The `cd4` data base contains information about CD4 (T-cells) counts on 369 HIV positive men. These subjects were followed as part of a large cohort from approximatively 3 years before seroconversion to roughly 5 years after seroconversion. Observations are taken every six months on average for a total of 2376 observations (average of 6.5 measures per patient). The dataset contains the following variables:

- `age`: age at seroconversion
- `packs`: smoking (packs/day)
- `rdu`: recreational drug use (yes/no)
- `snp`: number of sexual partners (centered?)
- `CESD`: depression - CESD scale
- `tss`: time since seroconversion
- `cd`: CD4 counts
- `id`: subject ID

The goals of the study were:

1. Estimate average pattern of CD4 decline.
2. Estimate pattern of CD4 decline for individuals.
3. Identify factors which predict CD4 changes.
4. Understand heterogeneity of outcome across men.

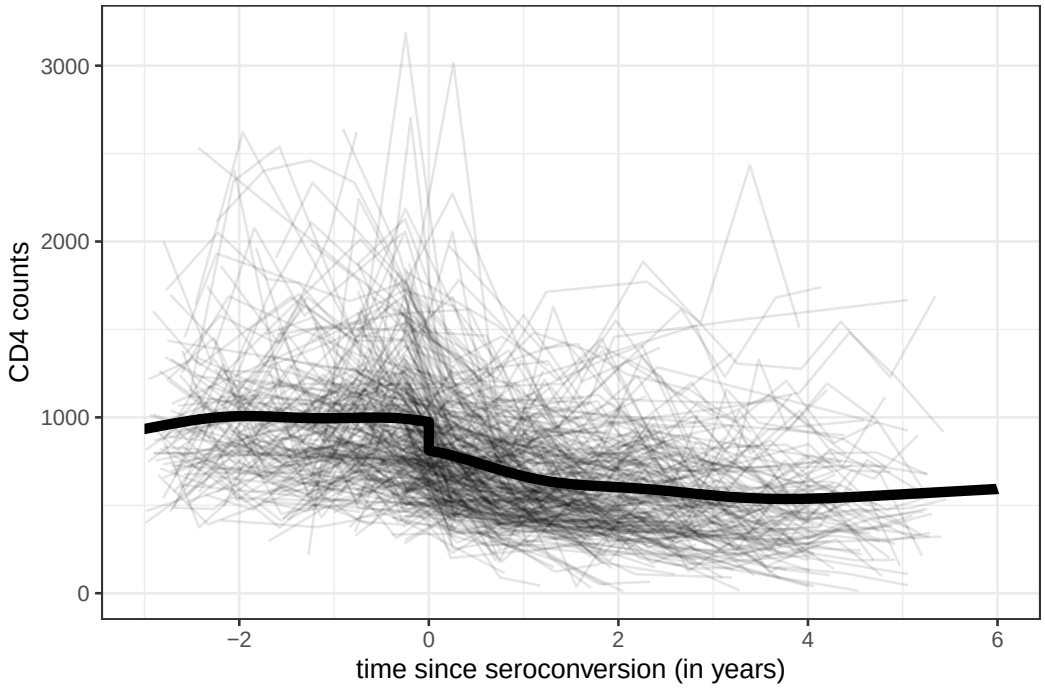

Figure 1: Spaghetti plot of CD4 trajectories.

Figure 1 shows the time evolution of the individuals trajectory with a smoothed mean estimator. In the context of the workshop, we will fit linear mixed models with as sole fixed effect various functions of time, with random intercepts for individuals and a first-order autoregressive structure for the errors.

1. How could we model the following changes in CD4 counts over time? Create the auxiliary variables for the mean model.
   (a) CD4 is stable, drops at seroconversion, then is stable again
   (b) CD4 changes linearly with time

Compiled 18/11/2020 at 15:05

   (c) CD4 changes linearly, drops at seroconversion then declines linearly again (with the same slope before and after).
   (d) CD4 changes linearly, drops at seroconversion then declines linearly again (different slopes)
   (e) CD4 changes linearly until seroconversion, then declines more steeply after
   (f) CD4 is stable, jumps at seroconversion, then drops linearly after seroconversion
   (g) CD4 is stable, then drops linearly after seroconversion (no jump)
   (h) CD4 is stable, then drops nonlinearly (quadratically) after seroconversion with a jump at seroconversion
   (i) CD4 is stable, then drops nonlinearly (quadratically) after seroconversion with no jump at seroconversion.
2. Fit the various representations of time
3. Write down the mean model equation before, at, and after seroconversion.
4. Determine which model fits best evolution of CD4 counts using information criteria.
5. Using the best fitting model, does the effect of time vary across subject?
6. Compare the regression coefficients for time from the model with random intercepts and slopes to that with random intercept and serial correlation.
7. Is the serial correlation needed in the model with random intercepts and random slopes?
8. Is there evidence that smokers have a different change over time?

Compiled 18/11/2020 at 15:05