



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 II: Statistical methods in challenging situations

Análisis de supervivencia en investigación cardiovascular (II): metodología estadística en situaciones complejas

Xavier ROSSELLO^{a,b,c,d,*} and Maribel GONZÁLEZ-DEL-HOYO^e

^a *Servei de Cardiologia, Institut d'Investigació Sanitària Illes Balears (IdISBa), Hospital Universitari Son Espases, Palma de Mallorca, Balearic Islands, Spain*

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

^c *Facultad de Medicina, Universitat de les Illes Balears (UIB), Palma de Mallorca, Balearic Islands, Spain*

^d *Medical Statistics Department, London School of Hygiene & Tropical Medicine, London, United Kingdom*

^e *Servei de Cardiologia, Hospital Universitari Vall d'Hebron, Vall d'Hebron Institut de Recerca, Universitat Autònoma de Barcelona, Barcelona, Spain*

Received 7 June 2021

Accepted 21 June 2021

*** Corresponding author:** Servei de Cardiologia, Hospital Universitari Son Espases - IdISBa, Ctra. de Valldemossa, 79, 07120 Palma, Balearic Islands, Spain.

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

ABSTRACT

This article is the second part of a series of two educational articles. In the first paper we described the basic concepts of survival analysis, summarising the common statistical methods and giving a set of recommendations to guide the strategy of survival analyses in randomized clinical trials and observational studies. Here, we introduce stratified Cox models and frailty models, as well as the immortal time bias arising from a poor assessment of time-dependent variables. In order to address the issue of multiplicity of outcomes, we provide several modelling strategies to deal with other type of time-to-event data analyses, such as competing risks, multi-state models and recurrent-event methods. This review is illustrated with examples from previous cardiovascular research publications, and each statistical method is discussed alongside its main strengths and limitations. Finally, we provide some general observations about alternative statistical methods with less restrictive assumptions, such as, the win ratio method, the restrictive mean survival time, and accelerated failure time model.

Keywords

Survival analysis

Cardiovascular disease

Methodology

RESUMEN

Esta revisión es la segunda parte de dos artículos sobre metodología estadística. En el primero, se describían los conceptos básicos del análisis de supervivencia, los métodos estadísticos más comúnmente utilizados, y se aportaba un conjunto de recomendaciones para ayudar a establecer una estrategia de análisis de supervivencia, tanto en el contexto de un ensayo clínico aleatorizado como en el de un estudio observacional. En este segundo artículo, se introducen el

1 modelo estratificado de Cox y el modelo de fragilidad, y se ilustra el sesgo de tiempo inmortal
2 secundario a una evaluación errónea de variables tiempo-dependientes. Para abordar el
3 problema de la existencia múltiples *eventos clínicos*, se introducen distintas aproximaciones
4 estadísticas, como el análisis de riesgos competitivos, los modelos multiestado y el modelo de
5 eventos recurrentes. Todos ellos se ilustran con ejemplos del campo cardiovascular,
6 resumiéndose las principales ventajas y limitaciones de cada uno de los métodos estadísticos.
7 Finalmente, se presentan algunas consideraciones generales sobre métodos estadísticos
8 alternativos, con asunciones menos restrictivas, como el método *win ratio*, el tiempo de
9 supervivencia medio restringido y el modelo de tiempo de evento acelerado.
10
11
12
13
14
15
16
17
18
19
20
21
22

23 **Palabras clave**

24
25
26 Análisis de supervivencia

27
28 Enfermedad cardiovascular

29
30 Metodología
31
32
33
34
35
36
37

38 **Abbreviations**

39
40
41 CIF: cumulative incidence function

42
43 CPH: Cox proportional hazards

44
45 CVH: cardiovascular-related hospitalization

46
47 DAPT: dual antiplatelet therapy

48
49 HF: heart failure

50
51 MI: myocardial infarction

52
53 PH: proportional hazards

54
55 RCT: randomized clinical trial

56
57 RMST: restricted mean survival time
58
59
60
61
62
63
64
65

Abreviaturas

DAP: doble antiagregación plaquetaria

ECA: ensayo clínico aleatorizado

FIC: función de incidencia acumulada

HC: hospitalización cardiovascular

IAM: infarto agudo de miocardio

IC: insuficiencia cardiaca

RP: riesgos proporcionales

RPC: riesgos proporcionales de Cox

TSMR: tiempo de supervivencia medio restringido

Based on the fundamentals covered in the previous review article,¹ this second article explores more complex challenging situations in survival analyses. We now present some extensions to the Cox proportional hazards (CPH) model, such as the stratified and frailty models, and the use of time-dependent variables. We explore the problems faced by researchers in the cardiovascular field due to multiplicity of outcomes, and present some approaches to tackle this issue, such as the use of composite outcomes, competing risks, multistate models and recurrent-event methods. One increasingly popular topic is the use of the win ratio approach. To give a comprehensive overview of the most common statistical approaches in survival analyses, some other methods are briefly introduced, such as the restrictive mean survival time, and accelerated failure time model 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 knowledge is absolutely necessary.

EXTENSIONS OF THE COX PROPORTIONAL HAZARDS MODEL

Stratified Cox proportional hazards model

The CPH model is by far the most commonly used model in survival analysis. Some extensions to this model can be considered when it does not provide a good fit to our data. In the stratified CPH model, instead of assuming that the proportional hazards (PH) model holds for the overall cohort, we assume that the PH model holds within groups (or strata) of individuals. Study variables that are assumed to satisfy the PH assumption are included in the model, whereas the factor being stratified is not included, and is controlled by stratification. Hence, this method does not provide an estimate of the effect of the factor (or factors) defining the groups on the hazard (i.e., does not provide a hazard ratio of the stratifying variable) and is therefore not a suitable approach if the factor exhibiting non-proportionality is of primary

1 interest. To evaluate the effect of mineralocorticoid receptor antagonists in preventing sudden
2 cardiac death in patients with heart failure (HF) with reduced ejection fraction, a stratified CPH
3 model was needed to address the inevitably baseline differences across 11,032 patients
4 recruited from three placebo-controlled randomized trials (RCTs): RALES (Randomized
5 Aldactone Evaluation Study), EPHESUS (Eplerenone Post-Acute Myocardial Infarction Heart
6 Failure Efficacy and Survival Study), and EMPHASIS-HF (NCT00232180). Although all patients had
7 in common the fact that they had HF with reduced ejection fraction, subjects from EMPHASIS-
8 HF were at New York Heart Association class II, whereas those from RALES were at New York
9 Heart Association class III-IV, and subjects from EPHESUS had a recent myocardial infarction
10 (MI).² In other cases, CPH models are stratified by geographic region and baseline renal function
11 at baseline.³

27 **Frailty models**

30 In a stratified CPH, baseline hazards functions from different strata are unrelated. This
31 is based on the assumption that the study population is homogeneous across strata. However,
32 individuals may differ greatly within strata (e.g., in a RCT, with respect to the treatment effect,
33 whereas in an observational study, with respect to the influence of covariates in a given
34 association). The presence of unobserved individual-specific risk factors leads to unobserved
35 heterogeneity in the hazard, which is also referred to as frailty, or random effect. Importantly,
36 in many situations the population cannot be assumed to be homogeneous (e.g., a mixture of
37 subjects with different hazards). In this case, in contrast to the CPH model, a frailty model is
38 useful as it implies that baseline hazard functions are proportional to each other.

51 Frailty models are random effects models for time-to-event data,⁴ where the random
52 effect has a multiplicative effect on the baseline hazard function. In the context of survival
53 models this random effect is called “frailty” for historical reasons, as the term simply refers to
54 the fact that some individuals are intrinsically more “frail” than others. The classical example

occurs when a study involves the recruitment of patients from different hospitals. Survival times from subjects at the same hospital tend to be similar (e.g., due to treatment practices, level of tertiary activity, etc.) and there is a greater between-hospital variability than within-hospital variability. This clustered (or hierarchical) data needs a model accounting for the clustering. A natural way to model dependence of clustered event times is through the introduction of a cluster-specific random effect —the frailty.⁴ This random effect explains the dependence in the sense that had we known the frailty, the events would be independent. The use of frailty models is relatively popular. They have been applied in some studies using the EPICOR (Long-term Follow-up of Antithrombotic Management Patterns in Acute Coronary Syndrome Patients) registry,⁵ where in addition to adjusting for age, sex and other relevant confounders, the model had a random effect (shared frailty) at the hospital level.⁵

Time-dependent variables

Sometimes explanatory variables change over time in an individual (e.g., treatment, blood pressure, or smoking status). They are known as time-dependent, time-updated, or time-varying variables. If changes over time in these variables are not taken into account, the results yielded by a survival model may provide a bias known as “immortal time” bias, or survivorship bias.⁶

Immortal time refers to a period of follow-up where the event of interest cannot occur because the subject has not yet started the exposure.⁶ A subject is not literally immortal during this period, but remains event-free until classified as exposed. An incorrect consideration of this unexposed time period in the analysis will lead to immortal time bias.⁶ If the unexposed follow-up time is misclassified as exposed, patients in the exposed group are inherently given a survival advantage. Consequently, immortal time bias of this type results in spurious protective effects of the exposure. Classical immortal time bias examples in the literature, have been found in the Texas⁷ and Stanford Heart Transplant data,⁸ both studies concluding that a heart transplant

1 prolongs survival in those patients in a waiting list to be transplanted. However, data was poorly
2 analysed because heart transplant was not treated as a time-varying variable. The waiting time
3 of all patients who were alive until they received a transplantation was classified as exposed to
4 transplantation (instead of unexposed), and gave a survival advantage time to the transplanted
5 group. Deaths that occurred while waiting for transplantation were categorised into the non-
6 transplant cohort. By not being correctly classified, the immortal time produced an increase in
7 the mortality rate of the non-transplantation group, suggesting a benefit of transplantation.⁹
8 However, when adequately analysed, the major survival advantage of the intervention
9 disappeared when the follow up times were properly accounted for.¹⁰

21 When estimating the effect of time-dependent covariates, the follow-up period has to
22 be divided into subintervals (time until exposure, and time from exposure onwards), which
23 means that a subject might have more than one subinterval (e.g., more than one row in the
24 dataset, one line for each subinterval). Subjects enter the study alive, awaiting the exposure, are
25 censored when they get exposed, and start a new subinterval of time (new row in the dataset)
26 where the entry time is the censored time from the first subinterval, with a new covariate value
27 indicating post-exposure. Using the EPICOR cohort, Bueno et al.¹¹ evaluated the impact of dual
28 antiplatelet therapy (DAPT) duration in acute coronary syndrome patients, and the change from
29 DAPT to single antiplatelet therapy (SAPT) in mortality. DAPT was entered into the model as a
30 time-updated categorical variable (0 was being on DAPT, 1 meant a change to SAPT). For patients
31 who were always on DAPT, never receiving SAPT, the value of the time-updated variables was 0
32 throughout their follow-up. For those changing from DAPT to SAPT (usually after 1-year follow-
33 up), they provided follow-up time to both groups (time exposed to DAPT, and time exposed to
34 SAPT).

54 Two approaches can be used to estimate the impact of time-dependent covariates¹²:
55 the CPH model, which can accommodate time-dependent variables, and the landmarking
56 approach. The latter involves setting a landmark time point and using the value of the time-

1 dependent covariate at this landmark point as a time-fixed covariate. By using this approach,
2 subjects with an event before the landmark time point are excluded from the survival analysis,
3 which starts from the landmark time point onwards, in the subset of subjects at risk at that given
4 time.
5
6
7
8
9

10 **ALTERNATIVE MODELS FOR SURVIVAL DATA**

11
12 The CPH model is the most common approach used for the analysis of time-to-event
13 data. However, this regression model may not be appropriate in some situations, such as the
14 hazard ratio not being constant over time.^{13,14} Furthermore, one limitation of the CPH is that the
15 hazard ratio is a relative measure not quantifying absolute effects or associations. Other
16 approaches may overcome some of the limitations of PH analysis. But, when PH are satisfied,
17 the CPH model is the most statistically powerful method.¹³
18
19
20
21
22
23
24
25
26
27
28
29
30

31 **Restricted mean survival time**

32
33 Restricted mean survival time (RMST) is a measure of average survival from time 0 to a
34 specified time point, and may be estimated as the area under the survival curve up to that
35 point.^{13,14} Associations are expressed as the difference in RMST between groups at a suitable
36 follow-up time, which is easy to interpret by both clinicians and patients (e.g., if the outcome of
37 interest is mortality, the estimate would be loss of life expectancy). In addition, the difference
38 between RMST provides an absolute measure (e.g., in a two-arm RCT, RMST provides the
39 absolute benefit or harm). This approach does not require assumptions about hazards and has
40 the advantage of being valid under any distribution of survival time, or when it is expected for
41 an association to vary over time, such as an intervention with either early or late treatment
42 effects.
43
44
45
46
47
48
49
50
51
52
53
54
55
56

57 RMST analysis captures the entire survival history, does not change with extended
58 follow up time, and is routinely associated with a clinically meaningful time point.¹⁵ In heart HF
59
60
61
62
63
64
65

1 RCTs, RMST seems to add value to traditional PH analyses by providing clinically relevant
2 estimates of treatment effects, in line with the findings yielded by other statistical methods.¹⁶
3
4
5

6 **Accelerated failure time**

7
8
9 This approach is known as accelerated failure time model because the term “failure”
10 indicates the death or event; and the term “accelerated” indicates the factor for which the rate
11 of failure is increased. That factor is called “acceleration factor”.¹⁷ Instead of the hazard, the key
12 measure of the association between the study variable and survival time is the acceleration
13 factor, which is a ratio of survival times. Similar to the CPH model, the accelerated failure time
14 model describes the relationship between survival probabilities and a set of covariates,
15 estimating a relative (not an absolute) association. The accelerated failure time model provides
16 an estimate of the ratio of the median event times, which can be translated to clinicians as the
17 expected reduction in the duration of illness with treatment.¹⁷
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32

33 **MULTIPLICITY OF OUTCOMES IN SURVIVAL ANALYSIS**

34
35
36 Clinical studies may evaluate multiple outcomes to try to maximize the information
37 provided by clinical studies. In the cardiovascular field, the outcomes of interest might include
38 stroke, HF, MI, sudden death, cardiovascular death, or all-cause death. To avoid inflation of the
39 type I error rate by testing each outcome separately, one potential solution is to use a composite
40 endpoint by including all the outcomes based on the time-to-first-event principle. Composite
41 outcomes have several advantages,¹⁸ such as accounting for both fatal and nonfatal events, and
42 hence leading to higher event rates and power (thus requiring smaller sample sizes, or shorter
43 follow-ups).¹⁹ Nevertheless, they also present some weaknesses,²⁰ such as the underlying
44 assumption that each individual outcome involved in a composite is of similar importance to
45 patients.²¹ It is also common to have higher event rates and larger treatment effects associated
46 with less important components.²² Hence, the use of composite outcomes is not always optimal.
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

1 There are some situations that require more sophisticated statistical approaches than simply
2 using a composite outcome on a time-to-first event basis, such as: *a)* the use of a competing
3 risk assessment in the evaluation of non-fatal events, where the occurrence of fatal events can
4 bias the findings; *b)* the use of multistate models to take into account intermediate states (e.g.,
5 a HF hospitalization is common before a HF-related death)²³; *c)* the use of recurrent-event
6 methods to fully capture the burden of chronic diseases, which may involve several
7 hospitalizations over the follow-up period; and *d)* the win ratio approach to provide a
8 hierarchical assessment of the individual components of a composite outcome.
9
10
11
12
13
14
15
16
17
18
19
20

21 **COMPETING RISKS**

22 **The censoring assumption**

23
24
25
26
27
28 Uninformative or independent censoring is assumed for the most popular approaches
29 in survival analysis: those who are censored have the same hazard of the event of interest as
30 those who are not censored.^{24,25} In other words, the uncensored individuals who remain under
31 follow-up should be representative of the survival experience in the censored individuals.
32
33 However, if censoring occurs due to another known event taking place, the assumption of
34 uninformative censoring is violated. Competing risks occur when the event of interest is a
35 particular cause of failure (e.g., cardiovascular death), which can take place alongside other
36 causes of failure (e.g., non-cardiovascular death due to cancer). The competing risk may prevent
37 the event of interest from taking place: a subject who dies of cancer is no longer at risk of
38 cardiovascular death (figure 1).
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53

54 **Competing risk bias: impact on the cumulative incidence of events**

55
56
57 A competing risk bias happens when censoring is informative due to multiple causes of
58 failure. This bias has been reported in almost half of Kaplan-Meier analyses published in medical
59
60
61
62
63
64
65

journals.²⁶ If we estimate the survival probability of sudden death in patients with HF with reduced ejection fraction and censor the other causes of death (e.g., HF-related death), the cumulative incidence of events over time (which is 1 minus the survival probability) will overestimate the probability of death due to sudden death.² Indeed, by using the Kaplan-Meier estimator and censoring the other causes of death, we assume that those censored due to a HF-related death have the same future hazard of sudden death, as those who have not yet had any event. Since those who have already died from other reasons can never experience death from sudden death, this can never be true. By assuming that those already dead from other causes are still at risk of sudden death, and that they can be represented by those not yet experiencing any event, the Kaplan-Meier approach overestimates the probability of failure, and therefore underestimates the probability of surviving at a given time. Another classic example, for patients with implanted cardioverter-defibrillators, can be found elsewhere.²⁷

Addressing statistical analysis in competing risks

Two different hazard regression models are available in scenarios where competing risks are present^{28,29}: modelling the cause-specific hazard, or the subdistribution hazard function:

A) The cause-specific hazard function (use of cumulative incidence function [CIF]). The CIF estimates the incidence of an event of interest while allowing for a competing risk. Subjects experiencing the competing event are no longer considered at risk of the event of interest. In the simplest case, when there is only one event of interest and no competing risks, the CIF would equal the 1 – Kaplan-Meier estimate. The CIF takes into account both the probability of experiencing the event of interest, conditioned upon not experiencing either event (primary or competing) until that time. The sum of the CIF estimates for each outcome individually equals the CIF estimate of the composite outcome consisting of all competing events. Unlike the survival function in the absence of competing risks, the CIF function of

1 the event of interest, will not necessarily approach unity with time, because of the
2 occurrence of competing events that preclude the occurrence of the event of interest.²⁸ The
3
4 CIF can be interpreted as the instantaneous rate of the primary event in those subjects who
5
6 are currently event free.
7

8
9 *B) Fine and Gray model (use of subdistribution hazard function).* Fine and Gray modified the
10 CPH model to allow for the presence of competing risks.³⁰ The subdistribution hazard
11 function for a given type of event is defined as the instantaneous rate of occurrence of the
12 given type of event in subjects who have not yet experienced an event of that type. Hence,
13
14 in this model we are considering the rate of the event in those subjects who are either
15
16 currently event-free, or who have previously experienced a competing event (although it
17
18 feels unnatural to keep dead subjects at risk for other events). This differs from the risk set
19
20 for the cause-specific hazard function, which only includes those who are currently event
21
22 free. In this way, there is a subdistribution hazard function for each outcome (e.g., one for
23
24 sudden death, and another for HF-related death).
25
26
27
28
29
30
31
32
33
34

35 The CIF model estimates the impact of covariates on the cause-specific hazard function,
36 while the Fine-Gray subdistribution hazard model estimates the impact of covariates on the
37 subdistribution hazard function.²⁷ Because the CIF model relies on subjects actually at risk
38 (event-free subjects), hazard ratios from this model should be interpreted among individuals
39 who have not yet experienced the event of interest nor the competing event, and therefore, this
40 approach is optimal for answering etiological research questions. In contrast, by keeping at risk
41 those subjects who have experienced the competing risk, the subdistribution hazard model may
42 be of greater interest if the focus is on the overall impact of covariates on the incidence of the
43 event of interest, and is optimal to perform risk prediction and risk-scoring systems.³¹
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

MULTISTATE MODELS

Definition of absorbing and non-absorbing events

An absorbing event prevents the outcome of interest from subsequently taking place (e.g., a cardiovascular death prevents a cancer death). Sometimes, there is an intermediate event, which may occur before the absorbing event, known as non-absorbing event. These intermediate events are of particular interest when their occurrence substantially changes the likelihood of the outcome of interest to happen, and hence, may provide more detailed information on the natural history of the disease. The intermediate event can be interpreted as a deterioration or improvement step in the disease process. Solomon et al. illustrated this concept by assessing the influence of non-fatal hospitalizations for HF on subsequent mortality in patients with chronic HF.²³ In contrast to the relatively stable mortality risk observed over time in patients with HF from the CHARM (Candesartan in Heart failure: Assessment of Reduction in Mortality and morbidity) program, they found a higher likelihood to die in the immediate post discharge period of a HF hospitalization, which was directly associated with the duration and frequency of HF hospitalizations.²³ Having a HF hospitalization (non-absorbing event) changed the hazard of the outcome of interest (mortality).

Non-absorbing events can be modelled using multi-state models,³² where the focus is on the change of status over time (e.g., change from baseline status to HF hospitalization, and from there to cardiovascular death).³³

Multistate models: an extension of competing risks models

Multistate models provide a framework that allows for the analysis of the natural history of a disease. They are an extension of competing risks models (multistate model with one initial state and several mutually exclusive absorbing states), since they extend the analysis to what happens after the intermediate event.³⁴ In this review, only continuous-time models allowing

changes of state at any time will be considered. They are more realistic and can be seen as an extension of the standard survival model, as it describes how an individual moves between a series of discrete states in continuous time.

Multistate models are appropriate when a disease involves transitions between several well-defined distinct states. A two-state survival model is defined by a living state and a dead state. The two main features of the standard survival model are: *a)* there is *one event* of interest (the transition from alive to dead), which is unidirectional; and *b)* the timing of this event may be right-censored, in which case it is known that the event has not happened yet. A Kaplan-Meier curve can be thought of as a simple multistate model with two states, and one transition between those two states. The situation becomes more complex when non-absorbing events are included into the model. In the HF setting, HF hospitalization can be defined as a transient event (non-absorbing event). A three-state survival model would be defined by a HF-free state, a HF state, and a dead state. This sets three events: death from state 1 (HF-free state), death from state 2 (HF state), and transition from state HF-free to HF hospitalization (figure 2). The hazard rates defining movement from one state to another are defined as transition intensities, the instantaneous risk of moving from one state to another at a given time. These transition intensities are equivalent to the cause-specific hazards described for the competing risks approach, being this situation a particular case of multistate models. There are as many hazards to model as there are transitions.

To run a multistate model, a counting process data structure is used to frame the data. Hence, every time there is a transition, another row of information for that individual is needed in the dataset. On the other hand, in a traditional time-to-first event survival analysis, there is only one row of information per patient in the dataset, including the status and the survival time (time to event or time to censoring).

Recurrence of nonfatal HF or MI are non-absorbing events depending on time. Although they can be taken into account in standard survival models by including the non-absorbing event

as a binary time-dependent covariate for the risk of death, the best approach to tackle these patients' transitions in various states is by multistate modelling.³⁵ To improve the understanding of prognosis, a comprehensive model should include both death and non-fatal clinical events. CPH models are not strictly appropriate since observations are not independent. Multistate models overcome this limitation by assessing separately time-to-death and time-to-disease-related hospitalizations.³⁶

Multistate models in cardiology

Several studies have used multistate models in the cardiovascular field.³⁷ Beyond the classical example of chronic HF,³⁶ we can find other examples in patients with coronary multivessel disease. Using data from the BARI trial (NCT00000462), Zhang et al.³⁸ performed a multistate model, where the initial state were patients after randomization and before intervention, the intermediate state was non-fatal MI, and the final state was death. Standard survival analyses with Cox regression and Kaplan-Meier estimation for both, mortality and the composite outcome of death or non-fatal MI, showed no differences between coronary artery bypass grafting or percutaneous coronary angioplasty after a 10-year follow-up. Of note, this approach did not take into account the intermediate state (non-fatal MI). In contrast, multistate modelling broke the process into three transitions, and found significant differences in outcomes favouring coronary artery bypass grafting for patients in a transition path of non-fatal MI to death, whereas for patients without MI, there was no difference in terms of survival between patients who underwent coronary artery bypass grafting and those who underwent percutaneous coronary angioplasty. This study illustrates that the use of composite outcomes may not capture as much prognostic information as the use of a multistate model can do.

RECURRENT EVENT METHODS

Definition of recurrent event

Most studies evaluate time-to-first event endpoints, so that all subsequent equal events occurring after a first one are ignored in the analysis. In HF, this problem becomes even more important as “time-to-first” event analyses do not fully reflect the true burden of the disease³⁹: In patients admitted to Spanish Emergency Departments due to acute HF, 24% revisited the Emergency Department within 30 days, and 16% rehospitalized in the same follow-up period.⁴⁰ Of note, each subsequent HF hospitalization heralds a substantial worsening of the long-term prognosis.⁴¹ In contrast to conventional methods, the use of recurrent-event methods may capture the burden of disease (figure 3).

Recurrent event methods have been generally assumed to improve statistical precision and provide greater statistical power than more conventional time-to-first methods. For instance, in the CHARM-Preserved trial,⁴² a borderline result for time-to-first composite event analysis was achieved. However, *post-hoc* analyses with a recurrent event method, lead to a gain in statistical power and showed statistically evidence of efficacy.¹⁹

Statistical analysis in recurrent event methods

There are several statistical methods addressing the issue of recurrent events, but there is some controversy as to which of them is the most appropriate. There are two main approaches: through counts (or event rates), and through times between subsequent events. Non-informative censoring is assumed in both cases.

For the first approach based on measuring the number of events (e.g., number of hospitalizations due to worsening HF), there are two main methods.⁴³ The Poisson distribution is the most popular count model and can be used to determine if event rates differ between groups, whereas the negative binomial distribution is an alternative approach that allows for

different individual tendencies (frailties). The latter has been retrospectively used to evaluate recurrent hospitalizations in the EMPHASIS-HF trial.⁴⁴ In the TOPCAT (NCT00094302) trial, the pre-specified Poisson regression model was eventually replaced with a negative binomial model to allow for correlated events.⁴⁵

There are many methods for the second approach based on the time-to-event principle. The Andersen–Gill approach is an extension of CPH model, which analyses recurrent events as gap times (e.g., the times between consecutive events).⁴⁶ The Lin-Wei-Yang-Ying method is a modified Anderson-Gill model, with a robust variance’s estimator to account for the correlation between events, useful when covariates are considered time dependent. This approach was used for the primary outcome of total (first and recurrent) HF hospitalizations and cardiovascular deaths in the PARAGON-HF trial.⁴⁷ The Ghosh and Lin method offers a non-parametric estimate of the cumulative number of recurrent events through time, which incorporates death as a competing risk.⁴⁸

Final remarks about recurrent event methods

Recurrent event models seem to improve statistical precision and to provide greater statistical power than time-to-first event approaches.⁴⁹ However, the relative width of the 95% confidence interval associated with the recurrent-event analysis can sometimes be greater than that from the time-to-first analysis, suggesting a loss of precision.¹⁹ Using trial-based data (CHARM, TOPCAT, PARADIGM-HF [NCT01035255]), Clagget et al. found that the increasing heterogeneity of patient risk, a parameter not included in conventional power and sample size formulas, might explain the differences between time-to-first and recurrent-event analyses in terms of treatment effect estimation, precision, and statistical power.⁴⁹ In this study, they conclude that the greatest statistical gains from using recurrent-events methods occur in the presence of high patient heterogeneity and low rates of study drug discontinuation.⁴⁹

THE WIN RATIO METHOD

The win ratio was introduced in 2012 by Pocock et al.⁵⁰ as a new method for examining composite endpoints, and it is becoming progressively popular in cardiovascular RCTs.^{51,52} Unlike traditional methods evaluating composite endpoints, the win ratio accounts for relative priorities of their components, and even allows different types of components. For example, the win ratio can combine the time to death with the number of occurrences of a non-fatal outcome such as cardiovascular-related hospitalizations (CVHs) in a single hierarchical composite endpoint. It can also include quantitative outcomes such as quality-of-life scores.

Based on the principle of the Finkelstein–Schoenfeld test, the win ratio approach provides an estimate of the treatment effect (the win ratio) and confidence interval, in addition to a *P*-value.⁵³ In a simple two-arm RCT, the application of the win ratio can be summarised as: *a*) forming every possible patient-to-patient pair (every patient on the treatment arm is compared with every patient on the control arm); and *b*) within each pair, evaluating the component outcomes in descending order of importance until one of the pair shows a better outcome compared with the other. If the patient on the treatment has the better outcome it is called a “win”, if the control patient does better it is a “loss”, and, if none of these situations happen, then it is a “tie”.

This approach was used in the ATTR-ACT trial,⁵² which was a double-blind trial that randomized 441 patients with transthyretin amyloid cardiomyopathy to tafamidis (80 and 20 mg), or matching placebo for 30 months. In the primary analysis, the investigators hierarchically assessed all-cause mortality, followed by frequency of CVHs using the Finkelstein–Schoenfeld⁵³ and win ratio⁵⁰ methods. For each pair, they determined whether the patient receiving tafamidis “won” or “lost” compared with the patient receiving placebo. Their hierarchical assessment was to determine: *a*) who died first (the “loser”); and then, *b*) if neither died, who had the most CVHs (again the “loser”), both being assessed over their shared follow-up time. After adding up, they obtained a total of 8595 winners and 5071 losers. Hence, the win ratio was $8595/5071 = 1.70$,

with a 95% confidence interval of 1.26-2.29 and $P = .0006$.⁵⁴ By using traditional methods to evaluate a composite outcome of first CVH or death, we would have ignored repeated CVHs after the first CVH, as well as, any death happening after a CVH. The win ratio provides greater statistical power to estimate treatment differences by evaluating hierarchically each component of a composite outcome.

OTHER APPROACHES: LOGISTIC AND POISSON REGRESSION MODELLING

In addition to CPH models, survival data is often evaluated using logistic and Poisson regression models.^{55,56} The choice between those models is based on the design of the study and the nature of the research question.⁵⁷ Table 1 summarizes the main differences between these models in the setting of survival analysis.

CONCLUSIONS

In this second educational review, we put the focus on stratified CPH models, frailty models, and time-dependent variables. Competing risks, multistate models, recurrent event methods and the win ratio approach have been presented to tackle the issue of multiplicity of outcomes when the use of a composite outcome and a time-to-first event might not be optimal. The use of restrictive mean survival time and accelerated time model approaches have been also illustrated. Modelling adequately survival data is not a straightforward exercise. This review has offered practical advice on what should be considered before choosing the most appropriate model for survival data, as well as how to interpret the findings yielded by more complex statistical approaches.

FUNDING

None declared.

AUTHORS' CONTRIBUTIONS

X. Rossello and M. González-Del-Hoyo conceived the review. X. Rossello led the writing process, and M. González-Del-Hoyo has been in charge to find most of the examples to illustrate the methods. X. Rossello drafted the article, though both contributed substantially to its revision.

CONFLICT OF INTERESTS

None declared.

REFERENCES

1. Rossello X, González-Del-Hoyo M. Survival analyses in cardiovascular research, part I: the essentials. *Rev Esp Cardiol*. 2021. <https://doi.org/10.1016/j.rec.2021.06.003>.
2. 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.
3. Rossello X, Ferreira JP, Caimari F, et al. Influence of sex, age and race on coronary and heart failure events in patients with diabetes and post-acute coronary syndrome. *Clin Res Cardiol*. 2021. <https://doi.org/10.1007/s00392-021-01859-2>.
4. Aalen OO. Effects of frailty in survival analysis. *Stat Methods Med Res*. 1994;3:227-243.
5. Bueno H, Pocock S, Medina J, et al. Association between clinical pathways leading to medical management and prognosis in patients with NSTEMI. *Rev Esp Cardiol*. 2017;70:817-824.
6. Suissa S. Immortal time bias in pharmacoepidemiology. *Am J Epidemiol*. 2008;167:492-499.
7. Messmer B, Leachman R, Nora J, Cooley D. Survival-times after cardiac allografts. *Lancet*.

1969;293:954-956.

8. Clark DA, Stinson EB, Griep RB, et al. Cardiac transplantation in man. VI. Prognosis of patients selected for cardiac transplantation. *Ann Intern Med.* 1971;75:15-21.

9. Gail MH. Does cardiac transplantation prolong life? A reassessment. *Ann Intern Med.* 1972;76:815-817.

10. Mantel N, Byar DP. Evaluation of Response-Time Data Involving Transient States: An Illustration Using Heart-Transplant Data. *J Am Stat Assoc.* 1974;69:81-86.

11. Bueno H, Pocock S, Danchin N, et al. International patterns of dual antiplatelet therapy duration after acute coronary syndromes. *Heart.* 2017;103:132-138.

12. Putter H, van Houwelingen HC. Understanding Landmarking and Its Relation with Time-Dependent Cox Regression. *Stat Biosci.* 2017;9:489-503.

13. Gregson J, Sharples L, Stone GW, et al. Nonproportional Hazards for Time-to-Event Outcomes in Clinical Trials: JACC Review Topic of the Week. *J Am Coll Cardiol.* 2019;74:2102-2112.

14. McCaw ZR, Yin G, Wei L-J. Using the Restricted Mean Survival Time Difference as an Alternative to the Hazard Ratio for Analyzing Clinical Cardiovascular Studies. *Circulation.* 2019;140:1366-1368.

15. Royston P, Parmar MKB. Restricted mean survival time: an alternative to the hazard ratio for the design and analysis of randomized trials with a time-to-event outcome. *BMC Med Res Methodol.* 2013;13:152.

16. Perego C, Sbolli M, Specchia C, et al. Utility of Restricted Mean Survival Time Analysis for Heart Failure Clinical Trial Evaluation and Interpretation. *JACC Heart Fail.* 2020;8:973-983.

17. Patel K, Kay R, Rowell L. Comparing proportional hazards and accelerated failure time models: an application in influenza. *Pharm Stat.* 2006;5:213-224.

18. Heddle NM, Cook RJ. Composite outcomes in clinical trials: What are they and when should they be used? *Transfusion*. 2011;51:11-13.
19. Rogers JK, Pocock SJ, McMurray JJV, et al. Analysing recurrent hospitalizations in heart failure: A review of statistical methodology, with application to CHARM-preserved. *Eur J Heart Fail*. 2014;16:33-40.
20. Ferreira-González I, Busse JW, Heels-Ansdell D, et al. Problems with use of composite end points in cardiovascular trials: systematic review of randomised controlled trials. *BMJ*. 2007;334:786.
21. Stolker JM, Spertus JA, Cohen DJ, et al. Rethinking composite end points in clinical trials insights from patients and trialists. *Circulation*. 2014;130:1254-1261.
22. Ferreira-González I, Permanyer-Miralda G, Busse JW, et al. Methodologic discussions for using and interpreting composite endpoints are limited, but still identify major concerns. *J Clin Epidemiol*. 2007;60:651-657.
23. Solomon SD, Dobson J, Pocock S, et al. Influence of Nonfatal Hospitalization for Heart Failure on Subsequent Mortality in Patients With Chronic Heart Failure. *Circulation*. 2007;116:1482-1487.
24. Bland JM, Altman DG. Survival probabilities (the Kaplan-Meier method). *BMJ*. 1998;317:1572.
25. Rao SR, Schoenfeld DA. Survival Methods. *Circulation*. 2007;115:109-113.
26. Van Walraven C, McAlister FA. Competing risk bias was common in Kaplan-Meier risk estimates published in prominent medical journals. *J Clin Epidemiol*. 2016;69:170-173.e8.
27. Wolbers M, Koller MT, Stel VS, et al. Competing risks analyses: objectives and approaches. *Eur Heart J*. 2014;2936-2941.
28. Pintilie M. An introduction to competing risks analysis. *Rev Esp Cardiol*. 2011;64:599-605.

29. Latouche A, Allignol A, Beyersmann J, et al. A competing risks analysis should report results on all cause-specific hazards and cumulative incidence functions. *J Clin Epidemiol*. 2013;66:648-653.
30. Fine JP, Gray RJ. A Proportional Hazards Model for the Subdistribution of a Competing Risk. *J Am Stat Assoc*. 1999;94:496-509.
31. Austin PC, Fine JP. Practical recommendations for reporting Fine-Gray model analyses for competing risk data. *Stat Med*. 2017;36:4391-4400.
32. Andersen PK, Abildstrom SZ, Rosthøj S. Competing risks as a multi-state model. *Stat Methods Med Res*. 2002;11:203-215.
33. Meira-Machado LF, de Uña-Álvarez J, Cadarso-Suárez C, Andersen PK. Multi-state models for the analysis of time-to-event data. *Stat Methods Med Res*. 2009;18:195-222.
34. Putter H, Fiocco M, Gekus RB. Tutorial in biostatistics: Competing risk and multi-state models. *Stat Med*. 2007;26:2389-2430.
35. Upshaw JN, Konstam MA, Van Klaveren D, Noubary F, Huggins GS, Kent DM. Multistate model to predict heart failure hospitalizations and all-cause mortality in outpatients with heart failure with reduced ejection fraction. *Circ Hear Fail*. 2016;9:e003146.
36. Ieva F, Jackson CH, Sharples LD. Multi-state modelling of repeated hospitalisation and death in patients with heart failure: The use of large administrative databases in clinical epidemiology. *Stat Methods Med Res*. 2017;26:1350-1372.
37. Gasperoni F, Ieva F, Barbati G, et al. Multi-state modelling of heart failure care path: A population-based investigation from Italy. *PLoS One*. 2017;12:e0179176.
38. Zhang X, Li Q, Rogatko A, et al. Analysis of the bypass angioplasty revascularization investigation trial using a multistate model of clinical outcomes. *Am J Cardiol*. 2015;115:1073-1079.

39. Anker SD, McMurray JJV. Time to move on from “time-to-first”: Should all events be included in the analysis of clinical trials? *Eur Heart J*. 2012;33:2764-2765.
40. Rossello X, Bueno H, Gil V, et al. MEESSE-AHF risk score performance to predict multiple post-index event and post-discharge short-term outcomes. *Eur Hear Journal Acute Cardiovasc Care*. 2021;10:142-152.
41. Solomon SD, Anavekar N, Skali H, et al. Influence of ejection fraction on cardiovascular outcomes in a broad spectrum of heart failure patients. *Circulation*. 2005;112:3738-3744.
42. Yusuf S, Pfeffer MA, Swedberg K, et al. Effects of candesartan in patients with chronic heart failure and preserved left-ventricular ejection fraction: The CHARM-preserved trial. *Lancet*. 2003;362:777-781.
43. Glynn RJ, Buring JE. Ways of measuring rates of recurrent events. *Br Med J*. 1996;312:364-367.
44. Rogers JK, McMurray JJV, Pocock SJ, et al. Eplerenone in patients with systolic heart failure and mild symptoms: Analysis of repeat hospitalizations. *Circulation*. 2012;126:2317-2323.
45. Pitt B, Pfeffer MA, Assmann SF, et al. Spironolactone for Heart Failure with Preserved Ejection Fraction. *N Engl J Med*. 2014;370:1383-1392.
46. Andersen PK, Gill RD. Cox’s Regression Model for Counting Processes: A Large Sample Study. *Ann Stat*. 1982;10:1100-1120.
47. Solomon SD, McMurray JJV, Anand IS, et al. Angiotensin–Neprilysin Inhibition in Heart Failure with Preserved Ejection Fraction. *N Engl J Med*. 2019;381:1609-1620.
48. Ghosh D, Lin DY. Nonparametric analysis of recurrent events and death. *Biometrics*. 2000;56:554-562.
49. Claggett B, Pocock S, Wei LJ, et al. Comparison of time-to-first event and recurrent-event methods in randomized clinical trials. *Circulation*. 2018;138:570-577.

- 1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
50. Pocock SJ, Ariti CA, Collier TJ, Wang D. The win ratio: a new approach to the analysis of composite endpoints in clinical trials based on clinical priorities. *Eur Heart J*. 2012;33:176-182.
51. Redfors B, Gregson J, Crowley A, et al. The win ratio approach for composite endpoints: practical guidance based on previous experience. *Eur Heart J*. 2020;41:4391-4399.
52. Maurer MS, Schwartz JH, Gundapaneni B, et al. Tafamidis Treatment for Patients with Transthyretin Amyloid Cardiomyopathy. *N Engl J Med*. 2018;379:1007-1016.
53. Finkelstein DM, Schoenfeld DA. Combining mortality and longitudinal measures in clinical trials. *Stat Med*. 1999;18:1341-1354.
54. Pocock SJ, Collier TJ. Statistical Appraisal of 6 Recent Clinical Trials in Cardiology: JACC State-of-the-Art Review. *J Am Coll Cardiol*. 2019;73:2740-2755.
55. 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.
56. Miró Ò, Rosselló X, Gil V, et al. The Usefulness of the MEESI Score for Risk Stratification of Patients With Acute Heart Failure at the Emergency Department. *Rev Esp Cardiol (Engl Ed)*. 2019;72:198-207.
57. Cox DR. Regression Models and Life-Tables. *J R Stat Soc Ser B*. 1972;34:187-220.

FIGURE LEGENDS

Figure 1. A graphical representation of the competing risks model. A: a classic competing risk challenge with two fatal outcomes. B: a more challenging and realistic situation, where several risks are competing in patients with a severe disease.

Figure 2. A graphical representation of the illness-death model in a heart failure example. λ , transition intensity function (e.g., λ_{12} is the transition intensity function from state 1 to state 2).

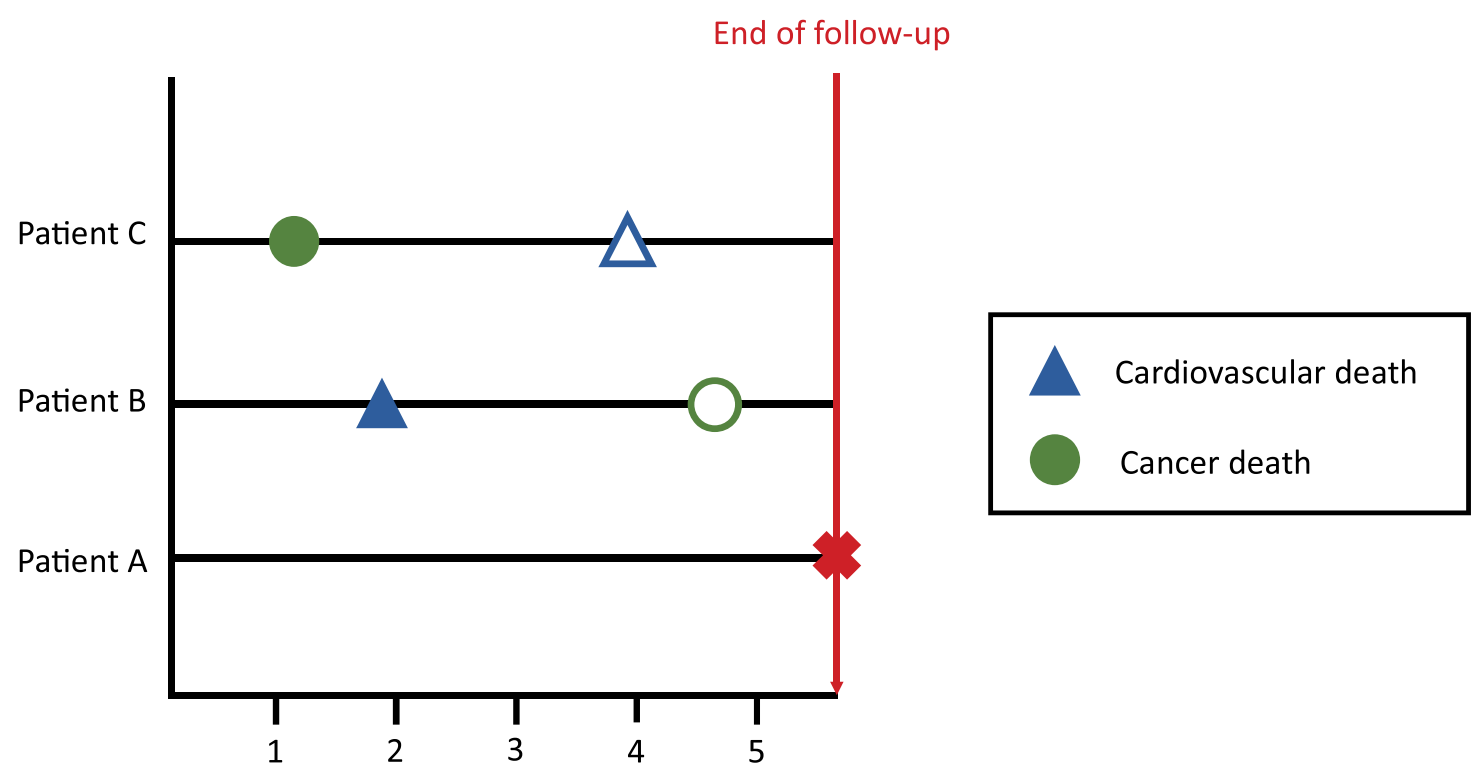
Figure 3. A graphical representation of the recurrent risks model. Several situations in the context of recurrent events are illustrated: patient A is hospitalized 4 times over follow-up (only the first event would be used in conventional survival analyses), patient B is hospitalized twice before dying (only the first hospitalization would be taken into account using traditional methods), and patient C dies during the follow-up without any previous hospitalization (using a composite endpoint, information about the disease burden would be lost in this situation).

Table 1. Summary of differences between the Cox proportional hazards, Poisson and logistic regression models

	Cox proportional hazards model	“Interval” Poisson regression model	Logistic regression model
Research question	How long until the event occurs, in a defined time endpoint?	How many times does the event occur in a defined time endpoint?	Does a subject reach the event in a defined timeframe?
Modelling	It models survival times	It models the rate at which the event occurs independently over time	It models whether an event occurs or not
Use of survival time	Analysis of individual survival times	Analysis of aggregated patient mortality rates	Analysis only of events, without taking into account when they happen. It does not use survival times
Outcome type	Time-to-event	Event-count data	Dichotomous event data
Association type	Hazard ratio	Incidence rate ratio	Odds ratio
Main assumption	Hazard function or death rate are proportional between groups	Rate of events or relative risk ratio remains constant over specific time intervals or are proportional to one another	It does not require the dependent and independent variables to be related linearly, but that the independent variables are linearly related to the log odds

Figure 1

Graphical representation of the competing risks model with two fatal outcomes



Reality checkpoint: Graphical representation of the competing risks model in patients with heart failure

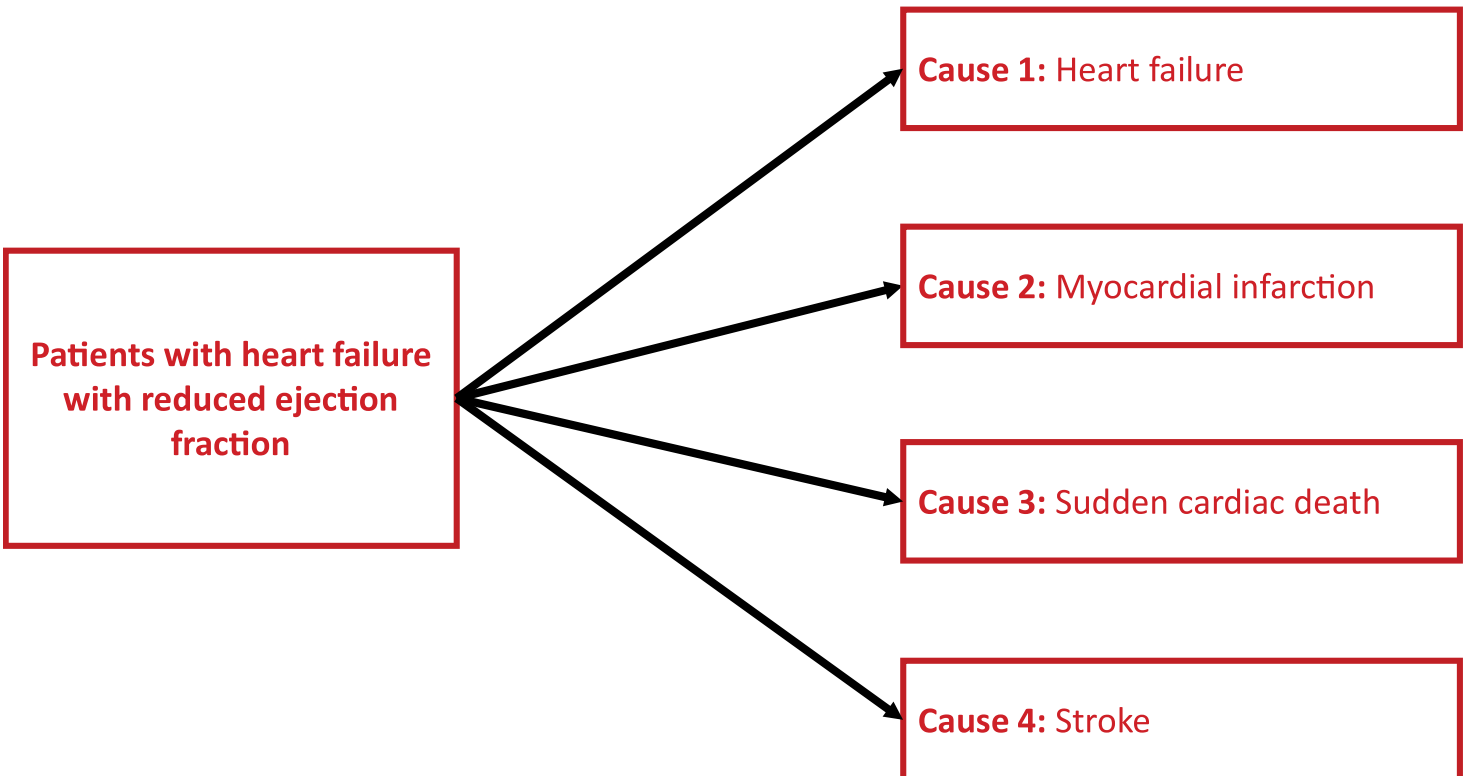


Figure 2

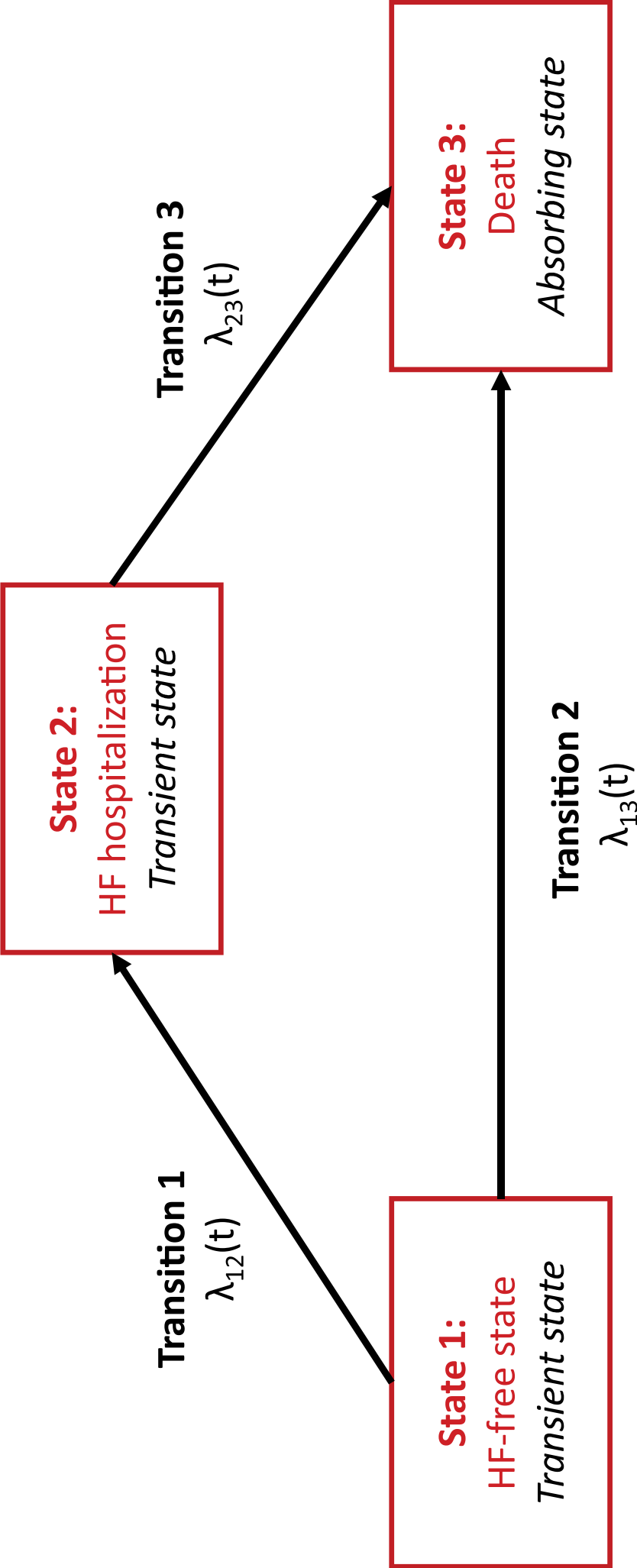


Figure 3

Graphical representation of the recurrent event model in patients with heart failure

