What is ...?

For further titles in the series, visit: www.whatisseries.co.uk

What are hazard ratios?

Martin Duerden BMedSci DRCOG MRCGP DipTher DPH Deputy Medical Director, Betsi Cadwaladr University Health Board, North Wales; Senior Clinical Lecturer, Centre for Health Economics and Medicines Evaluation, Bangor University

- **Hazard ratios** are commonly used when presenting results in clinical trials involving survival data, and allow hypothesis testing. They **should not be considered the same as relative risk ratios.**

- When hazard ratios are used in survival analysis, this may have nothing to do with dying or prolonging life, but **reflects the analysis of time survived to an event** (the event may, in some instances, include cure).

- **A hazard is the rate at which events happen**, so that the probability of an event happening in a short time interval is the length of time multiplied by the hazard. Although the hazard may vary with time, the assumption in proportional hazard models for survival analysis is that **the hazard in one group is a constant proportion of the hazard in the other group. This proportion is the hazard ratio.**

- When expressing the results of clinical trials, **the hazard ratio conveys no information about the size of clinical effect and should be considered alongside a measure of time,** usually the median time to the event under scrutiny comparing active treatment and control groups (the points at which half the subjects have experienced the event in each arm of the study).

What is ...?

What are **hazard ratios?**

Defining a hazard ratio

The hazard ratio is an expression of the hazard or chance of events occurring in the treatment arm as a ratio of the hazard of the events occurring in the control arm. The term hazard ratio is often used interchangeably with the term relative risk ratio to describe results in clinical trials. This is not strictly correct as there are subtle and important differences. It is useful to understand the meaning of the term and also be able to identify when it is used appropriately. Hazard ratios are increasingly used to express effects in studies comparing treatments when statistics which describe time-to-event or survival analyses are used. In most recent trial publications these have largely replaced direct comparisons of number of events (or 'rates') after a specific point in time, or at the end of a study, seen in tests such as the t-test.

For the technically minded, the hazard is usually denoted by *h(t)* and is the probability that an individual who is under observation at a time t has an event at that time. It represents the instantaneous event rate for an individual who has already survived to time *t*.

Suppose that *k* patients have events in the period of follow-up at distinct times, $t_1 < t_2 < t_3 < t_4 < t_5 < \ldots < t_k$. As events are assumed to occur independently of one another, the probabilities of surviving from one interval to the next may be multiplied together to give the cumulative survival probability. The probability of being alive at time t_j, $S(t_j)$, is calculated from the probability of being alive at t_{j-1}, $S(t_{j-1})$, the number of patients alive just before t_j, n_j, and the number of events at t_j, d_j (Equation 1).[1] In this equation $t_0=0$ and $S(0)=1$.[1]

Equation 1.[1]

$$S(t_j) = S(t_{j-1}) \left(1 - \frac{d_j}{n_j}\right)$$

There is a clearly defined relationship between *S(t)* and *h(t)*, which is given by the following calculus formula (Equation 2).[1]

Equation 2.[1]

$$h(t) = -\frac{d}{dt}[\log S(t)].$$

The hazard *h(t)* can be used for further statistical analysis, nowadays nearly always using computers. The hazard ratio can be calculated to compare groups and, strictly speaking, is the effect on the hazard of differences or 'covariates' (for example, drug treatment or control), as estimated by regression models which treat the logarithm of the hazard rate as a function of a baseline hazard, $h_0(t)$. One method, the Cox model, is the most commonly used multivariate approach for analysing survival time data in medical research. It is based on an assumption that the hazards remain proportionately constant and it is more correctly called the Cox proportional hazards model. Mathematically, the Cox model is expressed by the following equation (Equation 3).[2]

Equation 3.[2]

$$h(t) = h_0(t) \times \exp\{b_1 x_1 + b_2 x_2 + \cdots + b_p x_p\}$$

In this equation, the hazard function *h(t)* is dependent on, or determined by, a set of *p* covariates *(x1, x2, ..., xp)*, whose impact is measured by the size of the respective coefficients *(b1, b2, ..., bp)*.

Hopefully the following dialogue will make these concepts more accessible for most of us, who have more rudimentary mathematical skills.

Distinction from relative risk

In contrast to the hazard ratio, the relative risk ratio is a measure of how many events have occurred in a study expressed as a ratio of the proportion of events occurring in the treatment group compared with that in the control group. It is usually calculated at the end of the study and is quoted as having occurred over the average or median duration of the trial. One pitfall in therapeutic trials is picking a point in time to express the relative risk ratio of an event. This can be misleading as it could be used to select the point in time at which there was greatest separation between the treatment and the comparator arms. It should only be calculated at the end of the clinical trial, and the point at which the trial ends or is halted should be prespecified (rather than chosen selectively after looking at the results!). Using survival data and hazard ratios goes some way to preventing this type of selectivity (Box 1).[3]

Box 1. Examples of when to use survival data

Blood pressure

In a trial comparing blood pressure reductions caused by two drugs, it is assumed that the changes in blood pressure of the subjects caused by the different drugs are normally distributed (this is 'the sample' from a population). Calculations to determine whether the differences between the interventions are statistically different (the probability of the difference having occurred by chance) are based on statistical methods which can be applied to continuous variables.

The mean of the blood pressure differences are calculated, and the variance (and standard deviation) or range of blood pressure changes can also be deduced. Using these measures a statistical test such as a Student's t-test or analysis of variance (ANOVA)[3] can be carried out to determine the probability of the differences observed having occurred by chance. Conventionally it is accepted that if this probability is less than 0.05 ($p<0.05$) then the differences are statistically significant and the null hypothesis can be rejected – the treatments are not the same.

Aspirin and mortality

In a trial designed to observe whether aspirin reduces mortality, patients who had sustained a myocardial infarction are randomised to aspirin or to placebo. After several years have elapsed, the number who have died in each treatment group is analysed and compared. The question to be answered here is whether there is a relationship between aspirin use and the risk of a patient dying, or whether the aspirin does not affect mortality (the null hypothesis). One way to determine this is using tests on categorical data (either the patient dies or does not).

In this example the Chi-squared test of association[3] can be used to determine whether to reject the null hypothesis of no association. The results show that the proportion of patients given aspirin who die is less than the proportion that dies when given placebo. If the Chi-squared test gives a p-value of <0.05, then it is unlikely that this result has occurred by chance.

Statins and cardiovascular events

In a trial examining whether a statin prevents a cardiovascular event in patients who have been admitted to hospital with unstable angina, patients are randomised to the statin or to placebo on admission. In this instance the focus of the study is examining the time between randomisation and a subsequent event. It is unlikely that these times are normally distributed. In this type of trial it is better, and possibly more ethical, if the study does not wait until events have occurred in all subjects. Also, some patients may leave the study early and become lost to follow-up, so that the only information available regarding these patients will be that they were still without a further event at the last follow-up.

In this instance, it is preferred to analyse the data using a survival method such as Kaplan–Meier analysis.[3] The basic idea is that the trial is split up into distinct time intervals. In each time interval the probability of 'surviving' that time interval without an event is calculated and these probabilities are multiplied to give the probability of 'survival' up to a given time point. Survival probability curves are plotted for those given the statin and those given placebo and the hazard ratio between these survival curves is calculated. The p-value for this hazard ratio is <0.05, so it is unlikely that this difference in time to an event has occurred by chance and, therefore, it is decided that statins do prevent and delay cardiovascular morbidity after admission for unstable angina.

NB In the second example it can be seen that time-to-event data could also have been used as in the third example. These days most studies of this nature are conducted this way. Analysing data like this provides the added benefit of collecting information that allows assessment not just of whether a treatment prevents events but also by how much the time an event is delayed by treatment.

What is ...?

What are **hazard ratios?**

Survival data are not just about survival

The term hazard ratio is commonly used in medical literature when describing survival data. It is important to realise that survival data are not just used to describe the number of people who survive or die over a period of time. These data are used in medical research and statistics to describe how many people can reach a certain point in time without experiencing a hazard or event, which may not be death (for example, having an acute myocardial infarction) – or, conversely, determining the number that do – and are a useful descriptor. In some clinical trials; for example, looking at antibiotic response, survival data might be used to observe events such as recovery or cure.

There are a number of other good reasons for using survival statistics. One reason is that time to an event is rarely normally distributed, which can make conventional parametric statistical methods difficult or inappropriate. A good example of this is the measurement of progression-free survival time (or 'disease-free survival') in trials of cancer drugs; here the majority of events may occur quite early, possibly within months, but a few subjects have a prolonged remission and may not have progression of disease for some time; for example, a year or more.

Survival and censoring

Survival data can also be used to analyse clinical trials in which there are a high proportion of dropouts, either because of adverse events or due to other reasons such as low retention or 'compliance' in the trial. Such dropouts can be the cause of misleading results, can introduce bias and can make it difficult to fully understand the data. Survival analysis allows this information to be incorporated by the technique of censoring. It is unknown whether the person who drops out has an event or not. Censoring assumes that the subjects who drop out have the same hazard of an event as those that remain in the study. Usually this is a reasonable assumption, but on rare occasions it can also be misleading.

How is a hazard ratio calculated?

A hazard ratio is calculated from hazard rates, a precise description of a hazard rate is the 'conditional instantaneous event rate calculated as a function of time'. To understand this it helps to look at an example. If a group of 1,000 patients are given a treatment and in *month one*, 20 die; then the hazard rate for *month one* is 20/1,000. If in *month two*, 20 die; the hazard rate for *month two* is 20/980 and so on. In this case the hazard rate is the number of patients dying divided by the number still alive at the start of that interval.

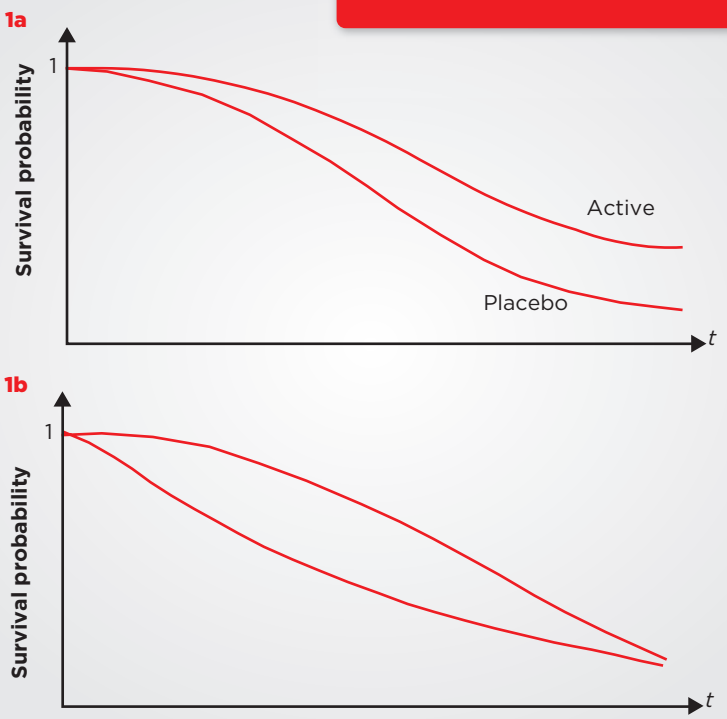
By looking at the hazard rate over small increments of time (giving an approximation of the instantaneous event rate) it is possible to compare the rate with the rate occurring in another group of patients given an alternative treatment, ideally within a randomised controlled trial. At different points in time the ratio of the hazard rates can be calculated. If the pattern of events is similar in each group it can be assumed that this ratio remains constant. Thus, the hazard ratio is the ratio of the hazard rates; that is, a ratio of the rate at which patients in the two groups are experiencing events. The log-rank test, which is often used for statistical analysis in these cases, tests the null hypothesis that this ratio is 1 (event hazard rates are the same).

To understand this further, as stated, a hazard ratio of 1 corresponds to equal treatments, a hazard ratio of 2 implies that at any time twice as many patients in the active group are having an event proportionately compared with the comparator group. A hazard ratio of 0.5 means that half as many patients in the active group have an event at any point in time compared with placebo, again proportionately.

Proportional hazards – not always the case

In many cases this assumption of 'proportional hazards' holds, but in some situations this may not be true. In Figure 1a, the assumption looks valid and the two hazard rates display the same basic attributes so that although the hazard rates are themselves not constant over time, a reasonable assumption would be that

Figure 1. Examples of hazard ratios plotted over time



their ratio is approximately constant. This does not follow in Figure 1b, which demonstrates a reason why the proportional hazard assumption can go astray: the short-term benefit of an active treatment does not maintain an effect in the longer term. For example, with some cancer treatments, such as interferon alfa in renal cell cancer, the effect of the active treatment is to create halting of tumour growth so that the event (progression of disease) in the active group is less than in the placebo group. However, after a period of time the event rate in the active group begins catch up with the event rate in the placebo group, as the disease escapes control.

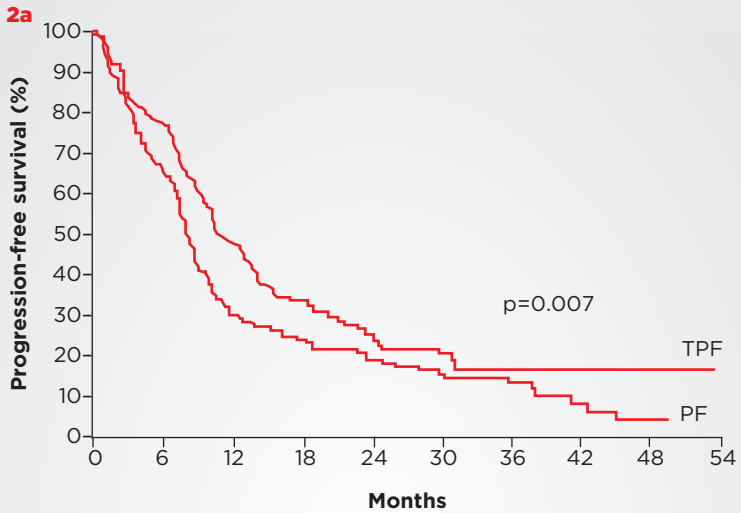
A hazard ratio of 2 – not twice as fast

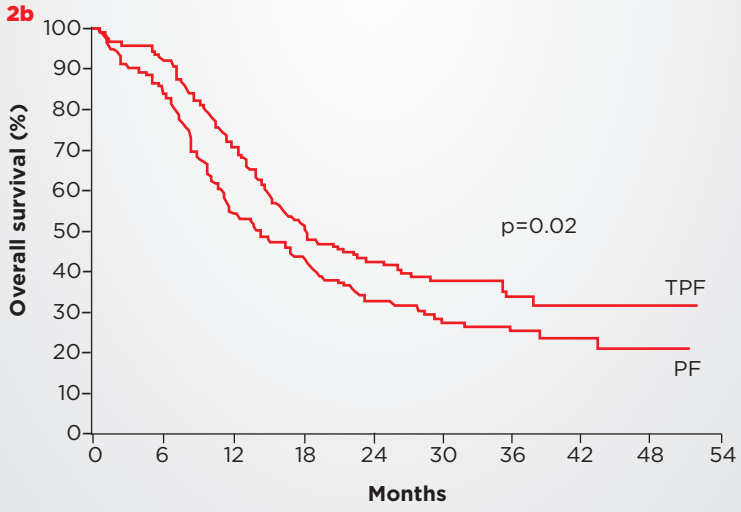
A hazard ratio of 2 could be misinterpreted by some as showing that patients in the placebo group progressed twice as fast as those in the control group. This is analogous to a relative risk of 2 doubling an event rate. Following this logic a misunderstanding would be to think that the median progression time was doubled by the treatment; that half as many patients were likely to have progressed by a particular day or that the treatment group was likely to have progressed half as quickly as the control group. This is a common pitfall and is incorrect as the hazard rates can only be inferred in a probabilistic sense (using statistics based on probabilities) from the occurrence of events in a population of at-risk individuals during a follow-up time interval. The correct interpretation is that a hazard ratio of 2 means that treatment will cause the patient to progress more quickly, and that a treated patient who has not yet progressed by a certain time has twice the chance of having progressed at the next point in time compared with someone in the control group.

In this example, the hazard ratio should be thought of as the odds that a patient will progress more slowly with treatment. It is a term that does not reflect a time unit of the study. This difference between hazard-based and time-based measures has been described as the distinction between the odds of winning a race and the margin of victory. This is why a hazard ratio should be regarded as the measure which allows

What are **hazard ratios?**

Figure 2. Effects of TPF and PF therapy on progression-free (a) and overall survival (b)[4]





PF: cisplatin and fluorouracil; TPF: docetaxel, cisplatin and fluorouracil

calculation for hypothesis testing, but it should be considered alongside a measure of time (and ideally while looking at survival curves on a graph) to describe the size of the treatment effect. In many survival analyses the best measure of time to consider is the median: the time at which 50% of participants will have experienced the event in question.

An example: hazard ratios in a study of head and neck cancer

An example of hazard ratios describing survival in a cancer study is shown in Figure 2 and Table 1.[4] In this case, survival analysis is used to describe true survival in people with advanced head and neck cancer (a term used to describe squamous cell cancer of the

throat, tongue, neck, sinus and so on). The safety and efficacy of types of 'induction chemotherapy' for patients with squamous cell carcinoma of the head and neck were evaluated, where induction chemotherapy is a treatment used in anticipation of radiotherapy. In this multicentre, randomised, Phase III, European study, 358 patients with previously untreated, unresectable, locally advanced stage III and IV tumours received either docetaxel, cisplatin and fluorouracil, or cisplatin and fluorouracil. Patients without further progression received radiotherapy within four to seven weeks. The primary endpoint, median *progression-free survival*, was significantly longer in the group receiving docetaxel, cisplatin and fluorouracil (11.0 months) than in the group receiving cisplatin and fluorouracil (8.2 months). The hazard ratio was 0.72 (95% confidence interval [CI] 0.57, 0.91; p=0.007). The median follow-up was 32.5 months.[4]

Figure 2b also shows a secondary endpoint, median *overall survival* (the point at which 50% of patients were still alive), which was significantly longer in the group receiving docetaxel, cisplatin and fluorouracil (18.8 months) than in the group receiving cisplatin and fluorouracil (14.5 months). The hazard ratio was 0.73 (95% CI 0.56, 0.94; p=0.02).

Table 1. PF and TPF in unresectable head and neck cancer[4]

Variable	PF (n=181)	TPF (n=177)	Hazard ratio (95% CI)	p-value
Progression-free survival				
Median duration – months	8.2	11.0	0.72 (0.57, 0.91)	0.007*
Rate – %				
At one year	31	48		
At two years	20	25		
At three years	14	17		
Overall survival				
Median duration – months	14.5	18.8	0.73 (0.56, 0.94)	0.02*
Rate – %				
At one year	55	72		
At two years	32	43		
At three years	26	37		

* The p-value was calculated with the use of an adjusted Cox proportional hazards model
CI: confidence interval; PF: cisplatin and fluorouracil; TPF: docetaxel, cisplatin and fluorouracil

Conclusion

In conclusion, hazard ratios are commonly used in survival analysis to allow hypothesis testing. They are similar to, but not the same as, relative risk ratios/reduction. As most clinical trials now use 'time to event' or survival methods of analysis, it is useful to be able to understand this distinction when reading these trials ●

References
1. Clark TG, Bradburn MJ, Love SB, Altman DG. Survival analysis part I: basic concepts and first analyses. *Br J Cancer* 2003; **89:** 232–238.
2. Bradburn MJ, Clark TG, Love SB, Altman DG. Survival analysis part II: Multivariate data analysis – an introduction to concepts and methods. *Br J Cancer* 2003; **89:** 431–436.
3. Altman D. *Practical Statistics for Medical Research.* London: Chapman & Hall, 1991.
4. Vermorken JB, Remenar E, van Herpen C *et al.* Cisplatin, fluorouracil, and docetaxel in unresectable head and neck cancer. *N Engl J Med* 2007; **357:** 1695–1704.

Further resources
1. Spruance SL, Reid JE, Grace M, Samore M. Hazard ratio in clinical trials. *Antimicrob Agents Chemother* 2004; **48:** 2787–2792.
2. Blagoev KB, Wilkerson J, Fojo T. Hazard ratios in cancer clinical trials – a primer. *Nat Rev Clin Oncol* 2012; **9**: 178–183.
3. Shaneyfelt T. Interpreting Hazard Ratios. University of Alabama at Birmingham, 3 May 2013. www.youtube.com/watch?v=z1b2hFzXsrU (last accessed 1 April 2014)
4. Sedgwick P. Hazard ratios. *BMJ* 2011; **343**: d5918.

What is ...?

What are **hazard ratios?**

First published 2009
Revised edition 2014
This publication, along with the others in the series, is available online at
www.whatisseries.co.uk

Published by Hayward Medical Communications, a division of Hayward Group Ltd, The Pines, Fordham Road, Newmarket CB8 7LG.
Tel: 01638 723560. Fax: 01638 723561.
email: hmc@hayward.co.uk

For reprints, permissions or to discuss training packages, email: **hmc@hayward.co.uk**