Report_01

Jan van Rongen

2021-10-07

Introduction

We analyse data from the State of Israel related to COVID vaccines and infections. This is c continuation on the technical documents that contain all code used. They were a reanalysis of a document circulating on Internet, but this document focusses on proper modelling.

The document itself is an RMarkdown worksheet that, when formatted in pdf, hides the code.

Data sources and cleaning

As before We use four data sources: three from the Israeli government site and one from a spreadsheet with population data. The data was imported manually on 2021-10-02.

The file `vaccinated-per-day-2021-09-28` is aggregated to a per week file. Numeric fields with `<5` or `<15` were converted to 3 and 8 resp. Weeks are identified by their first day. Four missing records for the 90+ age category were added to `cases-among-vaccinated-134.csv`

Using the `population` table we construct cumulative totals of fulle vaccinated, single vaccinated and not vaccinated.

One file is not yet used.

Goal and methods.

To analyse as best as possible two groups: fully vaccinated vs non vaccinated on rate of infection. Previous analysis (tech_doc_3) shows that a detailed breakdown in age groups does not help. Adults were vacinated immediately, under 20 much later. So we can restrict aourselves to these two groups.

Fully versus unvaccinated.

Fully vaccinated in week x: those who received exactly two doses and their last dose in or before week x-2. It is highly unlikely that they were infected fwice while fully vaccinated so we assume that did not happen.

Unvaccinated: again here we assume that we measure the first infection.

Disease prior to start of 2020-12-20 vaccination: we assume that the population that already had at least one infection is $W = 0.15$ of the total population. That should be a parameter that can be changed,

State TRansition diagram

We draw a State Transition Diagram of all possible states a person could go through, one of the many possibilitie for such a diagram. We have to make choices because nmost possible diagrams have transitions that we cannot calculate. This one might work. .

state transition diagram
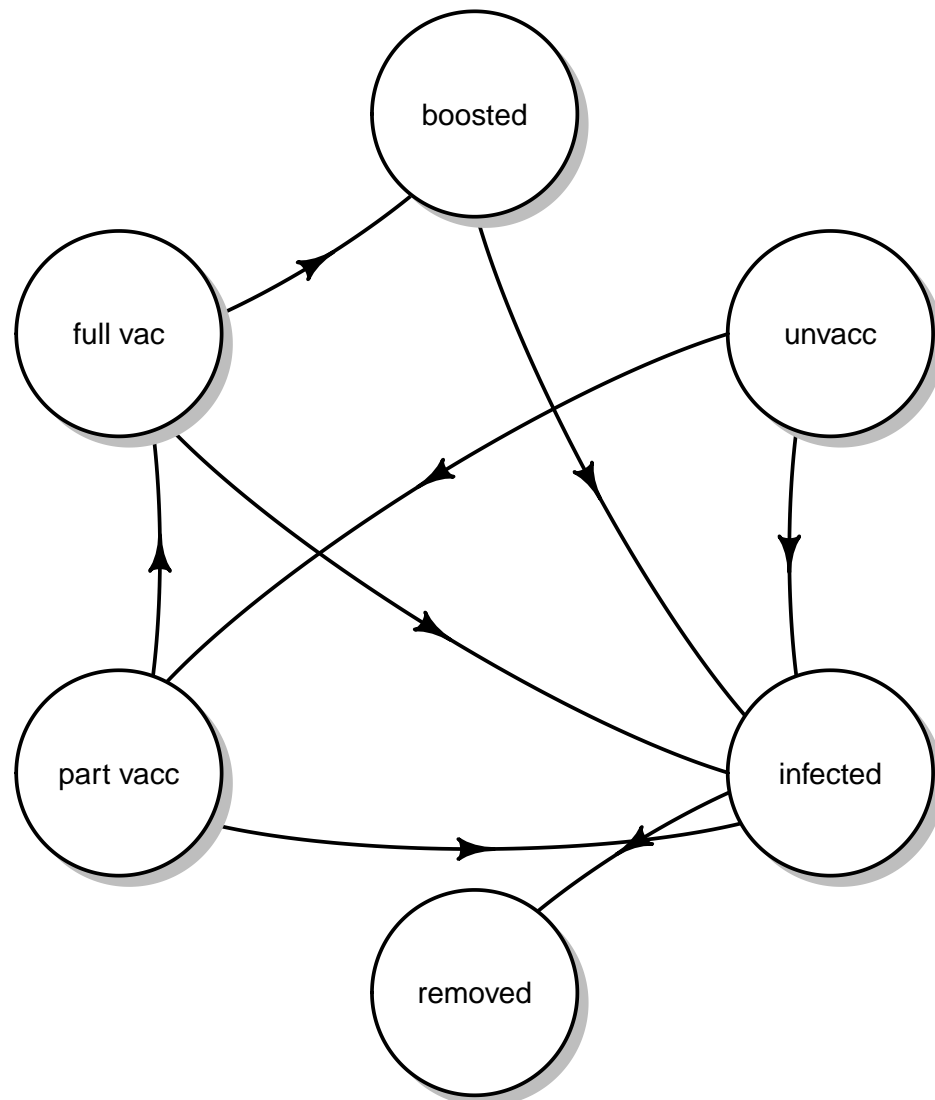
We do not know how many infected people are removed from the system, but in clinical trials in survival analysis people that reach an end state *are* removed from the system. So we will do that. Of course this is an approximation we use as there is nothing better.

Calculating Susceptible population.

From this diagram we can derive how to calculate the susceptible populations for the unvacced and fully vacced groups. We know how to calculate them from the weekly vaccinations, but they have to loose their infected members at the proper moment in time.

Algorithm

The assumption is that the effectiveness of the vaccin decreases over time. From the SIR model we can learn that for unvaccinated population S_n, the amount of new incidents is a linear function of the effective $R_e(t)$. A vaccine with initial effectveness β will have decaying effectiveness with a rate of $(\beta.e^{-\alpha.t})$ so the remaining effectiveness is one minus that. Taking the group of fully vaccinated people (wwithout booster) we model the risk ratio $RR_{f/n} = 1 - \beta.e^{-\alpha.t}$

This is in fact an exponential regression of $RR_{f/n} - 1$ against the function $-\beta.e^{-\alpha.t}$.

To get a truly least squares estimate (so that the errors heve zero sum) we use the R function `nls` from the base library and we estimate the confidence interval using the `invertR`package.
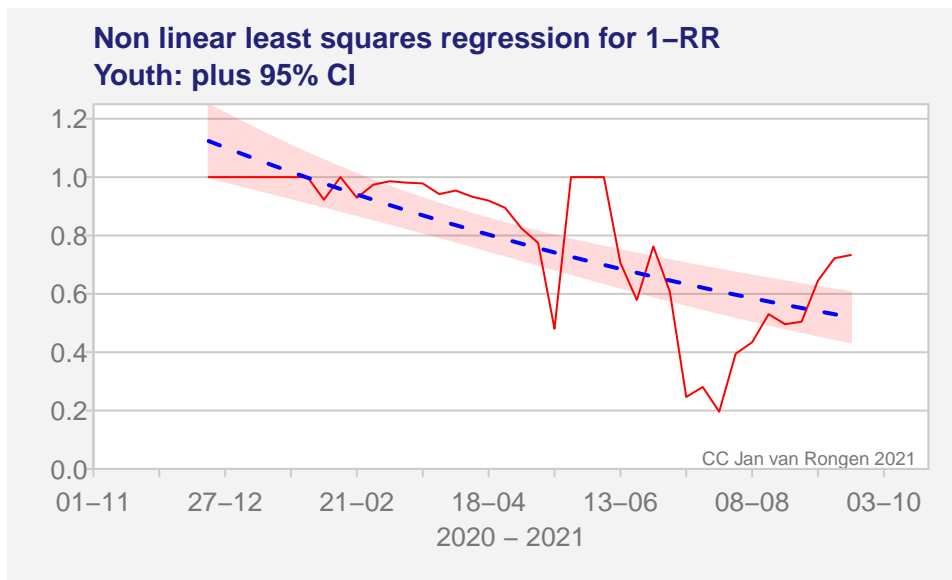
Get the data ready

We construct two data-frames for the Adults and the Youth Age groups with the relevant information. We assume that 15% of the population was infected before start vaccination.

And now the estimates.

Plus Youth

Non linear least squares regression for 1−RR
Youth: plus 95% CI

CC Jan van Rongen 2021
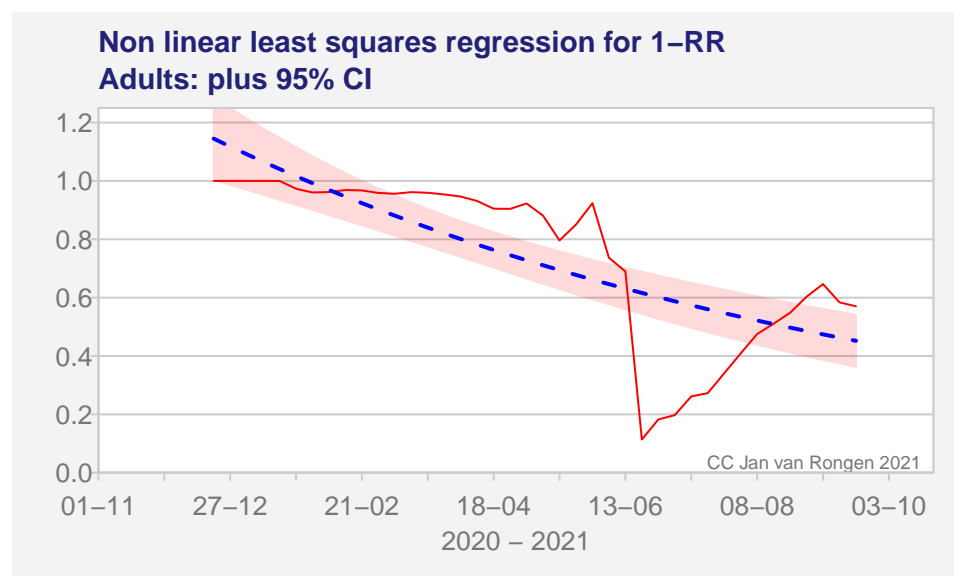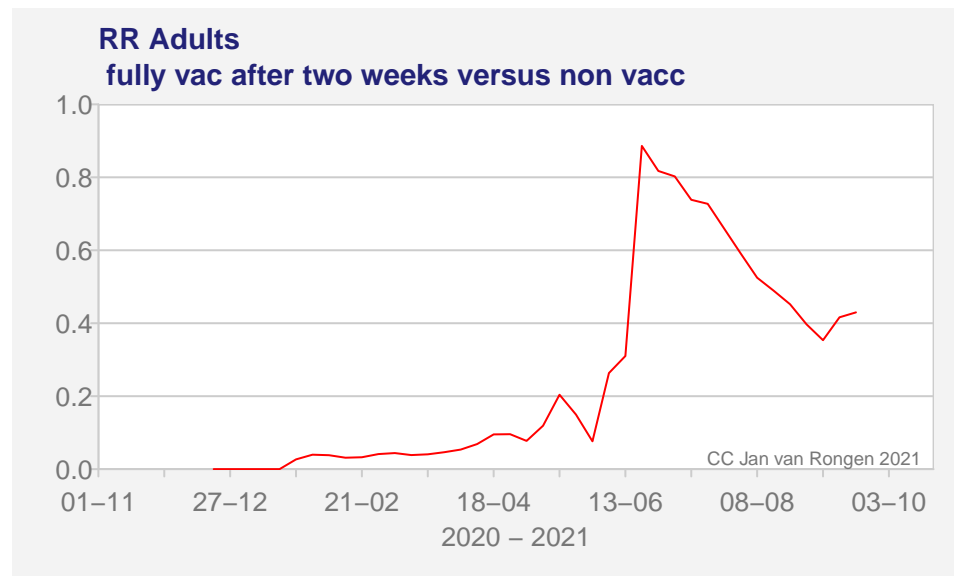
Discussion

The adjustments on the input data are more realistic than previous attempts. The effect is that the RR decrease by the adjustments for positives, but increases when the proportion of positives in 2010 is increased. The overall piture howvere is smoother than without those adjustments,

From here we might conclude that the decay has a half-time of 6-10 months, but the data clearly does not fit the exponential model very well. In fact you woud never think of such a refression if you would only see the original curve.

As the above model is a Cox regression in disguise, we must conclude that any survival-like model does not model this data.