Alternatives to Cox regression

In this course, I'm covering the two most commonly used survival analysis methods, Kaplan-Meier and Cox regression. These are popular largely because they are easy to run in standard software but mainly because you don't need to make any assumptions about the shape of the hazard function, the specific way that risk changes over time. There are, however, a number of other important types of methods for analysing such data. The example you've been using in this course concerns trying to predict mortality following admission for heart failure. Cox doesn't care about the distribution of survival times or what the hazard function looks like. This is why it's called "semi-parametric": it has some parameters – those of the predictors – but it has no parameters to describe the hazard function for patients with a value of zero for the predictors (i.e., patients with age zero and all the reference categories for the categorical variables). (For completeness, the simple proportion alive at a given time point and the Kaplan-Meier estimate are examples of non-parametric survival analysis.)

However, making assumptions about the **shape** of the hazard function – adding parameters to the model to describe the shape, making the model "fully parametric" – can lead to better prediction. More accurate prediction of a patient's survival time or risk of death within a given timeframe is vital for enabling the patient and his or her doctor and clinical team to make decisions regarding treatment. Risk models can put patients into, for example, low-, medium- or high-risk in a process called risk stratification, and high-risk patients can be offered different treatment plans from low-risk ones. It may be that you can do better than the Cox model in terms of risk prediction for a given data set and patient outcome. The Cox model was developed to look at the effect of covariates on the hazard function rather than to estimate survival times. A fully parametric can help here, especially if the Cox model assumptions are violated (more on the Cox assumptions later in the course).

There are **several such fully parametric models** such as Weibull, exponential, log-normal, and log-logistic models, where hazard function has to be specified. The **Weibull** distribution is used widely in medicine because of its flexibility: its hazard function can be increasing, decreasing, or constant over time. A special case of it is an exponential distribution, which is simple because it has only one parameter. This is because the hazard function is constant when the survival time is exponentially distributed. If you want a hazard that increases and then decreases over time, try either the log-logistic or the log-normal.

There are further extensions to the basic survival analysis approach, such as allowing for the fact that the values of some predictors change over time (Cox can deal with this) and handling multiple events (patient outcomes) in the same model. This is useful for disease recurrence, for example. Also, like so many statistical methods, survival analysis can be run in a **Bayesian** framework. Bayesian analysis is in general more complicated but very powerful. In essence, it involves mixing your data and your prior beliefs about what is related to what and deriving probabilities that something is true. What I'm teaching on this course and throughout this series of courses within the specialisation is called classical or “frequentist” statistics. In the classical framework, there’s no formal use of prior knowledge in the underlying maths. The answer you get is completely driven by the data. There’s a philosophical as well as a mathematical difference between the two, and much has been written about it – it’s a huge subject, way too big to go into here. It’s often claimed that there are two rival camps, but (happily) it’s also often claimed that there are no such camps and that many people, including me, use both methods (which often give similar results anyway in practice).