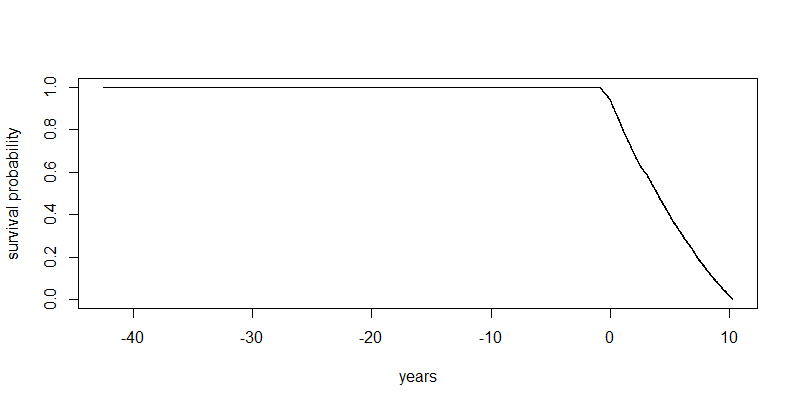
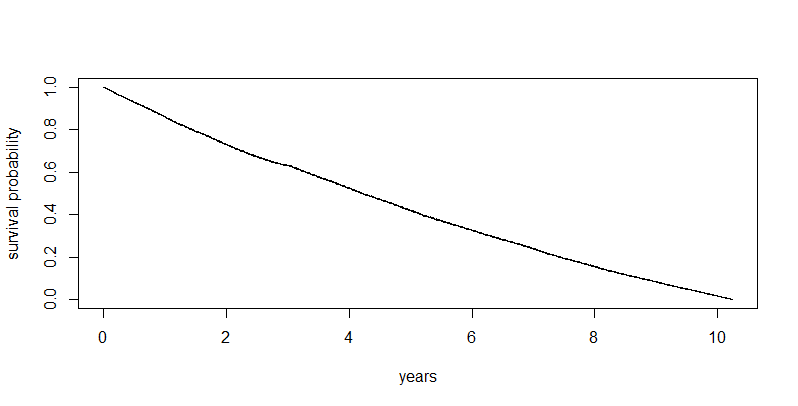
**Kaplan-Meier Plots**

* This is the plot produced before filtering the diagnosis cohort to only include patients with a time to censor greater than zero:
  + 0=event, 1=censored
  + KM <- survfit(Surv(time\_to\_censor, censor\_var) ~ 1, data=cohort\_diag)
  + Plot(KM)



* Same plot after filtering the time to censor variable:



* In the follwoing graph I have coded 0=censored, 1=event:
  + The rest of the graph is the same as above:

A line graph with numbers and a line

AI-generated content may be incorrect.

* Summary of the KM results (obviously because half of patients in the cohort did not get retinopathy there is no median value):

A black text on a white background

AI-generated content may be incorrect.

**Cox proportional hazards model**

* This was the first, full model with all covariates recommended in the meeting:

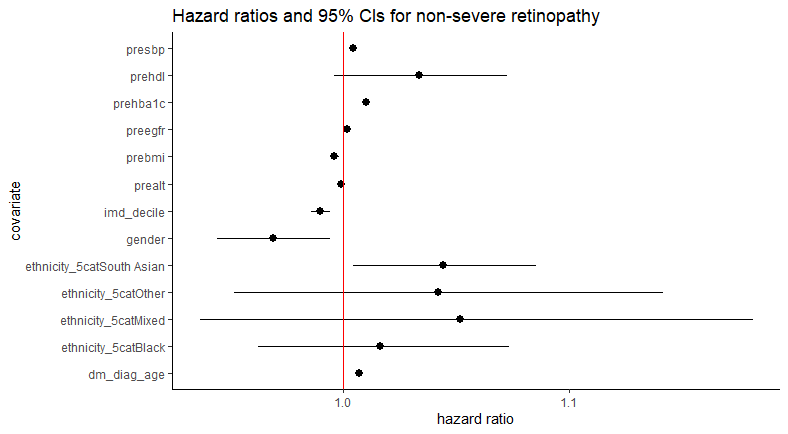
A screenshot of a computer

AI-generated content may be incorrect.

A screenshot of a computer

AI-generated content may be incorrect.

* Exponentiated coefficients (hazard ratios):
  + Age at diagnosis
    - For each additional year, there is 0.70% increase in hazard
  + Gender
    - Females have a 3.12% decrease in hazard compared to males
  + Ethnicity
    - Only South Asian was significant
    - In comparison to White ethnicity, South Asian patients had a 4.42% increase in hazard
  + IMD
    - For every one decile increase, there was 1.02% decrease in hazard
  + HbA1C
    - For every one mmol/mol increase, there was a 1.03% increase in hazard
  + BMI
    - For every one kg/m2 increase, there was a 0.38% decrease in hazard
  + EGFR
    - For every ml/min/1.73m2 increase, there was 0.16% increase in hazard
  + HDL
    - Not significant, but for every mmol/L increase there was 3.36% increase in hazard
  + ALT
    - For every U/L increase, there was a 0.09% decrease in hazard
  + SBP
    - For every mmHg increase, there was a 0.42% increase in hazard
* Plot of tidy table of Cox proportional model output



**Stepwise regression on the Cox Proportional Model**

* This was slightly tricky at first kept getting an error
* So ended up having to make a complete cases data frame which worked
* Results
  + initial AIC = 633878.6
  + Step 1:
  + Ethnicity
    - Taking out ethnicity reduced the AIC score to 633876
    - so made the model better
  + other variables
    - Taking out all the other variables increased the AIC score
    - So made the model worse
  + Ethnicity was then taken out
  + Step 2:
    - Continues checking the impact of taking out and putting back in variables
    - No changes are necessary
    - So the best model is one with ethnicity removed
* Stepwise model process table:

A screenshot of a computer program

AI-generated content may be incorrect.

A screenshot of a computer program

AI-generated content may be incorrect.

* The next set of results shows a summary of the Cox Proportional Hazard model without the ethnicity variable (i.e., the best model)
  + Most of the exponentiated coefficients are largely the same

A screenshot of a computer program

AI-generated content may be incorrect.

A screenshot of a computer program

AI-generated content may be incorrect.