

# L03 Reproduction numbers and their estimation<sup>1</sup>

Michael Höhle<sup>1</sup>

<sup>1</sup>Department of Mathematics, Stockholm University, Sweden



m\_hoehle

STA427 FS2021

Statistical Methods in Infectious Disease Epidemiology  
Epidemiology, Biostatistics and Prevention Institute  
University of Zurich, Switzerland



University of  
Zurich<sup>UZH</sup>

---

<sup>1</sup>LaMo: 2021-03-08 @ 22:33:20

# Outline

## 1 Effective Reproduction Number

## 2 Implementation and Results

- $R(t)$  for Switzerland

## 3 Outlook

# Overview

## 1 Effective Reproduction Number

## 2 Implementation and Results

- $R(t)$  for Switzerland

## 3 Outlook

# Outline

1 Effective Reproduction Number

2 Implementation and Results

3 Outlook

# Basic Reproduction Number

- In the previous lectures we have considered the **basic reproduction number**  $R_0$ , which was defined<sup>2</sup> as

$$R_0 = \beta/\gamma S(0)$$

- $R_0$  is the mean number of direct offspring one infectious case generates in a *completely susceptible population without intervention measures*
- However, once an epidemic unfolds susceptibles are depleted. Furthermore, interventions or behavioral changes aim at reducing  $\beta$ ,  $\gamma$  or the amount of susceptibles ( $\rightarrow$  vaccination).
- In order to assess public health interventions the basic reproduction number is not useful (Delamater, Street, Leslie, Yang, and Jacobsen 2019).

---

<sup>2</sup> $\beta$  was the infectious contact rate,  $\gamma$  the mean duration of the infectious period and  $S(0)$  is the number of susceptibles at time zero ( $S(0) \approx N$ ).

## Effective Reproduction Number (1)

- We define the time-varying effective reproduction number  $R(t)$  as a time varying quantity denoting be the average number of secondary cases generated at at calendar time  $t$
- If contact and recovery rates do not vary with time we get

$$R(t) = R_0 \times \frac{S(t)}{S(0)}$$

- Interventions could be reflected by extending this to

$$R(t) = (1 - c(t)) \times R_0 \times \frac{S(t)}{S(0)}$$

## Effective Reproduction Number (2)

- The derivations in this lecture will follow Fraser (2007), who does not explicitly consider depletion of susceptibles
- Let the **transmissibility** function  $\beta(t, \tau)$  be a function of calendar time  $t$  and time since infection  $\tau$
- Dependence on  $\tau$  could, e.g., be due to time-dependence of viral shedding in the host, but also changes in contact behaviour after infection, depletion of susceptibles
- Renewal equation

$$I(t) = \int_0^\infty \beta(t, \tau) I(t - \tau) d\tau \quad (1)$$

## Effective Reproduction Number (3)

- Case reproduction number

$$R_c(t) = \int_0^{\infty} \beta(t + \tau, \tau) d\tau$$

Interpretation: Average number of individuals an infected at time  $t$  can expect to infect

- Instantaneous reproduction number

$$R(t) = \int_0^{\infty} \beta(t, \tau) d\tau$$

Interpretation: Average number of individuals an infected at time  $t$  is expected to infect should conditions remain unchanged



## Effective Reproduction Number (4)

- Assuming the factorization

$$\beta(t, \tau) = \phi_1(t)\phi_2(\tau),$$

Fraser (2007) shows that by assuming  $\int_0^\infty \phi_2(\tau) d\tau = 1$  we get

$$\beta(t, \tau) = R(t)w(\tau) \tag{2}$$

- Here  $w(\tau)$  denotes the distribution of the new infections as a function of time since infection, i.e. the generation time distribution.

# Effective Reproduction Number (5)

Inserting (2) into (1) yields:

- Instantaneous reproduction number

$$R(t) = \frac{I(t)}{\int_0^\infty I(t-\tau)w(\tau)d\tau}$$

- Case reproduction number

$$R_c(t) = \int_0^\infty R(t+\tau)w(\tau)d\tau$$

# Discretised Instantaneous Reproduction Number (1)

- If only discrete observations for time intervals  $t_i \equiv [t_i, t_{i+1})$  are available, then one would use a discretized version of the instantaneous reproduction number

$$R(t_i) = \frac{l_i}{\sum_{j=0}^n w_j l_{i-j}}$$

- Note: the  $0 \leq w_i \leq 1$  now denote the probability mass function of a discretized version of the generation time, i.e.  $\sum_{i=0}^n w_i = 1$ .

## Discretised Instantaneous Reproduction Number (2)

- Casting this into a distributional framework, e.g.,

$$I_i | I_{i-1}, I_{i-1}, \dots \sim Po \left( R(t_i) \times \sum_{j=0}^n w_j I_{i-j} \right)$$

allows the consideration of  $R(t_i)$  as a parameter to be estimated in a statistical model

- This also allows the construction of confidence intervals or credibility regions for  $R(t_i)$ .

## Smoothed Instantaneous Reproduction Number

- Estimate  $R(t)$  using discrete data is subject to fluctuations and it is therefore in general not recommended to just use one time point (Cori, Ferguson, Fraser, and Cauchemez 2013)
- Suggested improvement: Compute instantaneous reproduction number over an interval of  $\tau$  days as follows

$$R_{\tau}(t_i) = \frac{\sum_{k=i-\tau+1}^i I_k}{\sum_{k=i-\tau+1}^i \sum_{j=0}^n w_j I_{k-j}}$$

- Note that this is slightly difference from just using the mean of the  $\tau$  computed  $R(t)$  values, i.e.

$$\overline{R}_{\tau}(t) = \frac{1}{\tau} \sum_{k=i-\tau+1}^i R(t_{i-k})$$

# Superspreading

- So far we have considered the *mean* number of offspring one infectious case generates
- This view ignores the variability in the offspring distribution, e.g., the larger the variation, the greater the probability of extinction
- If we can identify settings where super-spreading occurs, one can reduce the mean number of offspring substantially

# Outline

## 1 Effective Reproduction Number

## 2 Implementation and Results

- $R(t)$  for Switzerland

## 3 Outlook

# Outbreak simulation (1)

- Consider a growth model and denote by  $y_t$  the expected number of new symptom onsets we observe on day  $t$ .
- Let  $(w_1, \dots, w_M)'$ , denote the probability mass function of the serial interval distribution, i.e.  $P(GT = i) = w_i$  for  $i = 1, 2, \dots, M$ .
- Assume that the expected number of cases can be described by the homogeneous linear difference equation

$$y_t = \sum_{i=1}^M R_c(t-i) w_i y_{t-i}, \quad t = 2, 3, \dots$$

- Note: We use the case reproduction number and ignore terms on the RHS when  $t - M \leq 0$ .

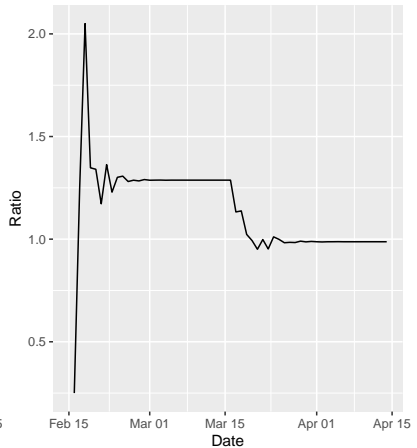
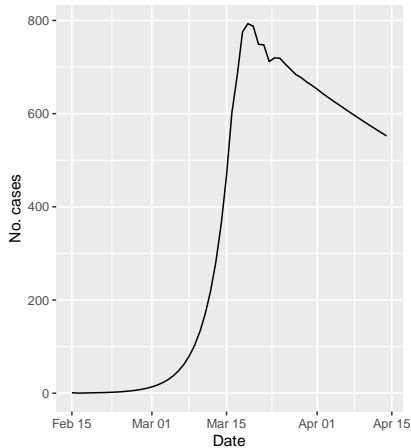


## Outbreak simulation (2)

- Somewhat arbitrarily we fix  $y_1 = 1$  and conceptually denote by  $t = 1$  the 15th of February 2020 in calendar time.
- To simulate a COVID-19 like outbreak with lockdown type intervention we use

$$R_e(t) = \begin{cases} 2.5 & \text{if } t \leq 2020-03-15 \\ 0.95 & \text{otherwise} \end{cases}$$

## Outbreak simulation (3)



## EpiEstim (1)

- An implementation of the instantaneous reproduction estimate is available in the R package `EpiEstim` (Cori, Ferguson, Fraser, and Cauchemez 2013)
- Input is a 'data.frame' containing the column 'dates' and 'I' (for the incidence each day)
- It also computes credibility regions for  $R(t)$  and allows one to address uncertainty in the estimation of the serial interval

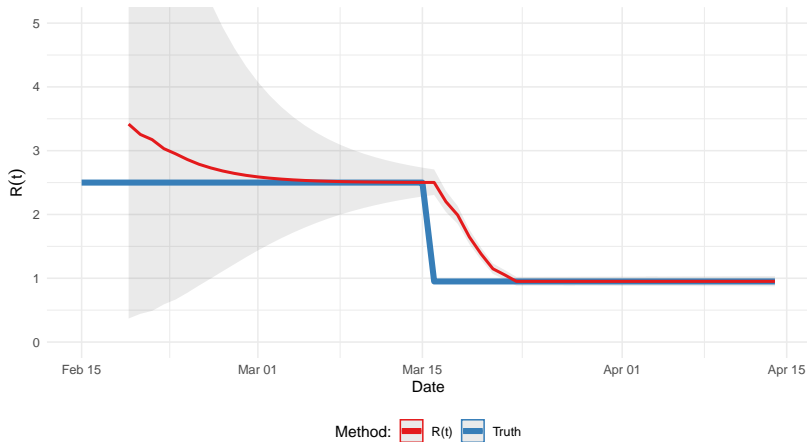
## EpiEstim (2)

```
library(EpiEstim)
# Rename data.frame columns to names handled by the EpiEstim pkg.
out_epiestim <- out %>% rename(I = y, dates = Date) %>% select(dates, I)

# Estimate the instantaneous reproduction number
res <- EpiEstim::estimate_R(out_epiestim, method = "non_parametric_si",
                           config=make_config(si_distr=GT_obj$GT,
                                                t_start=2:nrow(out_epiestim),
                                                t_end=2:nrow(out_epiestim))
)

# Convert result to a data.frame
rt_irt_df <- data.frame(Date=res$dates[res$R$t_end],
                        R_hat=res$R$`Mean(R)`,
                        lower=res$R$`Quantile.0.025`,
                        upper=res$R$`Quantile.0.975`,
                        Method="R(t)")
```

# EpiEstim (3)

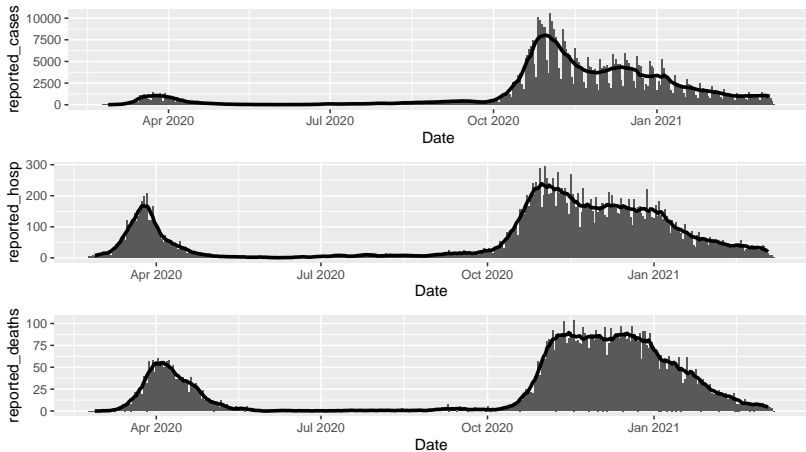


# $R(t)$ for Switzerland

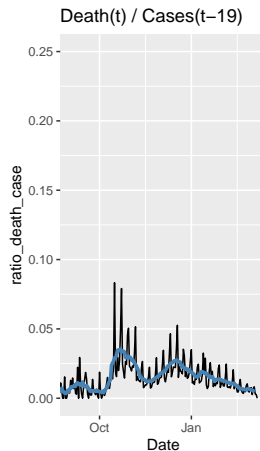
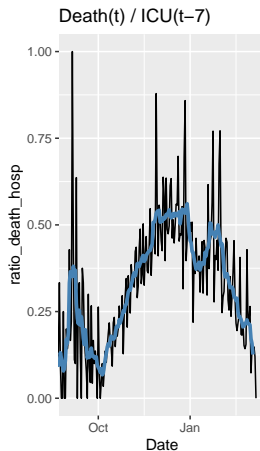
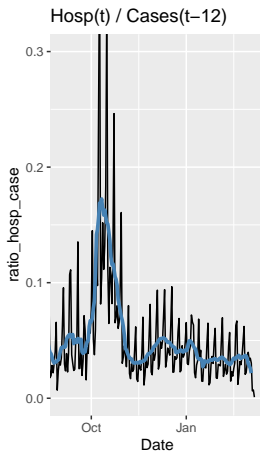
- Aim: Produce a crude  $R(t)$  for Switzerland based on the daily number of new cases as reported by the Swiss Federal Office of Public Health (FOPH)
- In order to do so, we first look a bit closer at the data as of 2021-03-08

reported_cases	reported_hosp	reported_deaths
562364	23665	9294

# COVID-19 Time Series for Switzerland (1)

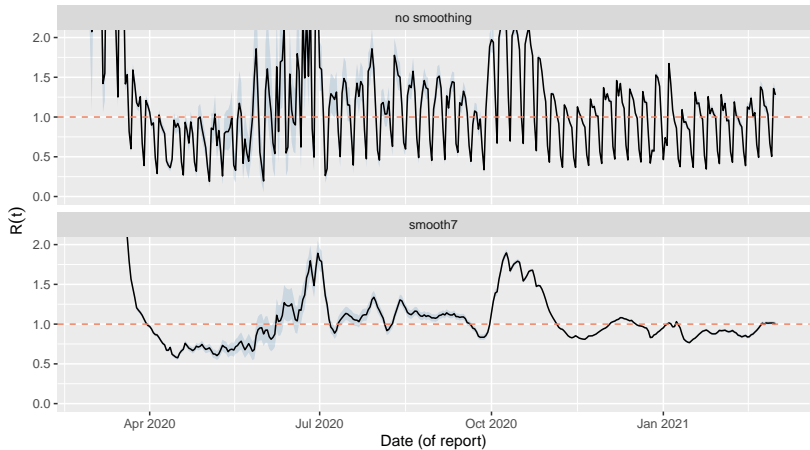


# COVID-19 Time Series for Switzerland (2)





# Coarse $R(t)$ Estimate (1)



## Coarse $R(t)$ Estimate (2)

### Limitations:

- All cases are considered as community transmissions, but
  - there is an increase in the overall number of tests
  - some settings are tested more pro-actively than others (e.g. hospital staff)
  - imported cases are ignored (negligible from mid March)
- Serial interval
  - The analysis used an estimate from Nishiura, Linton, and Akhmetzhanov (2020) with mean 4.8 days and a standard deviation of 2.3 days
  - Difference between generation time and serial interval (infection vs. symptom onset - see Svensson (2007))
  - However, in this coarse analysis the event is actually "Day of report".
  - No specific Swiss serial interval was available to me and no estimation uncertainty considered

## Coarse $R(t)$ Estimate (3)

- $R(t)$  always needs context, e.g., the absolute number of (reported) cases.
- The implication of  $R(t) = 1.3$  is very different when you have 10,000 cases or when you have 100.
- Especially when the case number is low, spikes due to outbreaks can lead to large  $R(t)$  fluctuations
- Be very careful when using  $R(t)$  graphs to assess interventions!

# Outline

- 1 Effective Reproduction Number
- 2 Implementation and Results
- 3 Outlook

# Outlook

- Not discussed in this lecture: COVID-19 analyses need to look at age groups, because of the strong age dependence in, e.g, mortality
- Good overview preprint about pitfalls when estimating  $R(t)$  by Gostic et al. (2020)
- Site estimating COVID-19  $R(t)$  for countries around the world:

<https://epiforecasts.io/covid>

## Literature I



Cori, Anne, et al. 2013. “A New Framework and Software to Estimate Time-Varying Reproduction Numbers During Epidemics”. American Journal of Epidemiology 178 (9): 1505–1512. doi:10.1093/aje/kwt133. eprint: <http://aje.oxfordjournals.org/content/178/9/1505.full.pdf+html>. <http://aje.oxfordjournals.org/content/178/9/1505.abstract>.



Delamater, P. L., et al. 2019. “Complexity of the Basic Reproduction Number ( $R_0$ )”. 25 (1): 1. doi:10.3201/eid2501.171901.



Fraser, Christophe. 2007. “Estimating Individual and Household Reproduction Numbers in an Emerging Epidemic”. PLOS ONE 2, no. 8 (): e758.



Gostic, Katelyn M, et al. 2020. “Practical considerations for measuring the effective reproductive number,  $R_t$ ”. medRxiv (): 2020.06.18.20134858. doi:10.1101/2020.06.18.20134858.

## Literature II



Nishiura, Hiroshi, Natalie M. Linton, and Andrei R. Akhmetzhanov. 2020. “Serial interval of novel coronavirus (COVID-19) infections”. International Journal of Infectious Diseases 93 (): 284–286. ISSN: 1201-9712, visited on 04/12/2020. doi:10.1016/j.ijid.2020.02.060.



Svensson, Å. 2007. “A note on generation times in epidemic models”. Math Biosci 208, no. 1 (): 300–311. doi:10.1016/j.mbs.2006.10.010.