Survival analysis

Survival data

Think of this as a ‘survival time to an event’ – the event being your endpoint/outcome.

Case study:

<https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0186284>

Background:

* 107 patients with a primary lung adenocarcinoma staged between I to III between 2005-2011
* Outcome of interest – recurrence event (
* CEA, NSE, CA125, HE4, CA19-9 (tumour markers) collected from blood tests within 1 week before the surgery (time of collection irrelevant to us)
  + We can use the status of these biomarkers (i.e. if they are elevated) to indicate if they affect the outcome

Methods of the study:

1. “Patients diagnosed with a suspected second primary lung cancer (of any histology) were censored at the time the suspected second lung cancer was diagnosed (n=7)”
   1. “… since it would not be possible to distinguish if eventual later recurrences originated from the first or second lung cancer”
2. When medical records revealed uncertainties regarding potential metastases or if a metastasis might be derived from another primary tumour the cases were censored at the time of the event (n=5)
3. Disease free survival was estimated by the Kaplan-Meier curves and two-sided log-rank tests or log-rank trend test were used to compare survival curves
4. Cox regression was used to estimate the effects of prognostic variables on Disease free survival

Making more sense out of this:

1. An individual needs to be censored if:
   1. Suspected second primary lung cancer (method (1)) – they technically haven’t had the primary outcome (recurrence)
   2. Uncertain metastases (method (2)) – they MAY NOT have the primary outcome (recurrence), but we can’t be sure they’ve had it
   3. Patients who haven’t had the primary outcome (recurrence)
   4. Patients who were lost to follow-up or withdrew (no primary outcome of recurrence)

![A picture containing chart

Description automatically generated]()