Technical Document vs 0.4

Jan van Rongen

2021-10-06

Introduction

We analyse data from the State of Israel related to COVID vaccines and infections. This is the technical document that contains all code used. This is a reanalysis of a document circulating on Internet.

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

This version is an appendum to the previous version. we look at the different age groups in one week of april. First repeat the initial data processing.

Data sources and cleaning

We use four data sources: three from the Israeli government site [1] 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 not yet used.

Definitions and methods

Fully vaccinated are people from the day of their second dose. *Single vaccinated* are people that had one dose but not two (from the date of first dose). All others are *unvaccinated*. *Infected* are people that (on a certain day) tested positive. All others are *not infected* on that day.

A *cross table* is a 2x2 table with two (0,1) categories. The entries are the number of people in that combination of categories. When the matrix is `|a b | c d|` then the *ralative risk* AKA *risk ratio* AKA *RR* is (a/(a+b))/(c/(c+d)). The *odds ratio* is also known as *OR* is a.d/b.c

In epidemiologie RR and OR are used a lot. Numbers can be quite large, so in programming we have to avoid numerical overflow or incorrect rounding.

Lookint at RR;s in one week pper age group.

RR per age group in week 2021–04–25