

## Dataset and variables

Partial results from a trial of laser coagulation for the treatment of diabetic retinopathy. (dataset: **diabetic.xlsx**).

This dataset can also be found in the `survival` package in R

```
data(diabetic)
View(diabetic)
```

A data frame with 394 observations on the following 8 variables:

`id`: subject id

`laser`: laser type: xenon or argon

`age`: age at diagnosis

`eye`: a factor with levels of left right

`trt`: treatment: 0 = no treatment, 1= laser

`risk`: risk group of 6-12

`time`: time to event or last follow-up

`status`: status of 0= censored or 1 = visual loss

### Description:

The 197 patients in this dataset were a 50% random sample of the patients with "high-risk" diabetic retinopathy as defined by the Diabetic Retinopathy Study (DRS). Each patient had one eye randomized to laser treatment and the other eye received no treatment. For each eye, the event of interest was the time from initiation of treatment to the time when visual acuity dropped below 5/200 two visits in a row. Thus, there is a built-in lag time of approximately 6 months (visits were every 3 months). Survival times in this dataset are therefore the actual time to blindness in months, minus the minimum possible time to event (6.5 months). Censoring was caused by death, dropout, or end of the study.

## Questions

- Run a univariate Cox proportional hazards regression model against variable `trt`. How do you interpret the result?
- Plot the observed vs predicted for `trt` model
- Run the model diagnostics for `trt` and `age`.
- Run a multiple Cox proportional hazards regression model against all explanatory variables with  $p < 0.2$  in the univariable analysis. Interpret the results.
- Run a Weibull model against `trt`. Interpret the results.
- Compare graphically the survival curves predicted by the Weibull model with the KM estimates.