What is relative survival and what is its role in haematology?

Perme, M.P.; Wreede, L.C. de; Manevski, D.

Best Practice & Research Clinical Haematology 36 (2023) 101474

Contents lists available at ScienceDirect

Best Practice & Research Clinical Haematology

journal homepage: www.elsevier.com/locate/issn/15216926

What is relative survival and what is its role in haematology?

Maja Pohar Perme [a,*], Liesbeth C. de Wreede [b,c], Damjan Manevski [a]

[a] Institute for Biostatistics and Medical Informatics, Faculty of Medicine, University of Ljubljana, Vrazov trg 2, 1000, Ljubljana, Slovenia
[b] Biomedical Data Sciences, Leiden University Medical Center, Einthovenweg 20, 2333 ZC, Leiden, the Netherlands
[c] Clinical Trials Unit, DKMS, Augsburger Strasse 3, 01309, Dresden, Germany

ARTICLE INFO

Keywords:
Survival analysis
Relative survival
Net survival
Mortality tables
Competing risks
Haematological diseases

ABSTRACT

In many haematological diseases, the survival probability is the key outcome. However, when the population of patients is rather old and the follow-up long, a significant proportion of deaths cannot be attributed to the studied disease. This lessens the importance of common survival analysis measures like overall survival and shows the need for other outcome measures requiring more complex methodology. When disease-specific information is of interest but the cause of death is not available in the data, relative survival methodology becomes crucial. The idea of relative survival is to merge the observed data set with the mortality data in the general population and thus allow for an indirect estimation of the burden of the disease.

In this work, an overview of different measures that can be of interest in the field of haematology is given. We introduce the crude mortality that reports the probability of dying due to the disease of interest; the net survival that focuses on excess hazard alone and presents the key measure in comparing the disease burden of patients from populations with different general population mortality; and the relative survival ratio which gives a simple comparison of the patients' and the general population survival. We explain the properties of each measure, and some brief notes are given on estimation. Furthermore, we describe how association with covariates can be studied. All the methods and their estimators are illustrated on a sub-cohort of older patients who received a first allogeneic hematopoietic stem cell transplantation for myelodysplastic syndromes or secondary acute myeloid leukemia, to show how different methods can provide different insights into the data.

1. Introduction

Survival analysis is one of the main tools for evaluating the burden of a lethal disease, as it examines the elapsed time from a starting time point (e.g. diagnosis, start of first-line treatment) until death [1]. However, it is common that not all deaths occurring in the patient cohort are due to the disease in question. This happens in particular when studying older patient cohorts with longer follow-up. In ageing populations, the probability of cancer patients dying due to other causes is increasing, especially for those patients who have survived the acute most dangerous phase after diagnosis. Moreover, an increasing number of malignancies are changing from fatal to chronic diseases due to extensive advances in therapy, implying that many patients will die with, not due to cancer.

With a substantial proportion of the population dying due to other causes, the patient's overall risk of dying may be less informative, and describing the disease-specific risk becomes more relevant. If the data contain information on the cause of death, one can

* Corresponding author.
E-mail address: maja.pohar@mf.uni-lj.si (M. Pohar Perme).

https://doi.org/10.1016/j.beha.2023.101474
Received 17 January 2023; Received in revised form 1 March 2023; Accepted 2 May 2023
Available online 8 May 2023

estimate the probability of dying due to the disease and the excess hazard of death due to the disease. However, the cause of death information may not be available [2] or one may not wish to rely on it [3,4], and the field of relative survival has been developed to avoid needing such information in the analysis. Instead, the idea of relative survival is to compare patients' mortality to that of the general population, thus enabling an indirect estimation of the disease-specific burden.

Relative survival is frequently used when analysing the survival of cancer patients. The current paper focuses on its application in the study of haematological malignancies but almost all of its content can be applied to study survival of other patient groups as well. Some applications in haemato-oncology compare the measures across the major haematological malignancies [5–7] (also focusing on older [8] or younger patients [9]), or consider leukemia [10], (non-)Hodgkin lymphoma [11–13], multiple myeloma [14,15], or outcomes of patients who were treated by an allogeneic hematopoietic stem cell transplantation (alloHCT), mainly for different types of leukemia [16–21].

Our motivational example will be the survival of patients who received a first alloHCT for myelodysplastic syndromes (MDS) or secondary acute myeloid leukemia (sAML). In this context, the traditional division of mortality into relapse-related and treatment-related mortality was a valid approximation until recently. However, with increased access to alloHCT for older patients and improved survival through time, there is more room for a third component, i.e. mortality due to causes also acting in the general population. In particular, when focusing on the survival probability in a good prognosis subgroup, e.g. patients alive without relapse/progression after two years follow-up, death due to other causes is an outcome that may arise with a substantial probability. As a consequence of this, overall survival is not a sufficiently specific measure for describing the disease burden.

The disease-specific burden may be expressed with several different measures: one may be interested in the probability of dying due to the disease and due to other (population) causes, in comparing the survival of groups of patients without being affected by the other-cause mortality risks, or simply comparing the patients' survival probability to that of the general population. The differences between the measures may be subtle. Substantial advances and hence changes in the methodology to be used have been made in the relative survival field in the past two decades [22–24], leading to some confusion and misconceptions in the literature (e.g. not differentiating between net survival and the relative survival ratio, or ambiguity on the estimators that have been used). We believe that by providing a clear overview of the possible measures and their estimators in this paper, future studies in the field of haematology can make a more deliberate choice between measures and report more explicitly about them, thus making interpretation and comparisons between studies easier.

In Section 2, we will first present the idea of relative survival and then look at the different questions that may arise in this context. We believe that it is crucial to first understand the study question, and only then choose among the available methods of estimation. In Section 3, we provide a detailed application of the introduced methods on the dataset of MDS/sAML patients followed after alloHCT. In Section 4 we provide the final discussion and main conclusions of the paper.

2. Materials and methods

We start by introducing the relative survival measures in Section 2.1 and then consider their estimation in Section 2.2. Throughout this paper, we assume that the core information we have on each patient is the follow-up time and censoring status. All patients are followed from a meaningful starting point (usually diagnosis or treatment) and the censoring status indicates whether they have been followed until their death (event) or their follow-up ended while they were still alive, commonly due to end of study or loss to follow-up.

While the most basic representation of the outcome is given in the left-most graph of Fig. 1, note that in fact various causes of death exist, and we wish to distinguish between them and thus enable a more meaningful estimation of the burden of the disease. The causes of death are split into two groups - those due to the disease of interest (excess death) and all others (population death). If cause of death information was available in the data, this would be the common competing risks setting (middle graph of Fig. 1). However, throughout this work, we will assume that cause of death information is not given in the data (we simply know who died and who did not). This data setting is the main topic of this paper and is referred to as the 'relative survival setting' (right-hand graph of Fig. 1).

2.1. The measures of interest

Different measures may be of interest, here we list the most commonly used ones.

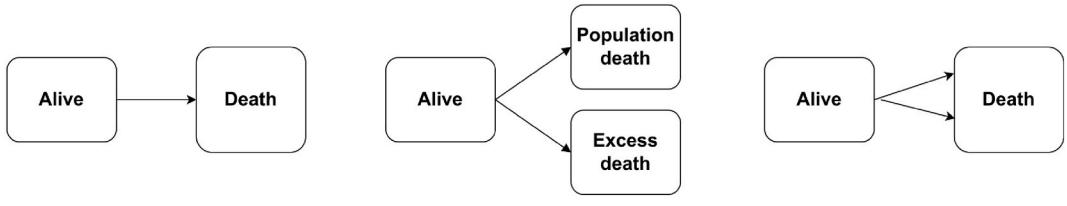

Fig. 1. Left: basic survival case (alive - death). Middle: competing risks scenario (alive - two observed causes of death). Right: relative survival scenario (alive - two unobserved causes of death).

2.1.1. Overall survival

The primary measure one considers in survival analysis is overall survival, i.e. the probability that a patient is still alive at a certain timepoint. In this case, death is the outcome event of interest without considering the fact that people may die of various causes. The survival probability monotonically decreases in time as more and more individuals experience their event. One can alternatively report the hazard of dying (instantaneous risk of death), which represents the rate (the 'speed') at which events occur at each time point. This is a more dynamic description of the risk of events in time which can increase or decrease in time. While the two measures (overall survival and hazard) may be insightful since they present the data from two different viewing points, both essentially represent the same information - knowing one, the other can be calculated.

2.1.2. Expected population survival

The relative survival field enables comparisons of an individual's (or group's) survival to the expected survival in the general population. One approach for performing such a comparison would be to gather a control sample of individuals from the general population with the same demographic variables (sex, age, etc.) as the patients under study. One would follow them in time and compare their survival to that of the patients. However, a more convenient (and inexpensive) approach is to use the general population mortality tables to this end. For every individual in the original data set one answers the question: if an individual is followed for a certain period of time, what is the probability that their counterparts in the general population (individuals with the same demographic characteristics, but without the disease) have survived this same time period? This is the so-called expected population survival for each individual. One can then average the individual expected probabilities to get what is referred to as the expected population survival of a group of patients (the opposite, i.e. one minus expected population survival is referred to as expected population mortality). As we will show in the following subsections, we use the population mortality tables not only to calculate the expected population mortality but also to indirectly extract disease-specific information from the observed data on patients.

2.1.3. Additive relative survival model

We now move on to the case where specific causes of death may be of our interest (see middle and right-hand graph in Fig. 1). Every individual in the general population at each time point has a certain hazard of dying that depends on their age and other demographic variables. This hazard is referred to as the population hazard, one can get it for each individual from the population mortality tables. In relative survival one assumes that the death hazard in the general population can be used to quantify the other-cause (non-disease-related) hazard in the patient group. The plausibility of the assumption will be further discussed in Section 2.2.1. If this holds, the so-called 'additive relative survival model' can be used which states that the overall (observed) hazard is the sum of the population hazard and the excess hazard due to the disease and its treatment. The model assumes that patients with the studied disease have a higher overall hazard of dying than the general population which is commonly true. This model is the basis for defining the endpoints of interest in relative survival. The excess hazard is also referred to as disease- or cause-specific hazard. In practice, the overall hazard is observed, the excess hazard is the quantity of interest, whereas the population hazard is essentially a nuisance parameter.

2.1.4. Crude mortality

While the hazard function is crucial for the theory, the summary measures are usually based on probabilities which are easier to interpret. As already noted, the overall survival is usually not very informative in terms of the disease burden since it only gives the percentage of individuals that survive for a certain time period (say 5 years). Therefore, the next logical question is: what proportion of patients has died due to the disease in question? The overall probability of dying (which equals one minus the survival probability) is thus split into the probability of dying due to the disease and probability of dying due to other causes. The two probabilities of dying are commonly referred to absolute risks or cumulative incidence functions. In the relative survival setting, the probability of dying due to the disease is of particular importance and is referred to as crude mortality.

The crude mortality provides information on the disease-specific hazard, but it is important to note that it may also strongly depend on the hazard of dying due to other causes. If the hazard of dying due to other causes is high (for example in old patients), this also affects the crude mortality, as some patients die of other causes before they could die due to the disease. When comparing crude mortality of different groups (for example different countries), one should therefore bear in mind that crude mortality is affected not only by the quality of treatment and care for the patients but also by the magnitude of the other-cause hazard. If in a certain group (for example a low-income country patient group) the other-cause hazard is substantially higher, fewer patients will die due to the disease and their crude mortality may be relatively low.

Following the same reasoning, the probability of patients dying due to other causes is lower than the population mortality of their counterparts. This is because many cancer patients die due to the disease which gives them less opportunity to suffer other-cause mortality compared to the general population, even though their population hazard is the same.

To summarize, crude mortality may be a very appealing measure when analysing a certain patient cohort. However, when considering the differences in treatment and care of various patient groups having different population mortality (e.g. patients from distinct countries with the same diagnosis), crude mortality may not be the best choice, as it is not clear whether an observed difference in crude mortality is due to differences in the excess or population hazard (the latter being typically not of interest).

2.1.5. Net survival

To report differences in the excess hazard, one thus needs a third measure which is referred to as net survival. Net survival is the survival function that depends on the excess hazard only and answers the question: how long would patients live in a hypothetical world where individuals can only die due to the disease? This hypothetical world is in itself not of interest; nevertheless, net survival

presents a useful direct measure of the quality of treatment and care. Its most important property is that it is not influenced by the population hazard, thus allowing comparisons between populations with different other-cause mortality (different countries, different periods of time, different age groups). The net survival measure is most commonly used in population-based cancer survival comparisons [25]. One minus net survival is not equal to crude mortality - some of the patients who would die due to cancer in the hypothetical world die due to other causes in the real world (analogous to the comparison between expected population mortality and the probability of dying due to other causes, see Section 2.1.4).

While the hypothetical world interpretation is key for understanding the properties of net survival, it makes more sense to interpret the results on the data in the "real world". To this end, we use the following result: since the excess hazard is the difference between the overall and the population hazard and the difference in hazards implies a ratio of survival functions, the net survival of each individual is the ratio between the patient's overall survival and their population survival. The net survival of a group can thus be interpreted as the average ratio between the patients' overall survival and the survival of their counterparts in the general population.

Note that the definition of net survival is problematic if the patients live on average longer than the general population. In this case, the additive relative survival model does not hold. Both net survival and crude mortality become impossible to interpret as probabilities.

2.1.6. Relative survival ratio

As the fourth and last measure, we consider the relative survival ratio. This measure simply compares the overall survival of a group of patients (see Section 2.1.1) to the expected population survival (see Section 2.1.2) of that group by calculating the ratio of the two measures.

This may seem very close to the net survival definition. But while net survival is the average of ratios, the relative survival ratio is a ratio of two averages (the overall and the population survival probabilities are averages of individual survival probabilities in those groups). Although clearly similar in terms of interpretation, the two measures are not mathematically the same and may yield quite different results. Historically, the relative survival ratio was understood to be the same as net survival, which is the root of much of the confusion in the literature.

To summarize, we can choose between the possible measures by understanding what is of interest.

- **Overall survival:** to know the probability that patients survive a certain time period (not interested in cause of death).
- **Crude mortality:** to know the probability of dying due to the disease in a certain time period.
- **Net survival:** to evaluate disease-specific mortality independently of the general population mortality and hence enable comparison between populations.
- **Relative survival ratio:** to compare the survival experience of the patients with the survival in a comparable group in the general population.

For a more detailed introduction and precise definitions of the measures, as well as an explanation of the relsurv package in R which can be used to do all analyses discussed in this paper, see Refs. [26–28].

2.2. Estimation methods

2.2.1. The relative survival data setting

Once the study question is clear and the measure(s) of interest chosen, an appropriate estimator must be used in the analysis. In the simplest setting where only death of any cause is considered, the data needed for each individual is the follow-up time and the censoring indicator, usually coded as 0 or 1, for censoring and death, respectively. When moving on to the relative survival case (with no cause of death information available), the only way forward is to bring the cause-specific information from elsewhere; namely by using the general population mortality tables. These are commonly available on the national level, for example at the HMD database which gives them in a uniform format [29]. Note that two basic criteria must hold when assuming the population hazard for all patients can be obtained from the general population mortality tables.

- First, if it were not for the disease in question, the mortality risk of the patients would be equal to that of the general population. This may be a credible assumption when there are no known strong risk factors that predispose persons both to the development of the disease in question and to other lethal diseases. This is certainly not true for lung cancer where smoking that in many cases led to cancer also greatly increases the risk of other morbidity. In the haematology field, the assumption is problematic when patient selection in a study is not diagnosis but treatment-based. In our example, alloHCT may not be accessible to all candidates and selection might be based on health-associated features. In such an application, a careful comparison of the included patient population and the general population is recommended if the data allow for that. Relative survival methods can still be used but the metrics might suffer from some bias. An adapted model for estimating the excess hazard when this assumption does not hold has been suggested [30].
- Second, the disease in question must contribute a negligible part to total mortality (even if all patients with the disease were excluded from the general population, the population mortality risk would remain practically equal). In practice, this assumption usually holds (the prevalence of haematological malignancies in the population is small). A recent population-based cancer study [31–33] has shown that this criterion is met with all specific cancer types, but is certainly not met when all cancers combined are considered in the analysis.

2.2.2. The estimators

Overall and population survival, crude mortality. The overall survival probability and overall hazard are standardly estimated using the Kaplan-Meier and the Nelson-Aalen estimator, respectively. The individual population hazards are directly obtained from the population mortality tables - for each individual, this is the value in the population tables reported for their demographic group (individuals of the same age and sex at the same calendar time). The population hazard for a whole group of patients is calculated as the average of the individual population hazards among those patients that are still at risk at a certain time. An estimator for the excess hazard is then the difference between the overall and the population hazard estimators. This procedure results in a non-parametric estimator of crude mortality [26].

Net survival. A bit more work is required to get an unbiased estimator of net survival. Simply using the above-mentioned estimator of the excess hazard gives the Ederer II estimator, which is a biased estimator of net survival. It is key to understand that the hazard function describes only the patients still at risk at a certain time point. In the estimation, one has to take into account that the risk group in the hypothetical world may not be the same as in the real world. The hypothetical world is defined as the scenario where individuals cannot die due to other causes, whereas in the real world, many, especially older patients, do. Therefore, the risk set in the real world after a certain period of time often includes too few old patients to represent the risk set in the hypothetical world. To correct for that, one must thus weigh the real-world risk set to make it more similar to that expected in the hypothetical world: the remaining older patients in the observed risk set get a higher weight to represent also those who have left the risk set because of dying due to other causes. The weights are inversely proportional to the population probability of still being in the sample. Since the number at risk in the observed data is too low, the number of events is too low as well and has to be weighted with the same weights. This weighted alternative to the Ederer II estimator has been referred to as the Pohar Perme estimator and shown to have the required statistical properties [22].

Relative survival ratio. The estimation of the relative survival ratio requires the overall and population survival, which can be estimated using the Kaplan-Meier estimator and obtained from the population mortality tables, respectively. This estimator is referred to as the Ederer I estimator. The Hakulinen estimator [34] is an alternative estimator that may help to remove a part of the bias in case of informative censoring in the overall survival, but may not be recommended in the case of non-informative censoring, where additional bias is introduced [35].

2.2.3. Comparisons

In practice, we often wish to compare net survival across different groups (e.g. countries, age groups). A log-rank type test can be used for performing such comparisons [36,37]. The null hypothesis is that net survival is equal for all groups throughout the follow-up time.

To compare net survival with respect to a continuous variable or to include more covariates in the comparison, one can turn to regression models. In survival analysis, regression modelling is commonly performed on the overall hazard, meaning that the effect of covariates is considered on the sum of the excess and population hazards. In relative survival, the goal is to focus specifically on the excess hazard. As the excess hazard is in a one-to-one relationship with net survival, modelling the excess hazard implies modelling the net survival.

The most common way to do that is to assume that the covariates are related to the excess hazard analogously to the Cox model. This Cox-type model assumes that the effect of covariates on the excess hazard (and thus net survival) is constant in time yielding hazard ratios, whereas the baseline excess hazard may vary in time. There are multiple ways of estimating such a model [28], depending on our assumptions of the baseline hazard. This can be either left unspecified and estimated using the EM algorithm as introduced in Ref. [38], or estimated using any flexible parametric baseline model through maximum likelihood [39–41]. In this work, we use the EM algorithm.

Instead of the excess hazard function, one could also model the crude mortality directly [42], but as in the competing risks data setting, this is rarely performed in practice.

3. Results

The practical use of the presented relative survival methods is shown using EBMT registry data [43] analysed in previous work [17]. The EBMT is an international (mainly Europe-based) organisation whose members collaborate in the field of autologous and allogeneic hematopoietic stem cell transplantation and other cellular interventions, mostly for malignant diseases. The data describe patients who received a first allogeneic hematopoietic stem cell transplantation (alloHCT) for MDS or sAML between January 2000 and December 2012. Full information on the data regarding selection criteria, patient characteristics and outcomes and a discussion of the suitability of relative survival methods for this cohort is available in the previous work [17].

Patients enter the dataset at alloHCT and are followed from this point onwards. A considerable number of patients die in the first two years after alloHCT due to relapse or complications of the treatment. For illustrative purposes, we focus on patients who have survived the first two years after alloHCT without a relapse of the underlying disease. As these patients have a better prognosis, it is sensible to evaluate the long-term survival of such a subgroup. We will refer to this time point (two years) as the landmark time and refer to the analysis using only the patients who survived and are relapse-free at the landmark time as the landmark analysis. The dataset contains data of 2578 patients from 21 countries with a median age of 56.8 years at alloHCT among which 58% were male and 42% female. From the original data, we use information on follow-up, relapse/progression (an intermediate event that may occur through time; for the remainder of this section we refer to relapse) and demographic covariates (age, sex, year of diagnosis, country). The maximum follow-up time is set at 8 years after the landmark.

We are interested in the effect of age on long-term survival. To study survival curves, we perform subgroup analysis. Here, we focus on patients older than 65 years at alloHCT for whom we expect a substantial number of deaths due to other causes. We compare them to the outcomes of patients aged less than or equal to 65 years. We start with overall survival (the non-parametric Kaplan-Meier estimate) shown in the left graph in Fig. 2 for the two subgroups. We see that the overall survival is lower in the older age group compared to the younger: at 8-year after the landmark, the survival estimates equal 54% (95% CI: [46%, 64%]) and 73% (95% CI: [70%, 76%]), respectively.

To get an idea of the cancer-specific mortality, we turn to absolute risks (crude mortality and probability of dying due to other causes). In Fig. 3, the absolute risks of dying due to the disease and due to other causes are shown for both age subgroups. The probability of dying due to the disease is larger for the older age group compared to the younger group (see left graph in Fig. 3). At 8 years after landmark, these probabilities equal 33.0% (95% CI: [23.5, 42.5]%) and 22.8% (95% CI: [20.0, 25.6]%), respectively. An even more apparent difference is obtained for the probability of dying due to other causes (right graph in Fig. 3); at 8 years, they equal 12.6% (95% CI: [11.8, 13.4]%) and 4.1% (95% CI: [4.0, 4.2]%) for the old and young age groups, respectively. The sum of the probabilities within each group gives the total probability of dying (33.0% +12.6%+54.4% = 100% for the older patients, where 54.4% is the overall survival estimate given above). The overall survival clearly indicates that older patients have higher mortality. With the analysis of absolute risks, we see that this is due to higher population mortality but also due to higher crude mortality. Furthermore, we observe that the probability of dying due to other causes at 8 years takes almost a two times as large proportion of the overall probability of dying in the older group - this proportion equals 27.6% ($(\frac{12.6\%}{12.6\%+33.0\%})$) in the older group compared to the 15.2% ($(\frac{4.1\%}{4.1\%+22.8\%})$) in the younger group. Thus population mortality plays a lesser role for the younger age group compared to the older one, both on an absolute and a relative scale.

In Fig. 3, the expected mortality in the general population is also shown for comparison. Note that, as mentioned in Section 2.1.4, this differs from the absolute risk of dying due to other causes in the patient population (see the solid and dotted curves in the right graph in Fig. 3). The general population mortality is the expected mortality of a group of individuals (from the general population) with the same demographic characteristics as the study patients. This measure does not depend upon the patient's overall survival and it may be considered as the mortality of a control group from the population. The probability of dying due to other causes is lower in the patients than in the general population not because they are fitter but because they die due to cancer first; their disease-related death 'protects' them from a later other-cause death.

Thus, the absolute risks are affected by both competing hazards. If one wishes to obtain disease-specific information, not affected by differences in general population mortality of the two age groups, net survival is a more suitable measure. The net survival estimates are shown in the right graph in Fig. 2. The net survival at 8 years after landmark equals 64% (95% CI: [54%, 76%]) and 77% (95% CI: [74%, 80%]) for the old and young age groups, respectively. We interpret these estimates as the average ratios of overall and expected survival across all patients. If we focus on older patients, their survival at 8 years is on average 64% of the survival of their counterparts in the general population. The net survival curves for the two age groups may also be compared using a formal statistical test based on a similar procedure as for the log-rank test [37]. This gives a p-value of 0.009. The net survival probability lies between the overall survival and 1 minus crude mortality probabilities: it is higher than overall survival since the excess hazard is smaller than the overall hazard, but lower than 1 minus crude mortality since part of the patients dying in the real world of population causes would die of the

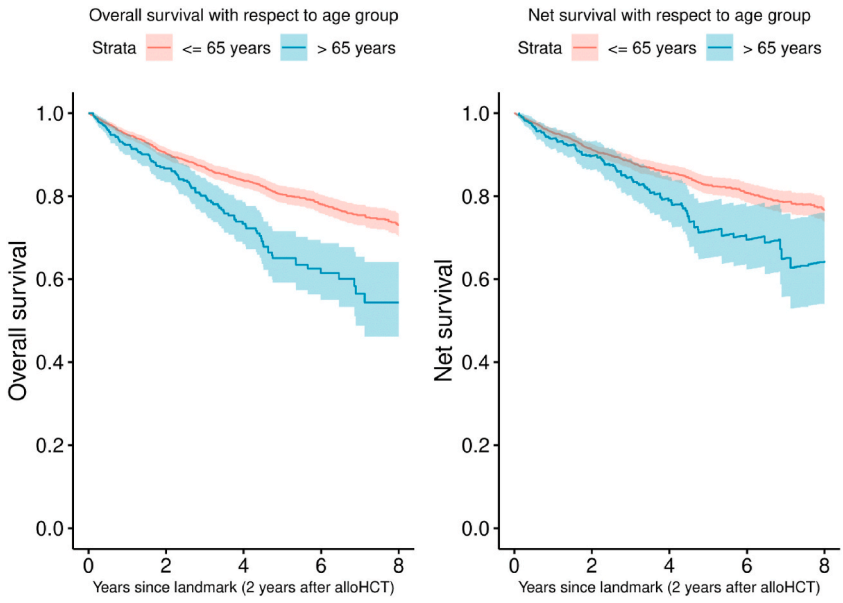

Fig. 2. The estimated overall survival (left) and net survival (right) with corresponding 95% confidence intervals for patients who are alive event-free at the 2-year landmark time. The results are shown with respect to the age group (65 years or younger and older than 65 years).

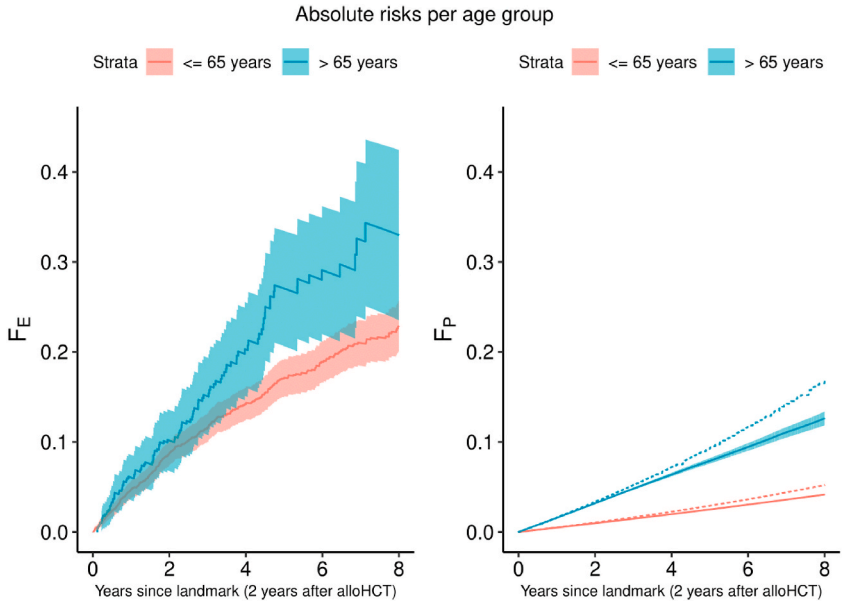

Fig. 3. The estimated absolute risks of dying for patients who are alive event-free at the 2-year landmark time. Absolute risks of dying due to the disease and due to other causes with corresponding 95% confidence intervals are shown in the left and right graphs, respectively. On the right graph, the expected population mortality in the matched general population is also shown (dotted curves). The results are shown with respect to age group; patients 65 years or younger (red) and older than 65 years (blue).

disease in the hypothetical world (see Section 2.1.5).

In the right graph in Fig. 2, we also observe that the net survival may increase within certain time intervals (e.g. group >65 years between 7 and 8 years after landmark). This should not be overinterpreted as it is a property of the estimator (there are no observed events but the population hazard increases, leading to a negative estimate of the excess hazard). Such an increase in net survival is common when the number of patients at risk is small and no events occur during a time period.

The relative survival ratio is calculated for the >65 age subgroup and shown in Fig. 4. It describes how the patients' survival compares to that of the matched general population. Since it is not defined as a probability (or a survival) function, its values can be outside of the [0,1] interval and the ratio may increase through time. The relative survival ratio is often confused with net survival. In practice, both measures often give similar but not equal values, as is the case in our example (Fig. 4). The relative survival ratio decreases throughout most of the time interval which means that the patient's survival is worse than the survival of the general population.

To study the effects of age and sex, which are often correlated, on the disease-specific hazard simultaneously, we fitted a semi-parametric Cox regression model. Results are shown in Table 1. Assuming a linear effect of age, the effect of age on the disease-specific hazard is statistically significant and the hazard ratio for age in years equals 1.03. Female patients have a slightly smaller

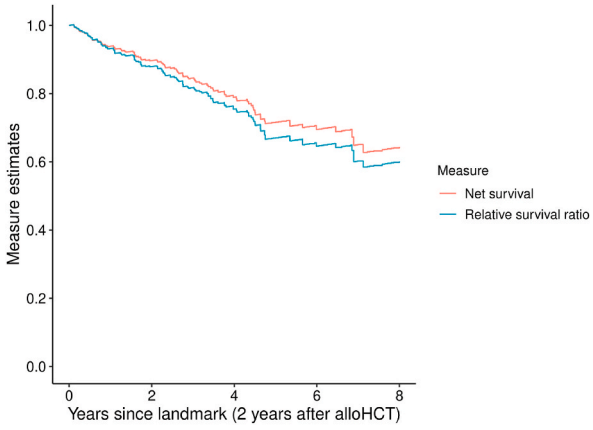

Fig. 4. The estimated net survival (red curve) and relative survival ratio (blue curve) for patients that are alive event-free at the 2-year landmark time. The results are shown only for patients older than 65 years.

Table 1
The estimated coefficients of the semi-parametric Cox-type model for the excess hazard with age and sex as covariates for patients who are alive event-free at the 2-year landmark time. Follow-up is artificially censored at 8 years since the landmark time.

Covariate	Hazard ratio	95% CI
age (in years)	1.03	[1.02, 1.04]
sex female (vs. male)	0.85	[0.70, 1.05]

excess hazard although this result is not statistically significant. As a comparison, one may also fit a semi-parametric Cox regression model on the overall hazard. We fitted such a model (results are available in the Supplementary material) which shows that in our case, the effect of the two covariates is similar on the overall and excess hazard.

Based on the Cox model we can also provide predictions. Fig. 5 shows the predicted excess hazard and net survival through time for two male patients aged 40 and 65 two years after the diagnosis (red and blue curves, respectively). Both graphs show essentially the same information but the hazards provide more insight into the dynamics of the process whereas the survival curves show cumulative results of it.

In the context of alloHCT, relapse is major intermediate event that influences the probability of dying. Multi-state models are a standard tool for calculating the probability of such events and their impact on different causes of death. This is illustrated on the EBMT data in the Supplementary material (Section S2).

The programming package relsurv [44] allows for performing the statistical analysis in R. A step-by-step illustration of the R code needed for the Results section is provided in Supplementary material.

4. Discussion

The core idea of the relative survival methodology is to extract disease-specific information from the observed data by merging the observed data with population mortality tables. Relative survival can yield important insights about patients's survival, whenever the patients' hazard of dying can be split into the hazard due to the disease and that due to other causes, with both causing a non-negligible proportion of deaths. This may occur with several haematological diseases, for example in Myeloma, Chronic Lymphocytic Leukemia or (good risk) AML, which all have an advanced median age of onset [5]. In this paper, we considered patients with MDS or sAML that have survived two years after alloHCT, thus entering a more chronic phase of the disease with a substantial proportion of (particularly older) individuals dying due to other causes and not due to the primary disease. In this situation, the risk of dying due to the disease becomes quite different from the overall risk of dying and relative survival methodology can bring valuable new information.

Two assumptions should be fullfilled to make it possible to use the population mortality data as a proxy for the other-cause hazard of the patients in the study. First, the studied disease should take a negligible part of the general population mortality; second, the population mortality tables should adequately describe the patients' hazard of dying had they not had the disease. If the assumptions do not hold, the estimated excess hazard cannot be fully attributed to the studied disease.

In this work, we have explained the different ways of analysing and interpreting the data using relative survival - various questions

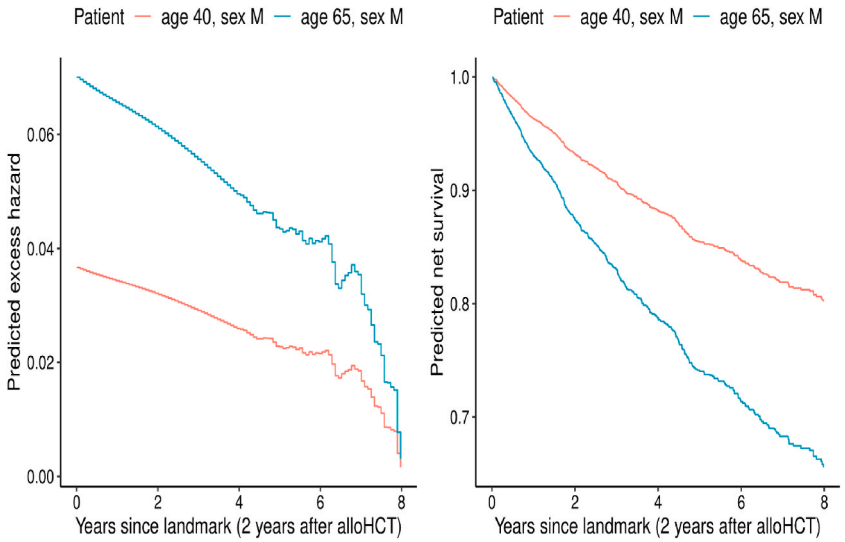

Fig. 5. Predicted excess hazard and net survival for two male patients aged 40 (red) and 65 (blue) based on the semi-parametric model provided in Table 1. When calculating the predicted excess hazard, an Epanechnikov kernel function based smoother has been used for smoothing the baseline excess hazard.

with subtle differences may be answered based on the data, thus elucidating the study topic from different viewpoints. The most commonly used measure in cancer survival is net survival, which describes the excess hazard due to the disease. Its independence from other-cause mortality makes it the measure of choice for comparing cohorts with different general population mortality. On the other hand, crude mortality gives insight into what one can expect in the real world - it depends on both disease-specific and other-cause hazard and describes the proportion of patients that die due to the disease in question. Both measures thus complement each other in describing the data. Historically, following the flawed assumption that the net survival and the relative survival ratio are the same quantities, inconsistent results have been obtained. The existing literature thus sometimes does not clearly inform the reader which measure has been used.

Relative survival is not a sensible approach when other-cause mortality is either negligible or predominant. In the first case, in situations where almost all deaths may be attributed to the studied disease (e.g. poor risk acute myeloid leukemia), the relative survival methods cannot provide any further understanding of the data since the overall survival probability and regression models on the overall hazard give very similar (if not identical) results as the relative survival methods. In the latter case, using relative survival may prove problematic. If most patients die due to other causes, very little information on the excess hazard is available in the data. This desired information is blurred by the high proportion of other-cause mortality, thus causing wide confidence intervals (that may be accompanied by unreasonably huge jumps in the estimated curves) and problems in regression model fitting even in large data sets. In particular, this occurs in cohorts comprising of old patients with high population mortality; but also in cohorts containing subsets of patients of very high age for whom the excess hazard after some time becomes impossible to study as they all die due to other causes. Forcing the analysis using such data may result in senseless results; one may remedy this problem by excluding the oldest patients from any long-term analysis.

Furthermore, there might also exist subsets of patients in the data that live longer than the general population, thus violating the key assumption of relative survival, i.e. that there exists an excess hazard (the overall hazard of the patients is higher than the population hazard). This may be the case in the long run, i.e. at later time points, due to a selection of fitter patients who survived the first perilous period, or due to increased health surveillance or a healthier lifestyle among (former) patients than in the general population. The described methodology is not appropriate in this case and unintuitive results can be obtained (i.e. negative hazards or increasing survival curves) if used nevertheless. Instead, one should rather turn to different models, for example, the multiplicative model (where the population hazard is multiplied by an excess term to obtain the overall hazard) [45] or individual relative survival [46].

Rather than focusing on the excess hazard, one could also evaluate the life expectancy of the cohort by comparing it to the general population, e.g. using the standardized mortality ratio (SMR) or absolute excess risk (AER). When considering long-term survival, the years lost/saved measure can also be used. A years difference measure may be defined in various ways, for example in the competing risks context [47] or by comparing the overall and population survival [48]. An overview of the years lost/saved measures and their estimation has been given in a recent paper [49].

To conclude, we believe that relative survival may often provide useful information in haematology studies as has been the case in the previous literature. Choosing the right measures has been an ongoing challenge in the practical setting and we hope that this paper will improve the way in which relative survival methodology is used.

Funding

This work was supported by the Slovene Research Agency ARRS (Methodology for data analysis in medical sciences, P3-0154, and project J3-1761).

Practice points

Analyzing cause of death is a common goal when considering the survival of patients with haematological malignancies.

Relative survival methods allow such analyses in the case when cause of death is not given in the data.

Research agenda

Provide an overview of the methods used in relative survival with haematological applications in mind.

Illustrate the use of relative survival based on a set of patients who received a first allogeneic hematopoietic stem cell transplantation for myelodysplastic syndromes or secondary acute myeloid leukemia.

Declaration of competing interest

None.

Acknowledgments

We thank the EBMT Data Office in Leiden for providing the MDS dataset, the EBMT Chronic Malignancies Working Party for their support of the original study and Johannes Schetelig (Medical Department I, University Hospital Carl Gustav Carus, Technical University Dresden/DKMS Clinical Trials Unit, Dresden) for expert input about this dataset and fruitful exchanges about relative survival in haematology.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.beha.2023.101474.

References

[1] Ellis L, Woods LM, Estève J, Eloranta S, Coleman MP, Rachet B. Cancer incidence, survival and mortality: explaining the concepts. Int J Cancer 2014;135(8):1774–82.
[2] Mariotto AB, Noone AM, Howlader N, Cho H, Keel GE, Garshell J, et al. Cancer survival: an overview of measures, uses, and interpretation. J Natl Cancer Inst Monogr 2014 Nov;2014(49):145–86.
[3] Begg CB, Schrag D. Attribution of deaths following cancer treatment. J Natl Cancer Inst 2002 Jul;94(14):1044–5.
[4] Percy C, Stanek E, Gloeckler L. Accuracy of cancer death certificates and its effect on cancer mortality statistics. Am J Publ Health 1981 Mar;71(3):242–50.
[5] Pulte D, Jansen L, Brenner H. Changes in long term survival after diagnosis with common hematologic malignancies in the early 21st century. Blood Cancer J 2020 May;10(5):1–8.
[6] Monnereau A, Troussard X, Belot A, Guizard AV, Woronoff AS, Bara S, et al. Unbiased estimates of long-term net survival of hematological malignancy patients detailed by major subtypes in France. Int J Cancer 2013 May;132(10):2378–87.
[7] Delgado J, Pereira A, Villamor N, Lopez-Guillermo A, Rozman C. Survival analysis in hematologic malignancies: recommendations for clinicians. Haematologica 2014 Sep;99(9):1410–20.
[8] Krok-Schoen JL, Fisher JL, Stephens JA, Mims A, Ayyappan S, Woyach JA, et al. Incidence and survival of hematological cancers among adults aged >=75 years. Cancer Med 2018 Apr.
[9] Desandes E, Lacour B, Belot A, Molinie F, Delafosse P, Tretarre B, et al. Cancer incidence and survival in adolescents and young adults in France, 2000-2008. Pediatr Hematol Oncol 2013 May;30(4):291–306.
[10] Ssenyonga N, Stiller C, Nakata K, Shalkow J, Redmond S, Bulliard JL, et al. Worldwide trends in population-based survival for children, adolescents, and young adults diagnosed with leukaemia, by subtype, during 2000-14 (CONCORD-3): analysis of individual data from 258 cancer registries in 61 countries. The Lancet Child & Adolescent Health 2022 Jun;6(6):409–31.
[11] Favier O, Heutte N, Stamatoullas-Bastard A, Carde P, van't Veer MB, Aleman BMP, et al. Survival after Hodgkin lymphoma: causes of death and excess mortality in patients treated in 8 consecutive trials. Cancer 2009 Apr;115(8):1680–91.
[12] de Vries S, Schaapveld M, Janus CPM, Daniels LA, Petersen EJ, van der Maazen RWM, et al. Long-term cause-specific mortality in hodgkin lymphoma patients. JNCI Journal of the National Cancer Institute 2020 Dec;113(6):760–9.
[13] Amaador K, Kersten MJ, Visser O, Posthuma EFM, Minnema MC, Vos JMI, et al. Conditional relative survival in Waldenstrom's macroglobulinaemia: a population-based study in The Netherlands. Br J Haematol 2022 Mar;196(5):1205–8.
[14] Eloranta S, Smedby KE, Dickman PW, Andersson TM. Cancer survival statistics for patients and healthcare professionals - a tutorial of real-world data analysis. J Intern Med 2021 Jan;289(1):12–28.
[15] Ludwig H, Bolejack V, Crowley J, Blade J, Miguel JS, Kyle RA, et al. Survival and years of life lost in different age cohorts of patients with multiple myeloma. J Clin Oncol: Official Journal of the American Society of Clinical Oncology 2010 Mar;28(9):1599–605.
[16] van Gelder M, de Wreede LC, Bornhauser M, Niederwieser D, Karas M, Anderson NS, et al. Long-term survival of patients with CLL after allogeneic transplantation: a report from the European Society for Blood and Marrow Transplantation. Bone Marrow Transplant 2017 Mar;52(3):372–80.
[17] Schetelig J, de Wreede LC, van Gelder M, Koster L, Finke J, Niederwieser D, et al. Late treatment-related mortality versus competing causes of death after allogeneic transplantation for myelodysplastic syndromes and secondary acute myeloid leukemia. Leukemia 2019 Mar;33(3):686–95.
[18] Manevski D, Putter H, Pohar Perme M, Bonneville EF, Schetelig J, de Wreede LC. Integrating relative survival in multi-state models-a non-parametric approach. Stat Methods Med Res 2022 Jun;31(6):997–1012. Available from: http://journals.sagepub.com/doi/10.1177/09622802221074156.
[19] Bhatia S, Francisco L, Carter A, Sun CL, Baker KS, Gurney JG, et al. Late mortality after allogeneic hematopoietic cell transplantation and functional status of long-term survivors: report from the Bone Marrow Transplant Survivor Study. Blood 2007 Nov;110(10):3784–92.
[20] Martin PJ, Counts GW, Appelbaum FR, Lee SJ, Sanders JE, Deeg HJ, et al. Life expectancy in patients surviving more than 5 years after hematopoietic cell transplantation. J Clin Oncol: Official Journal of the American Society of Clinical Oncology 2010 Feb;28(6):1011–6.
[21] Robin M, de Wreede LC, Wolschke C, Schetelig J, Eikema DJ, Van Lint MT, et al. Long-term outcome after allogeneic hematopoietic cell transplantation for myelofibrosis. Haematologica 2019 Sep;104(9):1782–8.
[22] Pohar Perme M, Stare J, Estève J. On estimation in relative survival. Biometrics 2012;68(1):113–20.
[23] Pokhrel A, Hakulinen T. How to interpret the relative survival ratios of cancer patients. Eur J Cancer 2008;44:2661–7.
[24] Hakulinen T, Seppä K, Lambert PC. Choosing the relative survival method for cancer survival estimation. Eur J Cancer 2011;47:2202–10.
[25] Allemani C, et al. Global surveillance of trends in cancer survival 2000-14 (CONCORD-3): analysis of individual records for 37513025 patients diagnosed with one of 18 cancers from 322 population-based registries in 71 countries. Lancet 2018 Mar;391(10125):1023–75. Available from: https://www.sciencedirect.com/science/article/pii/S0140673617333263.
[26] Pohar Perme M, Pavlič K. Nonparametric relative survival analysis with the R package relsurv. Journal of Statistical Software, Articles. 2018;87(8):1–27. Available from: https://www.jstatsoft.org/v087/i08.
[27] Pohar M, Stare J. Relative survival analysis in R. Comput Methods Progr Biomed 2006;81(3):272–8.
[28] Pohar M, Stare J. Making relative survival analysis relatively easy. Comput Biol Med 2007;37:1741–9.
[29] Human Mortality Database. University of California, berkeley (USA), and max planck institute for demographic research (Germany). Available at: www.mortality.orgorwww.humanmortality.de; 2021.
[30] Rubio FJ, Rachet B, Giorgi R, Maringe C, Belot A. On models for the estimation of the excess mortality hazard in case of insufficiently stratified life tables. Biostatistics 2019 May;22(1):51–67.
[31] Hinchliffe SR, Dickman PW, Lambert PC. Adjusting for the proportion of cancer deaths in the general population when using relative survival: a sensitivity analysis. Cancer Epidemiology 2012 Apr;36(2):148–52.
[32] Talback M, Dickman PW. Estimating expected survival probabilities for relative survival analysis–exploring the impact of including cancer patient mortality in the calculations. Eur J Cancer 2011 Nov;47(17):2626–32. Oxford, England: 1990.
[33] Dickman PW, Lambert PC, Coviello E, Rutherford MJ. Estimating net survival in population-based cancer studies, letter to editor. Int J Cancer 2013;133(2):519–21.
[34] Hakulinen T. Cancer survival corrected for heterogeneity in patient withdrawal. Biometrics 1982;38:933–42.
[35] Rebolj Kodre A, Pohar Perme M. Informative censoring in relative survival. Stat Med 2013;32:4791–802. https://doi.org/10.1002/sim.5877.
[36] Grafféo N, Castell F, Belot A, Giorgi R. A log-rank-type test to compare net survival distributions. Biometrics 2016;72(3):760–9. Available from: http://www.cmi.univ-mrs.fr/~castell/Publis/Logrank.pdf.
[37] Pavlič K, Pohar Perme M. On comparison of net survival curves. BMC Med Res Methodol 2017;17(1):79. https://doi.org/10.1186/s12874-017-0351-3.
[38] Pohar Perme M, Henderson R, Stare J. An approach to estimation in relative survival regression. Biostatistics 2009;10(1):136–46.
[39] Nelson C, Lambert PC, Squire IB, Jones DR. Flexible parametric models for relative survival, with application in coronary heart disease. Stat Med 2007;26:5486–98.

[40] Lambert PC, Royston P. Further development of flexible parametric models for survival analysis. STATA J 2009;9(2):265–90. Available from: https://ideas.repec.org/a/tsj/stataj/v9y2009i2p265-290.html.
[41] Charvat H, Belot A. Mexhaz: an R package for fitting flexible hazard-based regression models for overall and excess mortality with a random effect. J Stat Software 2021 Jul;98:1–36. https://doi.org/10.18637/jss.v098.i14.
[42] Klein JP, Andersen PK. Regression modeling of competing risks data based on pseudovalues of the cumulative incidence function. Biometrics 2005;61:223–9.
[43] The EBMT Patient Registry. Webpage, www.ebmt.org/ebmt-patient-registry; 2022 [Available from: https://www.ebmt.org/ebmt-patient-registry.
[44] Pohar Perme M, Manevski D. Relsurv: relative survival. R package version 2.2-8. Available from: https://cran.r-project.org/package=relsurv; 2022.
[45] Andersen PK, Borch-Johnsen K, Deckert T, Green A, Hougaard P, Keiding N, et al. A Cox regression model for relative mortality and its application to diabetes mellitus survival data. Biometrics 1985;41:921–32.
[46] Stare J, Henderson R, Pohar M. An individual measure of relative survival. J Roy Stat Soc C 2005;54(1):115–26.
[47] Andersen PK. Decomposition of number of life years lost according to causes of death. Stat Med 2013;32:5278–85.
[48] Andersen PK. Life years lost among patients with a given disease. Stat Med 2017;36(22):3573–82. Available from: https://onlinelibrary.wiley.com/doi/abs/10.1002/sim.7357.
[49] Manevski D, RužićGorenjec N, Andersen PK, Pohar Perme M. Expected life years compared to the general population. Biom J 2023.