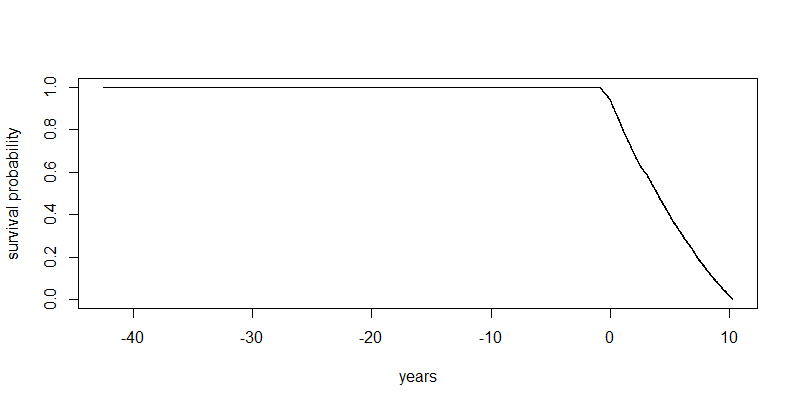
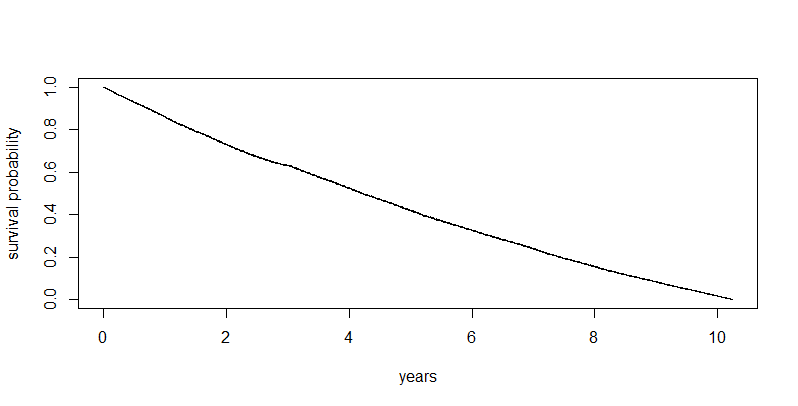
**Survival analysis**

* 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.

**Covariate details**

* Age at diagnosis
  + Self explanatory coding and units
  + <https://diabetesjournals.org/care/article/31/10/1985/25135/Timing-Is-Everything-Age-of-Onset-Influences-Long>
    - According to research there’s an inherent susceptibility to diabetic retinopathy with earlier-onset T2 diabetes
* Gender
  + 1 = male, 2 = female
  + <https://jech.bmj.com/content/78/8/479.abstract>
  + <https://onlinelibrary.wiley.com/doi/full/10.1155/2023/5919468>
    - Men with diabetes tend to be at a greater risk of complications (incl. retinopathy) irrespective of diabetes duration
* Ethnicity (5 category)
  + 0 = White, 1 = South Asian, 2 = Black, 3 = Other, 4 = Mixed
  + <https://www.sciencedirect.com/science/article/pii/S0039625712000070>
  + <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0032182>
    - The prevalence of DR is higher in people of South Asian descent and often also other minority ethnic groups compared to White people, also possible disparities in screening uptake
* IMD
  + Deciles
    - 1 = least deprived, 10 = most deprived
  + Quintiles
    - 1 = least deprived (deciles 1 and 2), 5 = most deprived (deciles 9 and 10)
  + <https://www.tandfonline.com/doi/full/10.2147/OPTH.S73272>
    - Deprivation can be a risk factor for late presentation of proliferative diabetic retinopathy
    - T2 is also more prevalent overall in
    - This paper also found that South Asian ethnicity was a risk factor for proliferative retinopathy, but wasn’t significant after controlling for HbA1c
  + <https://diabetesjournals.org/care/article/31/11/2131/26810/Screening-Uptake-in-a-Well-Established-Diabetic>
    - Again possibly caused by screening uptake disparities – higher deprivation is associated with poorer attendance at retinal screening
* Baseline HbA1C
  + mmol/mol
  + average blood glucose over the past 90 days
  + HbA1c over 48mmol/mol is diagnostic of T2
  + <https://link.springer.com/article/10.1007/s40200-022-00986-5>
  + <https://d1wqtxts1xzle7.cloudfront.net/100096241/14149-libre.pdf?1679340900=&response-content-disposition=inline%3B+filename%3DAssociation_of_HbA1c_levels_with_diabeti.pdf&Expires=1756898097&Signature=d7Nnu0KdgfEKbfupTlAMzOUeWwd7Tyrzrx~ko7KVZh17h0nWw5kPb4wrz0hlek0nhGiv1VnwuVXeWeaFehG6mddBJsZ9ay01F-5Vejff3~S9~qSMwnrFx4qj0ZC23EHHdMWBkxOxSOJwGTI6B5xH3tjQ2hwxe4IWIq3BJ7-~xtoLcbgPvGR5ziabAMfWMrzlt5EWPCxOn8RSWv750WjDnkcGjGj1CsQx9cjZYtQM5g0AdfuP-q7Qml6P~VcmCy5BVUpO7Daf~ugA0Gvzmn~F1tvGVB-TAkQEQMPKPi8rTkbOeoBw8krmIascms0WeF7YoikXjMU9BDY5lF0oeKdbEg__&Key-Pair-Id=APKAJLOHF5GGSLRBV4ZA>
    - Poorly controlled HbA1c is associated with DR – higher levels of HbA1c are linked to DR
* Baseline BMI
  + kg/m2
  + A function of height and weight
  + Higher BMI is a strong risk factor for T2
  + <https://ec.bioscientifica.com/view/journals/ec/10/7/EC-21-0172.xml>
    - Generally increased BMI/obesity is associated with increased prevalence of DR
  + <https://link.springer.com/article/10.1007/s40200-018-0339-5>
    - Although there is a slight debate over this, such as one paper finding an inverse relationship
  + <https://www.frontiersin.org/journals/endocrinology/articles/10.3389/fendo.2023.1108731/full>
    - However Mendelian randomization has found genetically predicted BMI is associated with an increased risk of DR
* Baseline EGFR
  + ml/min/1.73m2
  + Estimated Glomerular Filtration Rate – kidney function
  + Higher than 90ml/min/1.73m2 is generally good
  + Lower numbers are bad, and are sometimes associated with diabetes
  + <https://link.springer.com/article/10.1007/s11255-025-04547-6>
    - Systematic review and meta analysis found a significant association between reduced eGFR and DR
      * Seems to be a relatively robust finding
* Baseline HDL
  + mmol/L
  + high density lipoproteins – “good” cholesterol
  + Low levels of HDL cholesterol are consistently associated with increased risk of T2
  + <https://www.sciencedirect.com/science/article/pii/S0168822719301482>
    - One novel study found a relationship between *high* HDL and DR
    - Previously most studies show a relationship between low HDL and DR
  + <https://link.springer.com/article/10.1186/s12902-024-01599-0>
    - Also another study found an inverted U shape relationship
  + <https://diabetesjournals.org/care/article/35/11/2201/30665/Low-HDL-Cholesterol-and-the-Risk-of-Diabetic>
    - And another found that HDL was not an independent risk factor for retinal microvascular disease
  + Basically the relationship is somewhat unclear
* Baseline ALT
  + U/L
  + Alanine aminotransferase – enzyme in the liver
  + Increased levels can indicate liver cell damage
  + Data indicates moderate associations between ALT and risk of T2
  + <https://www.tandfonline.com/doi/full/10.2147/DMSO.S502129>
    - Has been found to not be an independent risk factor for DR
  + <https://www.nature.com/articles/s41598-022-13164-7>
    - Some studies have found an association between abnormal renal and liver function and diabetic retinopathy
  + <https://onlinelibrary.wiley.com/doi/full/10.1111/jdi.13489>
    - But non-alcoholic fatty liver disease was found to not have an association with DR by a meta-analysis
  + Somewhat unclear relationship
* Baseline SBP
  + mmHg
  + Systolic blood pressure
  + Elevated SBP is associated with future onset of T2
  + <https://onlinelibrary.wiley.com/doi/full/10.1155/2022/7876786>
    - Elevated SBP is pretty robustly associated with DR

**Cox proportional hazards model**

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

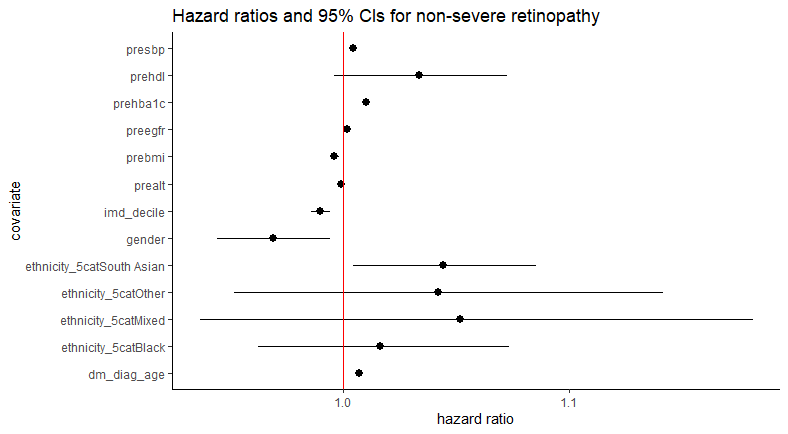
A screenshot of a computer

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.

**Results with IMD quintile and gender as categorical variables**

* The following results/plot were from when the IMD quintile variable was coded as ordered = TRUE, hence why it has analyzed whether there is a linear, quadratic, cubic etc. relationship:

A screenshot of a computer

AI-generated content may be incorrect.

A graph of a hazard ratio

AI-generated content may be incorrect.

* The following results/plot are from when the IMD quintile variable was coded as ordered = FALSE, hence why now it is treating the categorical variable usually with 1 being the reference category:

A graph of a hazard ratio

AI-generated content may be incorrect.

A screenshot of a computer

AI-generated content may be incorrect.

**Stepwise model selection results**

* Again using sequential selection, the model thought to be the best was the one that excluded ethnicity:

A screenshot of a computer program

AI-generated content may be incorrect.