Cf. "Rosuvastatin to Prevent Vascular Events in Men and Women with Elevated C-Reactive Protein" by Ridker, Danielson, Fonseca, Genest, et al. n engl j med 359;21 www.nejm.org november 20, 2008 2195

The website `www.biostat.mcgill.ca/hanley/c634` has a dataset that JH has created to reflect as closely as possible the two cumulative incidence curves, and the censoring pattern, in panel A of Figure 1 [Cumulative Incidence of Cardiovascular Events According to Study Group: cumulative incidence of the primary end point (nonfatal myocardial infarction, nonfatal stroke, arterial revascularization, hospitalization for unstable angina, or confirmed death from cardiovascular causes]. For each subject (row), in addition to an arbitrary serial number (ordered from first in to last in) it contains three data items: `tx` = 0/1 for Placebo/Rosuvastatin; `last.day` = minimum of (a) number of days until the subject had one of the events included in the primary endpoint, (b) number of days the subject had been followed from entry until the March 30, 2008 cutoff for the analysis and was still event-free at that time, (c) number of days until lost to follow-up or to a non-cardiovascular death, but still primary-event-free, at some date before that cutoff; `terminated.by.event` = 1 if an event included in the primary endpoint occurred on that last day and = 0 if not (i.e., if event-free at that last day). The website also has (annotated) `R` code that manipulates this file.

1 K-M, CI, NNT, etc

i. Obtain and plot the two K-M-based cumulative incidence curves on a single graph. (they are easier to read, and make more use of the white space, than the (decreasing from 1) "survival" curves.[1]).

ii. Obtain and report the upper and lower 95% limits for the 2-year cumulative incidence for each arm (limits are among the `R` objects returned by `survfit`).

iii. Calculate the NNTs based on the risk differences at 1, 2 and 4 years. Can you think of how to get confidence intervals for these NNTs? [it's the same challenge as getting an interval to accompany the difference of the two point estimates in part (ii)]

[1] see Pocock SJ, Clayton TC, Altman DG. Survival plots of time-to-event outcomes in clinical trials: good practice and pitfalls. The Lancet. 2002; 359(9318): 168689.

2 Event rates, rate ratios and rate differences, model-based CIs, etc

i. Obtain point estimates for the two event rates, their difference, and their ratio (these are "average" rates, averaged over the observed p-t distribution over the 4.5 years, and thus weighted towards the earlier years).

ii. Use these rates (incidence densities), and the relationship between incidence and cumulative incidence, to arrive at model-based cumulative incidence estimates at 1, 2, 4 and 5 years (5 is an extrapolation beyond the longest observed duration). Compare the NNTs derived from these with those based on the "observed" (K-M-based, and thus 'model-free' or 'non-parametric') cumulative incidence estimates.

3 Year-specific event rates, rate ratios and rate differences, and summaries of these

i. The `R` code allows you to create a file that "slices" each subject's follow-up into the amounts spent in each 1-year window, and from it to calculate year- and arm-specific numbers of events and person-years of observation. [If using some other software,[2] and unable to use it to create these time-slice data, you can use the file (on the website) created by the `R` code]. From these 10 cell-specific statistics, calculate year- and arm-specific event rates, and year-specific rate ratios and rate differences.

ii. How similar are the five rate ratios? rate differences? Use the formulae in chapter 8 of Rothman 2002 to obtain a summary ratio and a summary difference between arms. Do the calculation by your statistical package rather than by a calculator (cf. annotations in `R` code). *Rate ratios that don't differ (are homogeneous) over time are referred to as proportional rates (or if the time slices are a bit narrower, proportional hazards).*

iii. Do the year by year rates within the placebo arm show any general pattern? (again, see the annotations and remarks in the R code)

iv. (For placebo group) How different are the 5-year risks (cumulative incidence) when calculated using the same "average" rate, over the 5 years, vs. using the 5 yearly rates?

[2] The module `stsplit` in `Stata` can do this. In `SAS`, use one `output` statement per slice.