4 Survival analysis

Also known as 'Event History Analysis', 'Hazard Modelling' or 'Duration Analysis', survival analysis can be used to answer a wide range of research questions, which have a **time** or **duration** dimension.

The term survival analysis comes from biomedical studies where researchers were originally interested in studying mortality and the probability of individuals to survive past a given time point (for example the diagnosis of a disease or treatment for a disease).

More broadly, 'survival' is now used to indicate whether someone has experienced an event of interest by a given time point.

In survival analysis, we study the duration of time from when a participant enters a study (baseline) until the event occurs, the study ends, or the participant drops out of the study.

A measure of survival time needs to take into account:

- the baseline or time origin, for example, once the participants reach a specific age or when they receive a medical diagnosis;
- Which units you are measuring the duration in (e.g. seconds, months, years); and
- Whether or not someone experiences the event of interest before the end of the study period

Censoring

Within a study, for some participants the time of event will not be known because either: the analysis is carried out (or the study ends) before they have had the event, or the

participant drops out of the study before experiencing the event. This is called being 'censored'.

As we do have some information on censored participants; we know that their survival time is greater than their last observed follow-up time, we want to include them in our analyses.

In survival analysis we are interested in estimating the proportion of individuals experiencing a particular event by a specified time.

The risk of an individual experiencing the event at a given time point (assuming they hadn't already experienced it) is known as the **hazard rate**

The hazard rate may be different in different groups of people and may vary according to different exposures.

Hazard ratios can be used to compare hazard rates between two groups.

The **survivor rate** is the probability that the event **does not occur** before a time point.

The survivor rate at the start of the period is always 100%. The survival curve illustrates the probability of a participant 'surviving' at subsequent time points.

Lower survival rates indicate a greater probability of experiencing the event earlier. For example: the relative probability of 'survival' in employment by tenure in London compared to Glasgow. Hazard ratios can be used to assess the relative probability between two groups of experiencing an event.

Using appropriate statistical models, and similar to regression analysis, you can include multiple covariates.

To learn more about these analysis methods and how they can be performed in different software packages, check out our guidance and suggestions for further reading on the CLOSER website and Learning Hub.

Suggested citation: Kaye, N., Hardy, R., & O'Neill, D. (2021). *Analysing longitudinal data*. CLOSER Learning Hub, London, UK: CLOSER