an article published by Elsevier in Revista Española de Cardiología on 29 September 2021,available online: https://doi.org/10.1016/j.rec.2021.06.003

Focus on: Survival analyses in cardiovascular research, a practical guide

Enfoque: Análisis de supervivencia en investigación cardiovascular, una guía práctica

Survival analyses in cardiovascular research, part I: the essentials

Análisis de supervivencia en investigación cardiovascular (I): lo esencial

Xavier ROSSELLO[a,b,c,*] and Maribel GONZÁLEZ-DEL-HOYO[d]

[a] *Cardiology Department, Institut d'Investigació Sanitària Illes Balears (IdISBa), Hospital Universitari Son Espases, Palma de Mallorca, Spain*

[b] *Centro Nacional de Investigaciones Cardiovasculares (CNIC), Madrid, Spain*

[c] *Facultad de Medicina, Universitat de les Illes Balears (UIB), Palma de Mallorca, Spain*

[d] *Cardiology Department, Hospital Universitari Vall d'Hebron, Vall d'Hebron Institut de Recerca, Universitat Autònoma de Barcelona, Barcelona, Spain*

*** Author for correspondence:** Cardiology Department, Hospital Universitari Son Espases - IdISBa, Ctra. de Valldemossa 79, 07120 Palma de Mallorca, Islas Baleares, Spain.

E-mail address: fjrossello@cnic.es (X. Rossello).

ABSTRACT

This review is a practical guide to the essentials of survival analysis and their reporting in cardiovascular studies, although most of its key content can be extrapolated to other medical fields. This is the first in a series of two educational papers setting the grounds to address the most relevant statistical issues in survival analyses, which will smoothly drive the reader from the most basic analyses to the more complex situations. The focus will be on the type and shape of survival data, and the most common statistical methods, such as non-parametric, parametric and semi-parametric. Their adequacy, interpretation, advantages and inconveniences are illustrated by examples from the cardiovascular field. This article ends with a set of recommendations to guide the strategy of survival analyses for a randomized clinical trial and observational studies. Other topics, such as competing risks, multi-state models or recurrent-event methods will be addressed in the second article.

Keywords

Survival analysis

Cardiovascular disease

Methodology

RESUMEN

Esta revisión establece una guía práctica que comprende los conceptos básicos de los análisis de supervivencia y su aplicación en el estudio de las enfermedades cardiovasculares, si bien gran parte del contenido puede extrapolarse a otras ramas de la medicina. Este es el primero de dos artículos académicos que sientan las bases para abordar las principales cuestiones metodológicas empleadas en estudios de supervivencia, y guían al lector, desde los análisis más básicos hasta los más complejos. Esta revisión se centra en el tipo y forma de los datos de supervivencia, así como en los métodos estadísticos más utilizados, como las pruebas no paramétricas, paramétricas y semiparamétricas. La interpretación y valoración de la idoneidad de dichos métodos, así como sus ventajas e inconvenientes, son ilustradas con estudios del ámbito de las enfermedades cardiovasculares. El artículo concluye aportando un conjunto de recomendaciones para guiar en la estrategia del análisis de supervivencia, tanto en el contexto de un ensayo clínico aleatorizado como en el de estudio observacionales. En la segunda revisión se abordarán temas como el modelo de riesgos competitivos, el modelo de eventos recurrentes y los modelos multiestado.

Palabras clave

Análisis de supervivencia

Enfermedad cardiovascular

Metodología

Abbreviations

ACS: acute coronary syndrome

CPH: Cox proportional hazards

MACE: major adverse cardiac events

PH: proportional hazards

RCT: randomized clinical trial

Abreviaturas

ECA: ensayo clínico aleatorizado

MACE: eventos adversos cardiovasculares mayores

RP: riesgos proporcionales

RPC: riesgos proporcionales de Cox

SCA: síndrome coronario agudo

INTRODUCTION TO SURVIVAL ANALYSIS

This article gives an overview of survival analysis in cardiovascular research and aims to provide a practical tool to understand the range of statistical approaches to tackle time-to-event outcomes. We briefly describe features of survival time data and give an overview of several analytical approaches. For a better understanding we will illustrate how these methods have been applied to data from cardiovascular studies. This review is primarily descriptive in content, so no prerequisite mathematical, or statistical background are absolutely necessary. Basic concepts are summarised in figure 1.

What is survival analysis?

Survival analyses are applicable when the measure of interest is time until an event occurs (time-to-event outcomes). Despite using the word "survival", the event or outcome can be fatal or nonfatal (e.g., cardiovascular mortality or myocardial infarction, respectively).[1] The time at which the event occurs is referred to as an event time, survival time or failure time. The last two terms arose from the statistical methods developed for data analysis in cancer trials and for quality testing in manufacturing, respectively. Despite the term, the event under study is not necessarily a failure with negative connotations, meaning it could well be a positive event, such as appropriate implantable cardioverter defibrillator discharges.[2]

The survival time for each individual is the time from the starting point until the occurrence of the event of interest.[1,3] In a randomized clinical trial (RCT), survival time is usually estimated from the randomisation (usually happening with the start of the intervention).[4] In an observational study, survival time is more commonly calculated from entry into the study, from a fixed particular date, such as the date when first exposed to a risk factor (e.g., chemotherapy) or to an index event (e.g., acute coronary syndrome [ACS]).[5] Sometimes it is also calculated from an age point.

Sources of survival data

There are two main sources of survival data in cardiovascular research: RCTs and observational studies.

In RCTs the aim is usually to provide reliable evidence of treatment efficacy (or effectiveness), and safety. The simplest example involves individuals meeting the entry criteria, who are randomised to the intervention or the control group and followed-up for a given time to collect events and time to events. For example, the REBOOT trial (NCT03596385) is randomising patients with acute myocardial infarction and left ventricular ejection function ≥ 40% to either receive standard treatment with betablockers or without them and will be followed-up over a median period of 2.75 years to assess differences in the incidence of the composite endpoint of all-cause death, nonfatal reinfarction, or heart failure hospitalization.

In prospective observational studies (cohort studies), individuals differing in a primary exposure factor are recruited to a cohort and followed-up for a range of different times, in order to compare time-to-event outcomes in those exposed against those not exposed.[6] The EPICOR (l Long-term follow-up of antithrombotic management patterns in acute coronary syndrome patients) registry is an international cohort study, which recruited post-discharged patients with an ACS who were followed-up over 2 years for a variety of outcomes.[7] Its entry time (time zero) was fixed at discharge after an index event (ACS), and the events collected over follow-up were other cardiovascular events (such a recurrent ACS) and fatal events. The EPICOR registry has been used to evaluate differences in all-cause mortality across a variety of exposures, such as the geographic origin of the participants,[8,9] or whether they underwent coronary revascularization.[5]

In addition to RCTs and observational studies, other data sources for survival data that can be used, such as registries of administrative data. In Spain, the Spanish Minimum Basic Data Set has been used to describe temporal trends and in-hospital complications of several

cardiovascular conditions.[10] Nevertheless, their cross-sectional nature makes any time-to-event analysis impossible to happen.

Features of survival data: uninformative censoring

A key analytical problem in longitudinal studies (either RCTs or observational) is that the time to the event of interest may be censored, which means that for some subjects the follow-up may not be complete and hence, the event is not observed to happen. There are generally three reasons why censoring may occur during the study period of a given RCT evaluating major adverse cardiac events (MACE): *a)* the event might not be recorded for patients who are still alive at the end of follow-up; *b)* subjects lost to follow-up after a certain date (e.g., migration); and *c)* those who died from a different cause (e.g., cancer). In any of these situations, the actual survival time (e.g., time to MACE) is unknown. It would be inappropriate to exclude such individuals from the analyses, since the fact that they did not experience MACE whilst they were in the study, provides some information about survival. Of note, we only observe a time up to which we know they have *not* had the outcome at the "right side" of the follow-up period (phenomenon known as *right-censoring)*.[1]

The concept of censoring makes survival methods unique. For each patient we have two pieces of data: *a)* a time that is either the patient's event time, or the time at which the patient was last followed-up; and *b)* an indicator that denotes whether the time is an event time or a censoring time. In other words, each patient has either an event time or a censoring time after which the patient would no longer be observed. It will be assumed throughout most of this review that censoring is uninformative about event times.[1] This means that the time at which an individual is censored, or the fact that they are censored is not related with our time-to-event outcome. If a patient enrolled in a RCT drops out of the study, due to reasons related to the

study (e.g., subjects in the intervention arm feeling better or having side-effects) the censoring then becomes informative.

Violating this statistical assumption challenges the principles of the main approaches for survival data.[11] Methods to address censoring due to a competing event (e.g., death from another cause) will be discussed in our second review.

Key concepts in survival distributions

Two relevant functions are needed to understand and describe survival distributions. The survival function, S(t), is defined as the probability that a person "survives" beyond any specified time (or t), whereas the hazard function, h(t), is defined as the instantaneous failure rate. The hazard function is a conditional failure rate (conditional to those who have survived until time t). For instance, in a study assessing MACE, the function at year 2 only applies to those who were event-free in year 2, and it does not take into account those who had MACE before year 2. Note that, in contrast to the survival function, which focuses on not failing, the hazard function focuses on failing, that is, on the event occurring.

The hazard ratio (HR) is a measure of effect/association widely used in survival analyses. In its simplest form, the HR can be interpreted as the chance of an event occurring in the treatment arm (or exposed group) divided by the chance of the event occurring in the control arm (or unexposed group), of a RCT (or an observational study). The HR summarises the relationship between the instantaneous hazards (or event rates) in the two groups. The HR is estimated using regression-type analyses.

Aims and approaches of survival analyses

Most clinical studies focus on evaluating the impact of an intervention (e.g., betablockers) or exposure (e.g., geographic origin) on an outcome of interest (e.g., MACE). In

the setting of survival analyses, the outcome of interest is the survival time, and the focus is on either comparing survival times between treatment groups, or assessing the association between survival times and exposed *vs.* non-exposed subjects.

Methods for survival data need to account for censoring and the fact that survival times are strictly non-negative and commonly skewed to the right. If censored observations were excluded from the analysis, the results would be biased, and the estimates would be unreliable. Three basic approaches are presented in this review:

- *Non-parametric methods***.** These relatively simple methods (e.g., Kaplan-Meier estimates) do not make assumptions about the distribution of survival times. They are excellent for univariate analyses (e.g., primary outcome in RCTs), but not sufficient to handle more complex problems, such as confounding in observational studies.
- *Fully parametric methods.* They are regression-type analyses for survival data (e.g., Weibull model), which are analogous to regression approaches for other types of response (e.g., linear regression for continuous data, or logistic regression for a binary response). Nevertheless, they rely on some assumptions about the patterns of survival times, which should be carefully investigated.
- *Semi-parametric methods.* They lead to another regression-type analysis for survival data, often called Cox regression. The way survival times are associated with exposures of interest is parametrized, though they leave part of the full distribution of the survival times unspecified.

NON-PARAMETRIC ANALYSIS OF SURVIVAL DATA

Non-parametric techniques do not make any assumptions about the distribution of the survival times, which makes sense since survival data has a skewed distribution. The most common non-parametric approach for modelling the survival function is the Kaplan-Meier estimate. This method uses the actual observed event and censoring times.

Why use non-parametric methods?

Non-parametric methods are an adequate starting point for most survival analyses. First, these methods allow to estimate survivor and hazard functions without having to make parametric assumptions. Second, they provide a very intuitive way of graphically displaying survival data, taking into account censoring data. Third, these techniques are optimal to compare survival time by groups of individuals (categorical variables). Finally, non-parametric methods can be used to inform about whether some model assumptions can be made in more complex approaches of survival data (e.g., proportional hazards assumption).

Estimating the survival function: the Kaplan-Meier method

The simplest way to obtain an estimate of the survival function is by calculating the cumulative survival experienced by the subjects included in a cohort. This can be achieved using the Kaplan-Meier method, which estimates the cumulative survival (and failure) probabilities every time a subject leaves the study. The Kaplan-Meier estimate is also known as the "product limit estimate". This approach is based on conditional probabilities and is thoroughly explained in figure 2.

Another non-parametric approach is the life table estimate of the survival function. For the Kaplan-Meier estimate a discrete time setting is assumed (all survival times are observed at an exact time). Sometimes this piece of information is less precise and survival times are observed within a time range. A typical example is the evaluation of the number of deaths per year in a study population. In this setting, the life-table method is still popular in that it provides a simple summary of survival data in large study populations within time intervals.[12]

Recommendations for displaying Kaplan-Meier plots

Kaplan-Meier plots are presented in almost every article reporting survival analyses. Because of their prevalence, it is important to provide some suggestions for their appropriate representation. A poorly graphed Kaplan-Meier curve can lead to a wrong interpretation by the readership. In a landmark paper published around 20 years ago, Pocock, Clayton and Altman examined the survival plots that were shown in 35 RCTs and suggested some key recommendations to consider when presenting Kaplan-Meier plots.[3] In most cases, survival plots are best presented going upwards (cumulative proportion experiencing the event) rather than downwards (cumulative proportion free of event). Usually, plots going up highlight relative differences, because the x axis can be adapted to the scale of the curves. Instead, plots going down (using the whole vertical axis from 0 to 100%) make differences look much less pronounced and mainly emphasise differences in absolute terms. Therefore, a break in the y-axis should be avoided. Overlooking such a break, may lead the reader to a wrong perception of the potential treatment effect (or association), which would seem larger than the real difference. Ideally, plots should include some measure of statistical uncertainty - the standard error (or 95% confidence interval) for the estimated proportion of patients with (or without) the event can be calculated at any timepoint. In the rare cases when this uncertainty is presented, there is an increasing size of the standard error (or 95% confidence interval) bar over time, which illustrates the decreasing number of subjects at risk in the follow-up. Finally, plots should extend in time as far as there are sufficient number of subjects in the study and the number of participants at risk remaining in the study should be shown below the x-axis. If there is a large decline in the number of subjects as the study goes on, then caution should be exercised when interpreting differences between the curves in the right-hand side of the graph. These recommendations are illustrated in figure 3 with examples from articles published in *Revista Española de Cardiología*.[13–17]

The log-rank test to compare survival times in two groups

The log-rank test is used to compare the survival distributions of two or more groups. This non-parametric test, which only assumes the censoring to be non-informative, is based on a very simple idea: comparing observed vs. expected events in each group. For the comparison of two groups (e.g., testing the null hypothesis that the survival experience is the same for patients randomised to betablockers and control patients), we calculate the number of expected events in each group with the usual formula for 2x2 tables, for each time interval. The test will be carried out by comparting the sum of these "table-specific" expected values with the total observed events. If the observed numbers statistically differ from those expected, we may assume that the difference in the two groups is not happening at random. This approach can be also applied to explanatory variables with more than two groups, although the test statistic is more complex to calculate manually and its statistical power decreases as the number of comparisons increases.

Limitations of non-parametric methods

Non-parametric methods for analysing survival data are widely used, though they are not exempt from major limitations. They do not allow using continuous factors as dependent variables (it should be categorised to estimate hazards for people falling into different categories). Following the same reasoning, analyses involving several exposure variables, which may be binary continuous or categorical, are not possible unless each patient is categorised into a single category (e.g., for 2 factors with 3 categories, there will be 6 groups). This approach usually ends with small groups limiting any formal comparison. Even more importantly, there is no way to adjust for potential confounders, but to look separately at groups defined by the confounder (hence these methods quickly become cumbersome and the groups too small for

meaningful analysis). Taking all these limitations into account, a more advanced approach by using regression modelling can be applied to handle survival data.

PARAMETRIC REGRESSION MODELLING

Regression modelling is about establishing a mathematical model for the survival times, which defines how survival times depend on individual exposures (or randomised allocations). It is analogous to the use of other regression-type analyses, such as linear regression to study the dependence of continuous response variables on explanatory variables, or logistic regression to study the dependence of a binary outcome on explanatory variables. Regression models for survival data can be parametric or semi-parametric[18–20] and both provide a measure of effect known as HR. The models presented in this section are described as parametric, because they assume a particular shape for the distribution of the survival times which depends on defined parameters. In other words, survival times are assumed to follow a known distribution.

Why use parametric regression modelling for survival data?

Choosing a theoretical distribution to approximate survival data is as much an art as a scientific task, and regression modelling is going to be always a simplistic way to summarise the reality. Parametric models assume that the survival and hazard functions have a specific distribution which is often too structured and sometimes unrealistic for use with real data.[21] It is hard to isolate all potential causes that can lead to an event at a particular time, and to account mathematically for all of them. Nevertheless, there are some potential scenarios where it is appropriate to use parametric models.

The exponential model

The most basic parametric model for survival time is the exponential distribution, which is a form of the survival distribution characterized by a constant hazard function. According to this model, having a failure (or event) is a random event independent of time. The exponential distribution, often referred to as a purely random failure pattern, is well-known for its ''lack of memory'' which means that a subject's probability of an event in the future is independent of how long the subject has gone without an event[1]. The exponential distribution is characterized by having a single parameter (λ), a constant hazard rate. Hence, a high λ value indicates high risk and short survival, whereas a low λ value indicates low risk and long survival.[18] Exponential models are rarely used in the cardiovascular field, but it has been applied in other areas, such as cancer[22] or smoking cessation.[23]

The Weibull model

In many scenarios, it would not be reasonable to assume a constant hazard rate over time,[18] as under the exponential distribution. An alternative approach is to consider the Weibull distribution to parametrize survival times. Unlike the exponential distribution, it does not assume a constant hazard rate (hazard function is increasing or decreasing over time) and therefore has more flexibility, and a broader application. This method can be used when the hazard function in monotonically increasing or decreasing, and therefore is characterized by two parameters: the scale parameter λ (the same than for the exponential distribution), and the shape parameter γ, determining the shape of the distribution curve.[24] By adding a shape parameter, the distribution becomes more flexible and can fit more types of data (with increasing, decreasing, or constant risk). A value of $\lambda < 1$ indicates that the failure rate decreases

over time, whereas if λ > 1 the failure rate increases with time. When the scale parameter (λ) equals 1, the failure rate is constant over time (Weibull reduces to an Exponential model).

The reality about parametric regression modelling

Parametric regression models produce smooth predictions by assuming a functional form of the hazard and by directly estimating absolute and relative effects.[24] In addition to the Exponential and Weibull models, there are other parametric models (e.g., Gompertz, Lognormal, etc.),[18] but they fail beyond the scope of this review because the use of parametric modelling for survival data is relatively rare in cardiovascular research. Nevertheless, some recent examples can be found. In two recent publications from major RCTs, the Weibull regression model has been used to evaluate time to discontinuation to treatment allocation for informative reasons (e.g., death).[25,26] In the ENDURANCE trial, which compared a small centrifugal-flow left ventricular assist device to an axial-flow left ventricular assist device in patients with advanced heart failure who were ineligible for heart transplantation, the Weibull model was used for the primary endpoint of survival free from disabling stroke or device removal for malfunction or failure.[27]

THE COX PROPORTIONAL HAZARDS MODEL

The Cox proportional hazards (CPH) model is the most popular approach to evaluate the relationship between covariates and survival[1,18] and was first proposed by Cox in 1972 to identify differences in survival due to treatment and prognostic factors in clinical trials.[19,28] Whereas non-parametric methods typically study the survival function, with regression methods the focus is on hazard function.

The CPH is referred to as a semi-parametric approach,[28] as under this model the baseline hazard function is not parametrized, but the effect of the explanatory variables on the hazard is

parametrized (using β coefficients). In other words, the baseline hazard is left unspecified (not written in terms of parameters to be estimated), which means that no assumptions are made regarding the baseline hazard function for each group (survival rates can vary between exposed and unexposed in an observational study, or intervention and control groups in a RCT[1]), whereas the ratio of their hazards is assumed to be constant. What is parametrized are the effects of the covariate variables upon survival, which are constant over time and additive in one scale. We can test the association of each of the independent variables with survival time adjusted for other covariates.

Model assumptions

Like any other statistical model, the CPH method relies on some assumptions. Nevertheless, the semi-parametric approach makes fewer assumptions than the alternative parametric methods:

- The main assumption of the Cox model is that the HRs (defined as the ratio of the hazards between 2 groups) are independent of time. In other words, the ratio of hazards between groups is constant over time. This is known as the proportional hazards (PH) assumption. If an explanatory variable has a strong association with survival at the beginning of the follow-up but a weaker effect later on, this would be considered a violation of the PH assumption.
- The assumption that we have correctly specified the form of the explanatory variable – this assumption is bound to the PH assumption, since this may hold for a specific form of the explanatory variable but not for another. To comply with this assumption, explanatory variables might benefit from some transformation. Some continuous variables can be remodelled either using a binary cut-off point, for example for body mass index and haemoglobin, or using its continuous nature only above a certain

threshold, for example for creatinine and blood glucose.[29] The most common situation is the log transformation of highly skewed variables.

- Censoring is uninformative.
- Observations are independent.

Based on the above, the Cox model can fit any distribution of survival data if the PH assumption is valid (actually, most HRs are fixed proportional). For this reason, the Cox model is used so widely.

Model checking

Before reporting any findings, we should check, as far as possible, that the fitted model is correctly specified. If not, our inferences could be invalid, and we may draw incorrect conclusions. Model checking is not an easy issue in survival analysis. In this section, the focus will be on assessing the PH assumption, though some methods for assessing whether the functional form for the explanatory variables is correct, are also provided. The last two assumptions are not formally testable.

There are three main ways of assessing whether the PH assumption can reasonably hold.[30] The first of these methods involves the use of Kaplan-Meier estimates. If a categorical predictor satisfies the PH assumption, then the plot of the survival function *vs.* the survival time will yield two parallel curves. Similarly, the graph of the log(-log[survival]) vs. log of survival time graph should result in parallel lines if the predictor is proportional. This method becomes problematic with variables having many categories and it does not work well for continuous variables, unless they are categorised. Furthermore, the situation becomes too complex in multivariate models with several explanatory variables, which should need to be simultaneously tested using combinations of the covariates.[31] Hence, this approach is quickly becoming

cumbersome and unrealistic for more than 2 or 3 variables. The second method to check for the PH assumption is by formally testing whether the effect of explanatory variables on the hazard changes over time (this can be done by including an interaction between time and the explanatory variable into the model). A significant interaction would imply the hazard function changes over time, and the PH model assumption would be violated. Scaled Schoenfeld residual plots provide the third useful way to assess this assumption. Schoenfeld residuals can essentially be thought of as the observed minus the expected values of the covariates at each failure time.[31,32] The plot of Schoenfeld residuals against time for any covariate should not show a time-dependent pattern.

In addition to checking for the PH assumption, other aspects of model fit can be assessed using residuals.[30] Martingale residuals can be used as a way of investigating the appropriate functional form for continuous variables. A Martingale residual is the difference between what happened to a person (whether they had the event or not) and what is predicted to happen to a person under the model that has been fitted.[33] A plot of the martingale residuals from a null model (model without explanatory variables) against a continuous variable can be used to indicate the appropriate functional form for the continuous variable when it is entered in the model. Deviance residuals and other approaches are beyond the scope of this review.

Final remarks about non-parametric, parametric and semi-parametric approaches

Kaplan-Meier plots are an excellent way of graphically illustrating the survival experience over time, particularly when the rate of the outcome is changing irregularly. Should the exact survival or censoring time of each patient be known, the plots are easy to produce by the researchers, and easy to interpret by the readership. The log-rank test is a simple tool providing a significance test for comparing survival times between groups. However, the log-rank test only provides a p-value and does not provide a measure of effect, unlike any

regression-type approach providing a HR. A further advantage of the HR approach (either parametric or semi-parametric) is that it provides tools to investigate confounding and effect modification. Hence, non-parametric approaches are used to explore data and provide crude estimates, but regression-type approaches are commonly used to provide more precise estimates.

Developing an analysis strategy is a complex task. In survival analysis, incorporating model checking for PH assumption into our strategy adds an extra layer of complexity. This can be done at different points in the model building process. Nevertheless, if the PH assumption does not hold, some other options can be applied (changing the functional form of the explanatory variables entered in the model, using stratified CPH models, or using another type of approaches that will be explained in the second part of this review).

STRATEGY OF ANALYSIS: SET OF RECOMMENDATIONS AND SUGGESTIONS

Different strategies can be used in the setting of survival analyses. The approach to the data mostly depends on the type of study and the research question. In this section, we provide some suggestions for RCTs focused on estimating a treatment effect, and for observational studies focused on evaluating associations between exposures and survival times. These recommendations are not the only way to perform valid analyses – they are just a summary of a method covering all key aspects.

Randomised controlled trial

In a large RCT we can fairly assume that there are variables confounding the association between the randomised intervention and the outcome. One sensible strategy to evaluate survival data would be to *a)* describe the numbers of participants in the allocation groups and summarize the number of events in the each group; *b)* provide Kaplan-Meier estimates of the

survivor curves in the treatment groups, and use the log-rank test to evaluate the alternative hypothesis (survivor curves are different between treatment groups); *c)* use plots to informally evaluate whether a PH model would be adequate to yield a reliable estimate for the association between treatment and survival time (plots appear parallel when hazards for both groups are proportional); *d)* if a PH model seems reasonable, fit a CPH model (or a parametric model, if appropriate) to estimate the HR with its 95% confidence interval, and corresponding *P*-value; and *d)* perform more formal assessments of the PH assumption (e.g., by testing for an interaction of the treatment effect with time, or by plotting Schoenfeld residuals).

Observational studies

The research question raised in an observational study determines the choice of explanatory variables to be used in a survival model. Sometimes the focus is on estimating an exposure "effect" –the interest is mainly in one particular exposure, but other variables are needed to control for potential confounding. On other occasions, the focus is on understanding the associations between a set of explanatory variables and the time-to-event. In this case, the interest is in the independent concomitant associations of several exposures on survival, perhaps to determine which has the greatest impact on survival. Strategies for addressing these two types of observational studies are presented in the next section.

Prediction modelling, where the aim is to use a set of variables to build a model to predict survival in individuals from a new cohort,[34,35] is not covered in this review. In this setting, the term "predictor" is better suited than the term "explanatory variable", and the strategy is focused on producing precise estimates for outcome prediction and assisting in risk stratification and clinical decision making. Further details about prediction modelling can be found elsewhere.[34,35]

In case of observational studies aimed to estimate an exposure "effect", we suggest the following 3-step process based on a preliminary approach, a main analysis, and a last set of checks for modelling assumptions. Preliminary analyses would include: *a)* using Kaplan-Meier plots and log-rank tests to assess univariate associations between each exposure and the outcome; *b)* using non-parametric and residual plots to make preliminary assessments of the PH assumption for each exposure; and *c)* evaluating the association between the main exposure and each potential confounder by a simple but visual method (e.g., using cross-tabulation). Assuming a PH model would be appropriate (either a parametric model, or most commonly a Cox model), the main analysis would therefore involve the following 7 steps: *a)* fitting a survival model using only the main exposure to estimate the univariate association; *b)* fitting further survival models using the main exposure and adding each potential confounder, one at a time; *c)* evaluating the impact of adjusting for each confounder on the estimated association between the main exposure and survival (e.g., change in magnitude and direction of HR); *d)* assessing whether any covariate modifies the association between the main exposure and the event (e.g., whether there are interactions); *e)* fitting a multivariate model for the main exposure, with adjustment for confounders (pre-specified based on clinical knowledge, or found in step *c* and for interactions (found in step *d*); *f)* forcing those remaining potential confounders, not included in the model, back into the model one by one to evaluate whether they are confounders in the presence of other covariates; and *g)* adding any relevant confounders found in step *f* to the final model. Finally, further checks for modelling assumptions and overall model fit should be made before formally reporting the findings (e.g., by testing for an interaction of the treatment effect with time, or by plotting Schoenfeld residuals of the final model).

In a situation where there is not a clear prior hypothesis about which explanatory variables may be associated with survival, sometimes it may be necessary to perform an exploratory analysis. If there are not too many variables, it would be sensible to include them all in a model and evaluate the associations between each variable and the outcome after the

full adjustment for all other covariates. Alternatively, a similar 3-step process, based on a preliminary approach, a main analysis, and a model check, would be useful. Preliminary analyses would include using Kaplan-Meier plots and log-rank tests to evaluate the association between each variable and the outcome. After this first approach, the main analysis needs to address the issue of selecting the "best" set of covariates. This approach would require from the following tasks: *a)* evaluating each variable separately in a sequence of PH models (Cox or parametric, when appropriate) in order to assess univariate associations between explanatory variables and the outcome; *b)* including all variables selected in the previous step in a single PH model, and then excluding each variable one by one, to assess whether the exclusion has a significant impact on the log likelihood (statistical significance can be tested using likelihood ratio tests); *c)* entering each of the variables that have been removed in the previous step back into the model one by one, to check whether they add anything to the model (using likelihood ratio tests), and *d)* repeating step *c* as many times as needed for all explanatory variables which remain out of the model in each "cycle". Finally, further checks on modelling assumptions should be performed, as explained for other types of analyses.

Other approaches for the last two scenarios would be to use automatic tools to select explanatory variables, such as forward or backward selection step processes. Although they are valid in some scenarios (particularly prediction modelling), one should bear in mind that regression modelling is as much an art as a scientific task, that requires medical knowledge, and statistical expertise and experience.

CONCLUSIONS

This is the first in a series of two educational articles reviewing the basic concepts of survival analyses, setting the grounds to understand, compare and apply the most relevant statistical models for survival analyses. Aiming to reach the readership without much background in statistics, a key feature of this review is the integration of the essentials of survival

analysis with practical examples from cardiovascular studies, showing the reader how to perform survival analysis and how to interpret the results. The main model assumptions are presented to provide the tools to apply the adequate statistical model according to the shape of the data. Additionally, recommendations to guide the strategy of analyses have also been provided, hoping that they might have an impact on the critical appraisal and statistical performance of the readership. The second article will tackle a variety of more complex statistical challenges that are often faced in survival analyses. These include competing risks, recurrent-event methods, multi-state models, or the use of restricted mean survival time.

FUNDING

None.

AUTHOR CONTRIBUTIONS

X. Rossello and M. González-del-Hoyo conceived the review. X. Rossello led the writing process, and M. González-del-Hoyo was in charge of finding most of the examples to illustrate the methods. X. Rossello drafted the article, though both authors contributed substantially to its revision.

CONFLICTS OF INTEREST

None declared.

REFERENCES

1. Rao SR, Schoenfeld DA. Survival Methods. *Circulation*. 2007;115:109-113.

2. McLeod CJ, Ommen SR, Ackerman MJ, et al. Surgical septal myectomy decreases the risk for appropriate implantable cardioverter defibrillator discharge in obstructive hypertrophic

cardiomyopathy. *Eur Heart J*. 2007;28:2583-2588.

3. Pocock SJ, Clayton TC, Altman DG. Survival plots of time-to-event outcomes in clinical trials: good practice and pitfalls. *Lancet*. 2002;359:1686-1689.

4. Rossello X, Ariti C, Pocock SJ, et al. Impact of mineralocorticoid receptor antagonists on the risk of sudden cardiac death in patients with heart failure and left-ventricular systolic dysfunction: an individual patient-level meta-analysis of three randomized-controlled trials. *Clin Res Cardiol*. 2019;108:477-486.

5. Bueno H, Rossello X, Pocock SJ, et al. In-Hospital Coronary Revascularization Rates and Post-Discharge Mortality Risk in Non–ST-Segment Elevation Acute Coronary Syndrome. *J Am Coll Cardiol*. 2019;74:1454-1461.

6. Pascual I, Hernández-Vaquero D, Alperi A, et al. Survival in elderly patients with transcatheter aortic valve implants compared with the general populationv. *Rev Esp Cardiol*. 2020;73:822-827.

7. Rossello X, Medina J, Pocock S, et al. Assessment of quality indicators for acute myocardial infarction management in 28 countries and use of composite quality indicators for benchmarking. *Eur Heart J Acute Cardiovasc Care*. 2020;9:911-922.

8. Rosselló X, Huo Y, Pocock S, et al. Global geographical variations in ST-segment elevation myocardial infarction management and post-discharge mortality. *Int J Cardiol*. 2017;245:27-34.

9. Bueno H, Rossello X, Pocock S, et al. Regional variations in hospital management and post-discharge mortality in patients with non-ST-segment elevation acute coronary syndrome. *Clin Res Cardiol*. 2018;107:836-844.

10. Abu-Assi E, Bernal JL, Raposeiras-Roubin S, Elola FJ, Fernández Pérez C, Íñiguez-Romo A. Temporal trends and prognostic impact of length of hospital stay in uncomplicated ST-segment elevation myocardial infarction in Spain. *Rev Esp Cardiol*. 2020;73:479-487.

11. Pintilie M. An introduction to competing risks analysis. *Rev Esp Cardiol*. 2011;64:599-605.

12. Rossello X, Gil V, Escoda R, et al. Editor's Choice– Impact of identifying precipitating factors on 30-day mortality in acute heart failure patients. *Eur Heart J Acute Cardiovasc Care*. 2019;8:667-680.

13. Miró Ò, Rosselló X, Gil V, et al. The Usefulness of the MEESSI Score for Risk Stratification of Patients With Acute Heart Failure at the Emergency Department. *Rev Esp Cardiol.* 2019;72:198-207.

14. Wiebe J, Dörr O, Liebetrau C, et al. Outcome After Long-segment Stenting With Everolimus-eluting Bioresorbable Scaffolds Focusing on the Concept of Overlapping Implantation. *Rev Esp Cardiol.* 2016;69:1144-1151.

15. Sabaté M, Cánovas S, García E, et al. In-hospital and Mid-term Predictors of Mortality After Transcatheter Aortic Valve Implantation: Data From the TAVI National Registry 2010-2011. *Rev Esp Cardiol.* 2013;66:949-958.

16. López-de-Sá E, Martínez A, Anguita M, Dobarro D, Jiménez-Navarro M. Aldosterone Receptor Antagonist Use After Myocardial Infarction. Data From the REICIAM Registry. *Rev Esp Cardiol.* 2011;64:981-987.

17. Ielasi A, Latib A, Chieffo A, et al. Very Long-term Outcomes Following Drug-eluting Stent Implantation for Unprotected Left Main Coronary Artery Stenosis: A Single Center Experience. *Rev Esp Cardiol.* 2013;66:24-33.

18. Royston P, Parmar MKB. Flexible parametric proportional-hazards and proportional-odds models for censored survival data, with application to prognostic modelling and estimation of treatment effects. *Stat Med*. 2002;21:2175-2197.

19. Cox DR. Regression Models and Life-Tables. *J R Stat Soc Ser B*. 1972;34:187-220.

20. Hosmer DW, Lemeshow S, *Applied Survival Analysis : Regression Modeling of Time to Event Data*.2nd Ed. New York: Wiley; 1998. p. 87-112.

21. Orsini N. Review of Flexible Parametric Survival Analysis Using Stata: Beyond the Cox Model by Patrick Royston and Paul C. Lambert. *Stata J*. 2013;13:212-216.

22. Walle AJ, Al-Katib A, Wong GY, Jhanwar SC, Chaganti RSK, Koziner B. Multiparameter characterization of L3 leukemia cell populations. *Leuk Res*. 1987;11:73-83.

23. Elketroussi M, Fan DP. Time trends of smoking cessation analyzed with six mathematical survival models. *Int J Biomed Comput*. 1991;27:231-244.

24. Du X, Li M, Zhu P, et al. Comparison of the flexible parametric survival model and Cox model in estimating Markov transition probabilities using real-world data. *PLoS One*. 2018;13:e0200807.

25. Badve S V., Pascoe EM, Tiku A, et al. Effects of Allopurinol on the Progression of Chronic Kidney Disease. *N Engl J Med*. 2020;382:2504-2513.

26. Spertus JA, Jones PG, Maron DJ, et al. Health-Status Outcomes with Invasive or Conservative Care in Coronary Disease. *N Engl J Med*. 2020;382:1408-1419.

27. Rogers JG, Pagani FD, Tatooles AJ, et al. Intrapericardial Left Ventricular Assist Device for Advanced Heart Failure. *N Engl J Med*. 2017;376:451-460.

28. Cox DR. Partial Likelihood. *Biometrika*. 1975;62:269-276.

29. Pocock SJ, Huo Y, Van de Werf F, et al. Predicting two-year mortality from discharge after acute coronary syndrome: An internationally-based risk score. *Eur Heart J Acute Cardiovasc Care.* 2019;8:727-737.

30. Harrell FE, Lee KL, Mark DB. Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. *Stat Med*. 1996;15:361-387.

31. Hess KR. Graphical methods for assessing violations of the proportional hazards assumption in Cox regression. *Stat Med*. 1995;14:1707-1723.

32. Schoenfeld D. Partial Residuals for The Proportional Hazards Regression Model. *Biometrika*. 1982;69:239-241.

33. Therneau TM, Grambsch PM, Fleming TR. Martingale-based residuals for survival models. *Biometrika*. 1990;77:147-160.

34. Moons KGM, Kengne AP, Woodward M, et al. Risk prediction models: I. Development, internal validation, and assessing the incremental value of a new (bio)marker. *Heart*. 2012;98:683-690.

35. Moons KGM, Kengne AP, Grobbee DE, et al. Risk prediction models: II. External validation, model updating, and impact assessment. *Heart*. 2012;98:691-698.

FIGURES

Figure 1. Basic concepts defining survival analyses. List of basic terminology in survival methods.

Figure 2. How to build a Kaplan-Meier curve and estimate their survival function. A: in this example, we created a basic data set of 10 patients with a follow-up of 30 days. All of them are "at risk" of presenting the event at time 0. Five patients have the event (days 5, 10, 25, 25 and 30), whereas five patients have censored data: 4 patients are lost to follow-up (2 at day 15 and 2 at day 25), and only 1 patient finishes the follow-up without an event (censored at day 30). B: estimation of the survival function using the Kaplan-Meier method. C: interpretation of cumulative survival probabilities. As it happens with this example, the Kaplan-Meier survival estimate sometimes takes the form of a step plot rather than a smooth curve, reflecting the fact that this is an empirical estimate of the survival experience of a cohort at every point in time. Importantly, censored observations do not cause a reduction in the cumulative survival, but adjust the number at risk for when the next death occurs.

Figure 3. A-E: five examples of different ways to display Kaplan-Meier curves. Reproduced with permission from Elsevier and *Revista Española de Cardiología*. The source of each plot is displayed below each Kaplan-Meier plot and full references are shown in the reference list.[13–17]

Figura 1

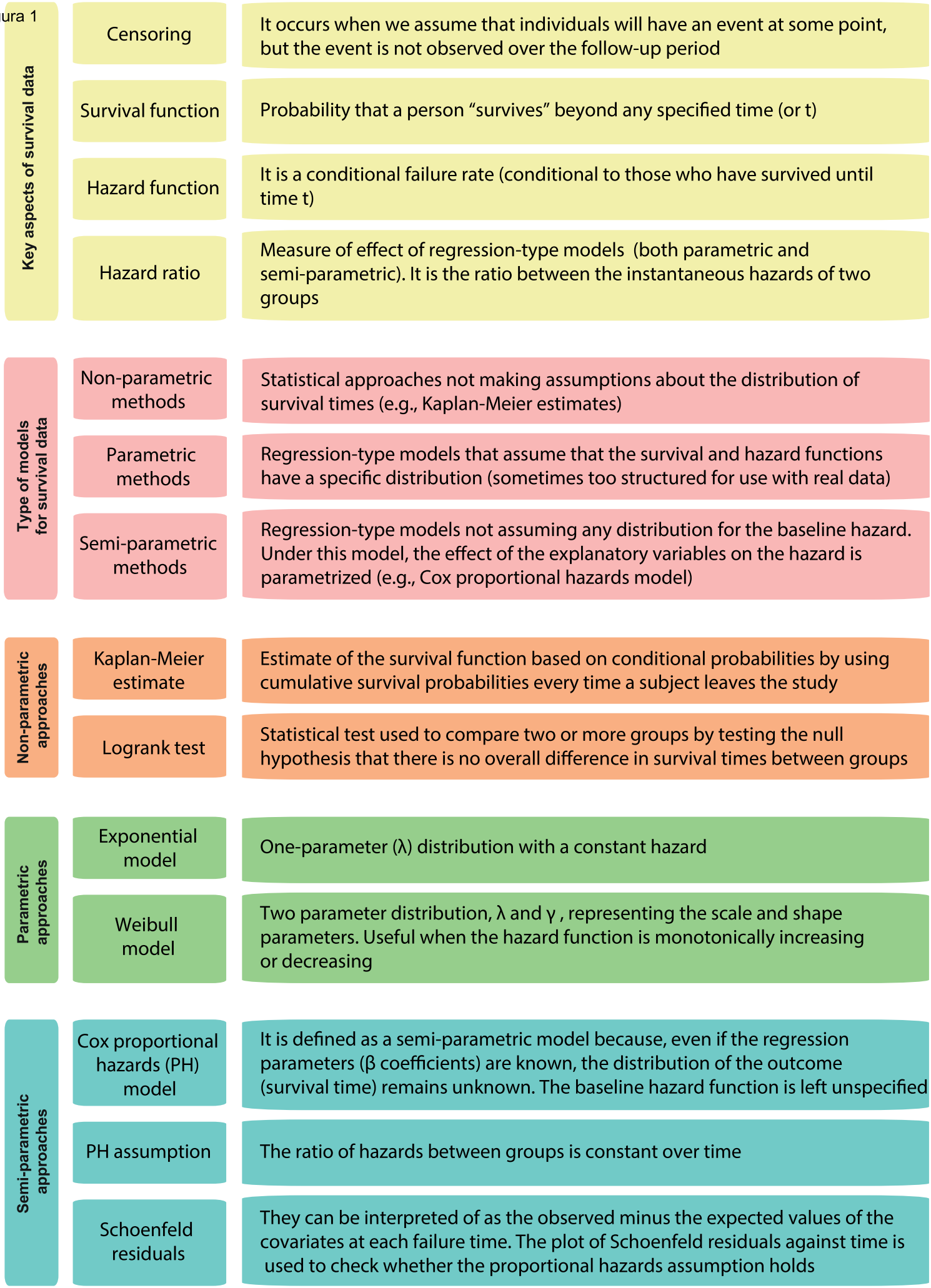
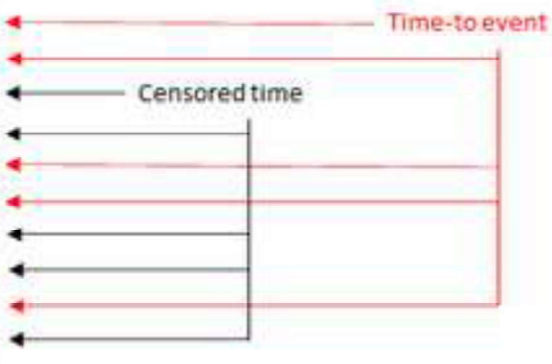

Key aspects of survival data
Censoring
It occurs when we assume that individuals will have an event at some point, but the event is not observed over the follow-up period
Survival function
Probability that a person "survives" beyond any specified time (or t)
Hazard function
It is a conditional failure rate (conditional to those who have survived until time t)
Hazard ratio
Measure of effect of regression-type models (both parametric and semi-parametric). It is the ratio between the instantaneous hazards of two groups
Type of models for survival data
Non-parametric methods
Statistical approaches not making assumptions about the distribution of survival times (e.g., Kaplan-Meier estimates)
Parametric methods
Regression-type models that assume that the survival and hazard functions have a specific distribution (sometimes too structured for use with real data)
Semi-parametric methods
Regression-type models not assuming any distribution for the baseline hazard. Under this model, the effect of the explanatory variables on the hazard is parametrized (e.g., Cox proportional hazards model)
Non-parametric approaches
Kaplan-Meier estimate
Estimate of the survival function based on conditional probabilities by using cumulative survival probabilities every time a subject leaves the study
Logrank test
Statistical test used to compare two or more groups by testing the null hypothesis that there is no overall difference in survival times between groups
Parametric approaches
Exponential model
One-parameter (λ) distribution with a constant hazard
Weibull model
Two parameter distribution, λ and γ , representing the scale and shape parameters. Useful when the hazard function is monotonically increasing or decreasing
Semi-parametric approaches
Cox proportional hazards (PH) model
It is defined as a semi-parametric model because, even if the regression parameters (β coefficients) are known, the distribution of the outcome (survival time) remains unknown. The baseline hazard function is left unspecified
PH assumption
The ratio of hazards between groups is constant over time
Schoenfeld residuals
They can be interpreted of as the observed minus the expected values of the covariates at each failure time. The plot of Schoenfeld residuals against time is used to check whether the proportional hazards assumption holds

Figura 2

Click here to access/download;Figuras (Figures);Fig 2 - REC - review 1.tif

A. List of patients with their observed follow-up time and status

Patient	Follow-up (days)	Event
A	5	1
B	10	1
C	15	0
D	15	0
E	25	1
F	25	1
G	25	0
H	25	0
I	30	1
J	30	0



B. Estimated survivor function (Kaplan-Meier estimate) based on conditionals probabilities

Periods of time (in days)	No. of patients at risk	No. of events	No. of censorings	Prob. of death	Survivor function
(0,5]	10	1	0	1/10=0.10	(1-1/10)=0.90
(5, 10]	9	1	0	1/9=0.11	(1-1/10)*(1-1/9)=0.80
(10, 15]	8	0	2	0	(1-1/10)*(1-1/9)*(1-0)=0.80
(15, 25]	6	2	2	2/6=0.33	(1-1/10)*(1-1/9)*(1-2/6)=0.53
(25, 30]	2	1	1	1/2=0.50	(1-1/10)*(1-1/9)*(1-2/6)*(1-1/2)=0.27

Key notes from the table:

1. Each row of the table represents a period finishing with at least one person leaving the study, either because of an evento or censoring
2. In each consecutive interval, the number at risk is reduced by the number of patients with the events or censored observations occurring in the previous Interval
3. The probability of dying is estimated for each interval separately.
4. The cumulative probability of surviving to the end of an interval is estimated using conditional probabilities because in order to have survived to the end of a period interval, a person must have survived all previous periods. This conditional probability is obtained by multiplying the sequence of the period-specific conditional Survival probabilities.

C. Interpretation of cumulative Survival plot using the Kaplan-Meier method

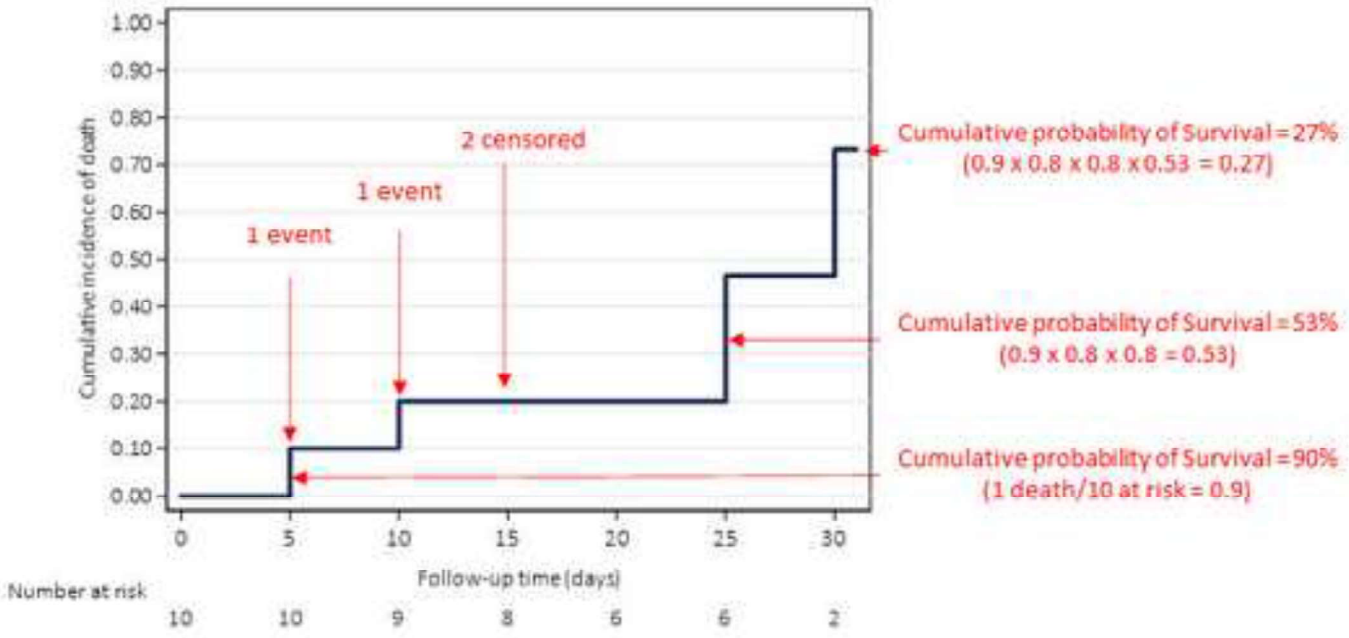
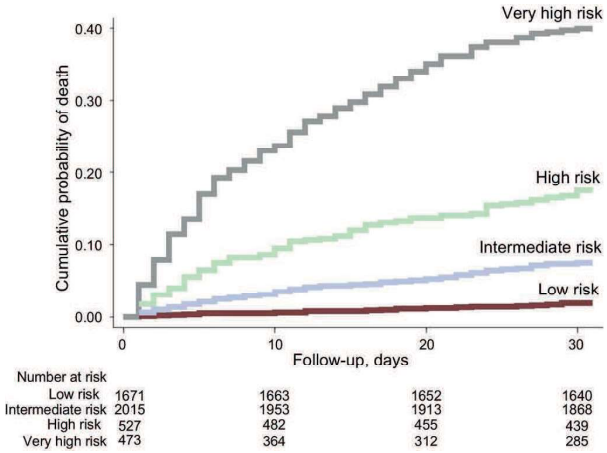


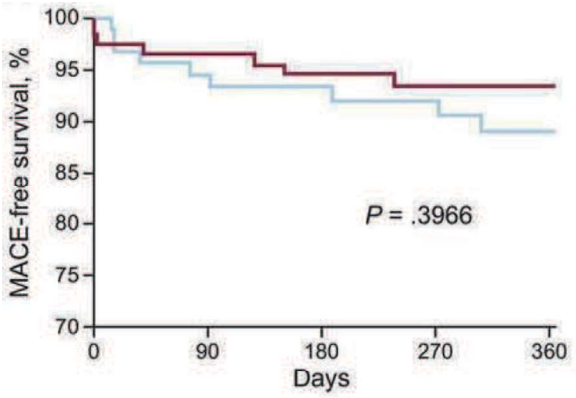
Figura 3

A **Survival plot going upwards with the number still at risk listed under the time axis. No uncertainty displayed at any time-point**



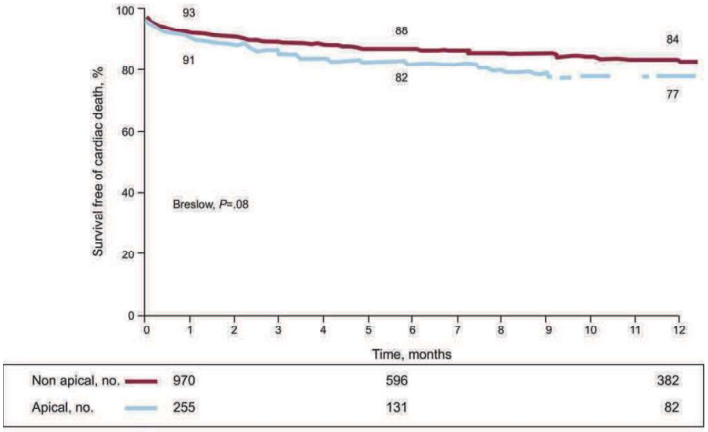
Source: Rev Esp Cardiol (Engl Ed) 2019;72:198-207

B **Survival plot going down and without the number still at risk listed under the time axis**



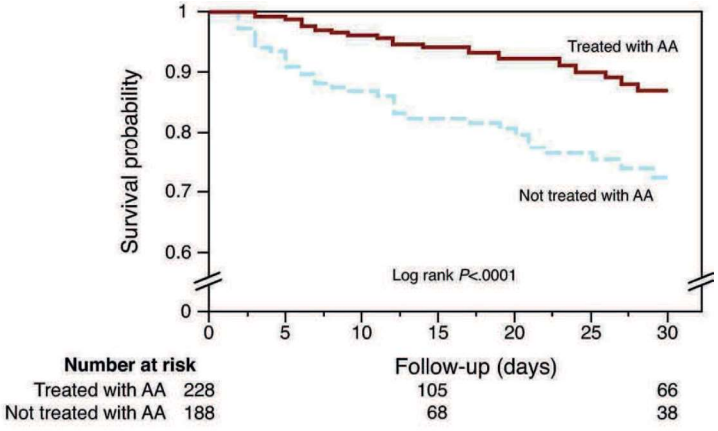
Source: Rev Esp Cardiol (Engl Ed) 2016;69:1144-1151

C **Survival plot going down and emphasising a small difference in absolute terms between groups. Displaying area of uncertainty would have shown great overlapping.**



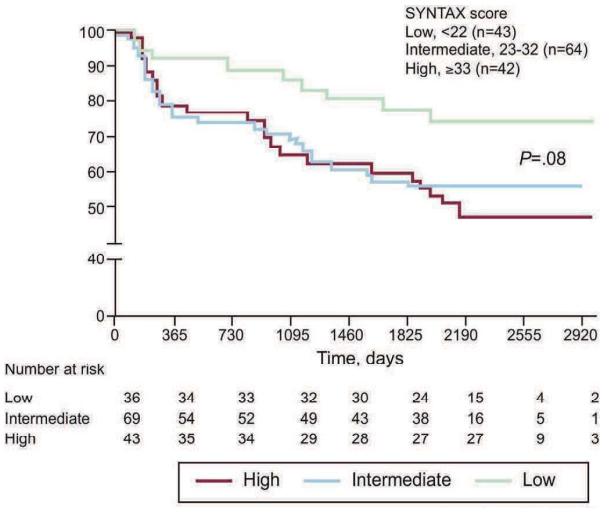
Source: Rev Esp Cardiol (Engl Ed) 2013;66:949-958

D **Survival plot going down and with a break in the y-axis emphasising the real difference in absolute terms**



Source: Rev Esp Cardiol 2011;64:981-987

E **Survival plot going down and with a break in the y-axis. The right-hand end of the plot offers little information given the numbers of participants remaining at risk (e.g. low and intermediate categories are crossing each other several times and separate where the plot is less informative)**



Source: Rev Esp Cardiol (Engl Ed) 2013;66:24-33