The clinical course of cirrhosis: the importance of multi-state models and competing risks analysis

Peter Jepsen[1, 2], pj@dce.au.dk

Hendrik Vilstrup[1], hendvils@rm.dk

Per Kragh Andersen[3], pkan@sund.ku.dk

1) Department of Hepatology and Gastroenterology, Aarhus University Hospital, Aarhus, Denmark
2) Department of Clinical Epidemiology, Aarhus University Hospital, Aarhus, Denmark
3) Department of Biostatistics, Institute of Public Health, University of Copenhagen, Copenhagen, Denmark.

Key words: Methodology; clinical epidemiology; disease progression; hepatocellular carcinoma; chronic liver disease.

Footnote page

Contact information: Peter Jepsen, Department of Hepatology and Gastroenterology, Aarhus University Hospital, Nørrebrogade 44, DK-8000 Aarhus, Denmark. Phone: +45 7846 3892. Fax: +45 7846 2860. E-mail: pj@dce.au.dk

List of abbreviations: Hepatocellular carcinoma : HCC

Financial support: None.

Abstract

Multi-state models are models of disease progression that for a patient group define multiple outcome events, each of which may affect the time to develop another outcome event. Multi-state models are highly relevant for studies of cirrhosis patients; both the classical perception of cirrhosis as either compensated or decompensated and the recent more complex models of cirrhosis progression are multi-state models. Therefore, researchers who conduct clinical studies of cirrhosis patients must realize that most of their research questions assume a multi-state disease model. Failure to do so can result in severely biased results and bad clinical decisions. The analyses that can be used to study disease progression in a multi-state disease model may be called *competing risks analysis*, named after the competing risks disease model which is the simplest multi-state disease model. In this review article we introduce multi-state disease models and competing risks analysis and explain why the standard armamentarium of Kaplan-Meier survival estimates and Cox regression sometimes gives bad answers to good questions. We also use real data to answer typical research questions about the course of cirrhosis and illustrate biases resulting from inadequate methods. Finally, we suggest statistical software packages that are helpful and accessible to the clinician-researcher.

Multi-state disease models are a class of disease course models. A multi-state disease model for a specific patient group defines multiple disease states, and the transitions between disease states are defined by the occurrence of outcome events, each of which may affect the time to develop another outcome event (1). Multi-state models are, therefore, highly relevant for studies of cirrhosis patients whose clinical course is fraught with complications—e.g., infections, ascites, variceal bleeding, hepatic encephalopathy, acute-on-chronic liver failure, hepatocellular carcinoma (HCC), and death (2, 3). It is increasingly clear that hepatologists must be familiar with these disease models and understand how to perform analyses within them if they want to attain a relevant understanding of their patients' disease, prognosis, and treatment priorities. We introduce multi-state disease models and competing risks analysis in this article.

Disease models

Central to any follow-up study of cirrhosis patients is its model of cirrhosis progression. The model can be shown as a set of disease states connected by arrows indicating the possible transitions between the states (Table 1). The chosen disease model should reflect the research question, and while there is no need to use an unnecessarily complicated model, it must be clinically realistic. It must also have clearly defined entry and exit times for all its disease states. The simplest possible model has only two disease states, 'alive' and 'dead' (Figure 1a), and it is suitable exclusively for analyses of all-cause mortality and for analyses of composite outcomes that include all-cause mortality, e.g. 'death or disease progression' as one single outcome. Such composite outcomes are used in studies of cancer progression and sometimes also in studies of cirrhosis progression. Often, however, the research interest lies in cause-specific

mortality, typically in the risk of death from cirrhosis. For such a study the researchers must consider that patients can also die from other causes, so it is necessary to use a disease model with three states, like the one in Figure 1b, i.e., a multi-state model. Another classic example of a multi-state model used in hepatology research is the model comprising three progressively advanced disease states: compensated cirrhosis, decompensated cirrhosis, and death (Figure 1c) (4). More detailed models of the clinical course of cirrhosis have also been proposed (2, 5, 6).

Having decided on the disease model, researchers must analyze their data accordingly. Analyses of the two-state disease model in Figure 1a are standard survival analysis, e.g. using the Kaplan-Meier technique, as detailed below. Analyses of the multi-state disease models in Figures 1b and 1c, on the other hand, must be *competing risks analysis* (7, 8). Competing risks analysis is a name shared by all analyses suitable for multi-state disease models. It is coined from the simplest possible multi-state disease model, shown in Figure 1b, which is often called the 'competing risks disease model'.

In the following sections we highlight the problems that we often encounter in analyses of data from multi-state models: handling censored observations, computing the cumulative risk of making a transition from one state to another, computing the probability of being in a specific state at a specific time, and examining prognostic factors. But first we will highlight a key issue: the distinction between *risks* and *rates*. Both describe the occurrence of an outcome event, but 'risk' describes the cohort members' probability of experiencing the outcome event, and 'rate' describes the speed with which the outcome event occurs within the cohort (Table 1). It is very important not to confuse the two, as we will explain.

Censoring

Censoring occurs when no outcome event is observed during a patient's follow-up. It is virtually always *right*-censoring, which means that follow-up ended *before* an outcome event occurred. In practice we never see *left*-censoring, which means that follow-up began *after* an outcome event occurred, and only rarely do we see *interval*-censoring, which means that an outcome event occurred between two observation periods. From here on we use 'censoring' as shorthand for 'right-censoring'.

The typical censoring event is an administrative one, such as an end-of-study date when all patients who remain in the cohort are censored. In order to get clinically meaningful and unbiased results from an analysis of data with censored observations, a censoring event must fulfill two criteria: First, if we are going to compute the *risk* of experiencing the outcome of interest, the non-existence of the censoring event must make sense from a clinician's perspective because the risk estimates will apply to a world in which the censoring event does not exist (1, 9). Second, the censoring event cannot be a marker of a good or bad prognosis. This is because the study sample (for example, a cohort of cirrhosis patients) is used to make inferences about cirrhosis patients in general, so it must be representative of cirrhosis patients in general *throughout* the follow-up period. Therefore the results from the analysis will be wrong if we censor those cirrhosis patients who have a particularly good or bad prognosis. The statistical term for this criterion is that censoring must be independent (1).

A common approach is to censor patients who make a different transition than the one of interest. This approach, however, is wrong—with one exception, which we describe in the next paragraph. First an example of wrongful censoring: We use a three-state

disease model for cirrhosis patients: compensated, decompensated, and dead (Figure 1c). Our aim is to describe the compensated patients' *risk* of decompensating. Should patients who die before they decompensate be censored at death? No! Per the criteria above, a censoring event's non-existence has to be clinically sensible, and the non-existence of death as an outcome of compensated cirrhosis does not make sense. How about censoring patients if they receive a liver transplant before decompensating? This can be tempting because transplantation has profound effects on the clinical course of compensated cirrhosis, and it might be perceived as unwanted noise in our study. But again, no! Censoring at transplantation is also wrong. First, the non-existence of liver transplantation is nonsense in a contemporary clinical context in most areas of the world. Second, our disease model dictates that cirrhosis patients must be either compensated, decompensated, or dead, and our cohort of uncensored patients will no longer be representative of *all* patients with compensated cirrhosis if it excludes those who are transplanted. Taken together, censoring at liver transplantation would violate both the criterion about non-existence being clinically sensible and the criterion about independent censoring. The best solution is to make 'transplanted' a distinct disease state and accept the more complex disease model (Figure 1d shows one possibility). The alternative solution is to simply ignore transplantations and just follow the patients until they decompensate, die, or experience a censoring event. This may well be a reasonable solution if there are so few transplantations that they have no discernable effect on the course of the total cohort with compensated cirrhosis. If more patients are transplanted, this solution will become clinically untenable because it will lump together transplanted and non-transplanted patients with widely different clinical courses. Clearly, there is no such thing as a universal model of cirrhosis progression that everyone has to use for

every purpose, but researchers should be careful to declare their disease model for every study.

The exception to the principle that it is wrong to censor patients who make a different transition from the one of interest is this: If the interest is in the *rate* (not the *risk*) of decompensation or another non-fatal outcome, censoring at death will not affect the results. This is so because an assumed constant rate of decompensation is estimated as the number of decompensations divided by the total observation time, and both measures are unchanged by censoring at death because follow-up would have stopped at death anyway. Moreover, with censoring at death, our cohort of uncensored patients—that is, compensated patients who are not dead—does meet the criterion of being representative of *all* compensated patients.

Cumulative risk

When clinicians inform patients about the outcome of a disease, they virtually always speak in terms of probabilities; the chance (in percent) of a positive outcome and the risk (in percent) of a negative outcome. Probabilities are far more intuitive than rates, or odds for that matter, and the probability of experiencing a specific outcome within a specific timeframe is called a 'cumulative incidence' or a 'cumulative risk', and often just 'risk'.

Cumulative risk can always be computed with the *cumulative incidence function* (10). As long as we keep to a two-state disease model, we can also compute cumulative risk as 1 minus the Kaplan-Meier estimate of survival probability, and this is convenient because the Kaplan-Meier method is simpler and more familiar to clinician-researchers. Unfortunately, many researchers also use the Kaplan-Meier method with a multi-state

disease model, e.g. to estimate the cumulative risk of decompensation. That is wrong! It gives us the cumulative risk of decompensation in a world where deaths do not occur, which is meaningless. Sometimes we see the cumulative risk of decompensation expressed as the 'cumulative chance of remaining free from decompensation', but the two are each other's complement and have exactly the same problems.

Example 1

This example, shown in Table 2, demonstrates how a two-state disease model (Figure 1a) can be analyzed correctly by both the cumulative incidence function and the 1 minus Kaplan-Meier method. We follow ten patients with compensated cirrhosis for up to ten years. The total follow-up time is 55 years and four patients die. The rate of death is 4 / 55 = 7.3 per 100 person-years, but we are interested in the *risk* of death, not the *rate*. The two methods arrive at the same estimate of cumulative risk (0.64, or 64%, at ten years) although the formulae are different.

Example 2

Like the previous example, the one in Table 3 follows ten patients with compensated cirrhosis for up to ten years, but we are now interested in decompensation and use a multi-state disease model (Figure 1c). The 10-year risk of decompensation is 37%, correctly computed with the cumulative incidence function. By using the 1 minus the Kaplan-Meier estimator we get *higher* (and wrong) risk estimates. The bias is rooted in the fact that the Kaplan-Meier method was never intended for analyses involving competing risks. It treats all non-decompensation events as censorings and consequently assumes that everyone who is censored would still be at risk of decompensation if only he or she could have been followed longer. That, however, is only true for patients who

experience a *true* censoring event, e.g. an end-of-study date. It is not true for patients who experience a competing outcome event; such patients are in fact no longer at risk of decompensation, whether the competing outcome is death or a non-fatal event like transplantation (cf. the disease model in Figure 1d). Despite their zero risk of decompensation, the Kaplan-Meier method assigns them the same risk of decompensation as those who continue to be under observation. The inevitable result is an inflated estimate of cumulative risk unless no one experiences the competing outcome event, but in studies of cirrhosis patients there are always many deaths. As shown in Table 3, the 1 minus Kaplan-Meier method gives us a 10-year risk of decompensation of 63%, a gross overestimate. Gooley et al. provide a more complete and reasonably accessible explanation of the bias from the Kaplan-Meier method (11).

Example 3

This example, shown in Figure 2, uses data on HCC risk in Danish patients with alcoholic cirrhosis (12). The computations are the same as in Example 2. Correctly calculated, the five-year HCC risk is 2.6%, but with the 1 minus Kaplan-Meier method we get a staggering 5.6%. Such a bias in real-life data serves as a warning against indiscriminate use of the Kaplan-Meier method for cumulative risks, as it may lead to wrong clinical decisions, in our example about HCC surveillance (13). Another example in this line is a recently published follow-up study of an Italian cohort of 494 cirrhosis patients, of whom 82% had chronic HCV infection (14). The authors correctly used the cumulative incidence function to show that the 20-year HCC risk was below 1% per year. This is markedly lower than 3–5% per year which is the perceived risk of HCC for patients with chronic HCV infection, judging from a 2011 review from a worldwide

authority (15). Probably, older estimates of HCC risk were upwards-biased because they were computed with the familiar but inappropriate 1 minus Kaplan-Meier method.

State occupation probability

When we study patients with compensated cirrhosis using the three-state disease model in Figure 1c, we now know how to compute the cumulative risks of the two competing outcomes, 'decompensation' and 'death before decompensation' (Figure 3, left and middle). However, neither risk estimate tells us what happens *after* decompensation, so our analysis is still clinically incomplete. One solution is to compute the risk of death starting from the time of decompensation (Figure 3, right). However, this solution has the serious downside that it does not provide us with the clinical information we need, namely an overall assessment of the outcome of a patient presenting with compensated cirrhosis. Rather, we end up with a fragmented state-by-state prognosis. The necessary overview is provided by an analysis of *state occupation probability*.

Analyses of state occupation probability exploit the fact that patients are always in exactly one disease state and give the distribution of patients' disease states during the follow-up. If the disease model has only three disease states, the proportion of patients in two of the states is computed from the cumulative risks, and then the proportion of patients in the third state is defined by subtraction. However, it is computationally easier—and with four or more disease states necessary—to use the *Aalen-Johansen estimator*, which is a generalization of the Kaplan-Meier estimator to multi-state models. It provides state occupation probabilities for any number of states and any number of connections between the states (16).

Example

For this example we use data from our study of Danish patients with alcoholic cirrhosis (2) and consider the 114 patients who presented with compensated cirrhosis. We define four disease states: 'compensated', 'decompensated', 'dead without decompensating', and 'dead after decompensating' (Figure 1e). Although the distinction between death before and after decompensation may be inconsequential for the patient, it is important to the clinician because it gives clues about the likely causes of death. Together, Figures 3 and 4 illustrate the patients' clinical course: Within five years, 49% decompensate (Figure 3, left), and 22% die before they decompensate (Figure 3, middle). The same 22% at five years are found in Figure 4 ("Death without decompensating"). Because 49% decompensate and 22% die without decompensating, there must be 100% – (49% + 22%) = 29% who remain compensated after five years. This is also shown in Figure 4 ("Compensated"). Now, to grasp the full picture of the clinical course of patients presenting with compensated cirrhosis we must also consider what happens *after* decompensation. It is evident that 66% die within five years of decompensating (Figure 3, right), but it cannot be inferred from Figure 3 that 36% of patients with compensated alcoholic cirrhosis first decompensate and then die, all within five years. Instead, this is shown in Figure 4 ("Dead after decompensating"). Figure 4 also gives us the proportion of patients who have decompensated but remain alive five years after presenting with compensated cirrhosis: it is 13% ("Decompensated"). This example shows how the analysis of state occupation probability is necessary to give us the risk estimates that we intuitively seek in our understanding of the patients' clinical course.

Prognostic factors

Researchers are interested in factors that change the risk or rate of a clinical event. Such factors are called *prognostic factors* in studies of the clinical course of a disease, and *risk factors* in studies of the development of a disease. We deal here with prognostic factors, because we describe studies of the prognosis of patients who have been diagnosed with cirrhosis. Prognostic factors (and risk factors) can be classified as *causal factors* or *predictive factors*. A causal factor is a potential *cause* of the event, whereas a predictive factor is *associated* with the event, but its effect may not be causal. It follows that all causal factors are predictive factors, but not all predictive factors are causal factors. Studies of causal factors differ from studies of predictive factors, and we describe them in the following paragraphs.

Causal factors

In studies of potential causes of an event, researchers evaluate and test a hypothesis, typically by estimating the strength of the association between the potential cause and the *rate* of the event. 'Rate' is the proper metric because it—unlike the *risk* of the event—is unaffected by the rate of competing events, which are unwanted noise in studies of causal factors. For the same reason analyses of causal factors are done in the same way whether or not the underlying disease model is a two-state or a multi-state model. Cox proportional hazards regression is the standard analytical tool because it is an easily interpretable analysis of rates with the possibility to adjust for confounders.

Example

We are interested in causes of HCC development because we wish to understand the pathogenesis. We hypothesize that estrogens protect against HCC development (17),

and we evaluate our hypothesis by examining the association between gender and the rate of HCC development. Specifically, we examine gender's effect on the transition rate from 'no HCC' to 'HCC' within a three-state disease model with states 'no HCC', 'HCC', and 'dead' (like Figure 1c). We use a Cox regression model including gender and potential confounders, e.g. cirrhosis etiology and severity, and the hazard ratio for gender gives us the confounder-adjusted effect of gender on the rate of HCC development. The effect of gender on the rate of the competing outcome event, death without HCC, is not important for our study. This example illustrates how each transition can be considered separately as long as we are interested in *rates*, not risks.

Predictive factors

In studies of predictive factors, researchers aim to identify one or more factors that predict a patient's *risk* of the event of interest. 'Risk' is the preferred metric because we want to predict what happens *in the clinic*, that is, in the presence of competing risks. In prediction studies we must use different methods depending on whether we have a two-state or a multi-state disease model.

With a two-state disease model, a predictor that increases the *rate* of the event also increases the *risk* of the event (1). Therefore it is safe, and standard practice, to use Cox regression although it is an analysis of rates. With a multi-state disease model, Cox regression on the event of interest fosters problems because a predictor that increases the *rate* of the event may not increase the *risk* of it (1). This can occur when the predictor also affects the rate of a competing event, as shown in the example below. The most commonly used solution to the problem with non-correspondence between rate and risk is to use Fine and Gray regression instead of Cox regression. Other solutions

involve pseudo-observations (18), or the combination of Cox regression models for all possible transitions, but we will not discuss these rarer alternatives here.

Fine and Gray regression

Fine and Gray regression is useful for studies of predictive factors within a multi-state disease model because it allows researchers to draw conclusions about predictive factors' effect on the *risk* of an event (8, 19). It estimates the *subdistribution hazard ratio* for a predictive factor (e.g., for men relative to women) (1, 9), but it is important to be aware that the subdistribution hazard ratio for men vs. women is not the same as the relative risk for men vs. women. In fact, the subdistribution hazard ratio does not translate to anything epidemiologically meaningful (Table 1), but it is nevertheless useful because the relative risk is > 1 if the subdistribution hazard ratio is > 1. This means that Fine and Gray regression can determine whether or not a predictive factor has an effect on the risk of an event, and it can also determine the direction of that effect, but not its strength.

It is also technically possible to use Fine and Gray regression in analyses of two-state disease models. In that case the method is identical to Cox regression, at least in theory. In practice, the results may vary marginally because of the way the two regression methods are implemented in statistical software packages. However, there is no advantage to using Fine and Gray regression when the well-known Cox regression will do, and we recommend against it.

Example

We are studying HCC surveillance among patients with alcoholic cirrhosis and want to narrow down the target population for surveillance. Our goal is to identify predictors of

HCC risk, but in this example—which uses manipulated data—we consider only one potential predictor: gender. Our disease model has three states: 'no HCC', 'HCC', and 'dead' (like Figure 1c), and we want to know whether gender is a predictor of the risk of making the transition from 'no HCC' to 'HCC'. Figure 5 shows men and women's hazard rates and cumulative risk of the two competing outcomes, HCC and death without HCC. We notice that male gender increases the *rate* of HCC development (Figure 5, top left), but not the *risk* of HCC development (Figure 5, bottom left). This occurs because male gender also increases the rate of death without HCC (Figure 5, top right), which is the competing event in our disease model.

The results of the Cox and the Fine and Gray regression models are consistent with the visual impression of the rates and risks. With Cox regression we get a hazard ratio of 1.36 (95% CI 1.01 to 1.82) for men vs. women, but we need risks, not rates, because we want to know whether men are more likely than women to develop HCC *in clinical practice*. With Fine and Gray regression we get a subdistribution hazard ratio of 1.16 (95% CI 0.87 to 1.56). This tells us that men are at greater HCC risk than women are, but men's risk is not necessarily 1.16-fold higher. More importantly, the confidence interval tells us that the gender difference in HCC risk can be due to sampling variation alone. In summary, based on these manipulated data we should offer men and women the same HCC surveillance schedule because they have the same *risk* of HCC. We could have made the poor decision to recommend HCC surveillance in men alone if we had relied on the analysis of rates.

Software

The statistical software package R (http://cran.r-project.org/) is not only free, it also has the best capabilities for analyzing multi-state disease models (20). The calculations and illustrations in this article are all done in R version 2.15.0 (21). Several user-contributed packages compute the cumulative incidence function, among them the 'survival' and 'cmprsk' packages. The 'survival' package also handles all aspects of survival analysis with two-state disease models (22). The 'cmprsk' package was designed for analyses of the competing risks disease model (Figure 1b) and also handles Fine and Gray regression (23). This package has been described in a dedicated book aimed at both scientists and statisticians (24), and clinician-researchers may find it illustrative and helpful although statisticians have voiced concerns over the notation used (25). Analyses of multi-state disease models with the Aalen-Johansen estimator can be done with the R packages 'msSurv' or 'mstate' (26, 27).

Stata users can download the 'stcompet' function designed to compute the cumulative incidence function. Currently, no Stata functions are designed for the Aalen-Johansen estimator, but the built-in Stata function 'stcrreg' does Fine and Gray regression. Users of SAS can use the %CIF macro to compute the cumulative incidence function (28), and in SAS version 9.4 the built-in PHREG procedure does Fine and Gray regression. There is no support for the Aalen-Johansen estimator.

Conclusion

Epidemiologic studies of outcomes other than all-cause mortality should generally be based on multi-state disease models and analysis methods suitable for such models. This is particularly true in studies of the clinical course of cirrhosis patients because they

have a high mortality, and the errors incurred by *not* using correct methods grows with the risk of death. Erroneous estimates can lead to bad decisions. The methods for analyzing outcomes within multi-state disease models are both reasonably simple and accessible in standard software packages. Table 4 presents an ultra-brief overview of our recommendations. Anything worth doing should be done right.

References

1. Andersen PK, Geskus RB, de Witte T, Putter H. Competing risks in epidemiology: Possibilities and pitfalls. Int J Epidemiol 2012;41:861-870.
2. Jepsen P, Ott P, Andersen PK, Sørensen HT, Vilstrup H. The clinical course of alcoholic liver cirrhosis: A Danish population-based cohort study. Hepatology 2010;51:1675-1682.
3. Rosselli M, MacNaughtan J, Jalan R, Pinzani M. Beyond scoring: A modern interpretation of disease progression in chronic liver disease. Gut 2013;62:1234-1241.
4. Garcia-Tsao G, Friedman S, Iredale J, Pinzani M. Now there are many (stages) where before there was one: In search of a pathophysiological classification of cirrhosis. Hepatology 2010;51:1445-1449.
5. Arvaniti V, D'Amico G, Fede G, Manousou P, Tsochatzis E, Pleguezuelo M, et al. Infections in patients with cirrhosis increase mortality four-fold and should be used in determining prognosis. Gastroenterology 2010;139:1246-1256.
6. D'Amico G, Garcia-Tsao G, Pagliaro L. Natural history and prognostic indicators of survival in cirrhosis: A systematic review of 118 studies. J Hepatol 2006;44:217-231.

7. Andersen PK, Abildstrøm SZ, Rosthøj S. Competing risks as a multi-state model. Stat Methods Med Res 2002;11:203-215.
8. Bakoyannis G, Touloumi G. Practical methods for competing risks data: A review. Stat Methods Med Res 2012;21:257-272.
9. Andersen PK, Keiding N. Interpretability and importance of functionals in competing risks and multistate models. Stat Med 2012;31:1074-1088.
10. Kalbfleisch JD, Prentice RL. The statical analysis of failure time data. 1st ed. New York: John Wiley & Sons, Inc., 1980.
11. Gooley TA, Leisenring W, Crowley J, Storer BE. Estimation of failure probabilities in the presence of competing risks: New representations of old estimators. Stat Med 1999;18:695-706.
12. Jepsen P, Ott P, Andersen PK, Sørensen HT, Vilstrup H. Risk for hepatocellular carcinoma in patients with alcoholic cirrhosis: A Danish nationwide cohort study. Ann Intern Med 2012;156:841-847.
13. EASL-EORTC clinical practice guidelines: Management of hepatocellular carcinoma. J Hepatol 2012;56:908-943.
14. D'Amico G, Pasta L, Morabito A, D'Amico M, Caltagirone M, Malizia G, et al. Competing risks and prognostic stages of cirrhosis: A 25-year inception cohort study of 494 patients. Aliment Pharmacol Ther 2014;39:1180-1193.
15. Sherman M. Hepatocellular carcinoma: Screening and staging. Clin Liver Dis 2011;15:323-334.
16. Andersen PK, Keiding N. Multi-state models for event history analysis. Stat Methods Med Res 2002;11:91-115.

17. Naugler WE, Sakurai T, Kim S, Maeda S, Kim K, Elsharkawy AM, et al. Gender disparity in liver cancer due to sex differences in MyD88-dependent IL-6 production. Science 2007;317:121-124.
18. Andersen PK, Perme MP. Pseudo-observations in survival analysis. Stat Methods Med Res 2010;19:71-99.
19. Fine JP, Gray RJ. A proportional hazards model for the subdistribution of a competing risk. J Am Stat Assoc 1999;94:496-509.
20. Scrucca L, Santucci A, Aversa F. Competing risk analysis using R: An easy guide for clinicians. Bone Marrow Transplant 2007;40:381-387.
21. R Development Core Team. R: A language and environment for statistical computing. Vienna, Austria, 2012.
22. Therneau T. A package for survival analysis in S (R package version 2.36-14), 2012.
23. Gray B. cmprsk: Subdistribution analysis of competing risks (R package version 2.2-4), 2013.
24. Pintilie M. Competing risks: A practical perspective. 1st ed. Chichester, UK: John Wiley & Sons, 2006.
25. Latouche A, Beyersmann J, Fine JP. Comments on 'analysing and interpreting competing risk data' by M. Pintilie. Stat Med 2007;26:3676-3679.
26. Ferguson N, Datta S, Brock G. msSurv: An R package for nonparametric estimation of multistate models. J Stat Softw 2012;50:1-24.
27. de Wreede LC, Fiocco M, Putter H. The mstate package for estimation and prediction in non- and semi-parametric multi-state and competing risks models. Comput Methods Programs Biomed 2010;99:261-274.

28. Lin G, So Y, Johnston G. Analyzing survival data with competing risks using SAS® software. http://support.sas.com/resources/papers/proceedings12/344-2012.pdf (accessed on 12 August 2014).

29. Last JM. A dictionary of epidemiology. 4th ed. New York: Oxford University Press, 2001.

Table 1. Glossary.

Disease state	The clinical course of a disease can be described by two or more disease states. Each state must have a clearly defined entry and exit time point, and a patient is always in exactly one disease state. The simplest possible disease model has only two disease states: alive and dead.
Competing risks	Patients face competing risks if they are at risk of *multiple* outcomes from their current disease state, as defined by the study's disease model. Conversely, if there is just one possible outcome from their current disease state (i.e., death from any cause), patients do not face competing risks. It follows that competing risks are the rule, not the exception, in clinical epidemiologic studies.
Censoring	The commonest type is right-censoring which means that follow-up stops before the patient experiences an outcome event. Left-censoring and interval-censoring are other types of censoring, but they are rarely used in studies of cirrhosis patients.
Cumulative risk (*or* cumulative incidence)	The probability of experiencing an outcome event, e.g. death, within a specified period. Ranges from zero to one. Example: In a study of all-cause mortality, ten patients are followed for up to ten years. Five patients die—after 1, 2, 3, 4, and 5 years, respectively—and the remaining five are censored alive after ten years. Because no one is censored before ten years, we can simply estimate the 10-year cumulative risk of all-cause mortality as 5 / 10 = 0.5 or 50%. With the Kaplan-Meier method we get the same estimate: (9/10) * (8/9) * (7/8) * (6/7) * (5/6) = 0.5.
Incidence rate	A measure of the *average* incidence of a clinical event, e.g. all-cause mortality, during the follow-up period. Computed as *total number of events / total follow-up time*. Ranges from zero to positive infinity. Example: In a study of all-cause mortality, ten patients are followed for up to ten years. Five patients die—after 1, 2, 3, 4, and 5 years, respectively—and the remaining five are censored alive after ten years. The total follow-up time is 1 + 2 + 3 + 4 + 5 + 5*10 = 65 person-years, and there are five deaths, so the incidence rate of all-cause mortality is 5 / 65 = 1 per 13 person-years.
Hazard rate	A measure of the *instantaneous* incidence of a clinical event, e.g., decompensation, at a particular time t during the study period. Defined mathematically as the limit, as Δt approaches zero, of the probability that a patient who has not experienced the outcome at time t will experience the outcome at time $t+\Delta t$, divided by Δt

	(29).
State occupation probability	Given patients' disease state(s) at the beginning of follow-up, the state occupation probability describes the distribution of patients' disease states at a specific time during the follow-up.
Aalen-Johansen estimator	A method to compute state occupation probabilities for multi-state disease models with any number of states and any number of connections between states.
Prognostic factor	A factor that modifies the risk or rate of an event. A prognostic factor may be a *causal factor* or a *predictive factor* (predictor). A causal factor is a potential cause of the event of interest. A predictive factor is associated with the event, but its effect may not be causal.
Relative risk (*or* risk ratio)	A measure of the association between an exposure and an outcome event. The ratio of cumulative risks between an exposed and an unexposed cohort.
Incidence rate ratio	A measure of the association between an exposure and an outcome event. The ratio of incidence rates between an exposed and an unexposed cohort.
Hazard ratio	A measure of the association between an exposure and an outcome event. Output from Cox proportional hazards regression. If the hazard rate of an event is constant during the study period, the ratio of hazard rates in an exposed vs. an unexposed cohort—the hazard ratio—equals the incidence rate ratio. In clinical practice the hazard rate varies during the study period, but even so the hazard ratio may be assumed constant. A prognostic factor's hazard ratio for outcome *j* has an intuitive interpretation: it is the factor's effect on the instantaneous rate of experiencing outcome *j* for patients who have not previously experienced an outcome event (9).
Subdistribution hazard ratio	A measure of the association between an exposure and an outcome event. Output from Fine and Gray regression and therefore used only in analyses involving competing risks. The ratio of subdistribution hazard rates between an exposed and an unexposed cohort. A prognostic factor's subdistribution hazard ratio for outcome *j* has a non-intuitive interpretation: it is the instantaneous rate of experiencing outcome *j* for patients who a) have not previously experienced an outcome event, *or* b) have previously experienced an outcome event other than *j* (9).

Table 2. Computation of the cumulative risk of death with the cumulative incidence function (top) and as 1 minus the Kaplan-Meier estimator (bottom). This example uses a two-state disease model, so there is only one possible event: death (Figure 1a). With the cumulative incidence function, the step size at t_1 is computed as the product of the conditional probability of death at t_1 given alive before t_1 (i.e., beyond t_0) and the probability of not having died before t_1 (i.e., beyond t_0, S_0). The steps are then added sequentially. The Kaplan-Meier method estimates the cumulative survival probability by computing the probability of surviving past each successive event time. The cumulative risk of death is, of course, the complement of the cumulative chance of surviving. It can be shown that the two methods are mathematically equivalent, as long as we use a two-state disease model.
t_0: beginning of time interval; t_1: ending of time interval; d: number of deaths at time t_1; n: number of patients followed beyond time t_0; S_1: Probability of being alive beyond time t_1 [the product sign, Π, means that whenever a death occurs S_1 is multiplied by $(n\text{-}d)/n$]; S_0: Probability of being alive beyond time t_0 (equal to S_1 from previous time interval).

Time interval start: t_0	Time interval stop: t_1	Censored at t_1	Death at t_1: d	N followed beyond t_0: n	Probability of survival beyond t_1 if followed beyond t_0: $(n\text{-}d)/n$	Probability of survival beyond t_1: $S_1 = \Pi\,(n\text{-}d)/n$	Probability of survival beyond t_0: S_0	Probability of death at t_1: d/n	Cumulative risk of death beyond t_1: $\Sigma\,(d/n) * S_0$
0	1	1	0	10	1	1	1	0	0 * 1 = 0
1	2	0	1	9	(9-1) / 9 = 0.89	0.89 * 1 = 0.89	1	1/9 = 0.11	0 + 1 * 0.11 = 0.11
2	3	1	0	8	1	1 * 0.89 = 0.89	0.89	0	0.11 + 0 * 0.89 = 0.11
3	4	1	0	7	1	1 * 0.89 = 0.89	0.89	0	0.11 + 0 * 0.89 = 0.11
4	5	1	0	6	1	1 * 0.89 = 0.89	0.89	0	0.11 + 0 * 0.89 = 0.11
5	6	0	1	5	(5-1) / 5 = 0.80	0.80 * 0.89 = 0.71	0.89	1/5 = 0.20	0.11 + 0.20 * 0.89 = 0.29
6	7	0	1	4	(4-1) / 4 = 0.75	0.75 * 0.71 = 0.53	0.71	1/4 = 0.25	0.29 + 0.25 * 0.71 = 0.47
7	8	0	1	3	(3-1) / 3 = 0.67	0.67 * 0.53 = 0.36	0.53	1/3 = 0.33	0.47 + 0.33 * 0.53 = 0.64
8	9	1	0	2	1	1 * 0.36 = 0.36	0.36	0	0.64 + 0 * 0.36 = 0.64
9	10	1	0	1	1	1 * 0.36 = 0.36	0.36	0	0.64 + 0 * 0.36 = 0.64

Time interval start: t_0	Time interval stop: t_1	Censored at t_1	Death at t_1: d	N followed beyond t_0: n	Probability of survival beyond t_1 if followed beyond t_0: $(n\text{-}d)/n$	Probability of survival beyond t_1: $S_1 = \Pi\,(n\text{-}d)/n$	Cumulative risk of death beyond t_1: $1 - S_1$
0	1	1	0	10	1	1	1 – 1 = 0
1	2	0	1	9	(9-1) / 9 = 0.89	0.89 * 1 = 0.89	1 – 0.89 = 0.11
2	3	1	0	8	1	1 * 0.89 = 0.89	1 – 0.89 = 0.11
3	4	1	0	7	1	1 * 0.89 = 0.89	1 – 0.89 = 0.11
4	5	1	0	6	1	1 * 0.89 = 0.89	1 – 0.89 = 0.11
5	6	0	1	5	(5-1) / 5 = 0.80	0.80 * 0.89 = 0.71	1 – 0.71 = 0.29
6	7	0	1	4	(4-1) / 4 = 0.75	0.75 * 0.71 = 0.53	1 – 0.53 = 0.47
7	8	0	1	3	(3-1) / 3 = 0.67	0.67 * 0.53 = 0.36	1 – 0.36 = 0.64
8	9	1	0	2	1	1 * 0.36 = 0.36	1 – 0.36 = 0.64
9	10	1	0	1	1	1 * 0.36 = 0.36	1 – 0.36 = 0.64

Table 3. Computation of the cumulative risk of decompensation with the correct cumulative incidence function (top) and as '1 minus the Kaplan-Meier estimator' (bottom). This example uses a three-state disease model, so there are two possible events: decompensation and death (Figure 1c). With the cumulative incidence function, the step size at t_1 is computed as the product of the conditional probability of decompensation at t_1 given alive and compensated beyond t_0 and the probability of being alive and compensated beyond t_0 ($S_{0, CR}$). The steps are then added sequentially. The Kaplan-Meier method estimates the cumulative risk of decompensation like in Table 2, wrongly counting deaths as censorings. The wrongful handling of patients who experience the competing outcome event (death) results in upwards-biased estimates of the cumulative risk of decompensation (0.63 vs. 0.37 at time 10).
t_0: beginning of time interval; t_1: ending of time interval; d_1: number of decompensations at time t_1; d_2: number of deaths at time t_1; d: number of events at time $t_1 = d_1+d_2$; *n:* number of patients followed beyond time t_0; $S_{1, CR}$: Probability of being event-free beyond time t_1 [the product sign, Π, means that whenever an event occurs $S_{1, CR}$ is multiplied by $(n-d)/n$]; $S_{0, CR}$: Probability of being event-free beyond time t_0 (equal to $S_{1, CR}$ from previous time interval); $S_{1, KM}$: Uninterpretable and biased estimate of the probability of not having decompensated beyond time t_1 [the product sign, Π, means that whenever a decompensation occurs $S_{1, KM}$ is multiplied by $(n-d_1)/n$].

Time interval start: t_0	Time interval stop: t_1	Censored at t_1	Decompensation at t_1: d_1	Death at t_1: d_2	Total events at t_1: d	N followed beyond t_0: n	Probability of no event beyond t_1 if followed beyond t_0: $(n-d)/n$	Probability of no event beyond t_1: $S_{1,CR} = \Pi\,(n-d)/n$	Probability of no event beyond t_0: $S_{0, CR}$	Probability of decompensation at t_1: d_1/n	Cumulative risk of decompensation beyond t_1: $\Sigma\,(d_1/n) * S_{0, CR}$
0	1	1	0	0	0	10	1	1	1	0	0 * 1 = 0
1	2	0	0	1	1	9	(9-1) / 9 = 0.89	0.89 * 1 = 0.89	1	0	0 + 0 * 1 = 0
2	3	0	1	0	1	8	(8-1) / 8 = 0.88	0.88 * 0.89 = 0.78	0.89	1/8 = 0.13	0 + 0.13 * 0.89 = 0.11
3	4	1	0	0	0	7	1	1 * 0.78 = 0.78	0.78	0	0.11 + 0 * 0.78 = 0.11
4	5	0	1	0	1	6	(6-1) / 6 = 0.83	0.83 * 0.78 = 0.65	0.78	1/6 = 0.17	0.11 + 0.17 * 0.78 = 0.24
5	6	0	0	1	1	5	(5-1) / 5 = 0.8	0.8 * 0.65 = 0.52	0.65	0	0.24 + 0 * 0.65 = 0.24
6	7	0	0	1	1	4	(4-1) / 4 = 0.75	0.75 * 0.52 = 0.39	0.52	0	0.24 + 0 * 0.52 = 0.24
7	8	0	0	1	1	3	(3-1) / 3 = 0.67	0.67 * 39 = 0.26	0.39	0	0.24 + 0 * 0.39 = 0.24
8	9	0	1	0	1	2	(2-1) / 2 = 0.5	0.5 * 0.26 = 0.13	0.26	1/2 = 0.5	0.24 + 0.5 * 0.26 = 0.37
9	10	1	0	0	0	1	1	1 * 0.13 = 0.13	0.13	0	0.37 + 0 * 0.13 = 0.37

Time interval start: t_0	Time interval stop: t_1	Censored at t_1	Decompensation at t_1: d_1	N followed beyond t_0: n	Probability of no decompensation beyond t_1 if followed beyond t_0: $(n-d_1)/n$	"Probability" of no decompensation beyond t_1: $S_{1, KM} = \Pi\,(n-d_1)/n$	"Cumulative risk" of decompensation beyond t_1: $1 - S_{1, KM}$
0	1	1	0	10	1	1	1 – 1 = 0
1	2	1*	0	9	1	1 * 1 = 1	1 – 1 = 0
2	3	0	1	8	(8-1) / 8 = 0.88	0.88 * 1 = 0.88	1 – 0.88 = 0.12
3	4	1	0	7	1	1 * 0.88 = 0.88	1 – 0.88 = 0.12
4	5	0	1	6	(6-1) / 6 = 0.83	0.83 * 0.88 = 0.73	1 – 0.73 = 0.27
5	6	1*	0	5	1	1 * 0.73 = 0.73	1 – 0.73 = 0.27
6	7	1*	0	4	1	1 * 0.73 = 0.73	1 – 0.73 = 0.27
7	8	1*	0	3	1	1 * 0.73 = 0.73	1 – 0.73 = 0.27
8	9	0	1	2	(2-1) / 2 = 0.5	0.5 * 0.73 = 0.37	1 – 0.37 = 0.63
9	10	1	0	1	1	1 * 0.37 = 0.37	1 – 0.37 = 0.63

* Death, wrongly treated as a censoring event by the Kaplan-Meier method.

Table 4. Overview of recommended methods, by research question and disease model.

Research question	Two-state disease model	Multi-state disease model
Cumulative risk	Kaplan-Meier estimator	Cumulative incidence function
State occupation probability	-	Aalen-Johansen estimator
Prognostic factor (causal)	Cox regression	Cox regression
Prognostic factor (predictive)	Cox regression	Fine and Gray regression

Figure legends

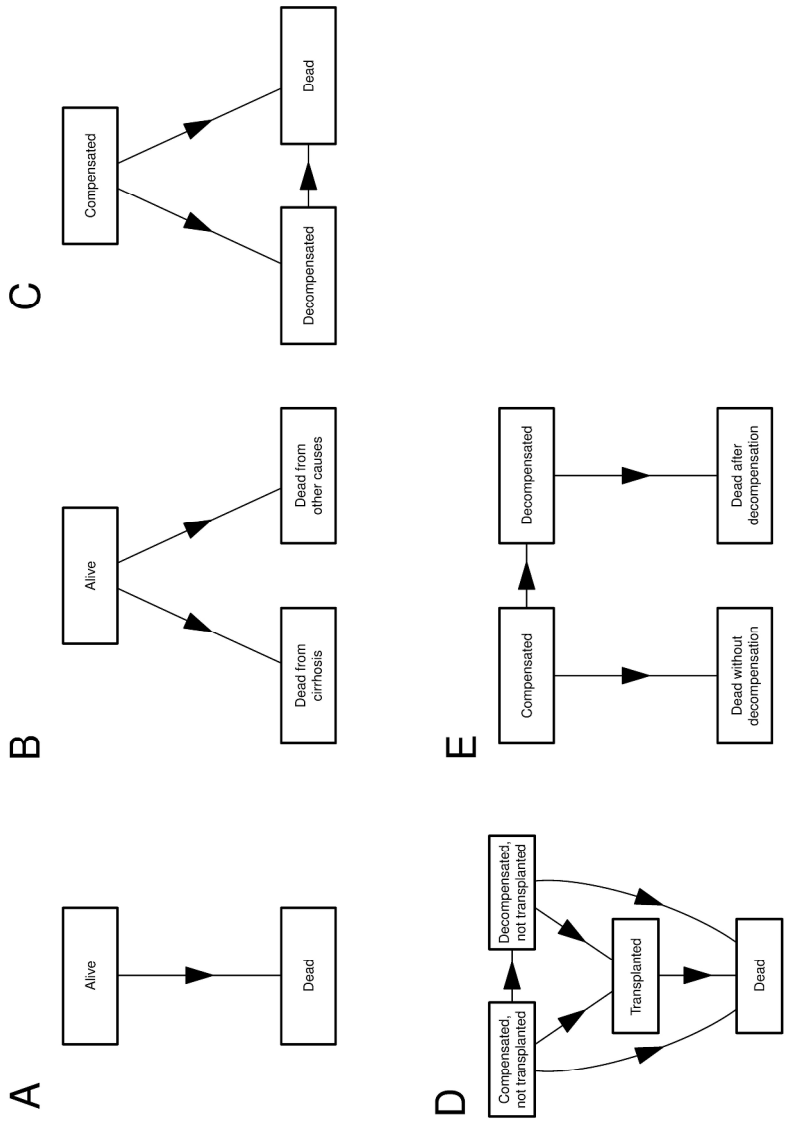
Figure 1: Disease models for liver cirrhosis. The model in panel A can be analyzed using standard survival analysis methods, whereas the competing risks disease model in panel B and the multi-state disease models in panels C, D, and E require different methods.

Figure 2: The cumulative risk of hepatocellular carcinoma (HCC) in patients with alcoholic cirrhosis computed correctly with the cumulative incidence function and wrongly as 1 minus the Kaplan-Meier estimator. The Kaplan-Meier method wrongly considers failure from a competing event (here: death without HCC) as a censoring event, and this results in upwards-biased estimates of the risk of experiencing the event of interest (here: HCC).

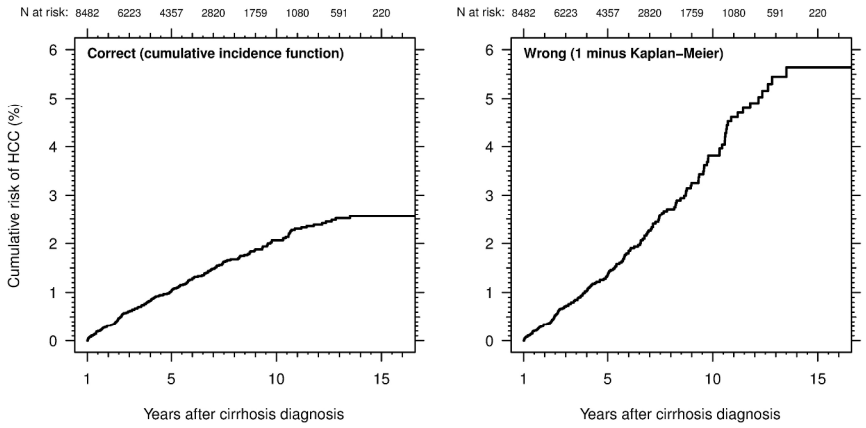
Figure 3: Cumulative risk of decompensation (left) and death before decompensation (middle) for 114 Danish patients with compensated alcoholic cirrhosis, and cumulative risk of death (right) from the time of decompensation for those 58 of the 114 patients who decompensated.

Figure 4: State occupation probabilities for 114 Danish patients followed from diagnosis of compensated alcoholic cirrhosis. The same data were used for Figure 3.

Figure 5: Analysis of made-up data on HCC and death in men and women with alcoholic cirrhosis. Panels show hazard rates and cumulative risks of HCC (top left and bottom left, respectively) and of death without HCC (top right and bottom right, respectively). Hazard rates are smoothed to facilitate the visual interpretation, but the smoothing is not important. What is important is that the *risk* of HCC is similar for men and women (bottom left) despite the fact that men have a consistently higher *rate* of HCC (top left). The risk of HCC is similar because men also have a higher rate of death without HCC (top right).

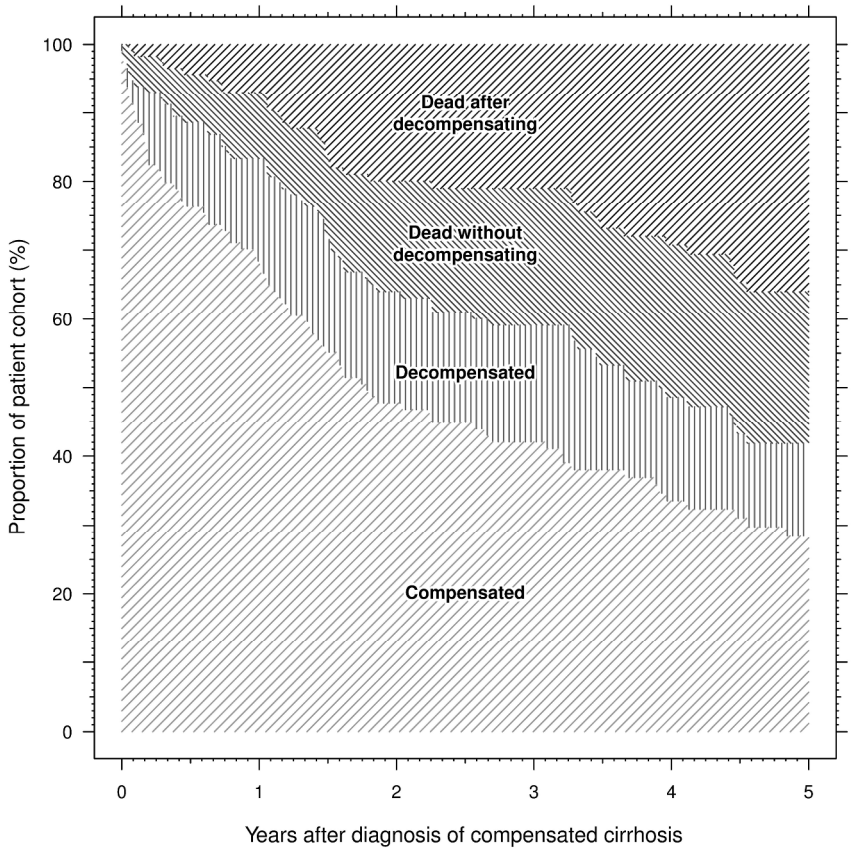


Disease models for liver cirrhosis. The model in panel A can be analyzed using standard survival analysis methods, whereas the competing risks disease model in panel B and the multi-state disease models in panels C, D, and E require different methods.
279x361mm (300 x 300 DPI)

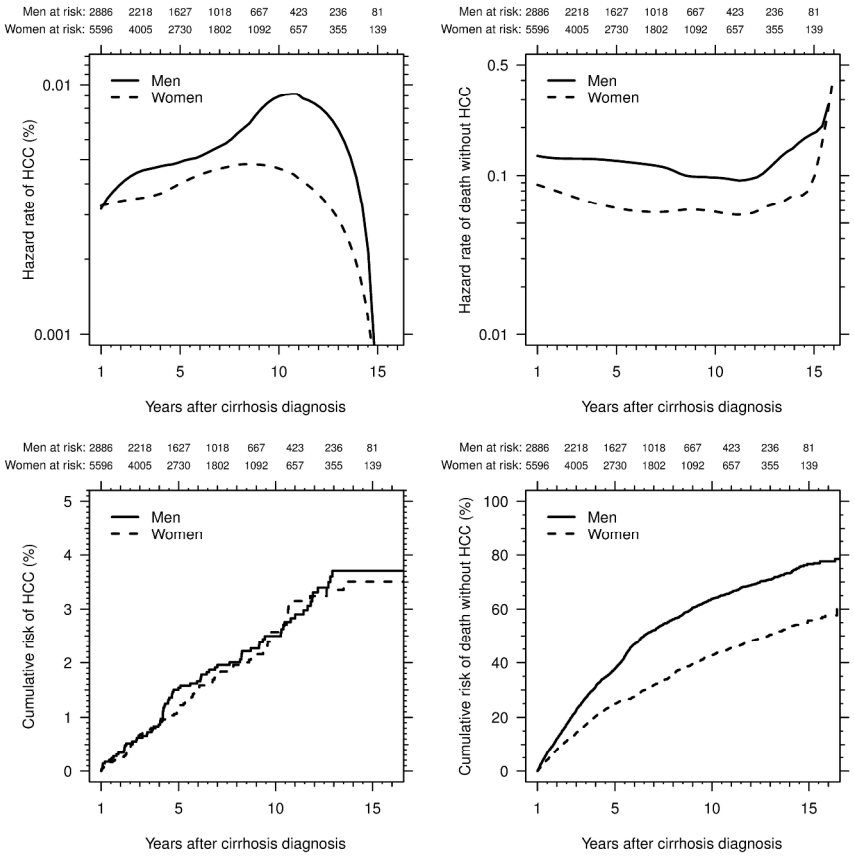


The cumulative risk of hepatocellular carcinoma (HCC) in patients with alcoholic cirrhosis computed correctly with the cumulative incidence function and wrongly as 1 minus the Kaplan-Meier estimator. The Kaplan-Meier method wrongly considers failure from a competing event (here: death without HCC) as a censoring event, and this results in upwards-biased estimates of the risk of experiencing the event of interest (here: HCC).
279x361mm (300 x 300 DPI)

Cumulative risk of decompensation (left) and death before decompensation (middle) for 114 Danish patients with compensated alcoholic cirrhosis, and cumulative risk of death (right) from the time of decompensation for those 58 of the 114 patients who decompensated.
279x361mm (300 x 300 DPI)



State occupation probabilities for 114 Danish patients followed from diagnosis of compensated alcoholic cirrhosis. The same data were used for Figure 3.
279x361mm (300 x 300 DPI)



Analysis of made-up data on HCC and death in men and women with alcoholic cirrhosis. Panels show hazard rates and cumulative risks of HCC (top left and bottom left, respectively) and of death without HCC (top right and bottom right, respectively). Hazard rates are smoothed to facilitate the visual interpretation, but the smoothing is not important. What is important is that the risk of HCC is similar for men and women (bottom left) despite the fact that men have a consistently higher rate of HCC (top left). The risk of HCC is similar because men also have a higher rate of death without HCC (top right).
279x361mm (300 x 300 DPI)