A short guide to Corvid, a SARS-CoV-2 epidemic simulator

Dennis Chao

Introduction

Corvid is an agent-based model to simulate SARS-CoV-2 transmission. The model is described in "Modeling layered non-pharmaceutical interventions against SARS-CoV-2 in the United States with Corvid", available on medRxiv (https://doi.org/10.1101/2020.04.08.20058487). The code is available at https://github.com/dlchao/corvid. The model is based on a previously published open-source influenza transmission model, FluTE.

Compiling and running corvid

Corvid is written in C++ and can be compiled using the Makefile in the "code/" directory. Just type "make corvid" or just "make" to make the executable. This just compiles the C++ and links the object files to produce the "corvid" executable.

Here is a simple config file, "config-seattle26":

```
## # basic configuration file for metro seattle
## label example-seattle-26
## datafile seattle
## R0 2.6
## runlength 180
## summaryfilename Summary-seattle26.txt
## logfilename Log-seattle26.txt
## tractfilename Tracts-seattle26.txt
## individualfilename Individuals-seattle26.txt
## individualfile 1
## randomnumberseed 1
## seedinfected 5
## seedinfecteddaily 0
## responsethreshold 1
```

This config file tells Corvid that you want to use the Seattle-based population, set R0=2.6, run the simulation for 180 days starting with 5 infected people, and don't add more infected people each day. It also specifies the random number seed, which you will need to vary if you want to see the effects of stochasticity on the outputs. The random number seed also affects the generation of the synthetic population, as described below. Just run "./corvid config-seattle26", and in about 10 minutes, it will output 4 files, whose names are specified in the config file. A full list of configuration parameters is in the README.

Corvid outputs

The "Summary" file has most of the results you will want. It starts with the Corvid version number, then outputs the label specified in the config file. This allows you to match the file to the config file that generated it. The summary file contains the parameters used for the run and basic outputs, like the number of people who were symptomatic each day and the attack rates. Note that the "Number symptomatic" is the number who are symptomatic each day (prevalence), not the number who become symptomatic each day (incidence). You can derive incidence from the "Cumulative symptomatic".

The "Tracts" file outputs the number of people in the synthetic population who live in each census tract. Corvid reads an input file that specifies the population size of each census tract, and it fills the tract with families of different size and age compositions randomly chosen from a distribution that reflects the US Census from the year 2000. Therefore, changing the random number seed will result in a slightly different population.

The "Log" file outputs the symptomatic and infection prevalence as well as the cumulative symptomatic for each census tract each day. Note that the symptomatic prevalence is not the symptomatic incidence (new cases).

The "Individuals" file is not normally generated because it can be large, but "config-seattle26" tells Corvid to output it by specifying "individualfile 1", where "1" means output the file and "0" or not specifying this parameter will cause it not to be output. The "individuals" file outputs one line per individual in the synthetic population, which can allow you to inspect household structures, determine where and when each person was infected, and build infection trees. The time of infection reported in the file is in units of half-days, so divide by 2 to convert to days.

Parsing Corvid output files

Here, we parse Corvid output files to generate summary statistics and plots. You can look at "corvid-guide.Rmd" to see the R code that produces all of the tables or figues in this pdf. The "config-seattle26" tells Corvid to use the Seattle-like population, based on 2000 Census data of metropolitan Seattle. The synthetic population has 563484 residing in 124 census tracts. Tracts have 1181 to 9003 people living in them, and 150 to 36823 people working in them. We can plot the number of people infected or symptomatic each day using the Summary file. The plots in this manuscript use Corvid version 0.5.

The Individuals file allows us to look at the time and cause of infection of each individual in the simulation. The average generation time was 8.38, median=8. The largest number of infections come from household members. School and workplaces contribute nearly equally. The rest of infections come from the general community. These proportions can be compared with those found across different agent-based models for R0=1.9 in Halloran et al 2008.

Table 1: Proportion of infections by settings for R0=2.6, %

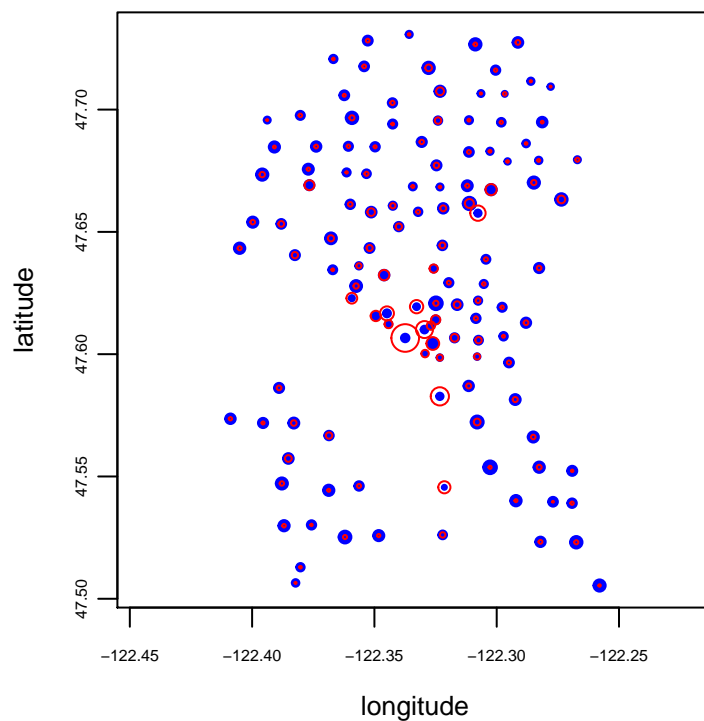| setting	infected26
school	17.9
work	16.2
home	30.2
community/neighbors	35.7
total	100.0

Figure 1: Map of the census tracts for the synthetic Seattle population. Dot size proportional to resident population. Red circle size is proportional to the number of people who work in each tract. Plotted using the 'seattle-tracts.dat' input file and the 'Tracts' output file.
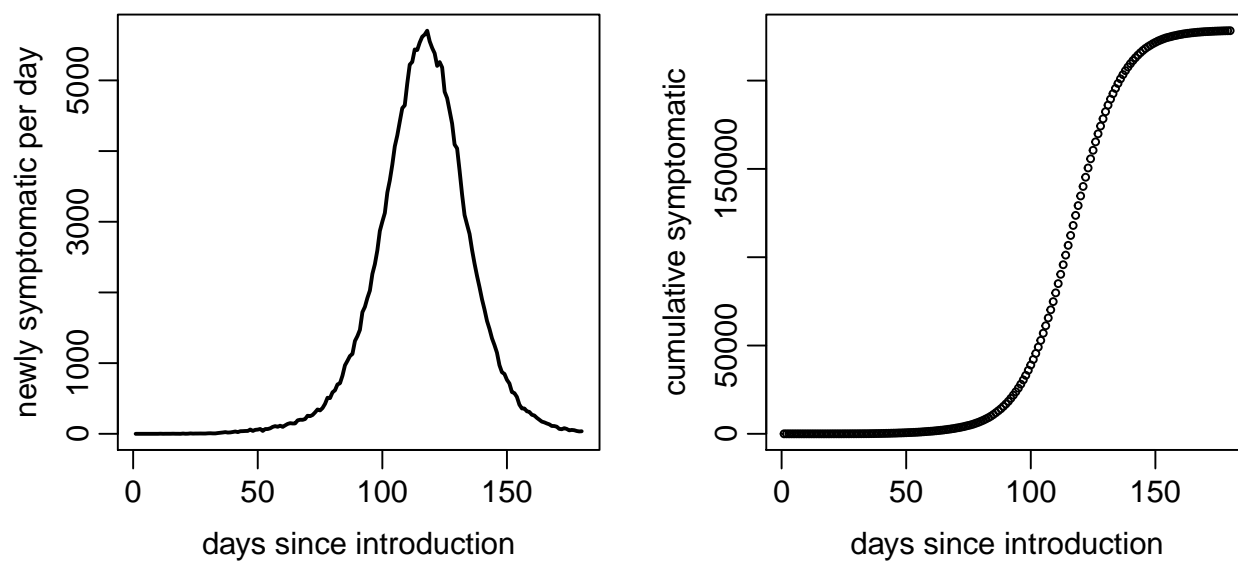

Figure 2: Symptomatic people over time in Seattle. On the left, the number of newly symptomatic people per day. On the right, the cumulative cases. Plotted from data in the Summary file.
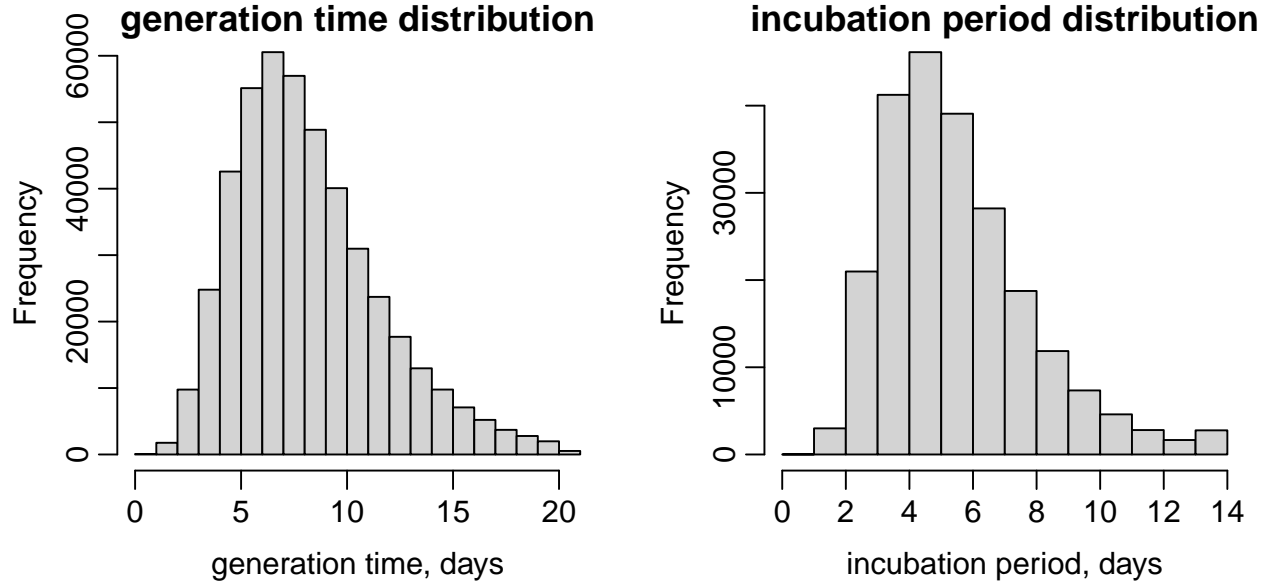
Figure 3: Generation time distribution in the Seattle model, R0=2.6. On the left, the distribution of generation times. On the right, incubation period distribution. Plotted using data from the Individuals output file.
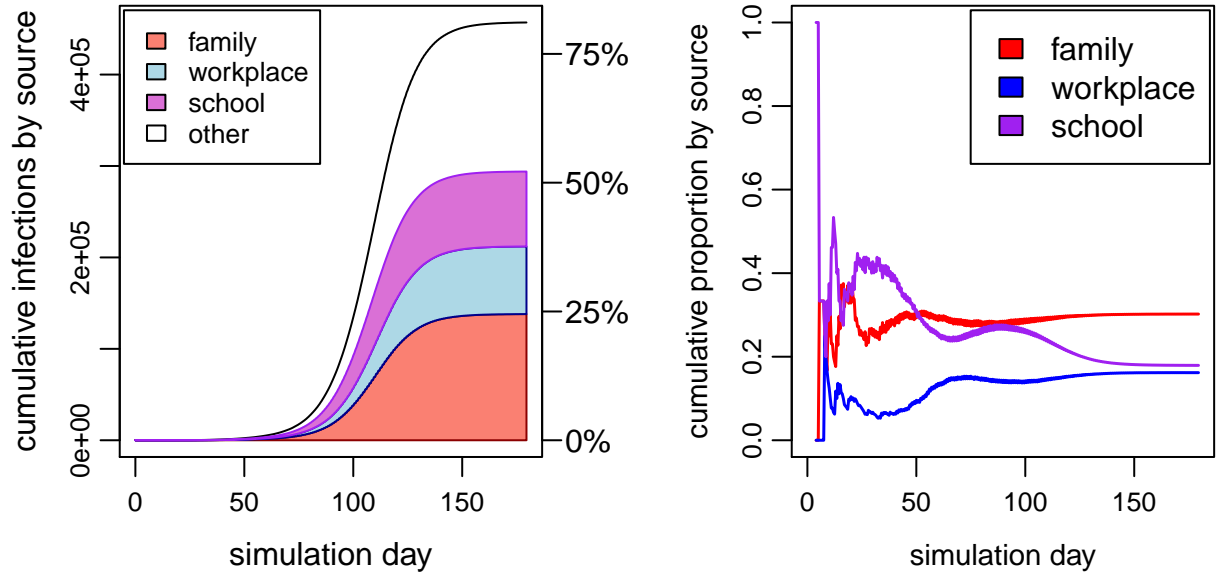

Figure 4: Settings of infections in the Seattle model (R0=2.6). On the left, cumulative infections in the Seattle population model by source of infection. On the right, cumulative proportion of infections by source. The sources plotted do not add up to one - the remainder of infections come from the general community.

4

Mapping locations of cases in Corvid

Day 15, 5 cases

Day 35, 99 cases

Day 50, 637 cases

Day 100, 38975 cases

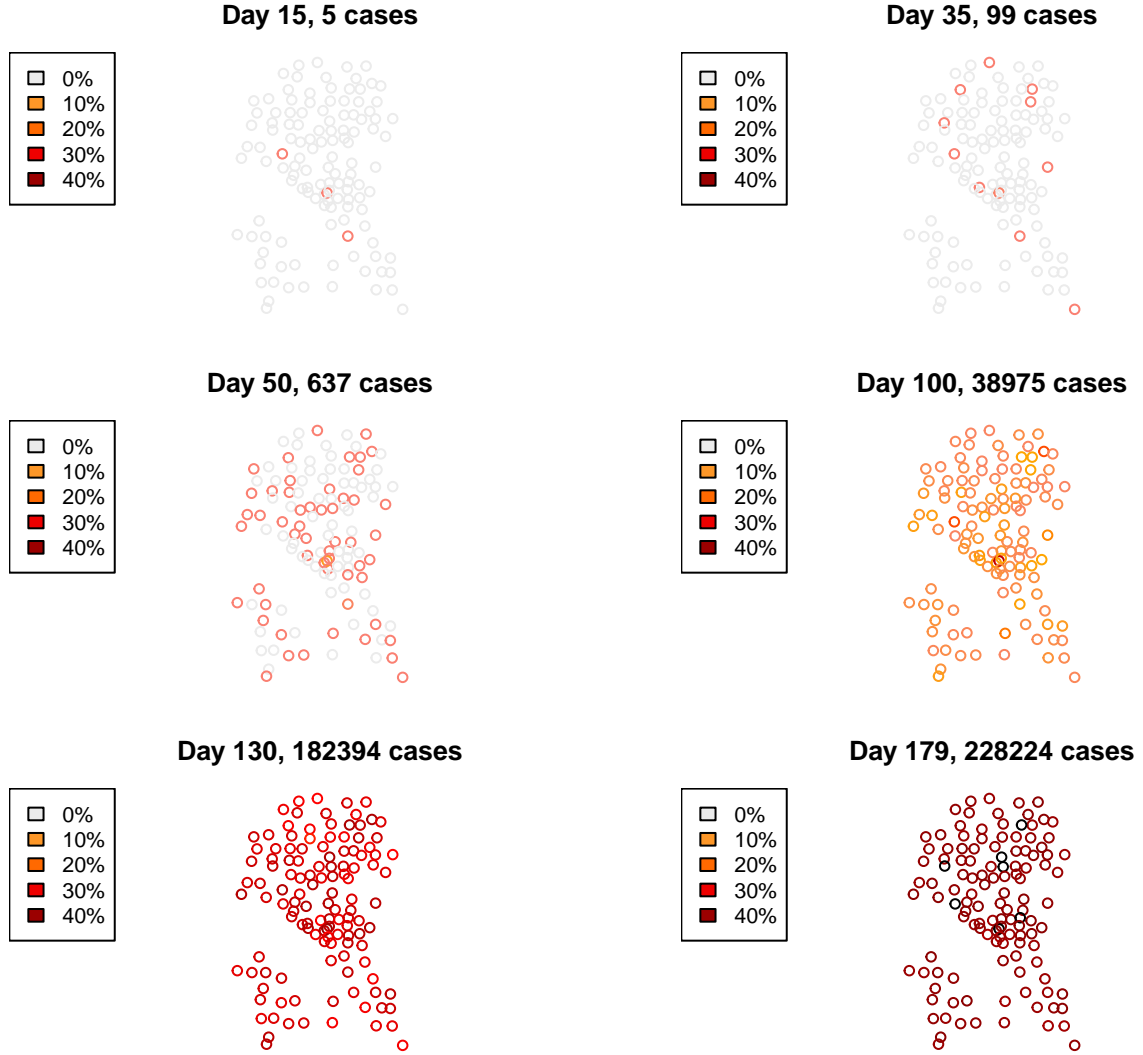**Day 130, 182394 cases**

Day 179, 228224 cases

Figure 5: Map of cumulative cases by census tract after epidemic starts with five infections on day 0. Color is based on cumulative symptomatic attack rate.

We can use the "Log" output files to see where and when cases occurred in the synthetic population. The "seattle-tracts.dat" input file has the lat/long of each tract, and the "Tracts" output file give the size of the synthetic population in each census tract. There are no "cases" on day 1 because the initially infected people are not symptomatic yet.

Simulating interventions

We can simulate interventions like liberal leave at workplaces (employees even more likely than usual to go home if sick), work from home policies (workers stay at home instead of going to work), voluntary isolation by sick people (isolation at home after symptom onset), self isolation by ascertained cases (isolation without transmission to household members, simulated by using an "antiviral"), and home quarantine of family members of sick people. Here's the configuration file "config-seattle26-workfromhome" for having a "work from home" policy start on day 60:

```
## # metro seattle with work-from-home
## label example-seattle-26-workfromhome
## datafile seattle
## R0 2.6
## runlength 180
## summaryfilename Summary-seattle26-workfromhome.txt
## logfilename Log-seattle26-workfromhome.txt
## tractfilename Tracts-seattle26-workfromhome.txt
## randomnumberseed 1
## seedinfected 5
## seedinfecteddaily 0
## responseday 60
## workfromhome 0.6
```

The "compliance" of work-from-home is set to 0.6, which is the fraction of workers who decide to work from home. Compliance for other interventions is set in the same way (except for school closures). There is no option (yet) to turn off most interventions after they are activated.

School closure is the only intervention with a limited duration. Here's the configuration file "config-seattle26-closeallschools" for shutting down all schools for 14 days starting on day 60:

```
## # metro seattle with total school closure on day 60
## label example-seattle-26-closeallschools
## datafile seattle
## R0 2.6
## runlength 180
## summaryfilename Summary-seattle26-closeallschools.txt
## logfilename Log-seattle26-closeallschools.txt
## tractfilename Tracts-seattle26-closeallschools.txt
## randomnumberseed 1
## seedinfected 5
## seedinfecteddaily 0
## responseday 60
## schoolclosurepolicy all
## schoolclosuredays 14
```
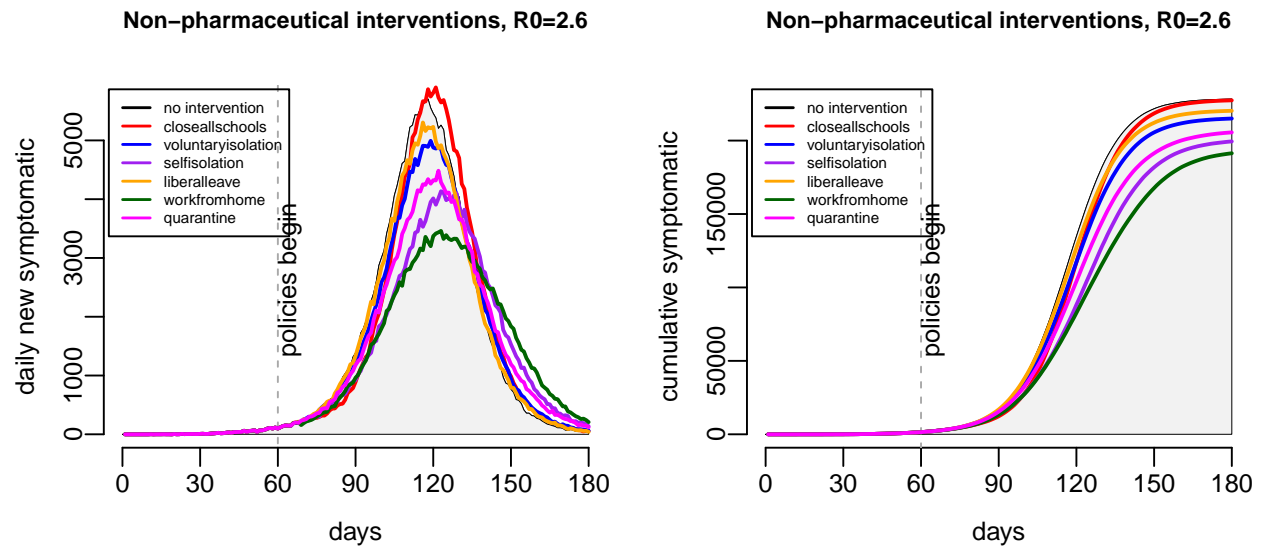
Non−pharmaceutical interventions, R0=2.6

Non−pharmaceutical interventions, R0=2.6

Figure 6: Non-pharmaceutical intervention simulations. R0=2.6 and compliance for the interventions is 60% (except for school closures, which is 100%).

You can also choose to start interventions after a certain fraction of the population has become ascertained, instead of using a pre-specified date. Here's the configuration file "config-seattle21-threshold" for shutting down all schools for 14 days after 10% of the population has become symptomatic:

```
## # configuration file for metro seattle where schools close after 10% of the population is symptomatic
## label seattle-21-threshold
## datafile seattle
## R0 2.1
## runlength 180
## summaryfilename Summary-seattle21-threshold.txt
## logfilename Log-seattle21-threshold.txt
## tractfilename Tracts-seattle21-threshold.txt
## individualfilename Individuals-seattle21-threshold.txt
## randomnumberseed 1
## seedinfected 5
## seedinfecteddaily 0
## ascertainmentfraction 1.0
## ascertainmentdelay 0
## responsethreshold 0.1
## responsedelay 0
## schoolclosurepolicy all
## schoolclosuredays 14
```

The "responsethreshold" is the cumulative fraction of the population that must be ascertained (after getting symptoms) before the epidemic is "declared". The response begins a set number of days (specified by "responsedelay") after the epidemic is declared. In this case, schools close for 14 days. In this example, we set the ascertainment rate to 100% with no ascertainment delay, so every symptomatic illness is immediately counted towards this threshold. Since epidemics with lower R0 progress more slowly, reaching the 10% threshold to trigger interventions takes longer for R0=2.1 than for 2.6 (see plot).

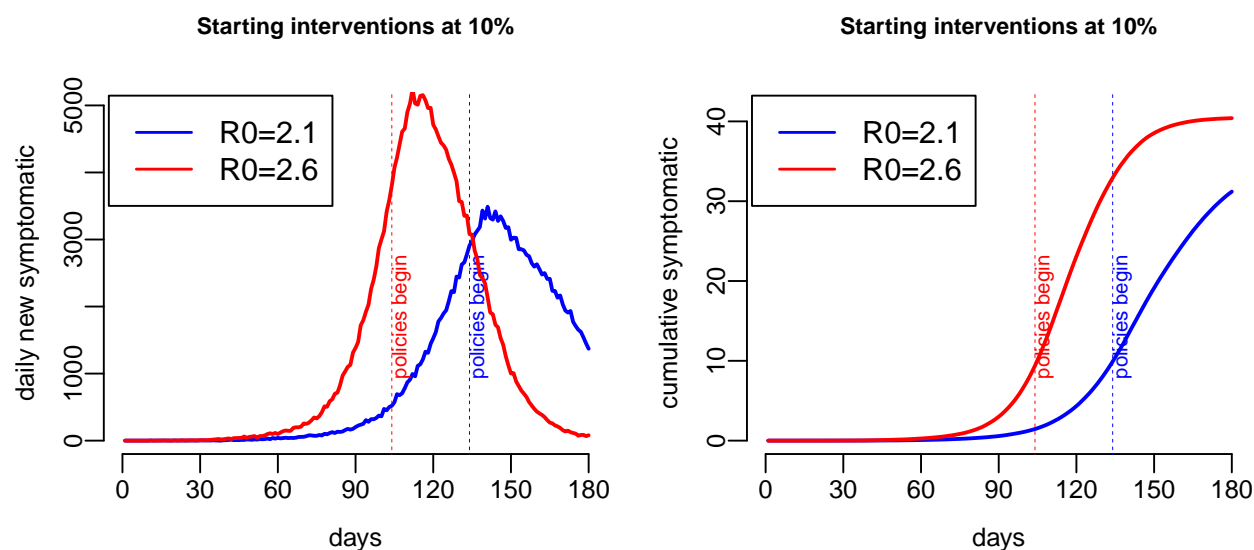Figure 7: Using thresholds to start responses. Schools close for 14 days after 10% of the population has become symptomatic. R0=2.1 or 2.6.

References

- Chao DL, Oron AP, Srikrishna D, Famulare M. Modeling layered non-pharmaceutical interventions against SARS-CoV-2 in the United States with Corvid. medRxiv. 2020 Apr 11. doi: 10.1101/2020.04.08.20058487 doi: 10.1101/2020.04.08.20058487.
- Chao DL, Halloran ME, Obenchain VJ, Longini IM Jr. FluTE, a publicly available stochastic influenza epidemic simulation model. PLoS Comput Biol. 2010 Jan 29;6(1):e1000656. doi: 10.1371/journal.pcbi.1000656
- Germann TC, Kadau K, Longini IM Jr, Macken CA. Mitigation strategies for pandemic influenza in the United States. Proc Natl Acad Sci U S A. 2006 Apr 11;103(15):5935-40.
- Halloran ME, Ferguson NM, Eubank S, Longini IM Jr, Cummings DA, Lewis B, Xu S, Fraser C, Vullikanti A, Germann TC, Wagener D, Beckman R, Kadau K, Barrett C, Macken CA, Burke DS, Cooley P. Modeling targeted layered containment of an influenza pandemic in the United States. Proc Natl Acad Sci U S A. 2008 Mar 25;105(12):4639-44. doi: 10.1073/pnas.0706849105.