

Nowcasting, short-term forecasting and R-values

Johannes Bracher (Karlsruhe Institute of Technology)

April 10, 2024

NOZGEKA infectious disease modelling short course



Welcome!



- ▶ This is the module **Nowcasting and Short-term forecasting** in the *NOZGEKA infectious disease modelling short course*.
- ▶ My name is **Johannes Bracher**.
 - ▶ Assistant Professor *Health Statistics* at Karlsruhe Institute of Technology since 04/2024.
 - ▶ Research topics: infectious disease modelling / forecasting, forecast evaluation, count time series modelling
 - ▶ Email:
johannes.bracher@kit.edu

Topics we will cover

- ▶ Some background on **disease surveillance data**.
- ▶ **Nowcasting:** what happened in the recent past, but has not yet appeared in our data?
- ▶ **R-values:** what is the current trend?
- ▶ **Short-term forecasting:** what is likely to happen next?
- ▶ **Open access literature** on the various topics will be referenced throughout the presentation. Some general resources you may find useful:
 - ▶ The Epidemiologist R Handbook,
<https://epirhandbook.com/>
 - ▶ Cori and Kucharski (2024),
<https://osf.io/preprints/osf/mg497>

The epidemic curve

- ▶ The most basic data on disease outbreaks is the *epidemic curve* – a histogram of case numbers by onset date (or another suitable / available date).
- ▶ Useful intro:
<https://www.cdc.gov/training/quickearns/epimode/>
- ▶ Example: COVID-19 cases in Germany according to RKI situation report:

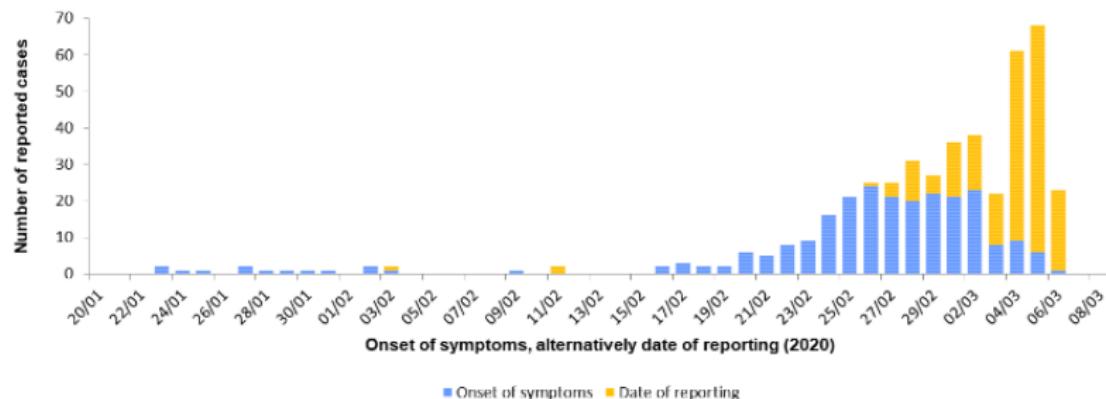


Figure 3: Epidemiological curve of the 447 electronically reported COVID-19-cases in Germany, by onset of symptoms and alternatively by date of reporting (08/03/2020)

Know your data

There are many different data types and it is important to know what you are dealing with!

► Data on disease / illness:

- ▶ **outbreak investigation data:** during an emerging outbreak, epidemiologists actively track cases.
- ▶ **sentinel data:** detailed data are routinely collected in some GP practices / hospitals etc. and extrapolated to the entire population.
- ▶ **mandatory surveillance data:** for identified cases for some diseases there is a legal obligation to report these to the authorities.

► Data on healthcare burden / mortality:

- ▶ Hospitalization / ICU counts.
- ▶ Deaths.

► Alternative / novel data sources:

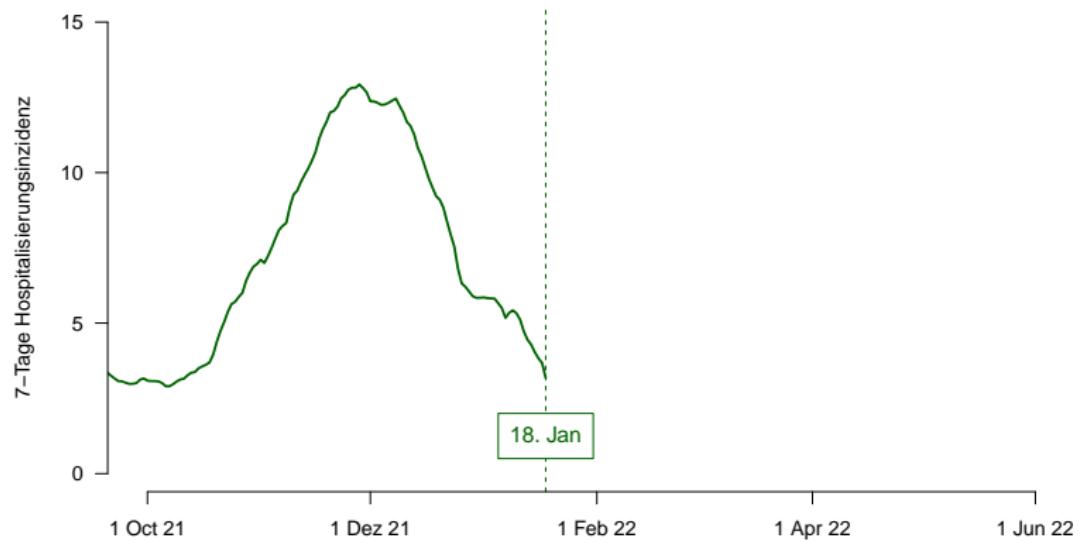
- ▶ Wastewater data
- ▶ Social media data
- ▶ ...

A statistician's view...

- ▶ In the following we focus on rather **generic statistical methods** to analyse infectious disease data.
- ▶ In practice you need to **know the particularities of your data**, but here we focus on the statistical fundamentals.
- ▶ But **keep a critical mind** when applying them in practice!

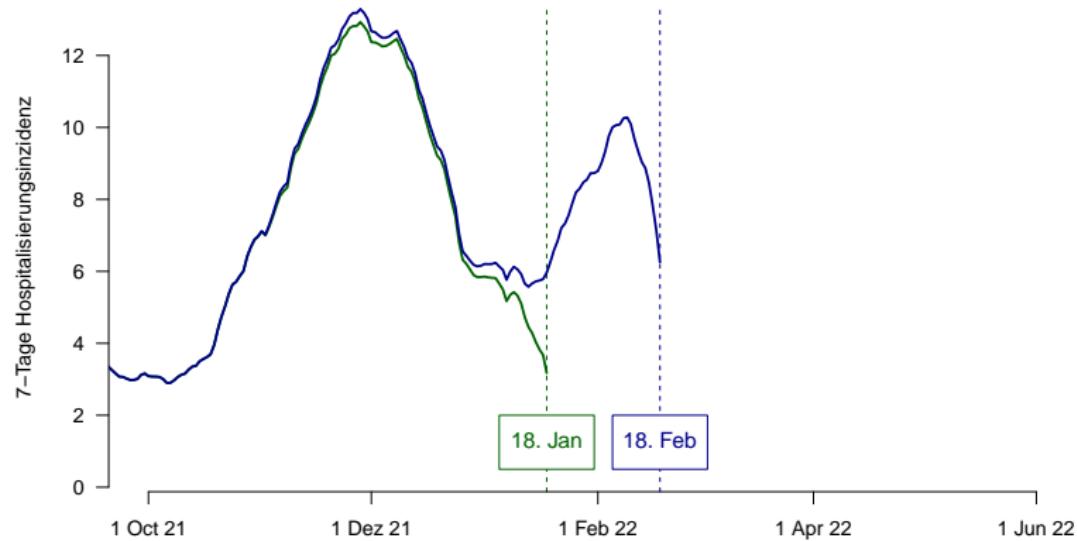
Part 1: Nowcasting

Illustrating the problem: German COVID-19 hospitalization incidences



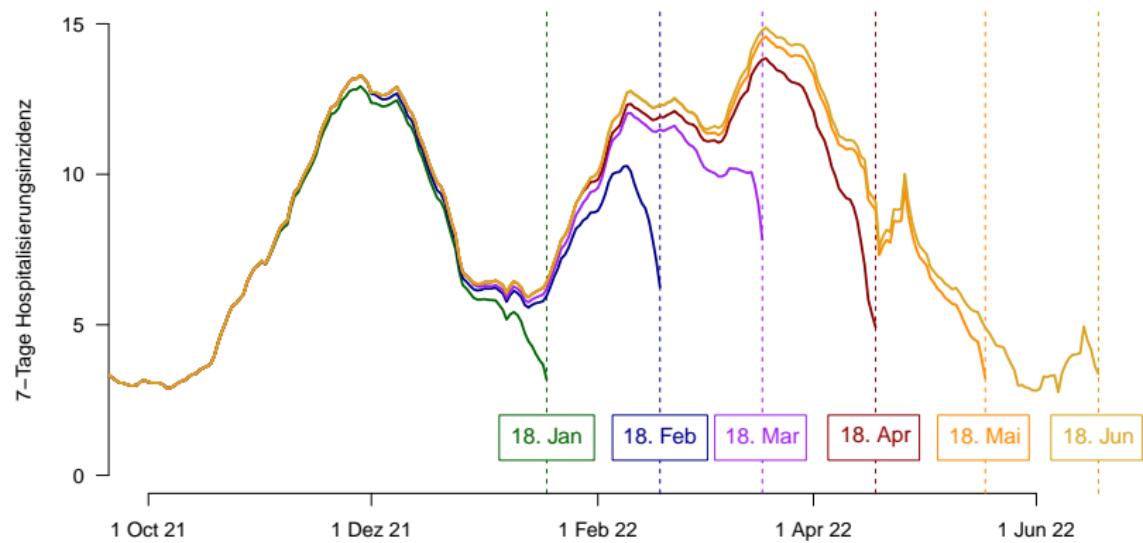
(sorry for German axis labels...)

Illustrating the problem: German COVID-19 hospitalization incidences



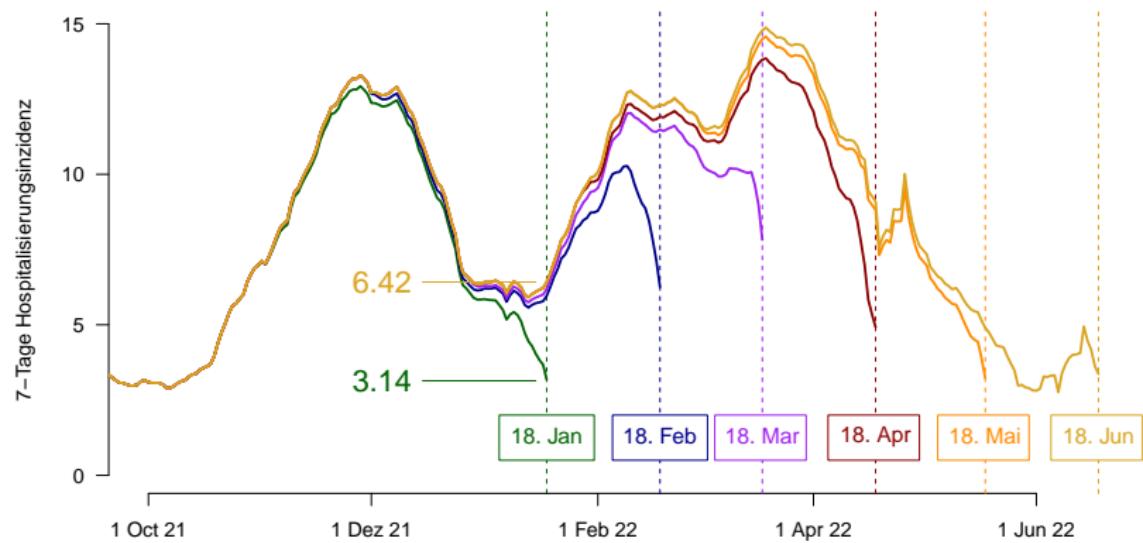
(sorry for German axis labels...)

Illustrating the problem: German COVID-19 hospitalization incidences



(sorry for German axis labels...)

Illustrating the problem: German COVID-19 hospitalization incidences



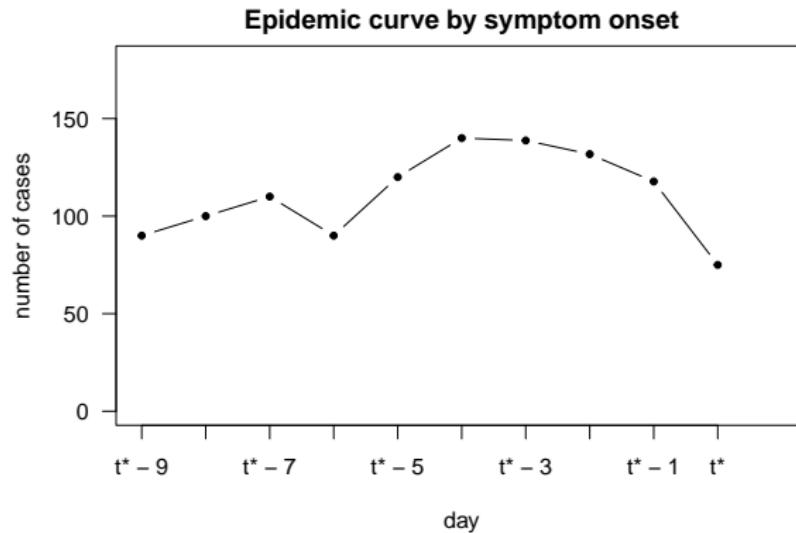
(sorry for German axis labels...)

Reporting delays

- ▶ Epidemiological time series are usually aggregated by symptom onset date / hospitalization date / date of death – some **epidemiologically meaningful date**.
- ▶ We call this the **reference date**.
- ▶ However, there is a **delay** between the reference date and the **reporting date**, which is when an event actually shows up in the data.
 - ▶ Note: we neglect that there may be another delay between reporting and the inclusion in the available data set.
- ▶ This leads to an **artificial dip** in the most recent values of an epidemiological time series.

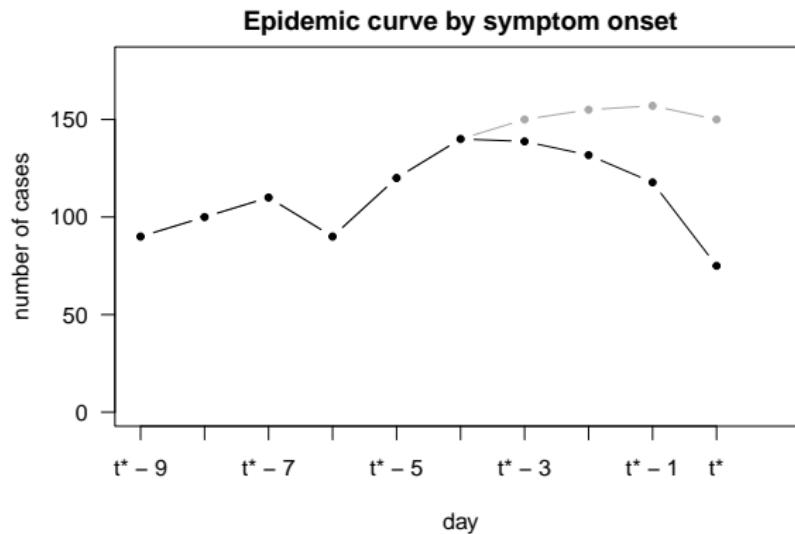
Illustrating the dip

- ▶ Imagine we are looking at the data on day t^* .
- ▶ This is what the data look like:



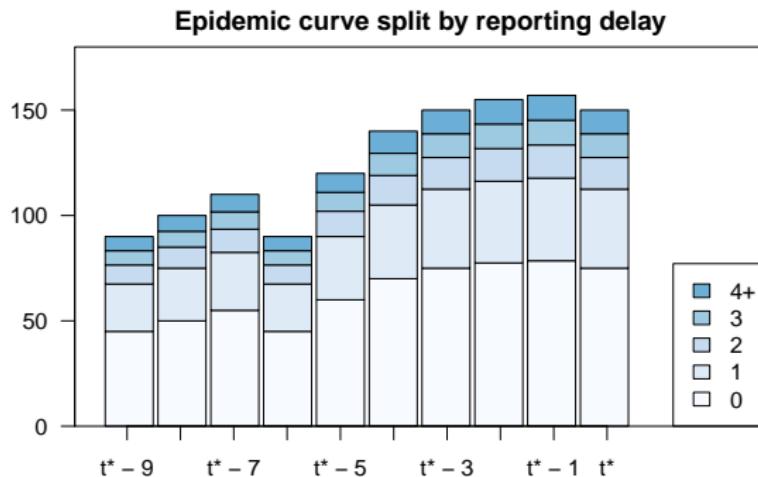
Illustrating the dip

- ▶ Imagine we are looking at the data on day t^* .
- ▶ But if we waited for a few days, they would look like this:



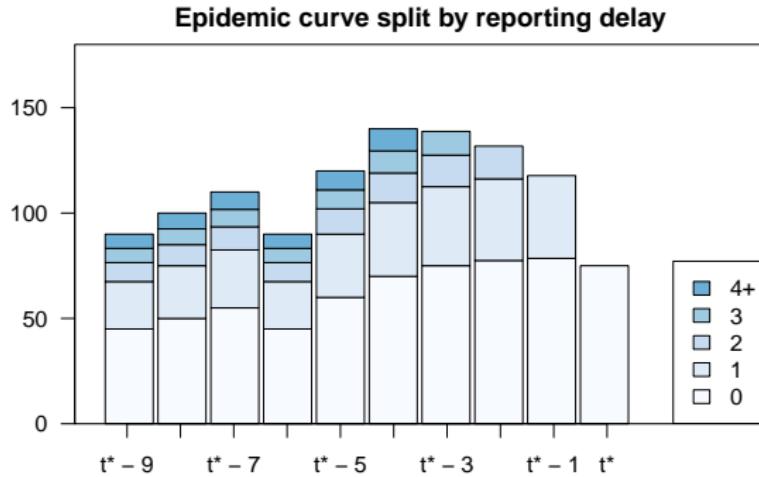
Illustrating the dip

- ▶ Imagine we are looking at the data on day t^* .
- ▶ It helps to split the data by the respective reporting delay:



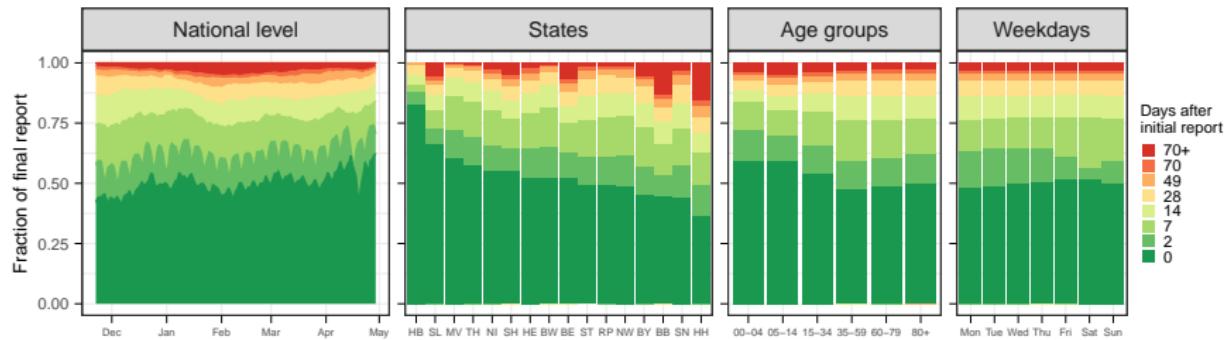
Illustrating the dip

- ▶ Imagine we are looking at the data on day t^* .
- ▶ This way we see that there's something missing for the last couple of days:



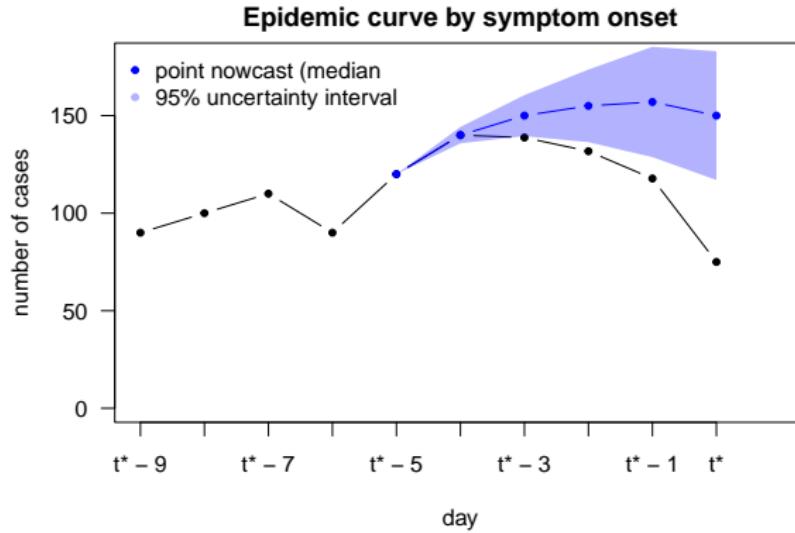
Reporting fractions for COVID hospitalizations in Germany

- ▶ Reporting fractions can vary over space and time, making them more difficult to handle.



Nowcasting

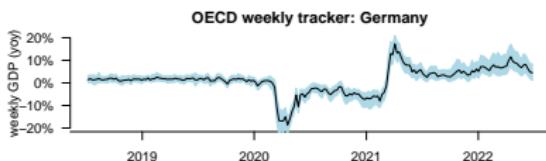
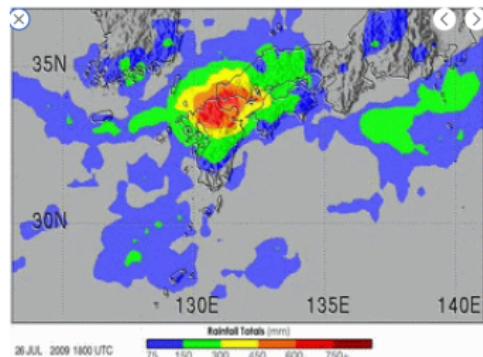
- ▶ The goal of **nowcasting** is to anticipate, what an epidemiological time series will look like after all delayed reports are in.



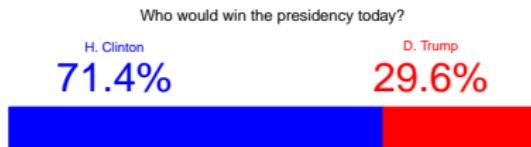
On a side note: nowcasting in different disciplines

GDP nowcast (OECD):

Precipitation nowcast:



Election nowcast
(fivethirtyeight.com, 8. Nov 2016):



Some notation

- ▶ We denote by
 - ▶ x_t the total number of cases with symptom onset on day t .
 - ▶ $x_{t,d}$ the number of cases with symptom onset on day t and reported on day $t + d$.
 - ▶ D an assumed maximum delay.
- ▶ We thus have

$$x_t = x_{t,1} + x_{t,2} + \cdots + x_{t,D} = \sum_{d=1}^D x_{t,d}$$

The reporting triangle

- ▶ The **reporting triangle** is a useful way of arranging the data needed for nowcasting.
- ▶ In the example below, the maximum reporting delay is 5 days.
- ▶ On day t^* , black cells are known, blue cells need estimation.

day	$d = 0$	$d = 1$	$d = 2$	$d = 3$	$d = 4$	$d = 5$	total
1	$x_{1,0}$	$x_{1,1}$	$x_{1,2}$	$x_{1,3}$	$x_{1,4}$	$x_{1,5}$	x_1
2	$x_{2,0}$	$x_{2,1}$	$x_{2,2}$	$x_{2,3}$	$x_{2,4}$	$x_{2,5}$	x_2
:	:						
$t^* - 5$	$x_{t^*-5,0}$	$x_{t^*-5,1}$	$x_{t^*-5,2}$	$x_{t^*-5,3}$	$x_{t^*-5,4}$	$x_{t^*-5,5}$	x_{t^*-5}
$t^* - 4$	$x_{t^*-4,0}$	$x_{t^*-4,1}$	$x_{t^*-4,2}$	$x_{t^*-4,3}$	$x_{t^*-4,4}$	$x_{t^*-4,4}$	x_{t^*-4}
$t^* - 3$	$x_{t^*-3,0}$	$x_{t^*-3,1}$	$x_{t^*-3,2}$	$x_{t^*-3,3}$	$x_{t^*-3,4}$	$x_{t^*-3,5}$	x_{t^*-3}
$t^* - 2$	$x_{t^*-2,0}$	$x_{t^*-2,1}$	$x_{t^*-2,2}$	$x_{t^*-2,3}$	$x_{t^*-2,4}$	$x_{t^*-2,5}$	x_{t^*-2}
$t^* - 1$	$x_{t^*-1,0}$	$x_{t^*-1,1}$	$x_{t^*-1,2}$	$x_{t^*-1,3}$	$x_{t^*-1,4}$	$x_{t^*-1,5}$	x_{t^*-1}
t^*	$x_{t^*,0}$	$x_{t^*,1}$	$x_{t^*,2}$	$x_{t^*,3}$	$x_{t^*,4}$	$x_{t^*,5}$	x_{t^*}

Nowcasting is essentially about multiplying

- ▶ Example: we wish to correct yesterday's case number.
- ▶ From experience we know that after one day, usually around half of the total cases have been reported.
- ▶ So we multiply the current number by 2. 😎
- ▶ Central assumption: The distribution of delays remains the same.

Practical steps in nowcasting

1. **Data management:** Keep storing “data snapshots”, i.e. data as available on each day / in each week.
 - ▶ ideally you can somehow make sure that “data releases” work similarly each time.
2. **Data preparation:** Generate reporting triangle.
 - ▶ this can be tedious :(
 - ▶ sometimes some pre-processing is necessary to avoid implausible entries (e.g., negative values)
3. **Descriptive analysis:** Analyse delay structure graphically (see slide ??) to check stability of delays.
4. **Nowcasting:** Run nowcasting model each day / week.
5. **Evaluation:** Go back to past nowcasts to check how well nowcasting worked.

A heuristic nowcasting procedure

A simple procedure is to fill in the reporting triangle step by step:

$$\hat{x}_{t^*,1} = x_{t^*,0} \times \frac{\sum_{t=1}^{t^*-1} x_{t,1}}{\sum_{t=1}^{t^*-1} x_{t,0}} = x_{t^*,0} \times \frac{\text{sum of light red cells}}{\text{sum of gray cells}}$$

day	$d = 0$	$d = 1$	$d = 2$	$d = 3$	$d = 4$	$d = 5$	total
1	$x_{1,0}$	$x_{1,1}$	$x_{1,2}$	$x_{1,3}$	$x_{1,4}$	$x_{1,5}$	x_1
2	$x_{2,0}$	$x_{2,1}$	$x_{2,2}$	$x_{2,3}$	$x_{2,4}$	$x_{2,5}$	x_2
\vdots							
$t^* - 5$	$x_{t^*-5,0}$	$x_{t^*-5,1}$	$x_{t^*-5,2}$	$x_{t^*-5,3}$	$x_{t^*-5,4}$	$x_{t^*-5,5}$	x_{t^*-5}
$t^* - 4$	$x_{t^*-4,0}$	$x_{t^*-4,1}$	$x_{t^*-4,2}$	$x_{t^*-4,3}$	$x_{t^*-4,4}$	$x_{t^*-4,4}$	x_{t^*-4}
$t^* - 3$	$x_{t^*-3,0}$	$x_{t^*-3,1}$	$x_{t^*-3,2}$	$x_{t^*-3,3}$	$x_{t^*-3,4}$	$x_{t^*-3,5}$	x_{t^*-3}
$t^* - 2$	$x_{t^*-2,0}$	$x_{t^*-2,1}$	$x_{t^*-2,2}$	$x_{t^*-2,3}$	$x_{t^*-2,4}$	$x_{t^*-2,5}$	x_{t^*-2}
$t^* - 1$	$x_{t^*-1,0}$	$x_{t^*-1,1}$	$x_{t^*-1,2}$	$x_{t^*-1,3}$	$x_{t^*-1,4}$	$x_{t^*-1,5}$	x_{t^*-1}
t^*	$x_{t^*,0}$	$\hat{x}_{t^*,1}$	$x_{t^*,2}$	$x_{t^*,3}$	$x_{t^*,4}$	$x_{t^*,5}$	x_{t^*}

colors: black: known. violet: already imputed. blue: missing.
red: imputed. gray: used for imputation.

This is similar to what is described in Lawless (1994),
doi.org/10.2307/3315820 (not OA); see also Supplement of Wolfram et al
2023, doi.org/10.1371/journal.pcbi.1011394 (OA)

A heuristic nowcasting procedure

A simple procedure is to fill in the reporting triangle step by step:

$$\hat{x}_{t^*-1,2} = (x_{t^*-1,0} + x_{t^*-1,1}) \times \frac{\text{sum of light red cells}}{\text{sum of gray cells}}$$

day	$d = 0$	$d = 1$	$d = 2$	$d = 3$	$d = 4$	$d = 5$	total
1	$x_{1,0}$	$x_{1,1}$	$x_{1,2}$	$x_{1,3}$	$x_{1,4}$	$x_{1,5}$	x_1
2	$x_{2,0}$	$x_{2,1}$	$x_{2,2}$	$x_{2,3}$	$x_{2,4}$	$x_{2,5}$	x_2
:							
$t^* - 5$	$x_{t^*-5,0}$	$x_{t^*-5,1}$	$x_{t^*-5,2}$	$x_{t^*-5,3}$	$x_{t^*-5,4}$	$x_{t^*-5,5}$	x_{t^*-5}
$t^* - 4$	$x_{t^*-4,0}$	$x_{t^*-4,1}$	$x_{t^*-4,2}$	$x_{t^*-4,3}$	$x_{t^*-4,4}$	$x_{t^*-4,5}$	x_{t^*-4}
$t^* - 3$	$x_{t^*-3,0}$	$x_{t^*-3,1}$	$x_{t^*-3,2}$	$x_{t^*-3,3}$	$x_{t^*-3,4}$	$x_{t^*-3,5}$	x_{t^*-3}
$t^* - 2$	$x_{t^*-2,0}$	$x_{t^*-2,1}$	$x_{t^*-2,2}$	$x_{t^*-2,3}$	$x_{t^*-2,4}$	$x_{t^*-2,5}$	x_{t^*-2}
$t^* - 1$	$x_{t^*-1,0}$	$x_{t^*-1,1}$	$x_{t^*-1,2}$	$x_{t^*-1,3}$	$x_{t^*-1,4}$	$x_{t^*-1,5}$	x_{t^*-1}
t^*	$x_{t^*,0}$	$\hat{x}_{t^*,1}$	$x_{t^*,2}$	$x_{t^*,3}$	$x_{t^*,4}$	$x_{t^*,5}$	x_{t^*}

colors: black: known. violet: already imputed. blue: missing.

red: imputed. gray: used for imputation.

A heuristic nowcasting procedure

A simple procedure is to fill in the reporting triangle step by step:

$$\hat{x}_{t^*,2} = (x_{t^*-1,0} + x_{t^*-1,1}) \times \frac{\text{sum of light red cells}}{\text{sum of gray cells}}$$

day	$d = 0$	$d = 1$	$d = 2$	$d = 3$	$d = 4$	$d = 5$	total
1	$x_{1,0}$	$x_{1,1}$	$x_{1,2}$	$x_{1,3}$	$x_{1,4}$	$x_{1,5}$	x_1
2	$x_{2,0}$	$x_{2,1}$	$x_{2,2}$	$x_{2,3}$	$x_{2,4}$	$x_{2,5}$	x_2
:							
$t^* - 5$	$x_{t^*-5,0}$	$x_{t^*-5,1}$	$x_{t^*-5,2}$	$x_{t^*-5,3}$	$x_{t^*-5,4}$	$x_{t^*-5,5}$	x_{t^*-5}
$t^* - 4$	$x_{t^*-4,0}$	$x_{t^*-4,1}$	$x_{t^*-4,2}$	$x_{t^*-4,3}$	$x_{t^*-4,4}$	$x_{t^*-4,5}$	x_{t^*-4}
$t^* - 3$	$x_{t^*-3,0}$	$x_{t^*-3,1}$	$x_{t^*-3,2}$	$x_{t^*-3,3}$	$x_{t^*-3,4}$	$x_{t^*-3,5}$	x_{t^*-3}
$t^* - 2$	$x_{t^*-2,0}$	$x_{t^*-2,1}$	$x_{t^*-2,2}$	$x_{t^*-2,3}$	$x_{t^*-2,4}$	$x_{t^*-2,5}$	x_{t^*-2}
$t^* - 1$	$x_{t^*-1,0}$	$x_{t^*-1,1}$	$x_{t^*-1,2}$	$x_{t^*-1,3}$	$x_{t^*-1,4}$	$x_{t^*-1,5}$	x_{t^*-1}
t^*	$x_{t^*,0}$	$\hat{x}_{t^*,1}$	$x_{t^*,2}$	$x_{t^*,3}$	$x_{t^*,4}$	$x_{t^*,5}$	x_{t^*}

colors: black: known. violet: already imputed. blue: missing.

red: imputed. gray: used for imputation.

Statistical background

- ▶ This essentially corresponds to a non-parametric estimation of the **discrete-time delay distribution**.
- ▶ The delay distribution is characterized by

$f(d)$ = probability that delay equals d

$F(d)$ = probability that delay is at most d ,

but can also be represented via terms of the form

$g(d) = \frac{f(d)}{F(d)}$ = probability that delay is d given it is at most d ,

which are easier and more efficient to estimate from the reporting triangle.

- ▶ Our procedure essentially does this iteratively.
- ▶ See Lawless (1994) for a detailed discussion.

Some practical aspects

- ▶ It is usually a good idea to only use a window of recent data to perform nowcasting (at least if there is reason to believe that delays change over time).
- ▶ Choosing the maximum delay D can be tricky.
 - ▶ Once D is chosen it is best to discard all reports with delay D for all dates to avoid artefacts.
- ▶ If you are dealing with daily data there are often weekday effects.
 - ▶ these can be hard to control as they are typically both in the “actual data” and the delay structure.
 - ▶ the heuristic method shown before can be misleading and biased for certain weekdays.

Some practical aspects (2)

- ▶ Delays may get longer on average if the reporting system is under stress.
- ▶ Delays may get shorter as a new reporting system gets more established.
 - ▶ Nowcasts will tend to **underestimate** the corrected values when delays get longer.
 - ▶ They will **overestimate** the corrected values when delays get shorter again.
- ▶ Important public holidays (in Germany: Christmas + New Year) can disturb nowcasts.
 - ▶ this affects both the nowcasts for these periods, but also the following time, as you keep using “weird data”.

Advanced methods: Bayesian hierarchical models

There is a wealth of more complex methods, most of them implemented in R:

- ▶ **epinowcast** (Sam Abbott et al, <https://package.epinowcast.org/articles/model.html>)
 - ▶ rich functionality, uses stan for Bayesian inference.
- ▶ **NobBS** (Sarah McGough et al, <https://cran.r-project.org/web/packages/NobBS/index.html>)
 - ▶ less functionality, more compact, uses JAGS for Bayesian inference. Paper:
<https://doi.org/10.1371/journal.pcbi.1007735>
- ▶ **nowcaster** (Rafael Lopes an Leo Bastos, <https://covid19br.github.io/nowcaster/>)
 - ▶ uses R-INLA for Bayesian inference.

Intuitive sketch of Bayesian hierarchical models

- ▶ Denote again by $x_{t,d}$ the number of events for time t reported with a delay d .
- ▶ Denote by λ_t the latent (unobservable) time series of infections in the population. For this we assume some simple renewal process like

$$\log(\lambda_t) = \underbrace{r_t}_{\text{growth rate at } t} \times \log(\lambda_{t-1}) + \epsilon_t, \epsilon_t \sim N(0, \sigma_\epsilon^2).$$

- ▶ The observed values are modelled as

$$x_{t,d} \sim \text{Poisson}(f(d) \times \lambda_t)$$

or a generalized version thereof.

- ▶ This model can be fitted and predictions can be generated using Bayesian inference.

Benefits of Bayesian hierarchical models

- ▶ These models **use the data more efficiently** as they “borrow strength across time points”.
 - ▶ predictions for each day also use data from the neighbouring days.
- ▶ These methods **work better for small numbers**.
 - ▶ zeros are not handled well by simple multiplication scheme.
- ▶ These models also return **uncertainty intervals**.
- ▶ These models **can be extended** to account for weekday patterns in disease incidence and reporting.

Keeping track of performance

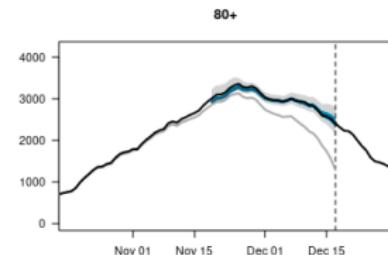
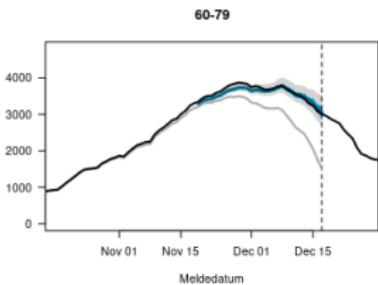
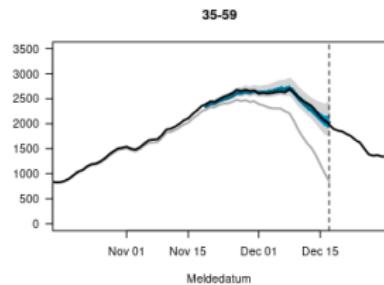
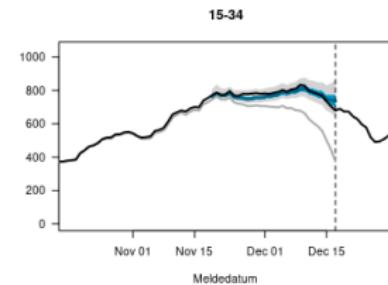
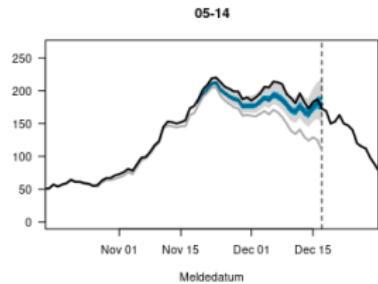
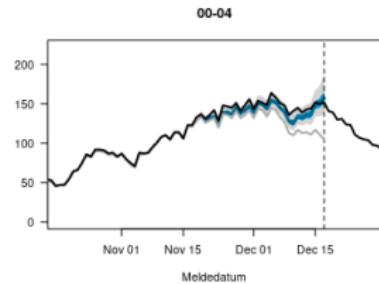
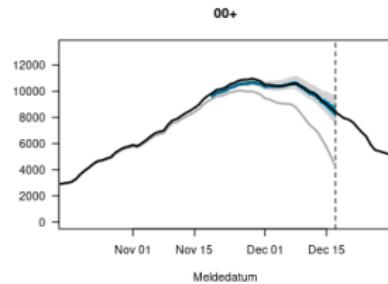
- ▶ It is important to keep track of whether or not your nowcasts yield acceptable corrections.
- ▶ We performed a comparative evaluation of eight different nowcasting methods for COVID-19 hospitalizations in Germany in Wolffram et al:

PLOS COMPUTATIONAL BIOLOGY

Collaborative nowcasting of COVID-19 hospitalization incidences in Germany

Daniel Wolffram , Sam Abbott, Matthias an der Heiden, Sebastian Funk, Felix Günther, Davide Hailer, Stefan Heyder, Thomas Hotz, Jan van de Kassteele, Helmut Küchenhoff, Sören Müller-Hansen, Diellë Syliqi, Alexander Ullrich, [...], Johannes Bracher [[view all](#)]

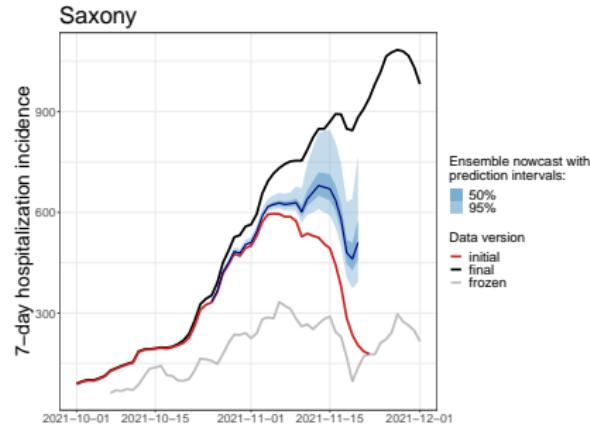
Diagnostic 1: inspection for individual dates



Nowcast intervals:
= 50%
■ 95%

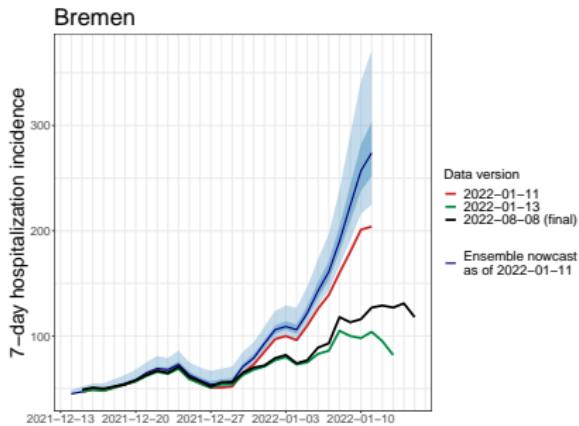
Data versions:
— 2021-12-17
— 2022-03-29

Spot problematic cases and identify reasons (I)



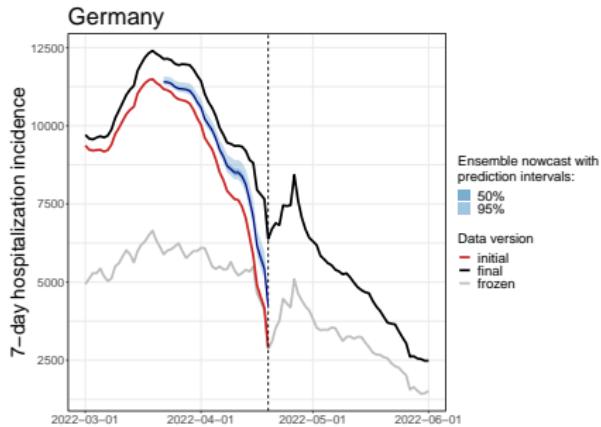
- ▶ In November 2021, hospitals in Saxony were overwhelmed, leading to a considerable lengthening of reporting delays.
- ▶ Consequently, the true number of hospitalizations was **strongly underestimated** in real time.

Spot problematic cases and identify reasons (II)



- ▶ In January 2022, a major reporting issue occurred in Bremen, requiring a major downward correction of the number of hospitalizations.
- ▶ Consequently, the true number of hospitalizations was **strongly overestimated** in real time.

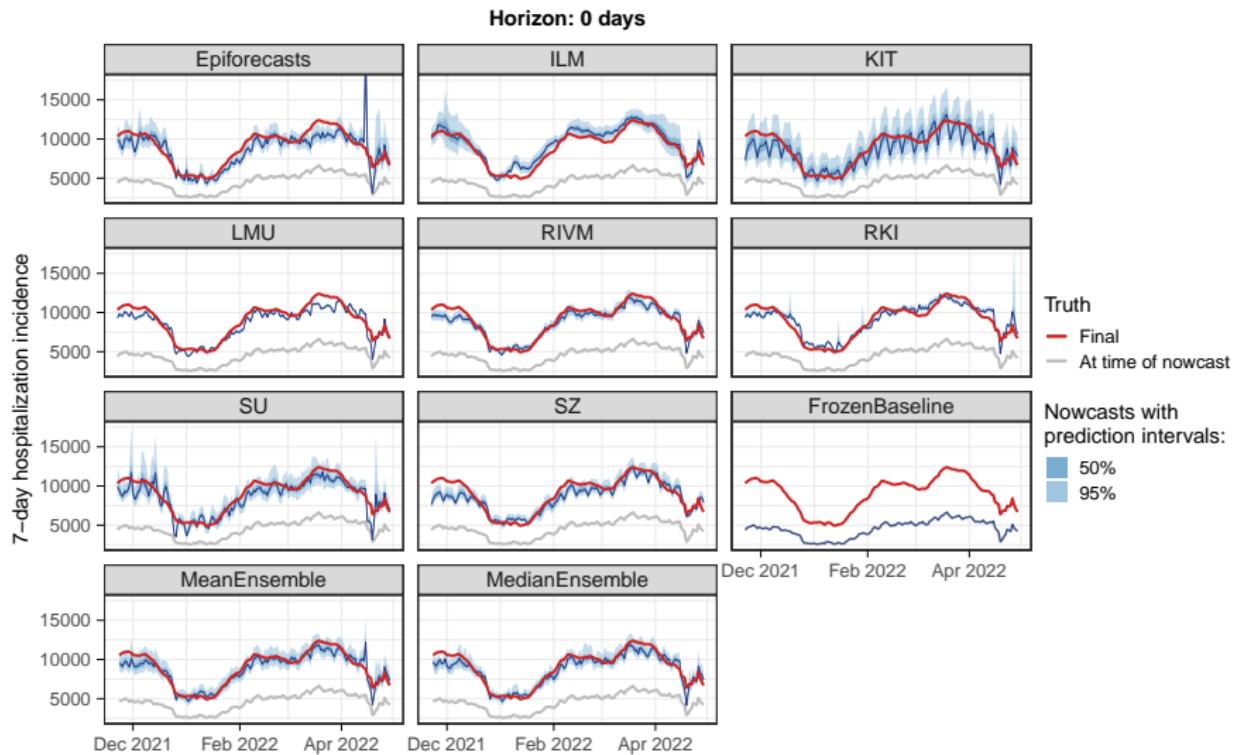
Spot problematic cases and identify reasons (III)



- ▶ In April 2022, the Easter break led to a considerable lengthening of reporting delays.
- ▶ Consequently, the true number of hospitalizations was **strongly underestimated** in real time.

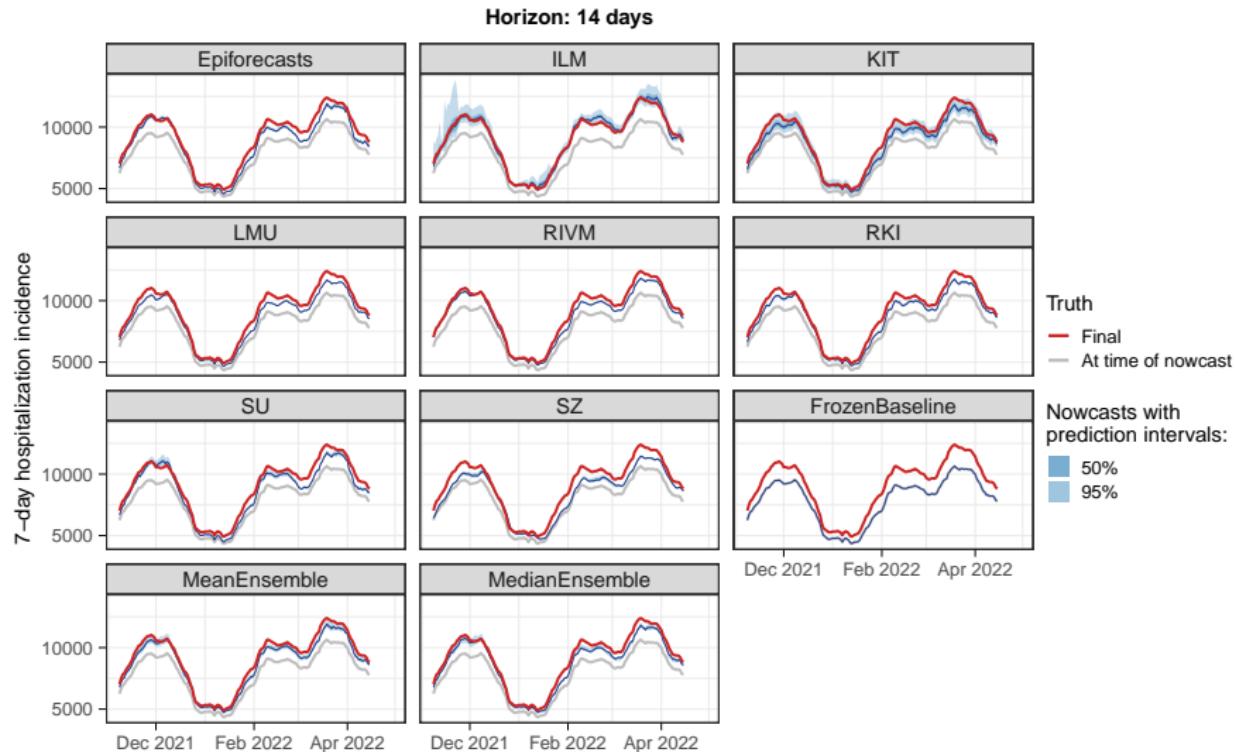
Diagnostic 2: Summary plots for fixed nowcast horizon

- ▶ This display makes systematic deviations easier to discern.



Diagnostic 2: Summary plots for fixed nowcast horizon

- ▶ This display makes systematic deviations easier to discern.

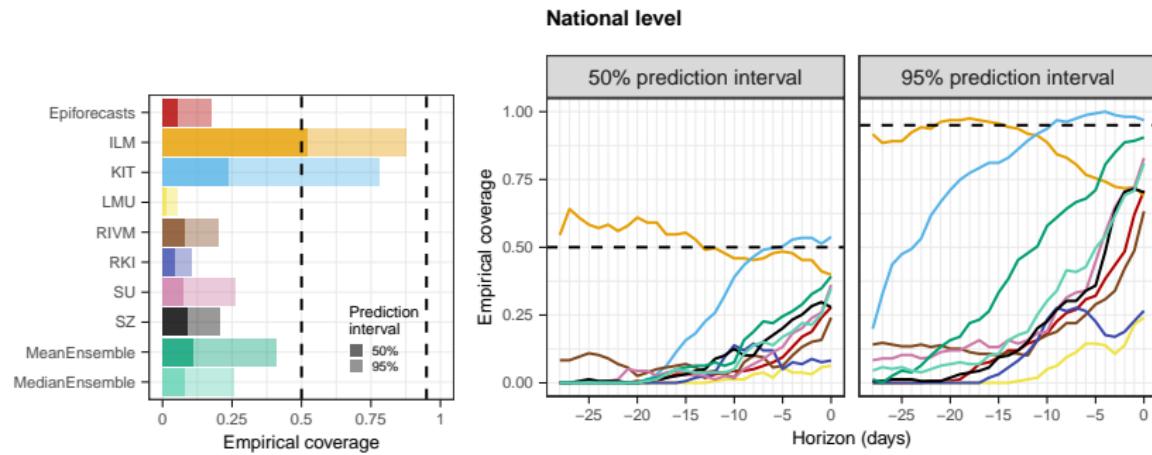


Importance of spotting problematic cases

- ▶ Unusual patterns / reporting quirks will mess up nowcasts when they occur – this is difficult to prevent.
- ▶ But they may even mess up things later – because the model fits can be perturbed when including these data.
- ▶ It may then be useful to exclude certain time periods in future analyses.
- ▶ Keeping track of these instances will also help you anticipate moments when nowcasts can't be trusted.

Diagnostic 3: Coverage percentages

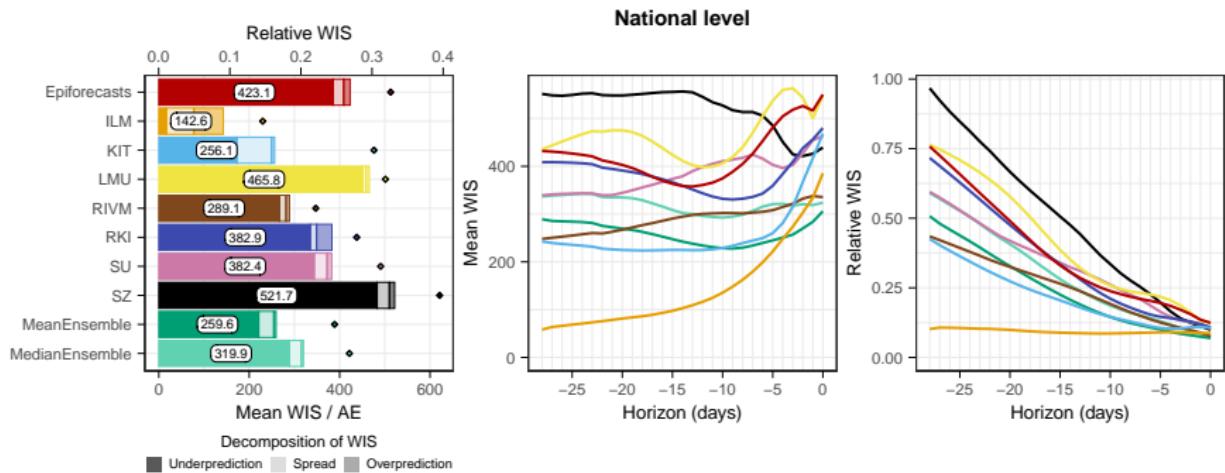
- ▶ How often do nowcast intervals contain true values?



- ▶ In this example all methods suffered from rather strong undercoverage (i.e., uncertainty intervals were too narrow).

Diagnostic 4: Comparative evaluation based on scores

- ▶ If different models are available we can compare the average “error” different models make.
- ▶ A commonly used score (\approx error measure) is the *weighted interval score* (lower values are better).



Bracher et al (2021),

<https://doi.org/10.1371/journal.pcbi.1008618>

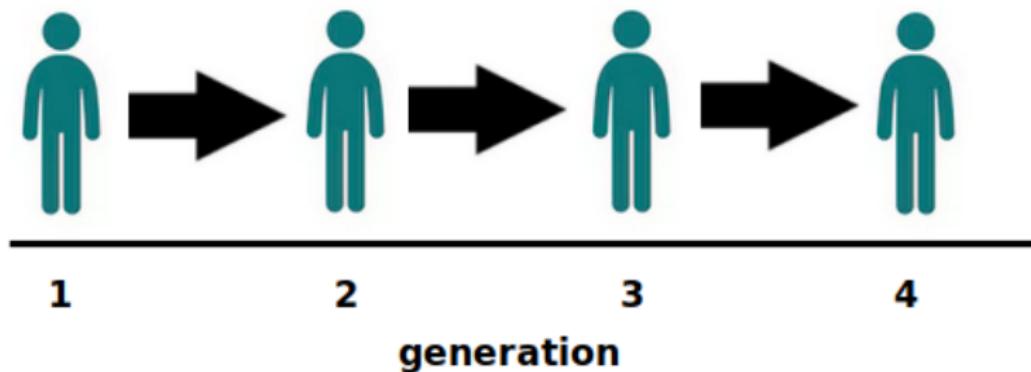
Takeaways from our comparative study

- ▶ All compared methods yielded practically useful nowcasts.
 - ▶ even quite simple approaches worked acceptably well.
- ▶ But they issued too narrow nowcast intervals which did not cover the true value as often as intended.
- ▶ The best results were achieved by averaging nowcasts from different models together (“ensembling”).

Part 2: Estimating effective reproductive numbers

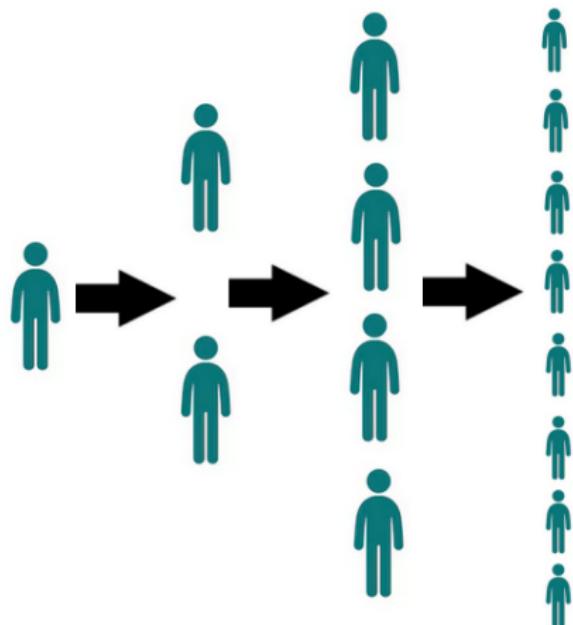
Intuition: the reproductive number R

- ▶ The reproductive number describes the **average number of secondary cases an infectious person generates.**
- ▶ Example: $R = 1 \Rightarrow$ the outbreak stagnates.



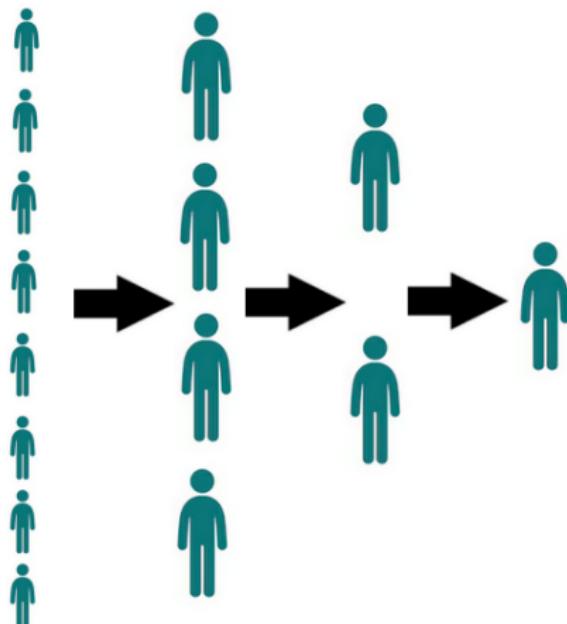
Intuition: the reproductive number R

- ▶ The reproductive number describes the **average number of secondary cases an infectious person generates.**
- ▶ Example: $R = 2 \Rightarrow$ the outbreak grows.



Intuition: the reproductive number R

- ▶ The reproductive number describes the **average number of secondary cases an infectious person generates**.
- ▶ Example: $R = 0.5 \Rightarrow$ the outbreak declines.



Some basic notions

- ▶ If $R > 1$, the epidemic grows exponentially, if $R < 1$ it declines.
- ▶ R depends on
 - ▶ the number of contacts a person has per time, c
 - ▶ the probability of transmission given contact, p
 - ▶ duration of infectiousness, D
 - ▶ the fraction of the population that is susceptible, s

and a simple model is

$$R_0 = c \times p \times D \times s.$$

Comparing subsequent generations

- ▶ To facilitate intuition, graphical displays often contain individual infectives arranged in generations.
- ▶ In some settings, like early phases of propagated outbreaks, generations can be discerned (roughly) from the epidemic curve.
- ▶ Example: Measles Cases by Date of Onset in Aberdeen, South Dakota, 1971

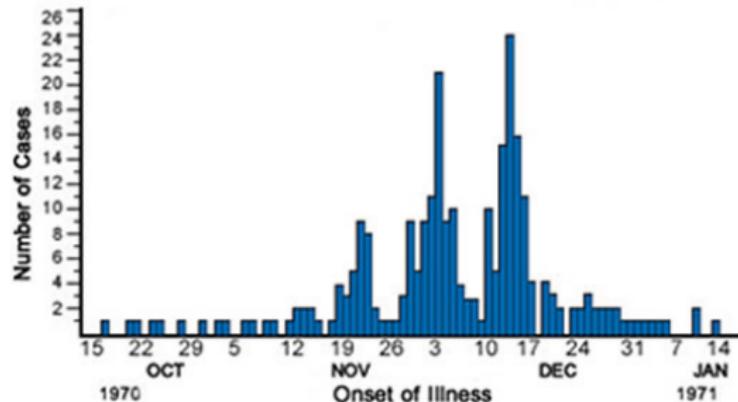


Figure source: US CDC,

<https://www.cdc.gov/training/quickearns/epimode/>

Generations and calendar time

- ▶ Usually we cannot observe infection events nor who infected whom – **we only observe numbers of infected and symptom onset times** (or some similar date).
- ▶ Reminder on terminology:

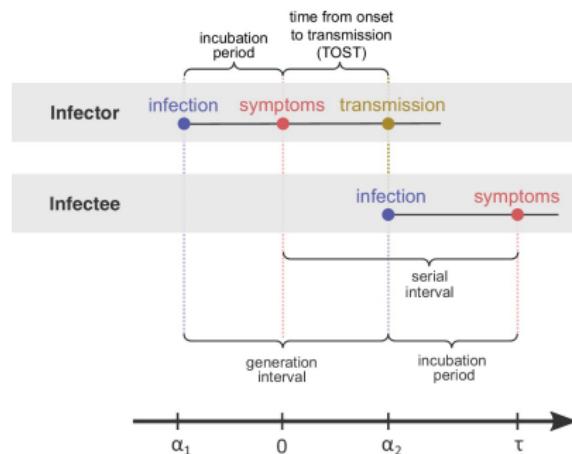


Figure from Sender et al (2022) eLife

<https://doi.org/10.7554/eLife.79134>

The renewal equation

- We need a model on the **population level and using calendar time**. A common approach is the **renewal equation**:

$$\underbrace{\mathbb{E}(X_t)}_{\text{expected infections on day } t} = R \times \sum_{d=1}^D \underbrace{X_{t-d}}_{\# \text{ infections at } t-d} \times \underbrace{w_d}_{\text{probability that generation time is } d}$$

- (w_1, w_2, \dots, w_D) is the generation time / serial interval distribution.
- I is the assumed maximum relevant generation time.
- Side note: We gloss over the distinction generation times / serial intervals, but there are some subtleties (see Britton and Scalia Tomba, <https://doi.org/10.1098/rsif.2018.0670>)

The many types of R values (I)

- ▶ Remember: R depends on
 - ▶ the number of contacts a person has per time, c
 - ▶ the probability of transmission given contact, p
 - ▶ duration of infectiousness, D
 - ▶ the fraction of the population that is susceptible, s
- ▶ c, p, D, s can vary over time, space, populations and the course of an outbreak
- ▶ R is thus **not an intrinsic property of a pathogen.**

The many types of R values (II)

- ▶ There are thus **many types of R-values** which are appropriate in different settings:
 - ▶ **basic** reproductive number R_0
 - ▶ time-varying **effective** reproductive number R_t
 - ▶ case reproductive number
 - ▶ instantaneous reproductive number
 - ▶ Other concepts:
 - ▶ “Angular reproduction numbers”
 - ▶ “Risk-averse reproductive numbers”
 - ▶ ...

The basic reproductive number R_0

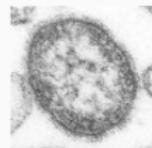
- ▶ The **basic reproductive number R_0** ("R naught") is the average number of secondary cases one infectious person generates **in a fully susceptible population**.
- ▶ Implicit assumption: "normal" behaviour, contact patterns.
 - ▶ i.e., we would set $s = 1$.
- ▶ R_0 determines whether a disease has a chance of establishing itself in a population.
- ▶ R_0 is relevant e.g., to determine which fraction of a population needs to be vaccinated to control an epidemic (textbook formula: $1 - 1/R_0$).
- ▶ Numerous estimation methods for different data types exist.
 - ▶ e.g., R package R0:
<https://cran.r-project.org/web/packages/R0/R0.pdf>

Empirical values of R_0 ^[citation needed]



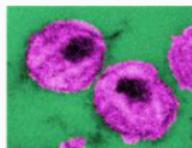
Malaria

R_0 10-100



Measles

R_0 5-18



HIV

R_0 2-12

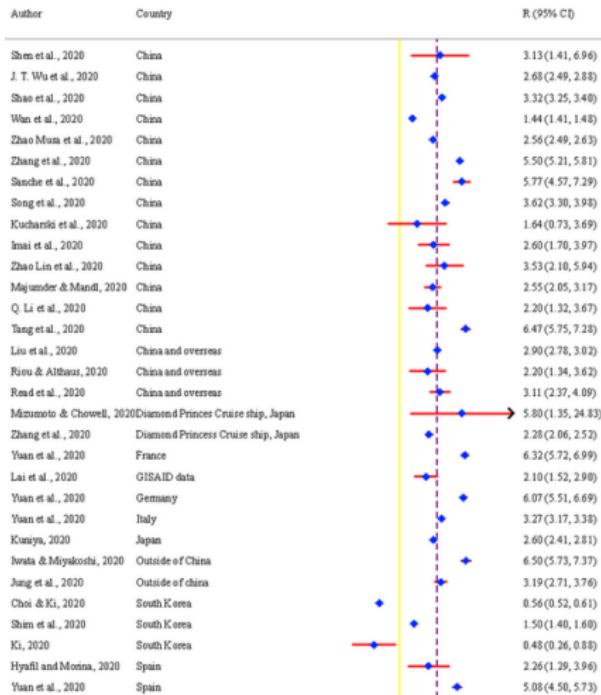


Ebola

R_0 1.5-2.5

Slide by S. Funk

R_0 of a pathogen is not a constant and not directly measurable



- ▶ Billah et al (2020) provide an overview of estimated R_0 of COVID-19 in various countries.
- ▶ These vary widely and many of them are associated with considerable uncertainty.

The effective reproductive number R_t

- ▶ As an outbreak unfolds, the effective reproductive number “at time t ” is a useful quantity to monitor.
- ▶ Example: R_t of COVID-19 in Germany

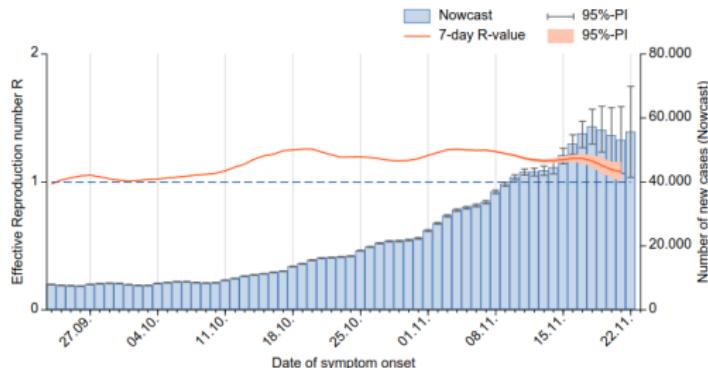
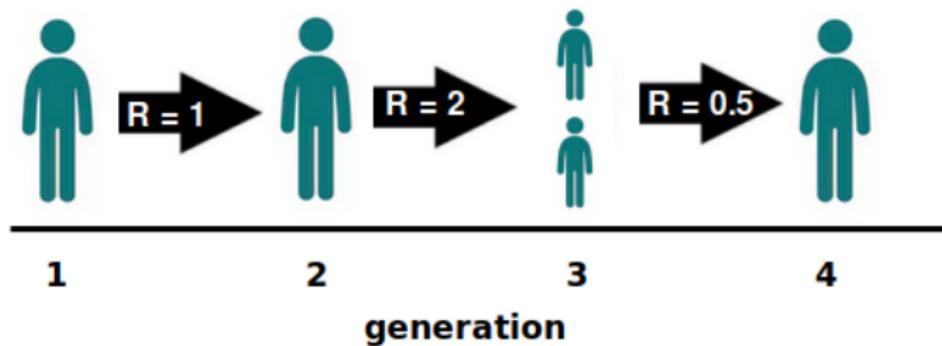


Figure 3: Estimated 7-day R-value (in orange) over the last 60 days, against the background of estimated number of COVID-19 cases according to illness onset [as of 26/11/2021, 12 AM, taking into account cases up to 22/11/2021].

- ▶ Recommended overview papers:
 - ▶ Gostic et al (2020),
<https://doi.org/10.1371/journal.pcbi.1008409>
 - ▶ White et al (2021), <https://doi.org/10.1093/aje/kwaa211>
 - ▶ Brockhaus et al (2023 – disclaimer: plugging my own work here),
<https://doi.org/10.1371/journal.pcbi.1011653>

But what do we mean by “at time t ”?

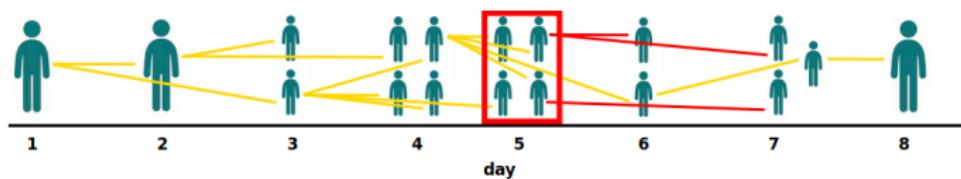
- If we think about R_t in terms of subsequent generations things are simple:



- But if we want to use data based on calendar time (as we usually do) things are less obvious.

Version 1: The case reproduction number R_t^{case}

- ▶ Intuition: the case reproduction number is **forward-looking** and compares the new infections on day t to the **following generation**.
- ▶ Example: $R_5 = 0.75$.



- ▶ Same problem as before: in practice we do not know who infected whom!

Version 1: The case reproduction number (II)

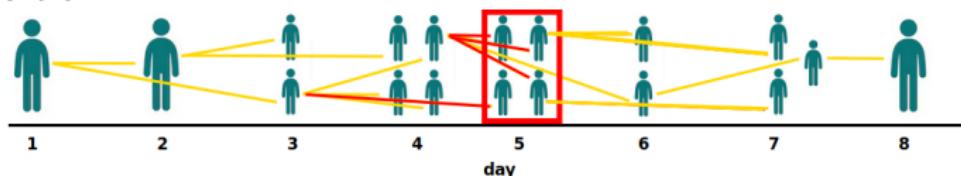
- To formalize R_t^{case} on the population level we use a **renewal equation** linking case incidences X_t on different days:

$$\underbrace{\mathbb{E}(X_t)}_{\text{expected infections on day } t} = \sum_{i=1}^D \underbrace{X_{t-d}}_{\# \text{ infections at } t-d} \times \underbrace{R_{t-d}^{\text{case}}}_{\text{"contagiousness" of these infectives}} \times \underbrace{w_d}_{\text{probability that generation time is } d}$$

- This is usually combined with a parametric distributional assumption for X_t (e.g., a Poisson distribution).
- implicit assumption: “contagiousness” of a given infective on day t depends on
 - how many days d ago they were infected
 - a value R_{t-d}^{case} **specific to their day of infection** $t - d$.

Version 2: The instantaneous reproduction number R_t^{inst}

- ▶ The instantaneous reproduction number is **backward-looking** and compares the new infections on day t to the **previous generation**.
- ▶ Intuition:



- ▶ Corresponding **renewal equation**:

$$\underbrace{\mathbb{E}(X_t)}_{\text{expected infections on day } t} = \underbrace{R_t^{\text{inst}}}_{\text{"contagiousness" at time } t} \times \sum_{d=1}^D \underbrace{X_{t-d}}_{\# \text{ infections at } t-d} \times \underbrace{w_d}_{\text{probability that generation time is } d}$$

- ▶ implicit assumption: “contagiousness” of a given infective on day t depends on
 - ▶ how many days i ago they were infected
 - ▶ a value R_t^{case} **specific to the current day** t .

The two renewal equations

- ▶ Forward-looking equation for case reproduction number:

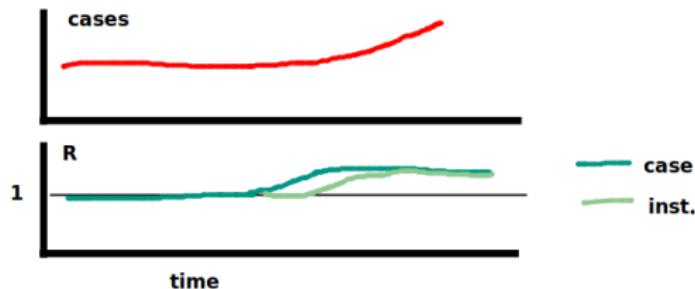
$$\underbrace{\mathbb{E}(X_t)}_{\text{expected infections on day } t} = \sum_{d=1}^D \underbrace{X_{t-d}}_{\# \text{ infections at } t-d} \times \underbrace{R_{t-d}^{\text{case}}}_{\text{"contagiousness" of these infectives}} \times \underbrace{w_d}_{\text{probability that generation time is } d}$$

- ▶ Backward-looking equation for instantaneous reproduction number:

$$\underbrace{\mathbb{E}(X_t)}_{\text{expected infections on day } t} = \underbrace{R_t^{\text{inst}}}_{\text{"contagiousness" at time } t} \times \sum_{d=1}^D \underbrace{X_{t-d}}_{\# \text{ infections at } t-d} \times \underbrace{w_d}_{\text{probability that generation time is } d}$$

Case vs instantaneous reproduction number (I)

- ▶ The case and instantaneous reproduction numbers are shifted by roughly the mean generation time m .
 - ▶ $R_{t-m}^{\text{case}} \approx R_t^{\text{inst}}$, i.e. R_{t-m}^{case} is “shifted left”.
 - ▶ The case reproduction number changes **before** changed dynamics case numbers become visible.
 - ▶ The instantaneous reproduction number changes later.



- ▶ So at time t knowing R_t^{case} would be “more interesting” than knowing R_t^{inst} .

Case vs instantaneous reproduction number (I)

- ▶ However, only the instantaneous reproduction number can be estimated properly in real time.
- ▶ Reason: estimation of the case reproduction number R_t^{case} requires data from after time t (as the renewal equation is “forward-looking”).
- ▶ Real-time monitoring of the effective reproductive number is thus mostly done using instantaneous reproduction numbers.

Estimating the instantaneous R_t with the R package EpiEstim (the “Cori method”)

- ▶ Ingredients:
 - ▶ Time series of daily cases by symptom onset.
 - ▶ Prior knowledge on generation time distribution (w_1, \dots, w_D).
- ▶ The following is based on the package documentation from <https://cran.r-project.org/web/packages/EpiEstim/vignettes/demo.html>
- ▶ The same method is also implemented as a web app (<https://shiny.dide.imperial.ac.uk/epiestim/> and an Excel spreadsheet (<https://shiny.dide.imperial.ac.uk/epiestim/>)).
- ▶ For an intro by Anne Cori see https://www.youtube.com/watch?v=T_7AHMB8pkQ

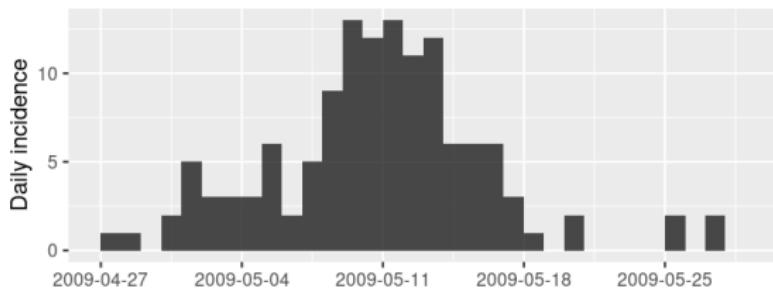
Example from EpiEstim documentation (II)

```
library(EpiEstim)
library(incidence) # for plotting

?EpiEstim
vignette("EpiEstim")
?EpiEstim::estimate_R

## load data on pandemic flu in a school in 2009
data("Flu2009")

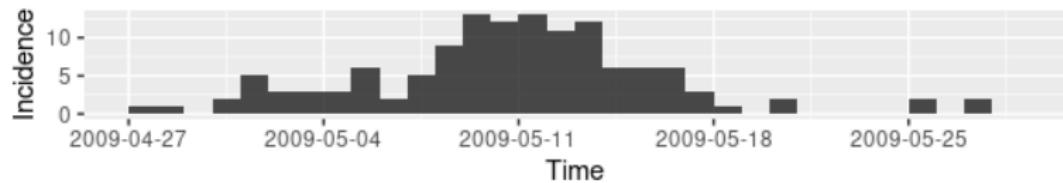
plot(as.incidence(Flu2009$incidence$I,
                  dates = Flu2009$incidence$dates))
```



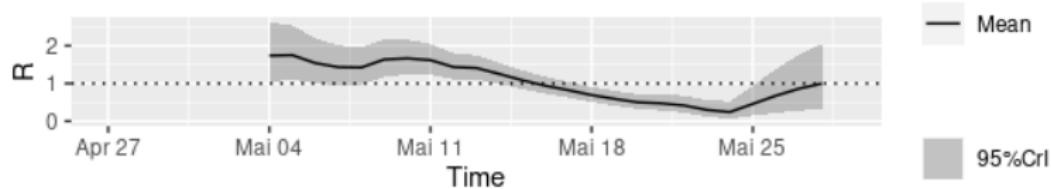
Estimation with suggested settings

```
# Version 1: estimating R for weekly windows (default),
# using a parametric serial interval distribution with
# mean 2.6, standard deviation 1.5
res_parametric_si <- estimate_R(Flu2009$incidence,
                                  method="parametric_si",
                                  config = make_config(list(
                                    mean_si = 2.6,
                                    std_si = 1.5)))
plot(res_parametric_si)
```

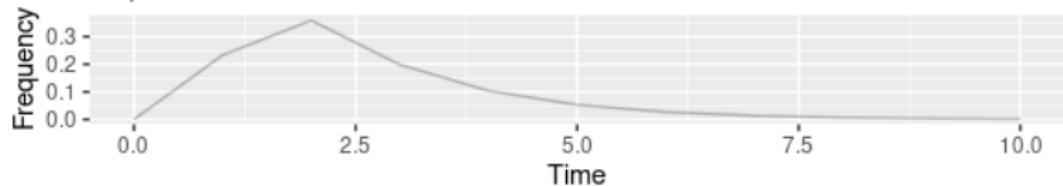
Epidemic curve



Estimated R



Explored SI distribution



What did we do here?

- ▶ EpiEstim fits the model

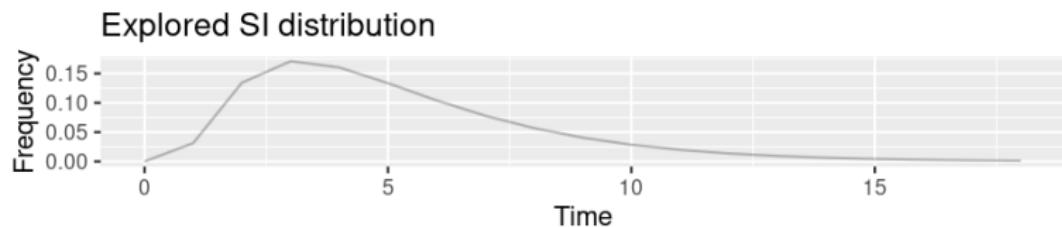
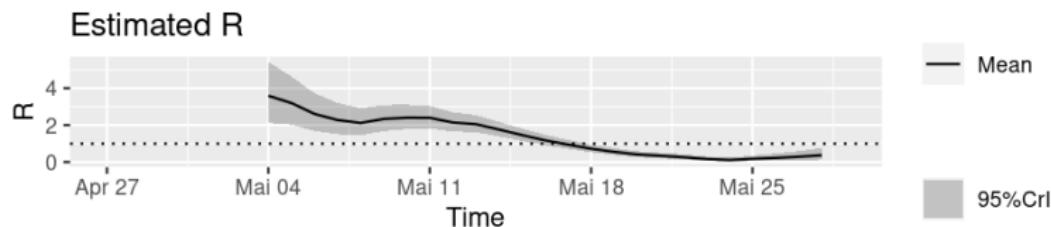
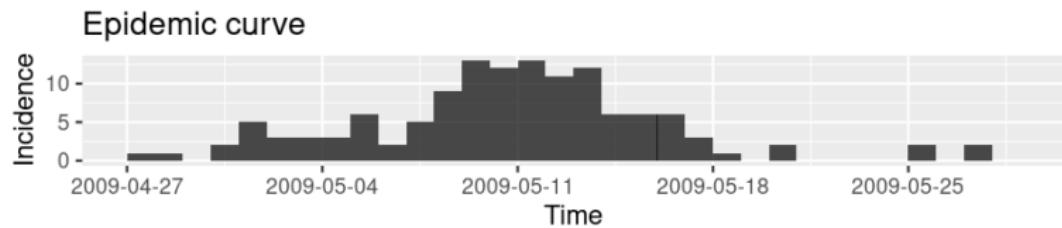
$$\underbrace{\mathbb{E}(X_t)}_{\text{expected infections on day } t} = \underbrace{R_t^{\text{inst}}}_{\text{"contagiousness" at time } t} \times \sum_{d=1}^D \underbrace{X_{t-d}}_{\# \text{ infections at } t-d} \times \underbrace{w_d}_{\text{probability that generation time is } d}$$

to **time windows** which by default are **7 days**.

- ▶ Technically, estimates are assigned to **the last day** of the respective interval. But they should be seen to **refer to the entire interval!**
- ▶ The **generation time (or serial interval) distribution** is parametric (a shifted, discretized Gamma distribution) with **mean and standard deviation provided by the user**.
- ▶ Technical remark: EpiEstim uses a conditional Poisson assumption for X_t (which could be criticized as it ignores overdispersion).

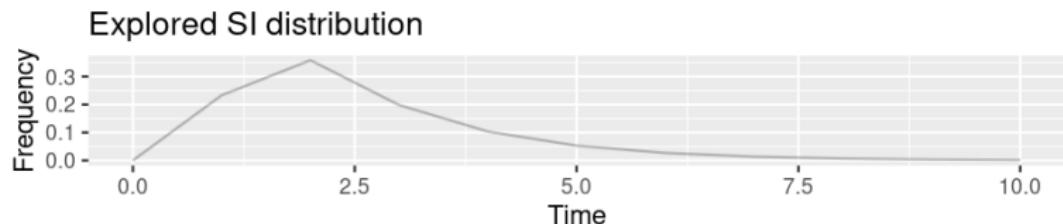
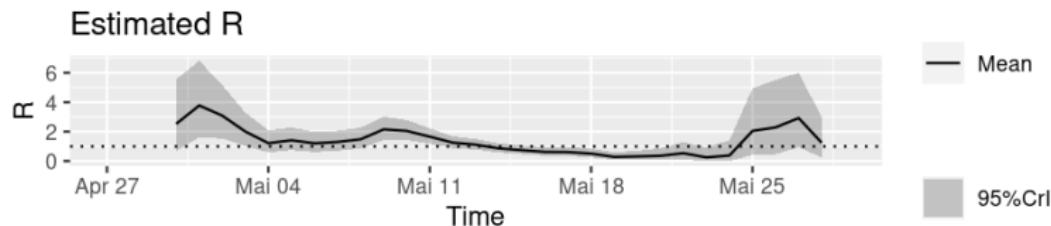
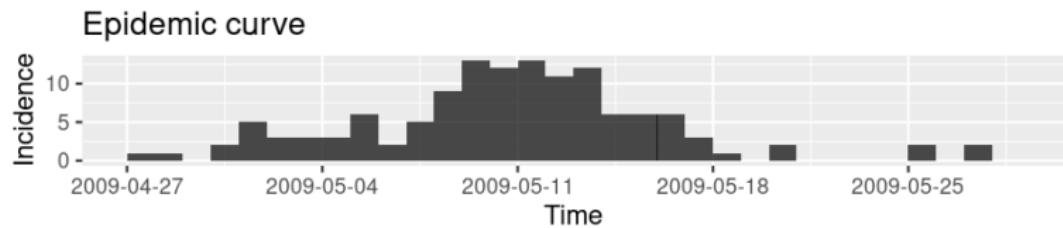
Increasing the mean serial interval

```
# Version 2: estimating R for weekly windows (default),
# using a parametric serial interval distribution with
# mean 5.2, standard deviation 3
res_parametric_si2 <- estimate_R(Flu2009$incidence,
                                    method="parametric_si",
                                    config = make_config(list(
                                        mean_si = 5.2,
                                        std_si = 3)))
)
plot(res_parametric_si2)
```

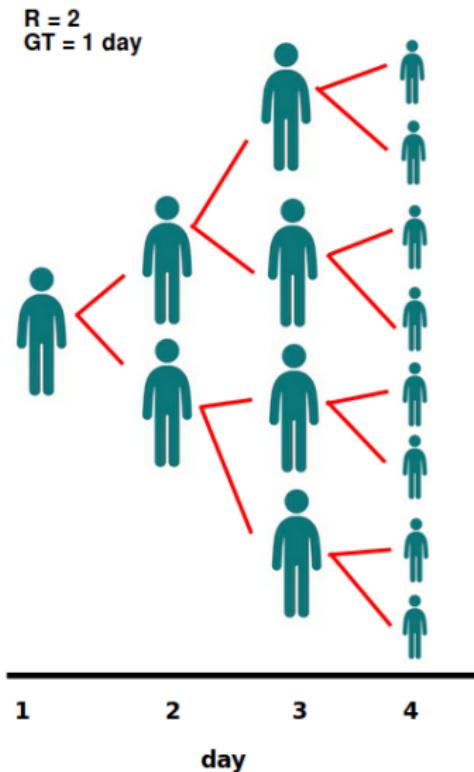


Shortening the mean serial interval

```
# Version 3: estimating R for weekly windows (default),
# using a parametric serial interval distribution with
# mean 1.3, standard deviation 0.75
res_parametric_si3 <- estimate_R(Flu2009$incidence,
                                    method="parametric_si",
                                    config = make_config(list(
                                        mean_si = 1.3,
                                        std_si = 0.75)))
)
plot(res_parametric_si3)
```

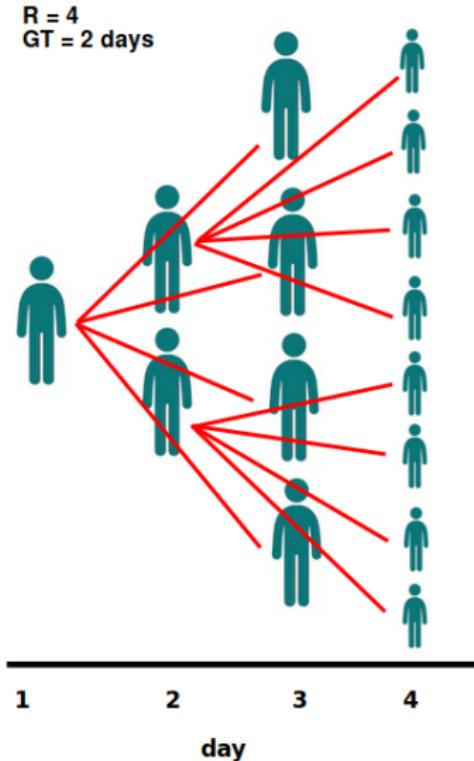


Role of the generation time / serial interval distribution



- ▶ The specification of the generation time distribution is a crucial step in the estimation of R_t :
 - ▶ longer average GTs lead to more extreme estimated R_t values (further away from 1, stronger amplitude).
 - ▶ shorter average GTs lead to less extreme R_t values.
- ▶ Details: Wallinga and Lipsitch (2007),
<https://doi.org/10.1098/rspb.2006.3754>

Role of the generation time / serial interval distribution

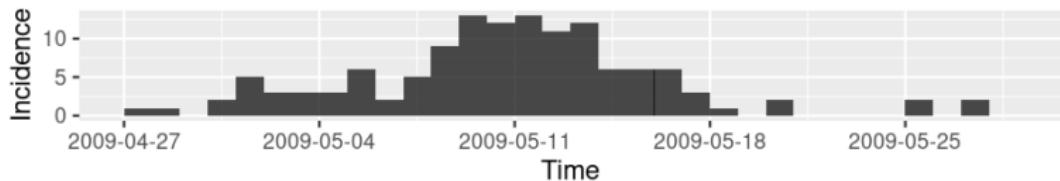


- ▶ The specification of the generation time distribution is a crucial step in the estimation of R_t :
 - ▶ longer average GTs lead to more extreme estimated R_t values (further away from 1, stronger amplitude).
 - ▶ shorter average GTs lead to less extreme R_t values.
- ▶ Details: Wallinga and Lipsitch (2007),
<https://doi.org/10.1098/rspb.2006.3754>

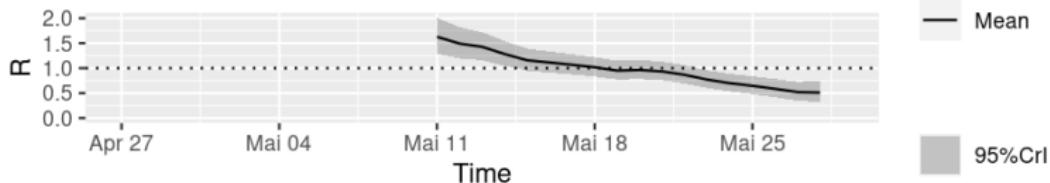
Increasing the estimation windows

```
# Version 4: estimating R for 14-day windows,
# using a parametric serial interval distribution with
# mean 2.6, standard deviation 1.5
T <- nrow(Flu2009$incidence)
t_start <- seq(2, T - 13)
# starting at 2 as conditional on the past observations
t_end <- t_start + 13
# adding 2 to get 3-day windows as bounds included in window
res_parametric_si4 <- estimate_R(Flu2009$incidence,
                                    method="parametric_si",
                                    config = make_config(list(
                                        t_start = t_start,
                                        t_end = t_end,
                                        mean_si = 2.6,
                                        std_si = 1.5)))
)
plot(res_parametric_si4)
```

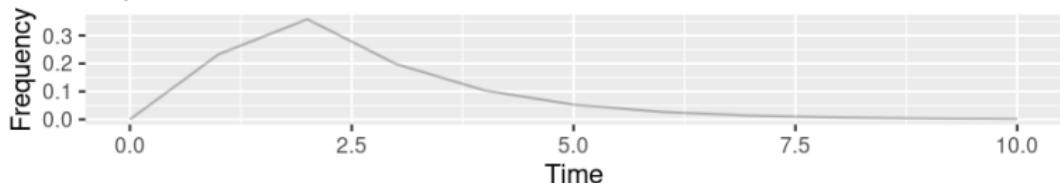
Epidemic curve



Estimated R



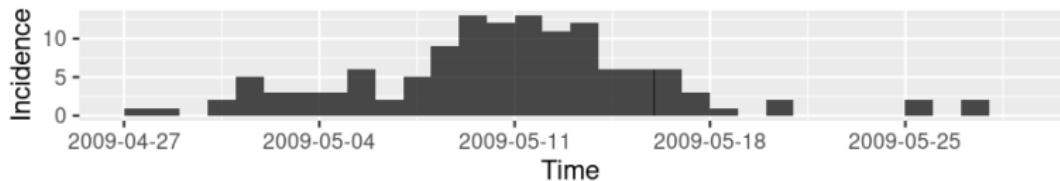
Explored SI distribution



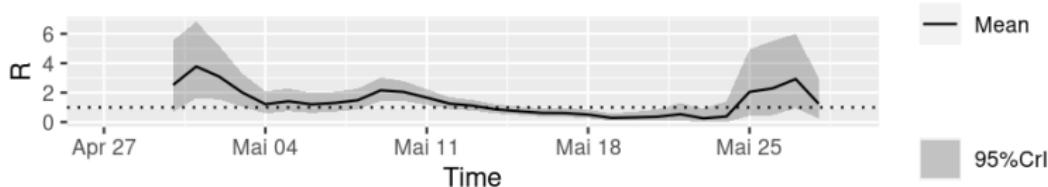
Shortening the estimation windows

```
# Version 5: estimating R for three-day windows,
# using a parametric serial interval distribution with
# mean 2.6, standard deviation 1.5
T <- nrow(Flu2009$incidence)
t_start <- seq(2, T - 2) # starting at 2 as conditional
                         # on the past observations
t_end <- t_start + 2 # adding 2 to get 3-day windows as
                     # bounds included in window
res_parametric_si5 <- estimate_R(Flu2009$incidence,
                                    method="parametric_si",
                                    config = make_config(list(
                                        t_start = t_start,
                                        t_end = t_end,
                                        mean_si = 2.6,
                                        std_si = 1.5)))
)
plot(res_parametric_si5)
```

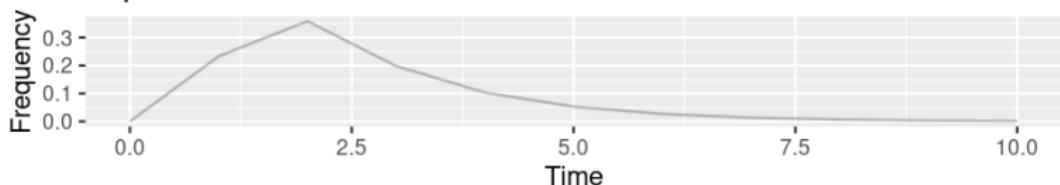
Epidemic curve



Estimated R



Explored SI distribution



Role of the estimation window

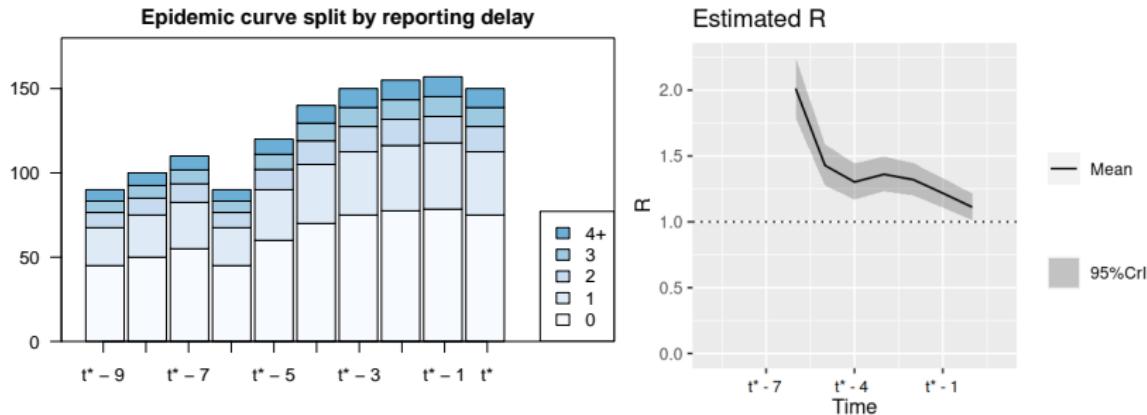
- ▶ Longer estimation windows lead to **smoother estimated curves**.
- ▶ Shorter estimation windows lead to **more wiggly estimated curves**.
- ▶ There is a tradeoff between detail and stability of the estimation.
- ▶ The interpretation of results changes as the values refer to **differently defined time periods**.
- ▶ Choosing **7-day windows** (the default) is a good idea as it also removes artefacts from weekly reporting patterns.
- ▶ Note: More sophisticated methods can avoid the specification of window sizes.

Some practical remarks

- ▶ In the early phases of an outbreak (when numbers are low) estimates can be somewhat unreliable.
- ▶ Estimated R_t values can be unreasonably large if surveillance efforts are increased / initiated at some point.
- ▶ Due to its statistical assumptions, the EpiEstim method may tend to produce too narrow uncertainty intervals, especially in settings with high incidence values (see Brockhaus et al 2023).

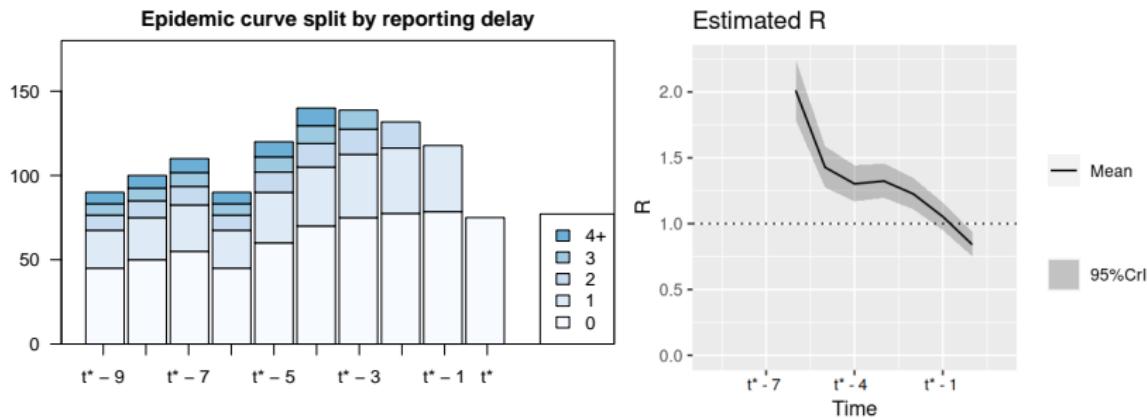
Nowcasting and effective reproductive numbers

- ▶ Nowcasting and R_t estimation are intimately linked.
Returning to the example from Slide ??.
- ▶ Estimates using incomplete real-time data:



Nowcasting and effective reproductive numbers

- ▶ Nowcasting and R_t estimation are intimately linked.
Returning to the example from Slide ??.
- ▶ Estimates using incomplete real-time data:



Nowcasting and effective reproductive numbers (continued)

- ▶ In practice, real-time R_t estimation requires an additional nowcasting step.
- ▶ Two-step procedures are easy to implement, but do not forward uncertainty from nowcasting step to R_t estimation.
- ▶ There are more methods which perform both steps jointly.
 - ▶ For an application of epinowcast see here:
[https://package.epinowcast.org/dev/articles/
single-timeseries-rt-estimation.html](https://package.epinowcast.org/dev/articles/single-timeseries-rt-estimation.html)

Advanced methods to estimate R_t

- ▶ There is an ever-growing number of R_t estimation methods:
 - ▶ EpiNow2: <https://epiforecasts.io/EpiNow2/>
 - ▶ EpiLPS:
<https://doi.org/10.1371/journal.pcbi.1010618>
 - ▶ EstimateR: <https://bmcbioinformatics.biomedcentral.com/articles/10.1186/s12859-023-05428-4>
 - ▶ ...

Know your data

- ▶ In practice, epidemiological data do not contain dates of infection, but dates of symptom onset, GP visit or something else.
- ▶ The data thus give a delayed (and blurred) picture of infection dynamics.
- ▶ This delay needs to be taken into account when interpreting R_t , especially when relating it to intervention measures or other events.
- ▶ Estimates need to be shifted by the assumed delay between infection and reference date in the data set.

Why are different estimates of the effective reproductive number so different?

- ▶ Different groups of researchers can make quite different assumptions / specifications on
 - ▶ generation times
 - ▶ incubation periods
 - ▶ smoothing parameters
 - ▶ data sources
 - ▶ statistical methods
- ▶ This can lead to quite different estimates.
- ▶ It is a good idea to critically check the assumptions underlying published R_t estimates.

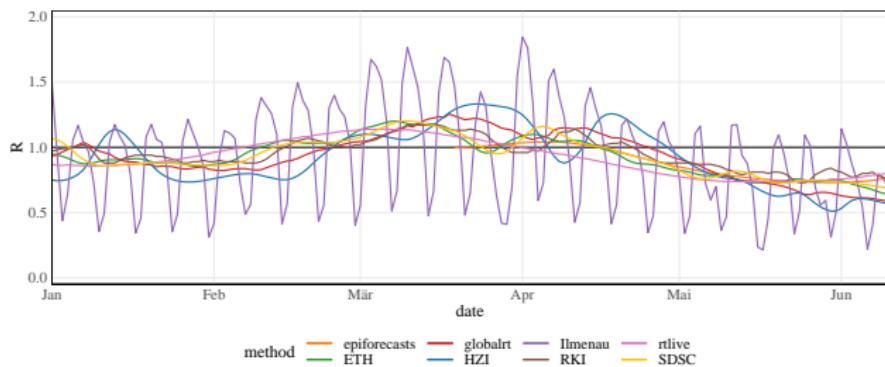
PLOS COMPUTATIONAL BIOLOGY

Why are different estimates of the effective reproductive number so different? A case study on COVID-19 in Germany

Elisabeth K. Brockhaus , Daniel Wolfram, Tanja Stadler, Michael Osthege, Tanmay Mitra, Jonas M. Littek, Ekaterina Krymova, Anna J. Klesen, Jana S. Huisman, Stefan Heyder, Laura M. Helleckes, Matthias an der Heiden, Sebastian Funk, Sam Abbott, Johannes Brächer 

Example: COVID-19 in Germany

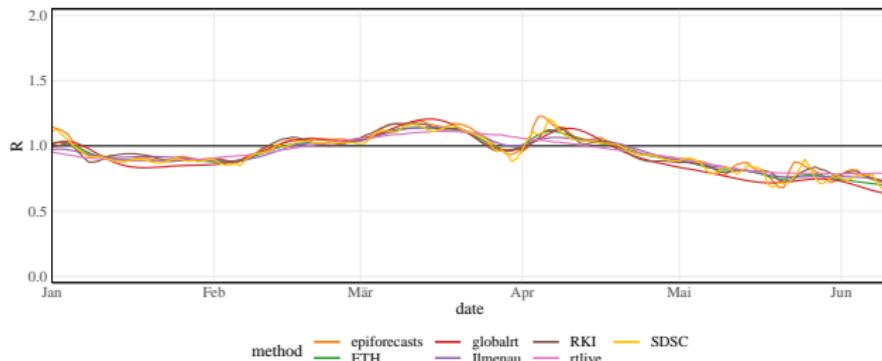
- ▶ Estimates as published by different research teams:



- ▶ Learning: Choice e.g., of generation time distribution is at least as influential as the statistical model used.

Example: COVID-19 in Germany

- ▶ Estimates from different statistical models after uniformization of auxiliary analytical choices:



- ▶ Learning: Choice e.g., of generation time distribution is at least as influential as the statistical model used.

Should we use growth rates rather than reproduction numbers?

- ▶ The daily epidemic growth rate r_t describes the dynamics of an epidemic via

$$\mathbb{E}(X_t) = X_{t-1} \times \exp(r_t) \approx (1 + r_t) \times X_{t-1}$$

- ▶ Unlike R_t , r_t can be estimated without assumptions on the generation time distribution.
- ▶ It is therefore sometimes argued that r_t rather than R_t should be used.
- ▶ Common reply: R_t and r_t are complementary in describing disease spread:
 - ▶ R_t describes the **strength** of spread: how much would we need to reduce contacts to stop the spread?
 - ▶ r_t describes the **speed** of spread: if the epidemic continues as currently, how long will it take until we reach a given number of daily cases?

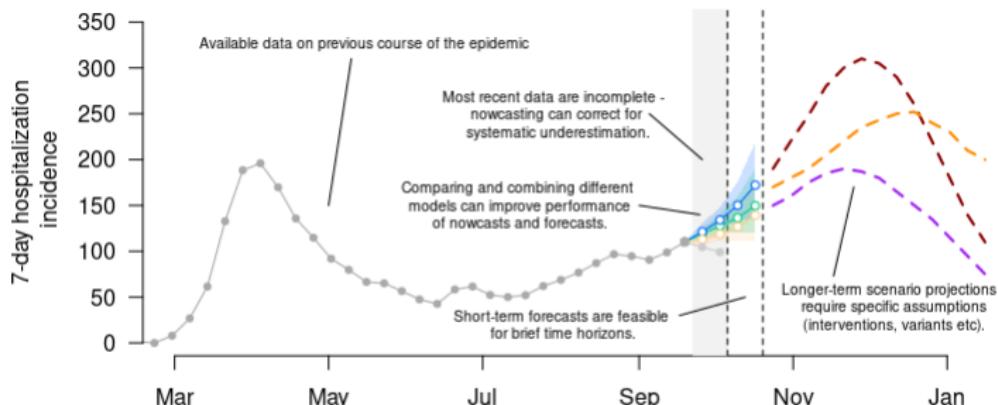
Part 3: Short-term forecasting

Short-term forecasting

- ▶ During an outbreak many questions require anticipation of future developments:
 - ▶ How many hospital beds are needed?
 - ▶ Where should ventilators be allocated?
 - ▶ How should vaccines be rolled out?
 - ▶ ...
- ▶ Consequently, an important task of epidemic modelling is prediction.
 - ▶ But modelling and forecasting are not synonymous – modelling has many other purposes (explanation, counterfactual reasoning, designing interventions...)
 - ▶ There is sometimes confusion about predictions, projections and scenarios.

Predictive epidemic modelling

- ▶ **Nowcasts** make statements about current trends based on partial data.
- ▶ **Short-term forecasts** are unconditional statements about the near future of an outbreak.
- ▶ **Scenarios or projections** make statements conditional on certain assumptions (e.g., on contacts, behaviour, biology).
- ▶ **Baseline projections** are projections under the assumptions that relevant aspects “remain the same”.



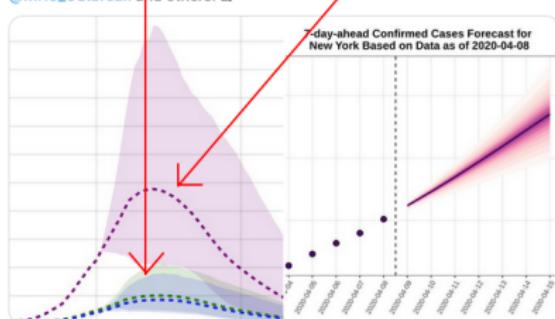
Different models can say very different things...



Nicholas G. Reich ✅ @reichlab · Apr 10

Replies to @reichlab

I've done some spot-checking of forecast accuracy of a few of the US models, and am planning a more thorough analysis of models that have been put out by @IHME_UW @Columbia @LosAlamosNatLab @alexvespi @MRC_Outbreak and others. 2/



Nicholas G. Reich ✅ @reichlab · Apr 10

These two #COVID19 models, both from very well-respected academic groups, are complete odds with one another. Their prediction intervals barely overlap! It highlights for me how little we really know about what will happen in the coming weeks. 7/

3

14

56

↑

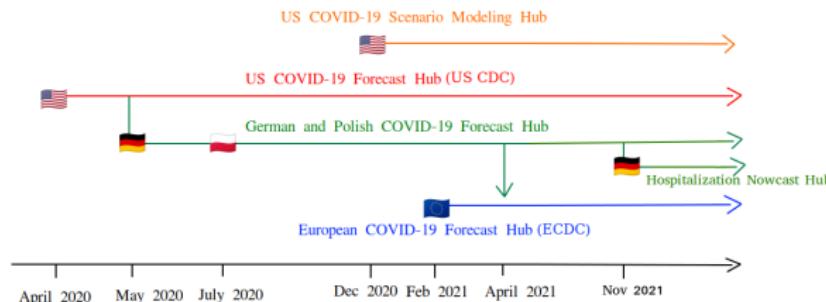
Collaborative forecasting: the Forecast Hub ecosystem

- ▶ Goal: An **open-science framework** to
 - ▶ make **standardized probabilistic** short-term forecasts of COVID-19 available in real time
 - ▶ **evaluate** them systematically and assess if potential of **combined** ensemble forecasts
 - ▶ **foster exchange** between disease modellers

Collaborative Hubs: Making the Most of Predictive Epidemic Modeling

Nicholas G. Reich, PhD, Justin Lessler, Sébastien Funk, PhD, Cécile Viboud, PhD, Alessandro Vespignani, PhD, Ryan J. Tilkshiran, PhD, Katriana Shea, PhD, Mélanie Schenke, PhD, Michael C. Runge, PhD, Ruth Rosenberg, PhD, Evan L. Ray, PhD, René Niehus, PhD, Helen C. Johnson, MRes, MSC, Michael A. Johansson, PhD, Harry Hochreiter, PhD, Lauren Gardner, MSE, PhD, Johannes Brächer, MS, PhD, Rebecca K. Borcherding, PhD, and Matthew Biggerstaff, ScD, MPH

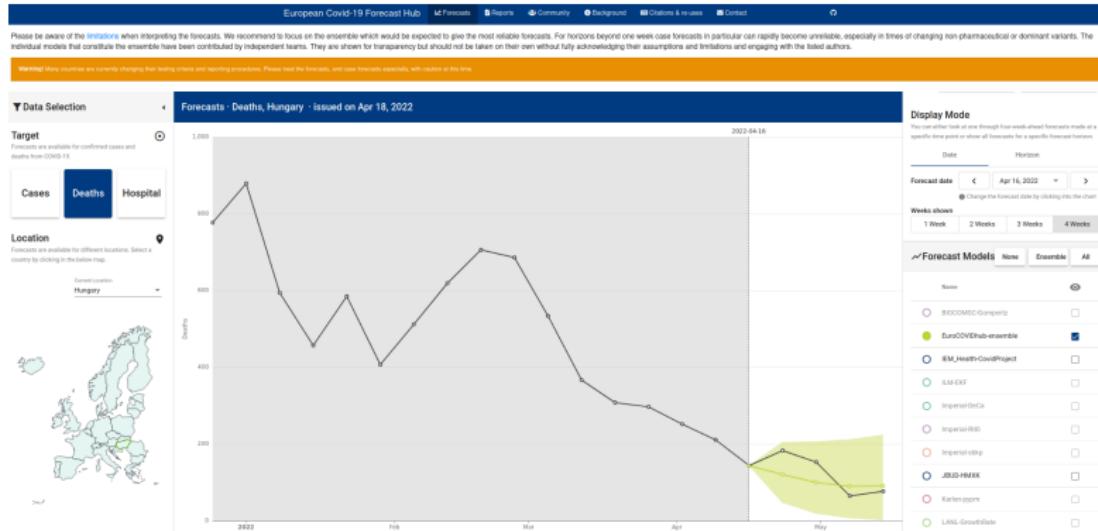
Reich et al (2022), American Journal of
Public Health



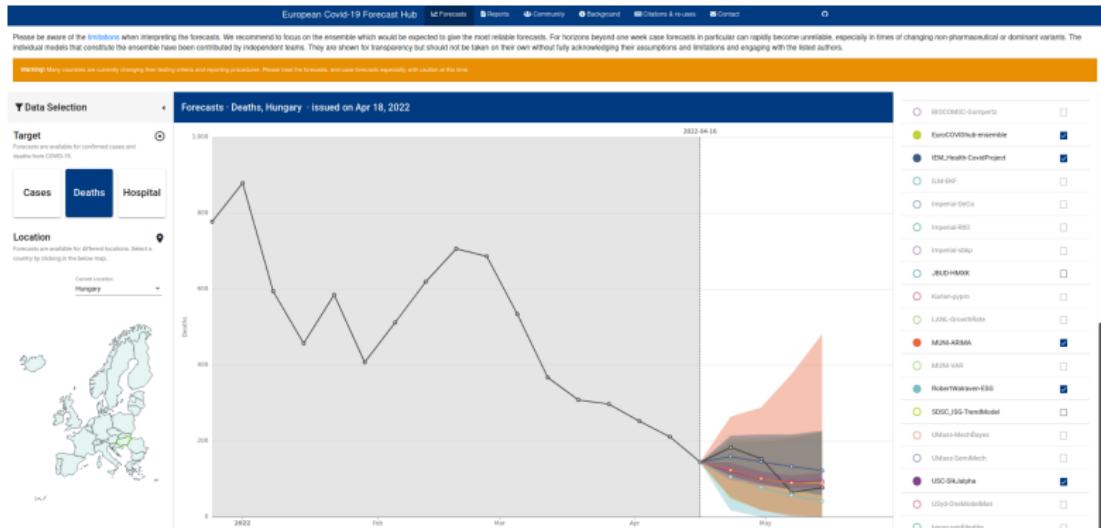
The public face (European COVID-19 Forecast Hub)



The public face (European COVID-19 Forecast Hub)

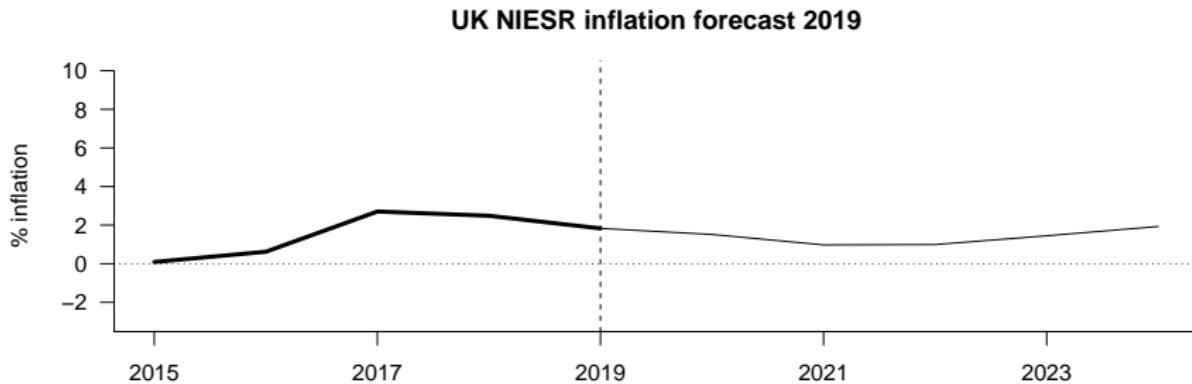


The public face (European COVID-19 Forecast Hub)



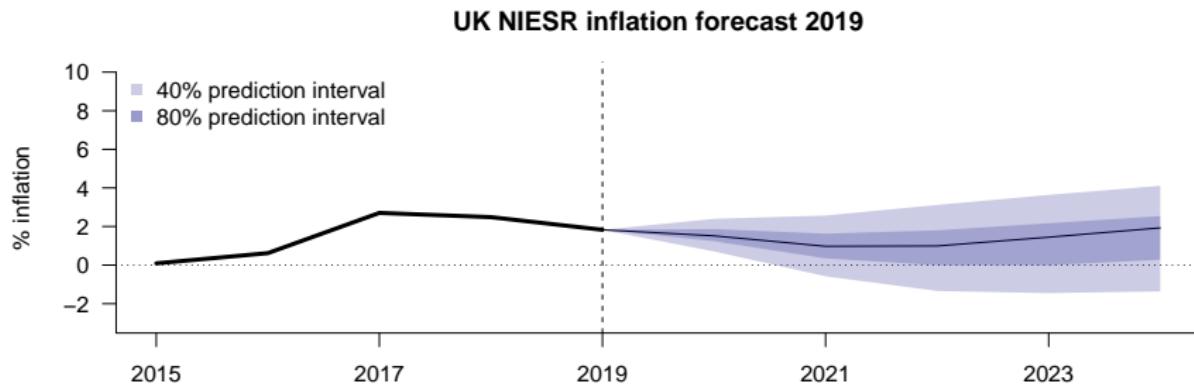
Why probabilistic forecasting?

- ▶ We do not only collect point predictions, but probabilistic forecasts.
- ▶ The characterization of remaining uncertainty is a central part of forecasting.
- ▶ Example: Inflation forecasts, National Institute for Economic and Social Research (UK):



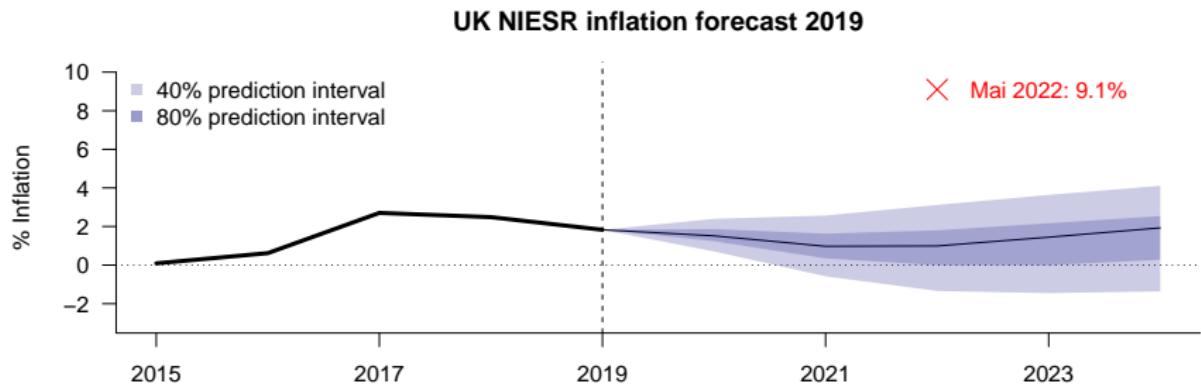
Why probabilistic forecasting?

- ▶ We do not only collect point predictions, but probabilistic forecasts.
- ▶ The characterization of remaining uncertainty is a central part of forecasting.
- ▶ Example: Inflation forecasts, National Institute for Economic and Social Research (UK):

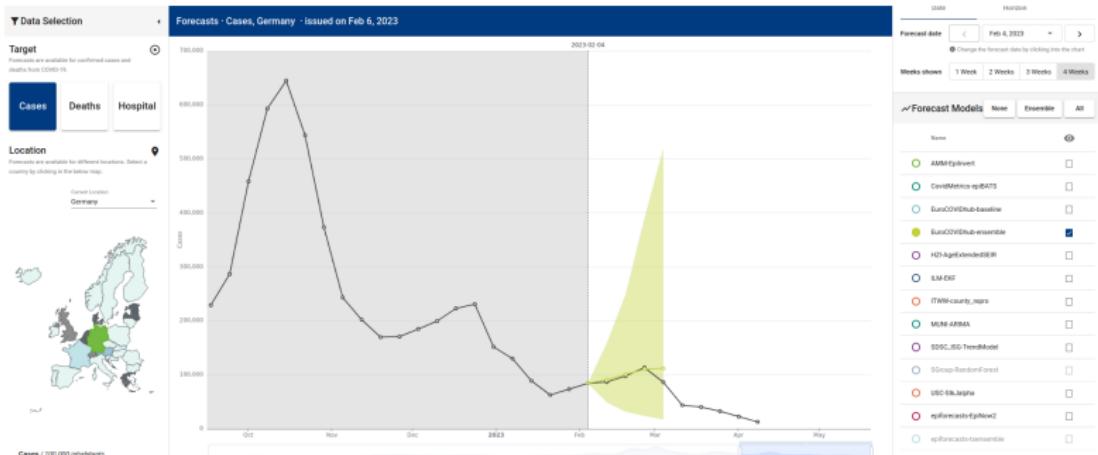


Why probabilistic forecasting?

- ▶ We do not only collect point predictions, but probabilistic forecasts.
- ▶ The characterization of remaining uncertainty is a central part of forecasting.
- ▶ Example: Inflation forecasts, National Institute for Economic and Social Research (UK):



Disease forecasts can be very uncertain!



Behind the scenes

Each Monday, teams submit forecasts in a standardized format to a public github repository (via PR):

- ▶ CSV files containing 23 predictive quantiles per target

The screenshot shows a GitHub repository page for 'covid19-forecast-hub-europe'. A specific commit titled 'actions-user Build ensemble' is highlighted. The commit message includes the text 'Build ensemble'. The commit was made by 'actions-user' and has been updated 5 days ago. It has 1 contributor. The file '2022-05-09-EuroCOVIDhub-ensemble.csv' is shown, containing 7683 lines (7681 sloc) and 472 KB. The file content is a CSV table with columns: forecast_date, target, target_end_date, location, type, quantile, and value. The data shows 7 rows of forecasts for various locations and targets on May 9, 2022.

1	forecast_date	target	target_end_date	location	type	quantile	value
2	2022-05-09	1 wk ahead inc case	2022-05-14	AT	point	NA	32137
3	2022-05-09	1 wk ahead inc case	2022-05-14	BE	point	NA	27929
4	2022-05-09	1 wk ahead inc case	2022-05-14	BG	point	NA	1899
5	2022-05-09	1 wk ahead inc case	2022-05-14	CH	point	NA	12597
6	2022-05-09	1 wk ahead inc case	2022-05-14	CY	point	NA	4161
7	2022-05-09	1 wk ahead inc case	2022-05-14	CZ	point	NA	4785

[https://github.com/covid19-forecast-hub-europe/
covid19-forecast-hub-europe](https://github.com/covid19-forecast-hub-europe/covid19-forecast-hub-europe)

- ▶ Open science: all output and codes publicly available under open licenses

Modelling approaches represented in the Forecast Hubs

A continuum of approaches

- ▶ **agent-based models**
- ▶ compartmental (SIR-type) models
- ▶ statistical (time series) models
- ▶ machine learning methods

- ▶ Human judgement methods

Detailed simulation at individual level:

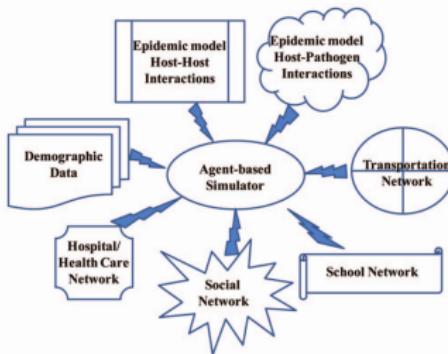


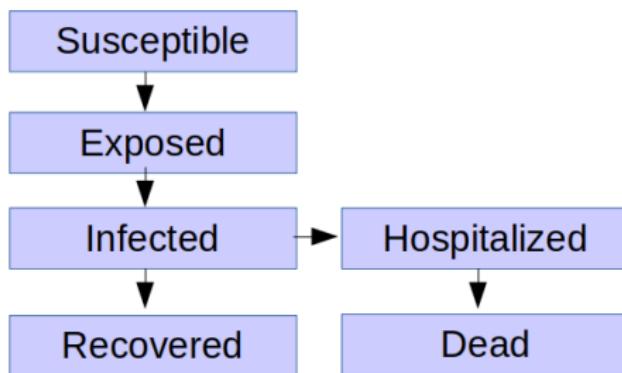
Figure from Siettos and Russo: Mathematical modeling of infectious disease dynamics. *Virulence* 4(4):295–306

Modelling approaches represented in the Forecast Hubs

A continuum of approaches

- ▶ agent-based models
- ▶ **compartmental
(SIR-type) models**
- ▶ statistical (time series) models
- ▶ machine learning methods
- ▶ Human judgement methods

▶ Decomposition of population into *compartments* (e.g. Susceptible, Infected and Recovered):



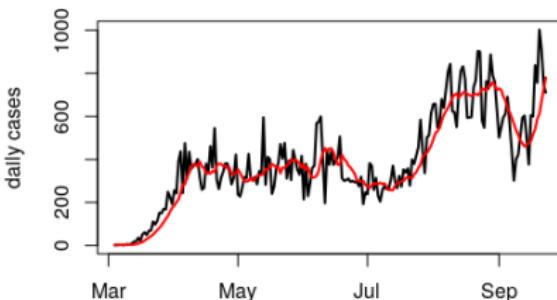
Modelling approaches represented in the Forecast Hubs

A continuum of approaches

- ▶ agent-based models
- ▶ compartmental (SIR-type) models
- ▶ **statistical (time series) models**
- ▶ machine learning methods

- ▶ Human judgement methods

▶ E.g. smoothing, autoregression:

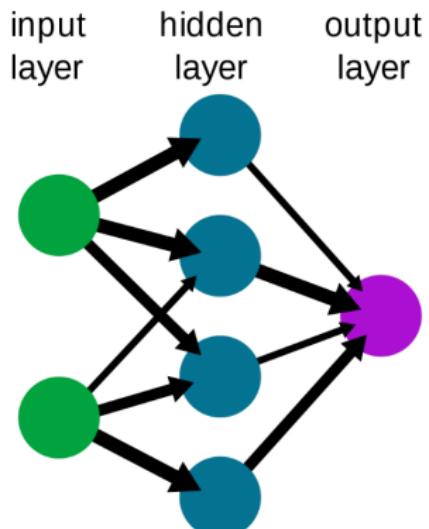


Modelling approaches represented in the Forecast Hubs

A continuum of approaches

- ▶ agent-based models
- ▶ compartmental (SIR-type) models
- ▶ statistical (time series) models
- ▶ **machine learning methods**
- ▶ Human judgement methods

E.g. neural networks:



Modelling approaches represented in the Forecast Hubs

A continuum of approaches

- ▶ agent-based models
- ▶ compartmental (SIR-type) models
- ▶ statistical (time series) models
- ▶ machine learning methods
- ▶ **Human judgement methods**

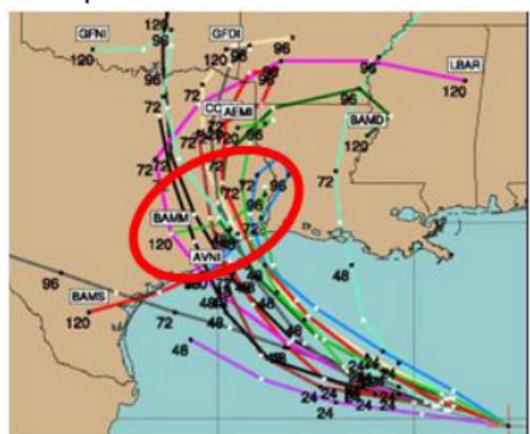
Surveys among experts or lay people:



Ensemble forecasts

- ▶ Statistical theory and experience
e.g., from meteorology show that combined ensemble forecasts are more reliable
 - ▶ different models exploit different mechanisms and data sources – the ensemble ties all together.
 - ▶ model uncertainty can be reflected.
- ▶ Increasingly common also in epidemiology (e.g., CDC FluSight)
- ▶ Intuition: We hope similarly many models will overpredict and underpredict...

Ensemble prediction of hurricane path:



A basic forecast for emerging diseases

- ▶ A simple (but quite common) forecasting approach is again based on the renewal equation
- ▶ This can also be read as a “baseline projection”: how quickly are cases going to rise / fall if the current trend continues?

$$\underbrace{\mathbb{E}(X_t)}_{\text{expected infections on day } t} = R \times \sum_{d=1}^D \underbrace{X_{t-d}}_{\# \text{ infections at } t-d} \times \underbrace{w_d}_{\text{probability that generation time is } d}$$

- ▶ A simple implementation is available in the R package `projections`, see
<https://www.repidemicsconsortium.org/projections/>

A basic forecast for seasonal diseases

- ▶ For seasonal diseases with stable patterns seasonal time series models are more appropriate than simple growth models.
 - ▶ R package `forecast`, <https://cran.r-project.org/web/packages/forecast/index.html>
 - ▶ R package `surveillance`,
<https://www.jstatsoft.org/article/view/v077i11>
- ▶ These models pick up the typical seasonal patterns of past years.

When will forecasts fail?

- ▶ It is important to make the key assumptions of forecasts explicit and to discuss their plausibility.
- ▶ Renewal equation models assume that the current rate of growth / decrease will not change in the short term.
 - ▶ i.e., contacts, transmissibility etc. will remain roughly the same.
- ▶ Seasonal models assume that the current season is not fundamentally different from previous ones.
 - ▶ e.g., post-COVID this is not the case for many respiratory diseases.

Evaluation of (probabilistic) disease forecasts

- ▶ As for nowcasts, it is important to monitor the performance of disease forecasts.
- ▶ There is a rich literature on best practices, see e.g.,
 - ▶ Funk et al (2019),
<https://doi.org/10.1371/journal.pcbi.1006785>
 - ▶ Bracher et al (2021),
<https://doi.org/10.1371/journal.pcbi.1008618>

Evaluation studies

RESEARCH ARTICLE | POPULATION BIOLOGY |

PNAS

Evaluation of individual and ensemble probabilistic forecasts of COVID-19 mortality in the United States

Daten Y, Crutcher J, Duan L, Liu Y, Valenza K, Lopez G, Johnstone Bracher G, Andrea Brennenstuhl, Alvaro J. Gómez Rueda, Emanuele Azzaro, Gerlach, Timmen, Graeling G, Keltz H, House, Yuen Hueing, Denisi Jaworska, Abdal H, Karthi Khandekar, Khoda Le, Arja Mütterönen, Jarad Niemi, Aman Shah, Ariane Stark, Yili Wang, Junhua Wattanachot, Martha W. Zem, Youyang Gu, Tariquddin Ishaq, Nayana Baruah, Ayush Deva, Mihir Kuburki, Sujanya Menju, Alpan Ravul, Sudham Shiro, Avinash Tewari, Jerome White, Niel F. Abernathy, Spencer Woody, Megha Dahani, Spencer Fox, Kelly Gaither, Michael Lachman, Lauren Ansol-Meyer, James G. Scott, Mauricio A. Vargas, Jeffrey C. Cope, J. van den Dethouwer, William P. Englund, Matthew W. Farting, Robert H. Huston, Brandon Jeffery, Oleg Linkov, Michael L. Mayo, Matthew D. Parizo, Michael A. Rosenthal, Benjamin D. Truett, Tylt Zhang, James Samuel Chen, Stephen Y. Farone, Jonathan Hess, Christopher P. Morely, Afaf Saadie, Doraiana Wong, Sabrina M. Corsetti, Thomas M. Baer, Martin C. Eisenberg, Karl Faibis, Yitao Huang, Emily T. Martin, Elisa McCalculus, Robert L. Myers, Tom Schwae, Daniel Shekhan, Graham Casey Gibson, Rose Yu, Liyan Gao, Yien Ma, Dongxia Wu, Xifeng Yan, Xiangyu Jin, Yaokang Wang, Yanping Chen, Uthong Gu, Yanting Zhou, Qianxun Gu, Jiezhichu Chen, Lingxiao Wang, Fan Au, Weiting Zhang, Dulan Zhu, Hannah Breden, Jeapline Leng, Steven McConnell, V.-P. Nagurni, Sonisha L. Ouerdi, Christopher Huhne, Louis Stroh, D. Turner, Zhifan Wu, Xueqiang Bai, Robert Wiklund, Qiwei Hong, Stanley Kong, Axel van de Walde, James A. Turlik, Michael Ben-Nun, Steven Eddy, Peter Bley, Lijun Kopecký, David Park, Bruce Hamery, Christina Kyriakides, Henrik Liss, John Millett, Michael Moloney, James Moran, Miroslav Nikolicic, Gábor Ozsán, Daniel Pach, Chris Schrader, Elizabeth Shulman, Daniel Segev, Ryan Spain, Chris Steedle, Barry Wilkinson, Alexander Wong, Sean Cawley, Guido Zusufa, Sean Moore, Rachel Orlman, Alex Perkins, David Kraus, Zhiheng Guo, Jiang Jiang, Wei Cap, Juan Lavista Ferres, Chaozhou Li, Tie-Yan Liu, Xing Xie, Zhou Zheng, Alessio Vespignani, Mattes Chrauzet, Jessica T. Davis, Kursang Mu, Ana Pastoriza, Pratik Rayas, Konrad Andrusz, Jackie Beck, Vinesh Fatin, Adrienne Georgescu, Retief Levy, Deeksha Sinha, Joshua White, Georgia Perakis, Mohammad Amine Benyoussef, David Nizet, Hédiou, Phoya Singh, Joann Sperber, Leanne Therapontos, Asterios Tsakris, Kaitlin Van Der Valk, Ali Jabbadzadeh, Devwani Shah, Nicolas Dela Perna, Lisa A. Cali, Kellie Sandus, Kara Wolfson, Diane Oftthus, Lauren Castro, Geoffrey Karché, June Michael, Debrah Miller, Marjorie Lutz, Cullen M. Lutz, Barbara Larway-Loves, Lauren Shiu, Katherine Talakauskas, Sheby Wilson, Elizabeth C. Lee, Karen Derr, Kanya H. Grantz, Alison L. Hill, Judith Karmazyn, Kathryn Karmazyn, Lindsay S. Karmazyn, Stephen A. Lauer, Joseph C. Lennartsson, Justin Lessler, Hannah R. Meredith, Javer Perez-Saez, Sam Shats, Clive P. Smith, Shaula A. Truelove, JOSH WEB, Maximilian Marchall, Lauren Gardner, Kristen Nixey, John C. Buratti, Ulvi Wale, Lei Guo, Zhiqiu Gu, Mengyan Xian, Xinyi Li, Guannan Wang, Yuxing Wang, Shen Yu, Robert C. Rastetter, Ryan Barber, Cormacala Gakidou, Samoni L. Hay, Steve Lim, Chia Murray, David Peleg, Heidi L. Durding, Prashanth Raghavan, Steven A. Spigel, Bradley T. Suchman, R. Adithya Prakash, Sajida Afridi, Jammy Cu, Alexander Rodriguez, Arka Tabassum, Jaya Xie, Pinar Yesilgül, John Asplund, Arden Baker, Bude Erol Oruc, Novakos Sertbas, Sercan O. Arsl, Mike Duzenberry, Arshad Fatheth, Eli Katal, Long T. Le, Chuan-Liang Li, Tomas Pfleider, Tariq Sale, Rajeshwar Sirisha, Thomas Tsai, Nade Yose, Insung Yoo, Leyisu Zhang, Sam Abbott, Nicosia A., Besse, Sebastian, Furu, G., Joffe Hellwinkel, Sophie R. Meakin, Katherine Sherratt, Mingyu Zhou, Rabi Kalentev, Thomas K. Karmazyn, Sami Pihl, Jeffrey Sharman, Michael L. Tamm, Dietrich Bertmann, O. Caglar Skali, Lennart, Subhashini Saha, Namita Tariq, Roushni Tariq, Sayar, Madeline Adams, Jagjeet Singh, Chhetri, Gurdeep S. Bal, Mary A. Ladd, Benjamin P. Juras, Peter Mueller, Jade Yuan, Xuefei Wang, Qizhou Wang, Shengping Zhang, Jingyu Zeng, Alton Green, Jacob Bres, Logan Brooks, Addison J. Hu, Maria Jafra, David McGrory, Balakumarareman Narasimhan, Colm Pollicino, Samarth Ravalas, Arun Kumar, Nishan Simon, Ryan J. Tofchiali, Jedidiah, Sebastian Funk, Antoni Myslinski, Karol Niedzielski, Jędrzej Nowosielski, Maciej Radwan, Franciszek Rakowski, Marcin Semejński, Ewa Szczeruk, Jakub Zieliński, Jan Kisielewski, Barbara Pabjan, Kirsten Holger, Yuli Kheifetz, Markus Scholz, Bilecik Przemysław, Marcin Bodchy, Maciej Filinski, Radosław Idzikowski, Tylt Krueger, Tomasz Ozanski, Johannes Bracher [see less](#)

nature communications

A pre-registered short-term forecasting study of COVID-19 in Germany and Poland during the second wave

J. Bracher , D. Wolfram, J. Deuschel, K. Görgen, J. L. Ketterer, A. Ullrich, S. Abbott, M. V. Barbarossa, D. Bertsimas, S. Bhatia, M. Bodchy, N. I. Bosse, J. P. Burgard, L. Castro, G. Fairchild, J. Fuhrmann, S. Funk, K. Gogolewski, O. Gu, S. Heyder, T. Hotz, Y. Kheifetz, H. Kirsten, T. Krueger, E. Krymova, M. L. Li, H. Meinke, I. J. Michaud, K. Niedzielski, T. Ozanski, F. Rakowski, M. Scholz, S. Soni, A. Srivastava, J. Zieliński, D. Zou, T. Gneiting, M. Schieme  & List of Contributors by Team — Show fewer authors



Predictive performance of multi-model ensemble forecasts of COVID-19 across European nations

Katherine Sherratt , Hugo Gruson, Rok Grah, Helen Johnson, Rene Niehus, Bastian Prasse, Frank Sandmann, Jannik Deuschel, Daniel Wolfram, Sam Abbott, Alexander Ullrich, Graham Gibson, Evan L. Ray, Nicholas G. Reich, Daniel Sheldon, Yijin Wang, Nutcha Wattanachot, Lijing Wang, Jan Trnka, Guillaume Obozinski, Tuo Sun, Dorina Thanou, Loïc Pottier, Ekaterina Krymova, Jan H. Meinke, Maria Vittoria Barbarosa, Neele Leithauser, Jan Mohring, Johanna Schneider, Jaroslaw Wlazlo, Jan Fuhrmann, Berit Lange, Isti Rodlath, Prashit Baccam, Heidi Gurung, Steven Stage, Bradley Suchoski, Jozef Budzinski, Robert Walraven, Inmaculada Villanueva, Vit Tuckin, Martin Smid, Milan Zajicek, Cesar Perez Alvarez, Borja Reina, Nikos I. Bosse, Sophie R. Meakin, Lauren Castro, Geoffrey Fairchild, Isaac Michaud, Dave Osthus, Pierfrancesco Alaimo Di Loro, Antonello Maruotti, Veronika Eclerova, Andrea Kraus, David Kraus, Lenka Pribylova, Bertsimas Dimitris, Michael Lingzhi Li, Soni Sakham, Jonas Dehning, Sebastian Mohr, Viola Priesemann, Grzegorz Redlarski, Benjamin Berger, Giovanni Ardenghi, Nicola Parolini, Giovanni Ziarrelli, Wolfgang Bock, Stefan Heyder, Thomas Hotz, David E. Singh, Miguel Guzman-Merino, Jose L. Aznar, David Morina, Sergio Alonso, Eric Alvarez, Daniel Lopez, Clara Prats, Jan Pablo Burgard, Arne Rodloff, Tom Zimmermann, Alexander Kuhlmann, Janez Zibert, Fulvia Pennoni, Fabio Divino, Marti Catala, Gianfranco Lovison, Paolo Giudici, Barbara Tarantino, Francesco Bartolucci, Giovanna Jona Lasinio, Marco Mingione, Alessio Farcomeni, Ajitesh Srivastava, Pablo Montero-Mano, Anrinudha Adiga, Benjamin Hurt, Bryan Lewis, Madhav Marathe, Przemysław Porebski, Sriwinaswan Venkatraman, Rafal P. Bartczuk, Filip Dreger, Anna Gambi, Krzysztof Gogolewski, Magdalena Grzelci-Słomka, Bartosz Krupa, Antoni Moszynski, Karol Niedzielski, Jędrzej Nowosielski, Maciej Radwan, Franciszek Rakowski, Marcin Semejński, Ewa Szczeruk, Jakub Zieliński, Jan Kisielewski, Barbara Pabjan, Kirsten Holger, Yuli Kheifetz, Markus Scholz, Bilecik Przemysław, Marcin Bodchy, Maciej Filinski, Radosław Idzikowski, Tylt Krueger, Tomasz Ozanski, Johannes Bracher [see less](#)

Takeaways (I)

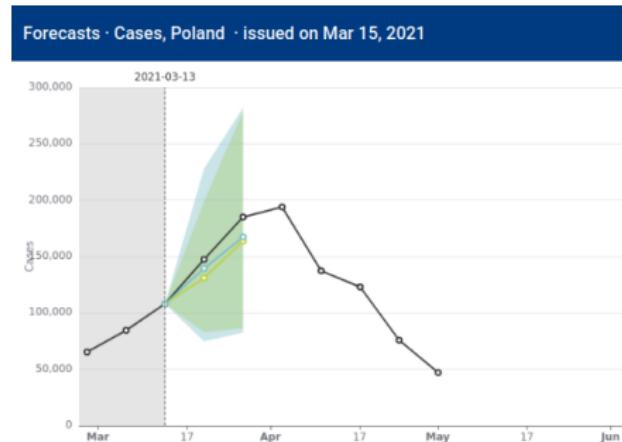
- ▶ Forecasts are highly heterogeneous – both in terms of central tendency and spread
- ▶ Predicting inflection points is very difficult – especially for cases
- ▶ Hardly any model reached nominal coverage of 50% and 95% prediction intervals (note: small n)



Blue: MOCOS model, Wroclaw University of Science and Technology; Green: Mean ensemble

Takeaways (I)

- ▶ Forecasts are highly heterogeneous – both in terms of central tendency and spread
- ▶ Predicting inflection points is very difficult – especially for cases
- ▶ Hardly any model reached nominal coverage of 50% and 95% prediction intervals (note: small n)



Blue: MOCOS model, Wroclaw University of Science and Technology; Green:
Mean ensemble

Takeaways (I)

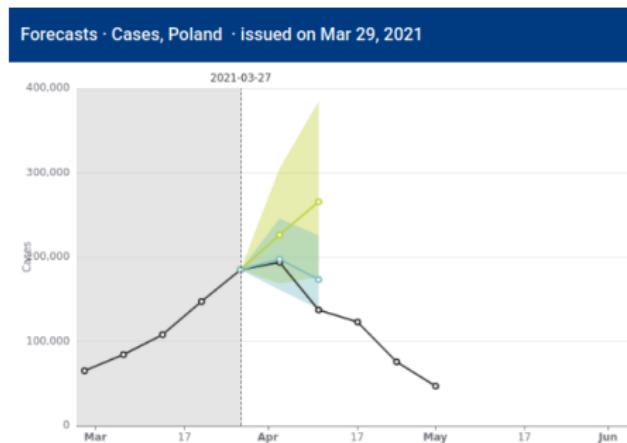
- ▶ Forecasts are highly heterogeneous – both in terms of central tendency and spread
- ▶ Predicting inflection points is very difficult – especially for cases
- ▶ Hardly any model reached nominal coverage of 50% and 95% prediction intervals (note: small n)



Blue: MOCOS model, Wroclaw University of Science and Technology; Green: Mean ensemble

Takeaways (I)

- ▶ Forecasts are highly heterogeneous – both in terms of central tendency and spread
- ▶ Predicting inflection points is very difficult – especially for cases
- ▶ Hardly any model reached nominal coverage of 50% and 95% prediction intervals (note: small n)



Blue: MOCOS model, Wroclaw University of Science and Technology; Green: Mean ensemble

Takeaways (I)

- ▶ Forecasts are highly heterogeneous – both in terms of central tendency and spread
- ▶ Predicting inflection points is very difficult – especially for cases
- ▶ Hardly any model reached nominal coverage of 50% and 95% prediction intervals (note: small n)



Blue: MOCOS model, Wroclaw University of Science and Technology; Green:
Mean ensemble

Takeaways (I)

- ▶ Forecasts are highly heterogeneous – both in terms of central tendency and spread
- ▶ Predicting inflection points is very difficult – especially for cases
- ▶ Hardly any model reached nominal coverage of 50% and 95% prediction intervals (note: small n)



Blue: MOCOS model, Wroclaw University of Science and Technology; Green:
Mean ensemble

Takeaways (I)

- ▶ Forecasts are highly heterogeneous – both in terms of central tendency and spread
- ▶ Predicting inflection points is very difficult – especially for cases
- ▶ Hardly any model reached nominal coverage of 50% and 95% prediction intervals (note: small n)



Blue: MOCOS model, Wroclaw University of Science and Technology; Green:
Mean ensemble

Takeaways (I)

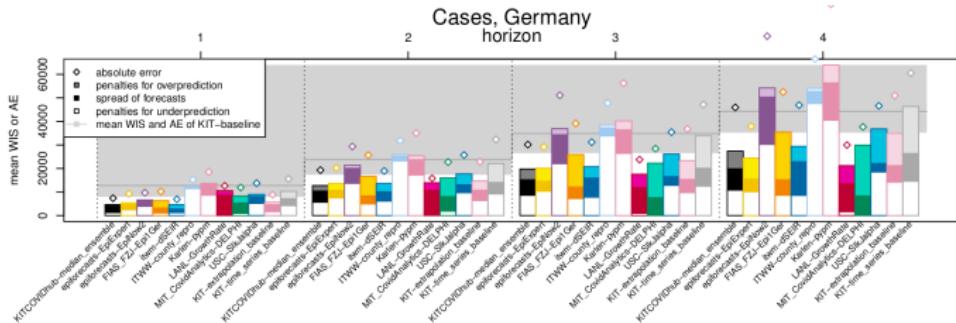
- ▶ Forecasts are highly heterogeneous – both in terms of central tendency and spread
- ▶ Predicting inflection points is very difficult – especially for cases
- ▶ Hardly any model reached nominal coverage of 50% and 95% prediction intervals (note: small n)



Blue: MOCOS model, Wroclaw University of Science and Technology; Green:
Mean ensemble

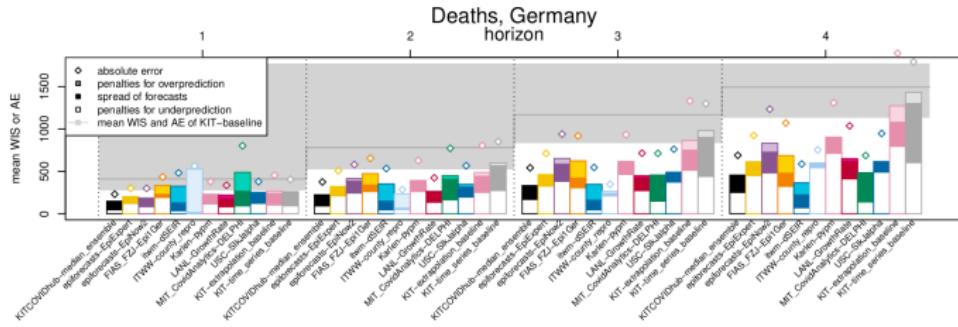
Takeaways (II)

- ▶ Simple baseline models can be tough to beat.
- ▶ Relative performance of different models varies substantially over time.
- ▶ Ensemble methods show good relative performance → **benefit of collaborative approach**
- ▶ Results Jan – Apr 2021:



Takeaways (II)

- ▶ Simple baseline models can be tough to beat.
- ▶ Relative performance of different models varies substantially over time.
- ▶ Ensemble methods show good relative performance → **benefit of collaborative approach**
- ▶ Results Jan – Apr 2021:



Acknowledgements

- ▶ These slides contain materials shared by the following colleagues:
 - ▶ Sebastian Funk, London School of Hygiene and Tropical Medicine.