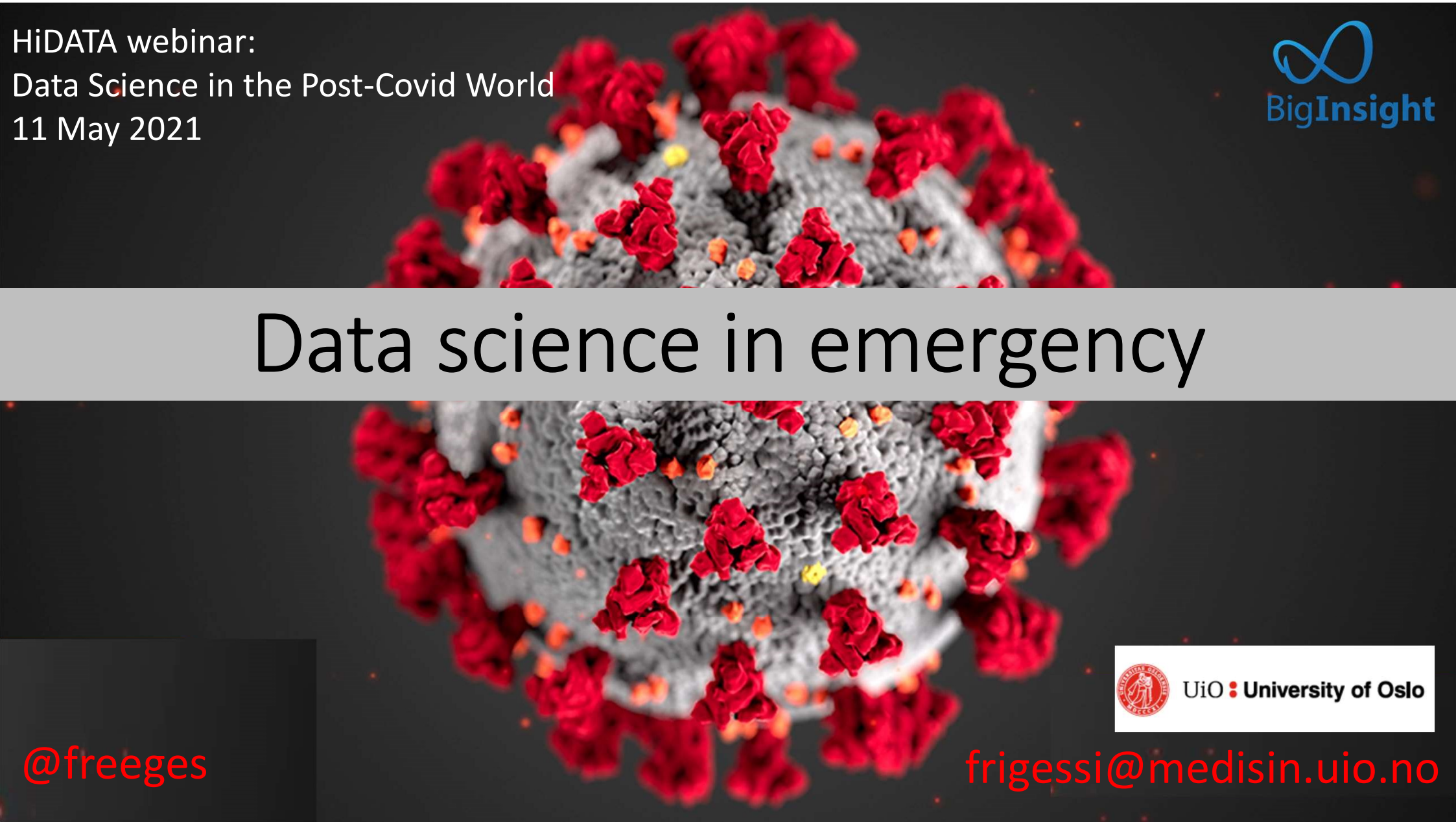
HiDATA webinar:
Data Science in the Post-Covid World
11 May 2021

BigInsight



Data science in emergency

@freeges

UiO : University of Oslo

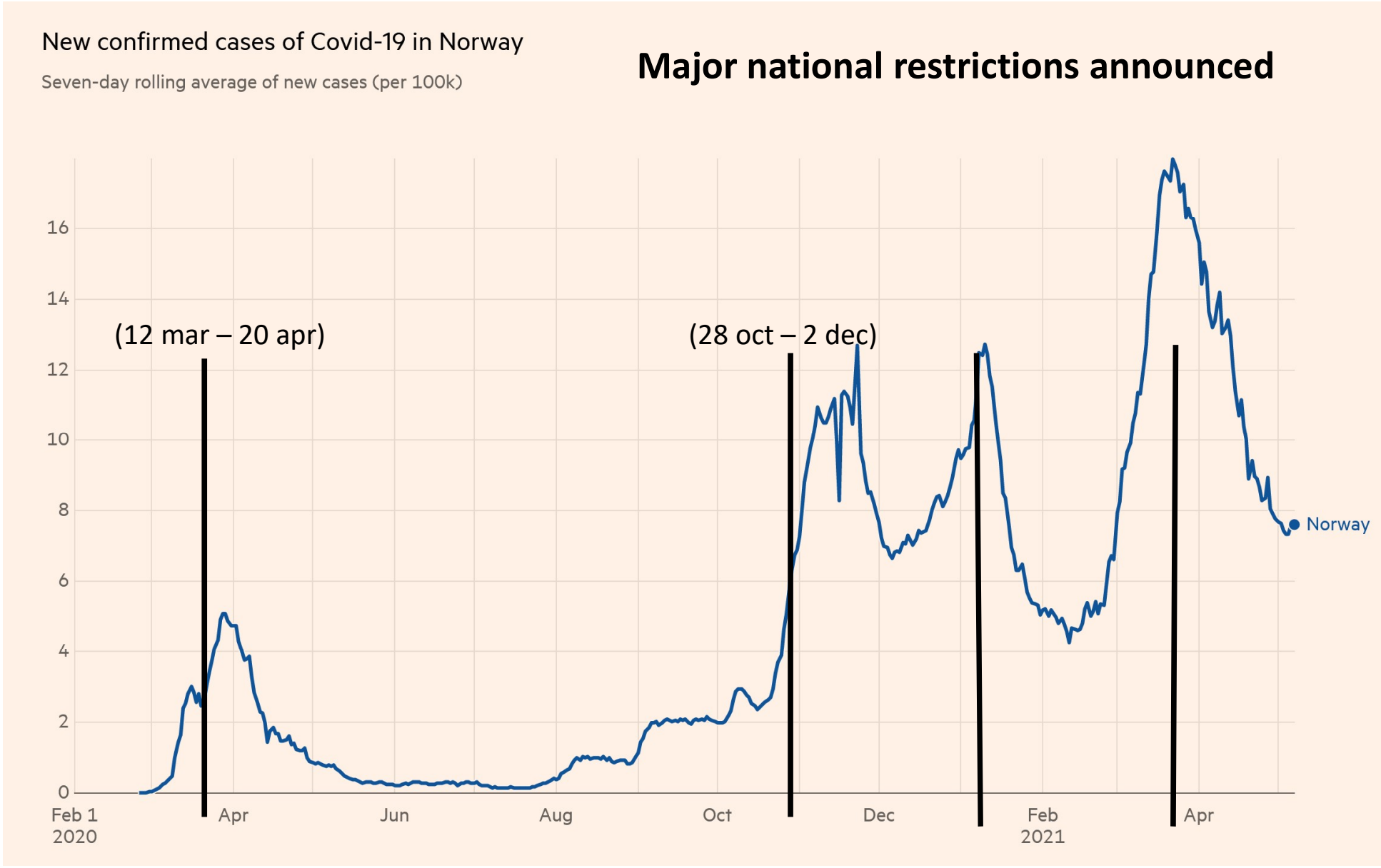
frigessi@medisin.uio.no

Norwegian strategy during Covid-19 emergency

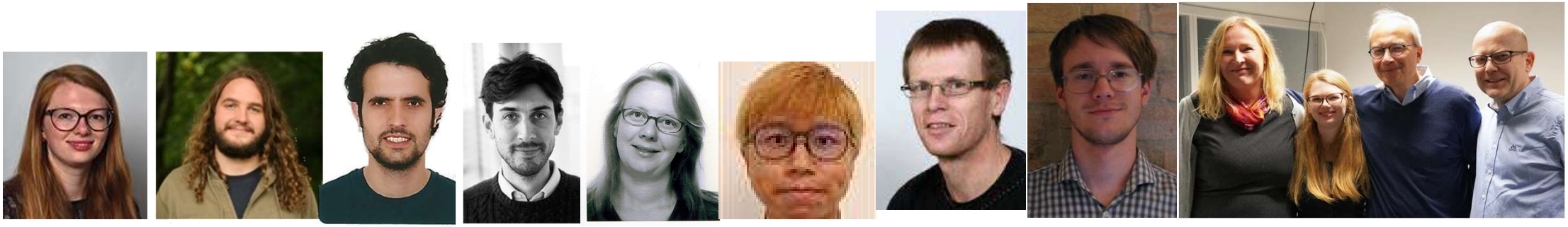
- **Keep control of the spread of the virus.**

- Control means:
  - the number of patients is manageable for the health system
  - the number of cases allows efficient testing, isolation, contact tracing and quarantine.

- Local interventions allowed and encouraged.


New confirmed cases of Covid-19 in Norway
Seven-day rolling average of new cases (per 100k)
Major national restrictions announced
(12 mar – 20 apr)
(28 oct – 2 dec)
16
14
12
10
8
6
4
2
0
Feb 1
2020
Apr
Jun
Aug
Oct
Dec
Feb
2021
Apr
Norway

NIPH COVID-19 modelling team



Birgitte Freiesleben de Blasio, Gunnar Øyvind Isaksson Rø, Alfonso Diz-Lois Palomares, Anja Brathen Kristoffersen , Francesco Di Ruscio, Solveig Engebretsen, Geir Storvik, Kenth Engø-Monsen.

,

Our tasks:

REQUESTS:

- Predict the number of covid-19 patients
- Predict number of patients requiring hospitalization, in particular ICU
- Estimate prevalence
- Estimate effective reproduction numbers

… at national and regional level

- National vaccination strategy
- National re-opening strategy

- 3 weeks
- 1 year

Data

- Total tested, total positive cases – excluding repeated tests
- Total imported positive cases
- Total new admitted patients to hospital
- Total new admitted to ICU, with and without ventilation
- Total deaths

All daily at municipal/hospital level.

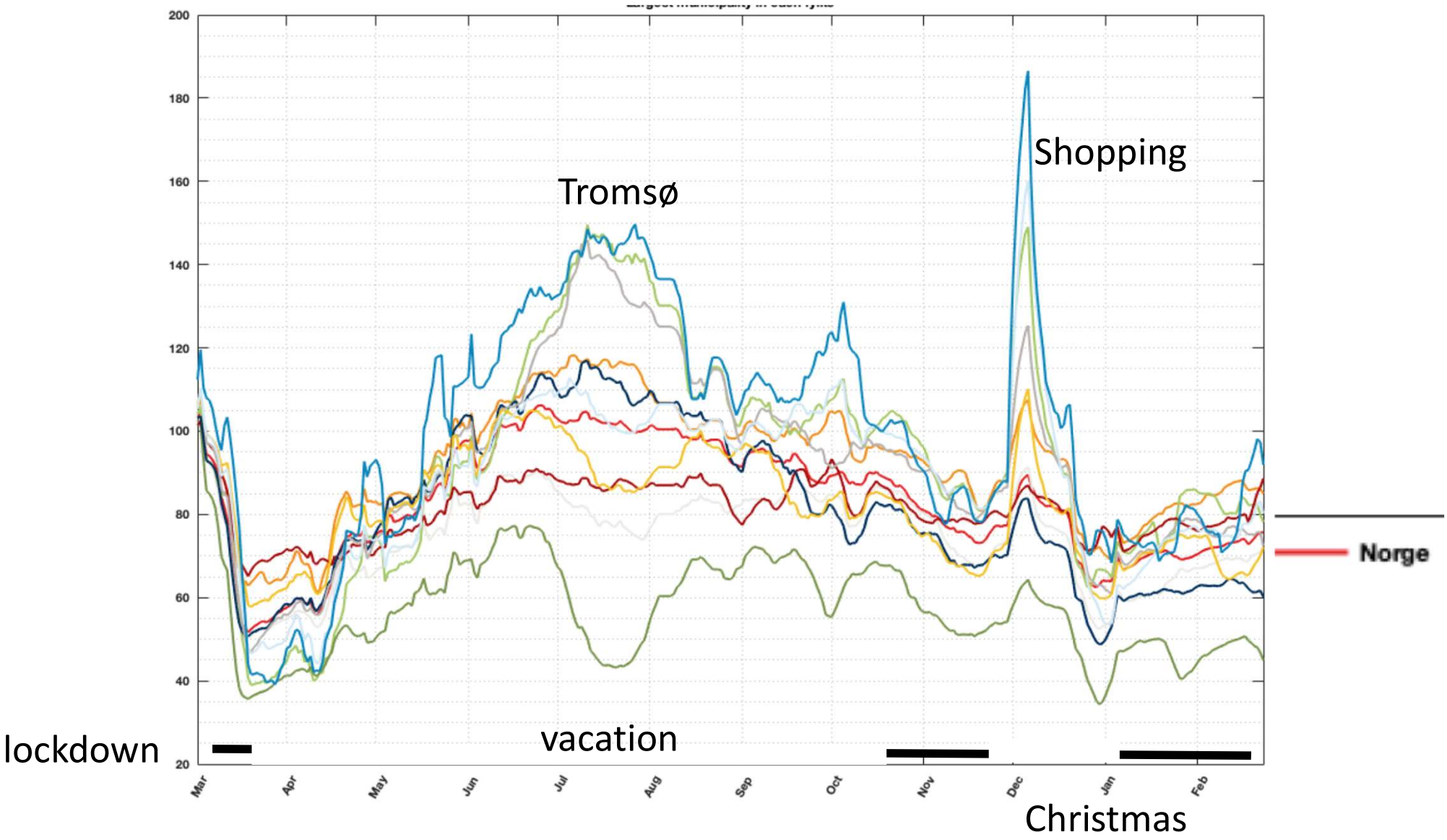
- Basic data pipelines were not ready at start
- Overlapping registries not always coherent
- Seasonal and organisational irregularities

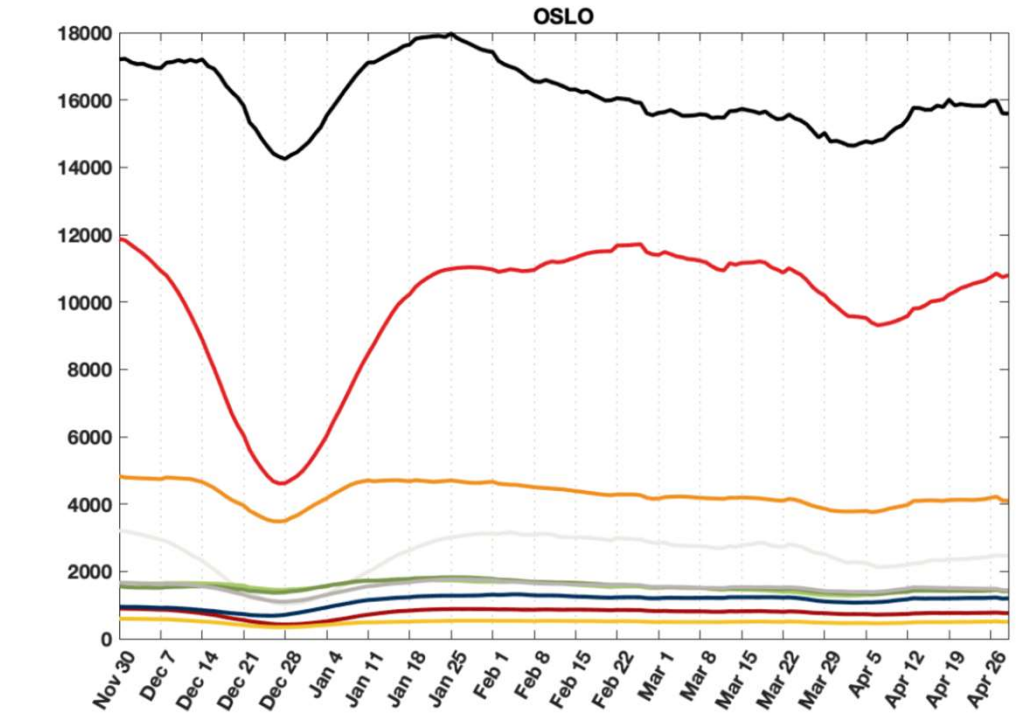
Mobility data

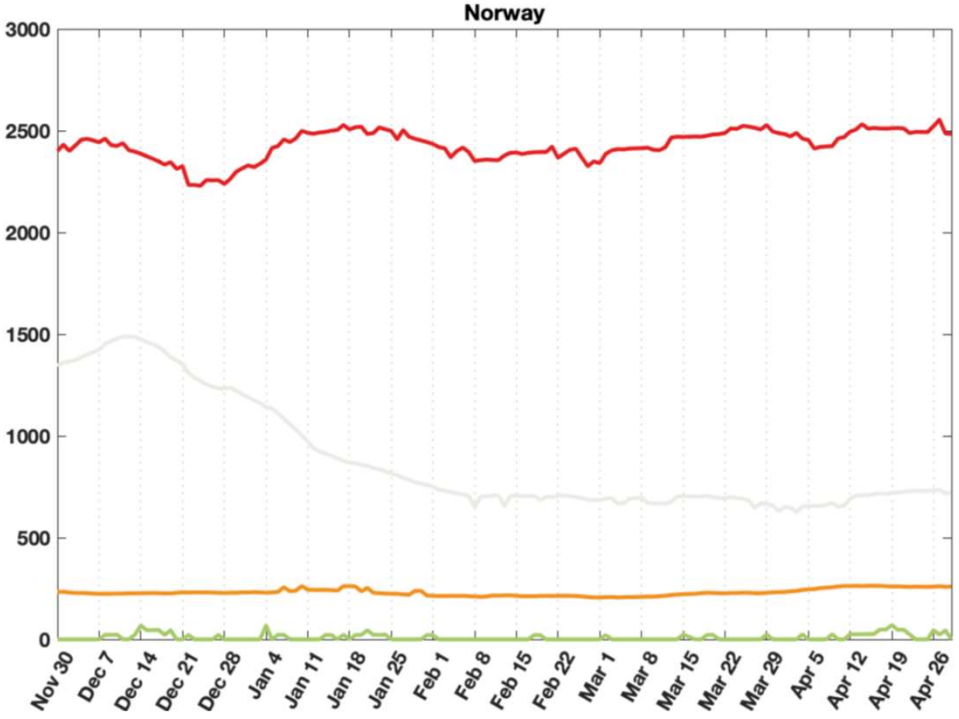
telenor mobile phone mobility, surrogate for real time mobility patterns

- Total movements between the 356 municipalities.
- 4 time intervals: 0-6, 6-12, 12-18, 18-24.
- Counts below 20 are censored.
- Telenor has ca. 48% market share – we scale up.

Date,	Time,	From,	To,	Count
20200224,	12,	1101,	1101,	8960
20200224,	12,	1101,	1103,	239
20200224,	12,	1101,	1108,	194
20200224,	12,	1101,	1111,	91


Tromsø
Shopping
Norge
lockdown
vacation
Christmas
200
180
160
140
120
100
80
60
40
20
Mar
Apr
May
Jun
Jul
Aug
Sep
Oct
Nov
Dec
Jan
Feb


OSLO
18000
16000
14000
12000
10000
8000
6000
4000
2000
0
Nov 30
Dec 7
Dec 14
Dec 21
Dec 28
Jan 4
Jan 11
Jan 18
Jan 25
Feb 1
Feb 8
Feb 15
Feb 22
Mar 1
Mar 8
Mar 15
Mar 22
Mar 29
Apr 5
Apr 12
Apr 19
Apr 26

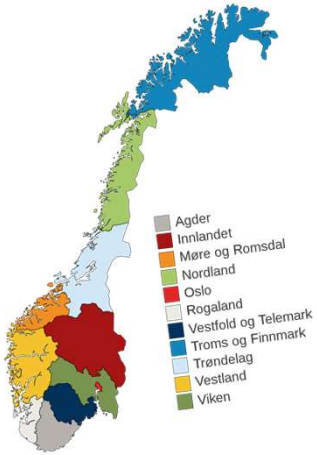

Norway
3000
2500
2000
1500
1000
500
0
Nov 30
Dec 7
Dec 14
Dec 21
Dec 28
Jan 4
Jan 11
Jan 18
Jan 25
Feb 1
Feb 8
Feb 15
Feb 22
Mar 1
Mar 8
Mar 15
Mar 22
Mar 29
Apr 5
Apr 12
Apr 19
Apr 26

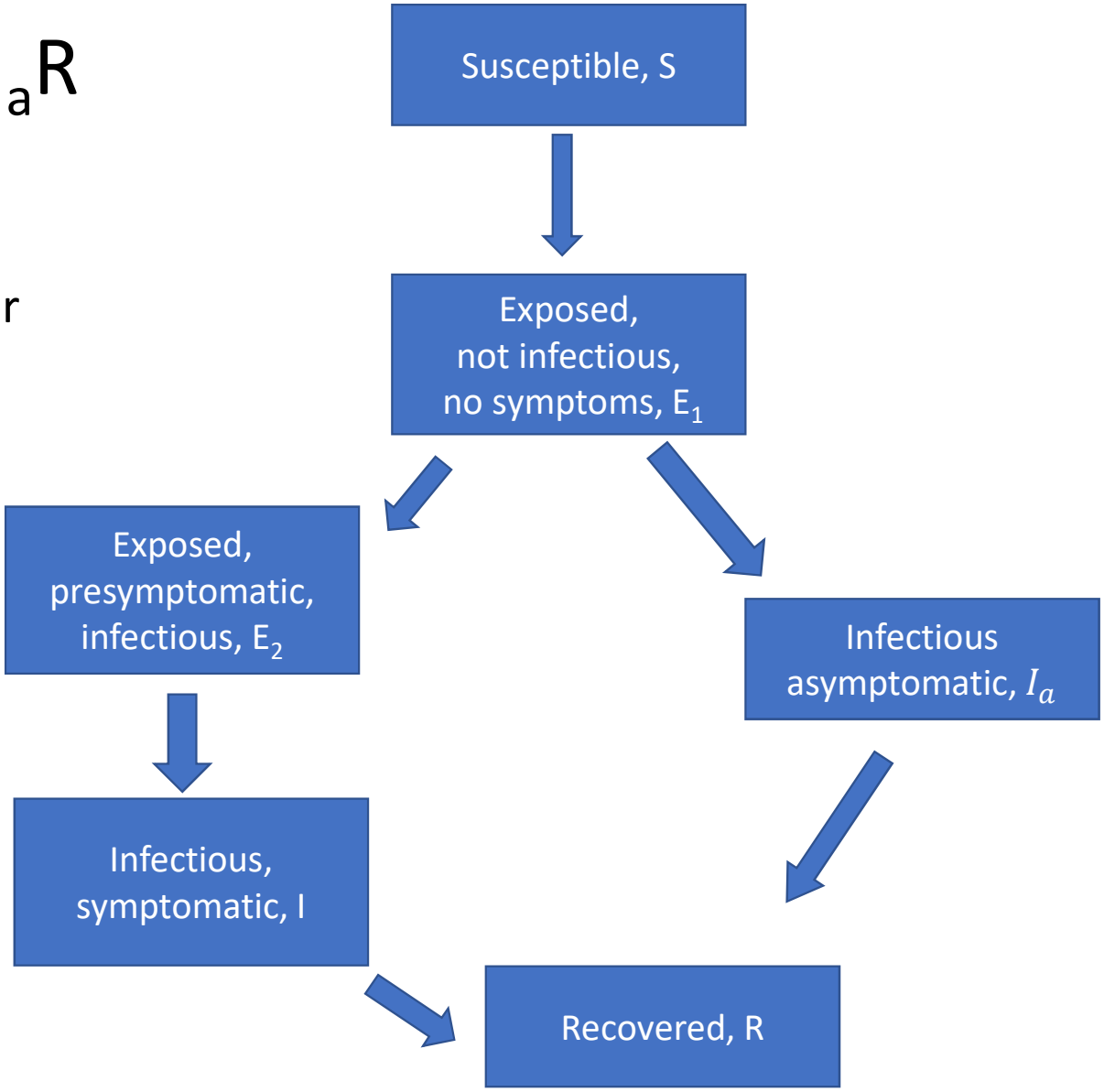

PL
LT
SE
NL
DK
LV
DE
ES
FI
REST


IN
BG
PK
MM

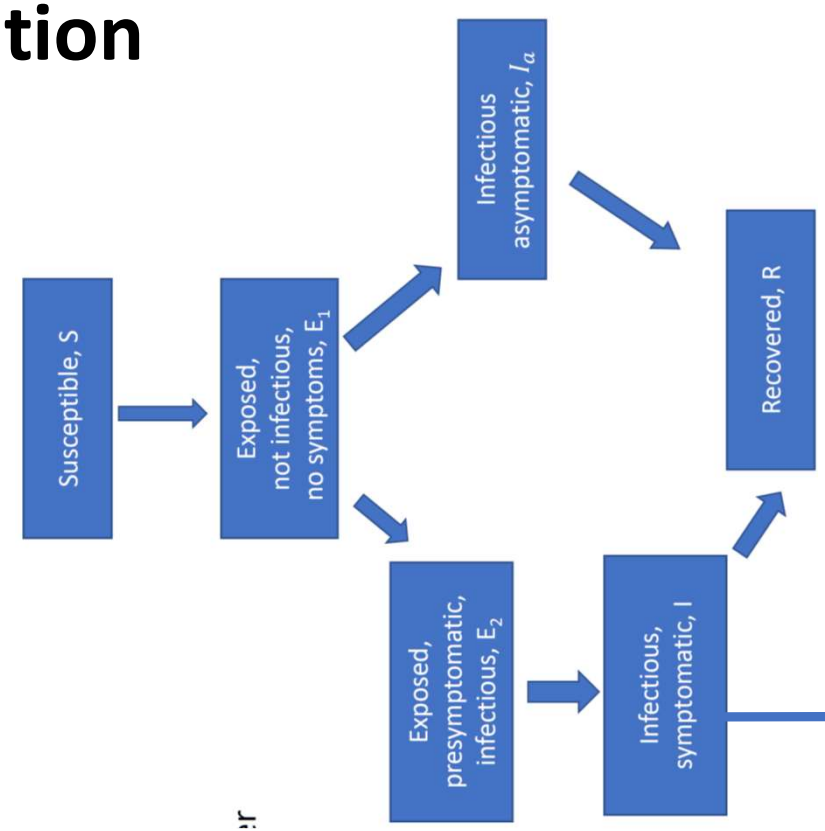
Regional stochastic $SEII_aR$

- One model per municipality (356)
- County level transmission parameter changing in time
- Other parameters at nation level





Hospitalisation



Hospital

Probability

Time gap

Science

CORONAVIRUS

Estimating the burden of SARS-CoV-2 in France

Henrik Salje[1,2,3]*, Cécile Tran Kiem[1,4]*, Noémie Lefrancq[1], Noémie Courtejoie[5], Paolo Bosetti[1], Juliette Paireau[1,6], Alessio Andronico[1], Nathanaël Hozé[1], Jehanne Richet[5], Claire-Lise Dubost[5], Yann Le Strat[6], Justin Lessler[3], Daniel Levy-Bruhl[6], Arnaud Fontanet[7,8], Lulla Opatowski[9,10], Pierre-Yves Boelle[11], Simon Cauchemez[1]†

8 July 2020

Probability of hospitalization per age group

Table S1

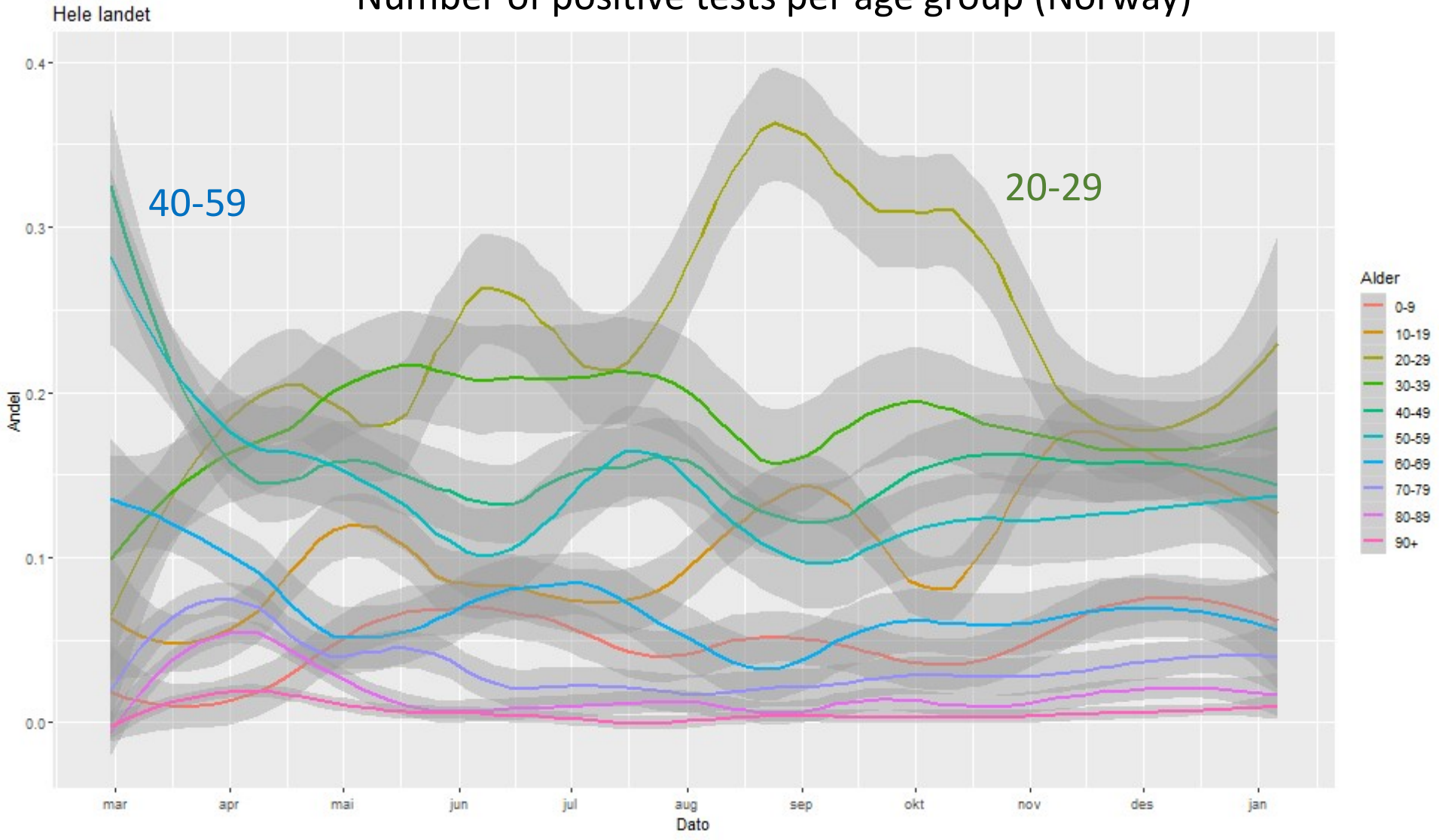
Age group	Percent infections hospitalized			Percent of hospitalized cases that go to ICU		
	Male	Female	Mean	Male	Female	Mean
<20	0.2 (0.08-0.2)	0.1 (0.07-0.2)	0.1 (0.08-0.2)	26.9 (23.1-31.1)	16.7 (14.3-19.3)	22.2 (19.1-25.7)
20-29	0.6 (0.3- 0.9)	0.5 (0.3-0.8)	0.5 (0.3-0.8)	14.0 (12.2-16.0)	8.7 (7.5-9.9)	11.6 (10.1-13.2)
30-39	1.2 (0.7-1.9)	0.9 (0.5-1.5)	1.1 (0.6-1.7)	19.2 (17.6-20.9)	11.9 (10.9-13.0)	15.9 (14.5-17.3)
40-49	1.6 (0.9-2.6)	1.3 (0.7-2.1)	1.4 (0.8-2.3)	26.9 (25.4-28.4)	16.6 (15.6-17.7)	22.2 (21.0-23.5)
50-59	3.2 (1.8-5.2)	2.6 (1.5-4.2)	2.9 (1.6-4.7)	33.4 (32.0-34.8)	20.7 (19.8-21.6)	27.6 (26.5-28.7)
60-69	6.7 (3.7-10.9)	5.1 (2.9-8.3)	5.8 (3.3-9.5)	37.3 (36.0-38.6)	23.1 (22.2-24.0)	30.8 (29.8-31.8)
70-79	11.0 (6.2-17.9)	7.8 (4.4-12.8)	9.3 (5.2-15.1)	30.2 (29.1-31.3)	18.7 (18.0-19.5)	24.9 (24.1-25.8)
80+	37.6 (21.1-61.3)	19.3 (10.9-31.6)	26.2 (14.8-42.7)	6.8 (6.5-7.2)	4.2 (4.0-4.5)	5.6 (5.3-5.9)
Mean	3.3 (1.8-5.3)	2.6 (1.5-4.3)	2.9 (1.7-4.8)	23.1 (22.6-23.6)	14.3 (13.9-14.7)	19.0 (18.7-19.44)

Table S1: Percent of infections that are hospitalized and end up in ICU by age and sex. Percentage of infections that are hospitalized and the percentage that end up in ICU, conditional on being hospitalized.

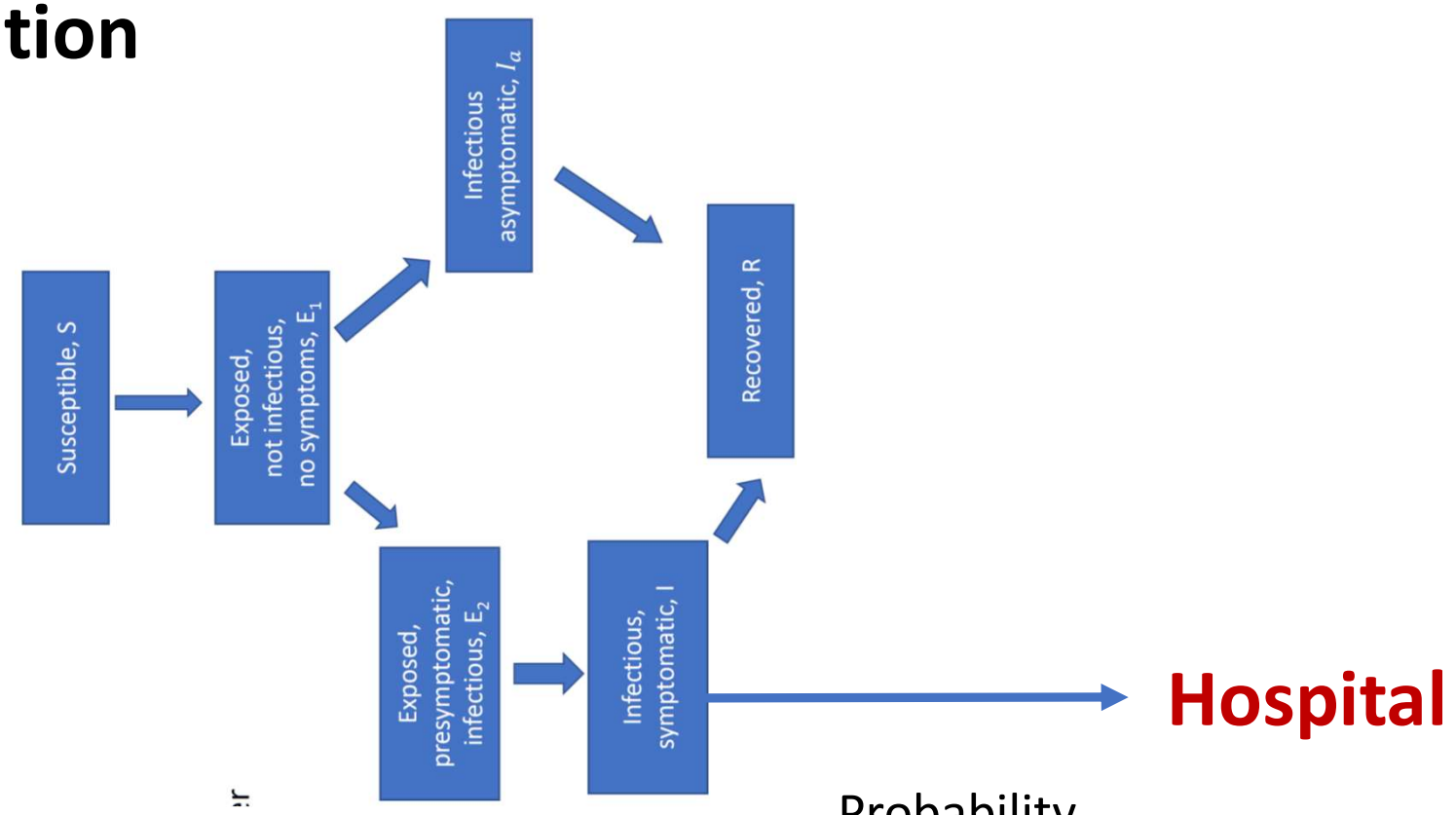
- Scaled to the **demography of each municipality** in Norway and corrected for percentage of elderly living in elderly homes.
- Assumes diseases burden follows demography
  Nation average

0-9 years	0.2%
10 - 19 years	0.2%
20 - 29 years	0.6%
30 - 39 years	1.3%
40 - 49 years	1.7%
50 - 59 years	3.5%
60 - 69 years	7.1%
70 - 79 years	11.3%
80+ years	27%

Number of positive tests per age group (Norway)

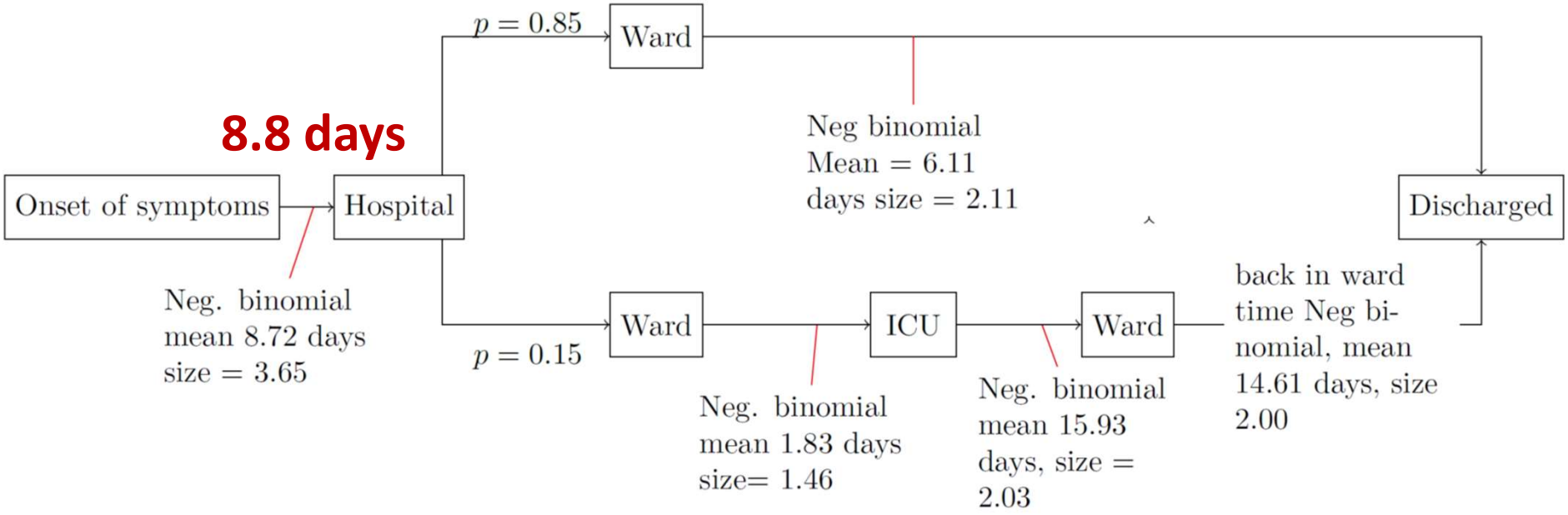


Hospitalisation



Hospitalisation lengths

Estimates from Norwegian registry data.



Up to 31 May 2020

Based on > 1500 hospitalisation cases, taking censoring into account.

Oslo

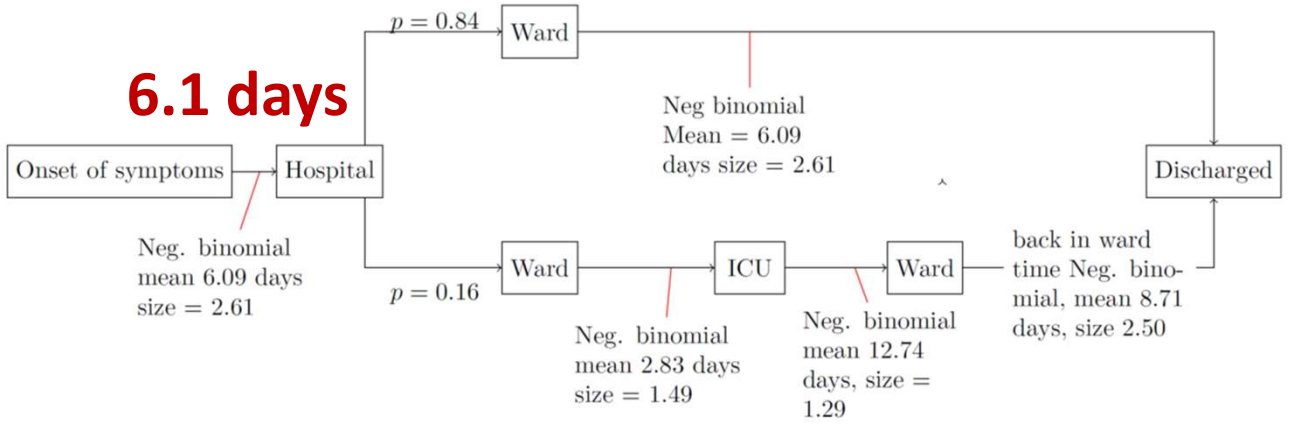


Figure 33: Hospital assumptions and parameters used between 1 January 2021 and 1 March 2021 for those living in Oslo

not Oslo

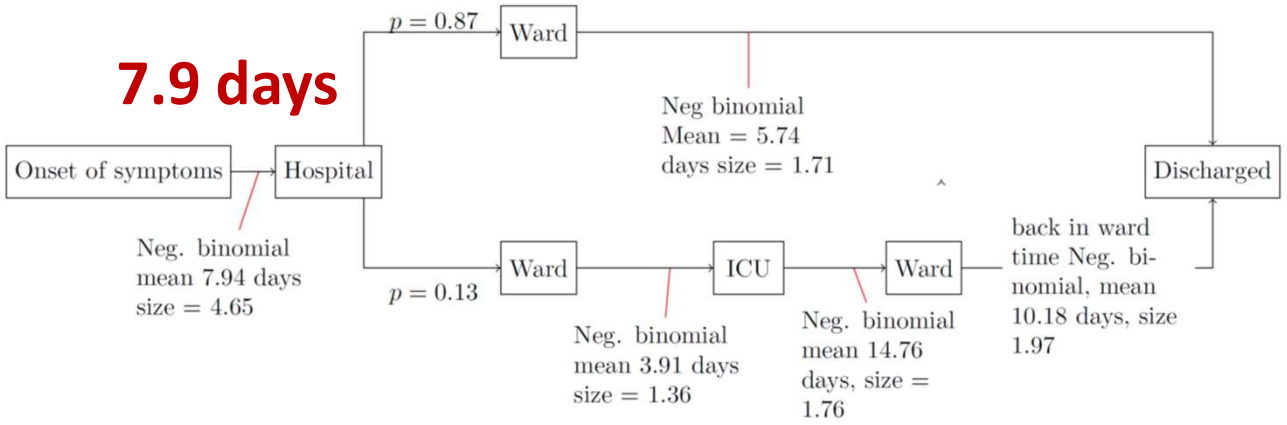
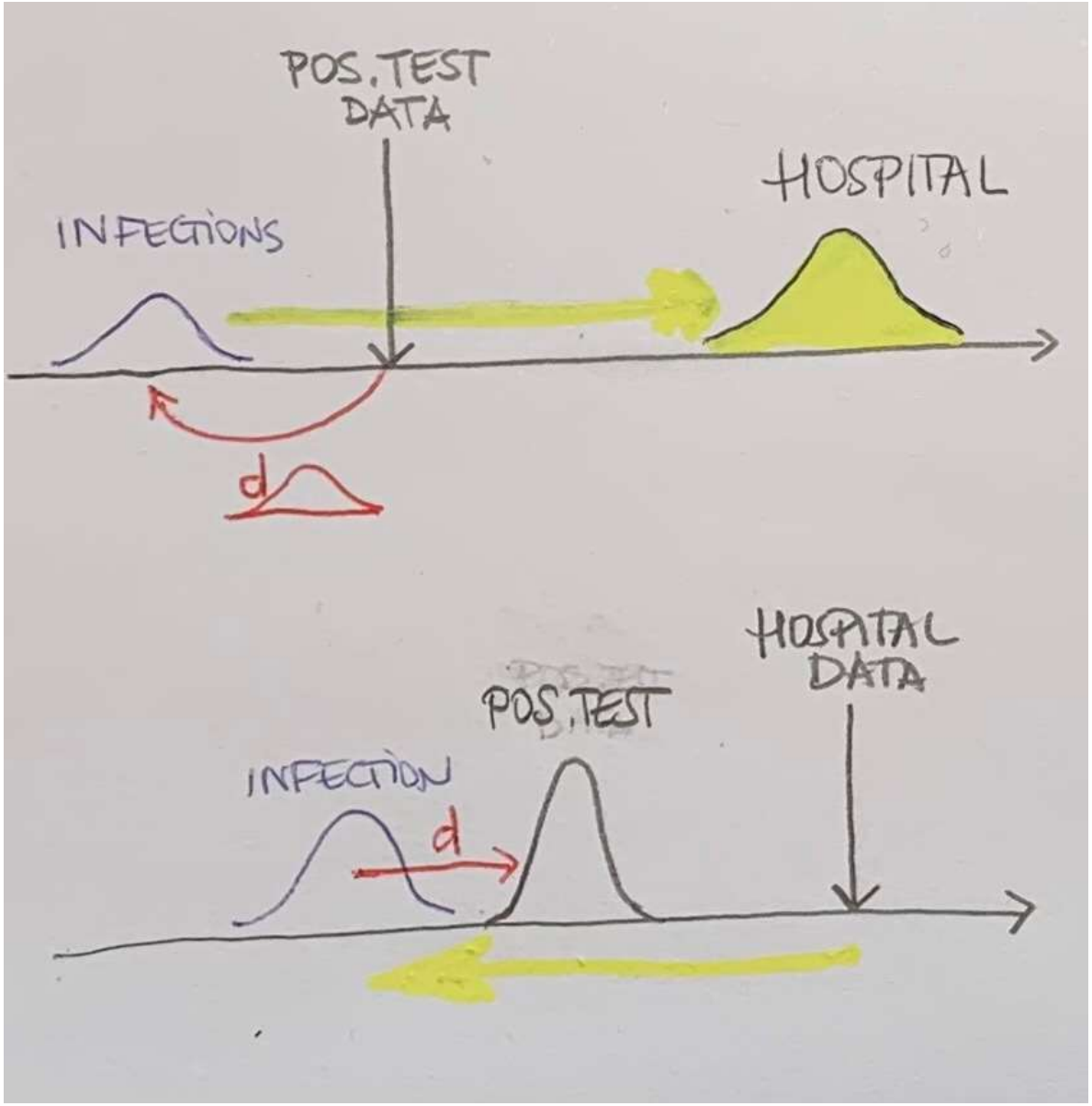


Figure 34: Hospital assumptions and parameters used between 1 January 2021 and 1 March 2021 for those not living in Oslo

From 1 Jan 2021

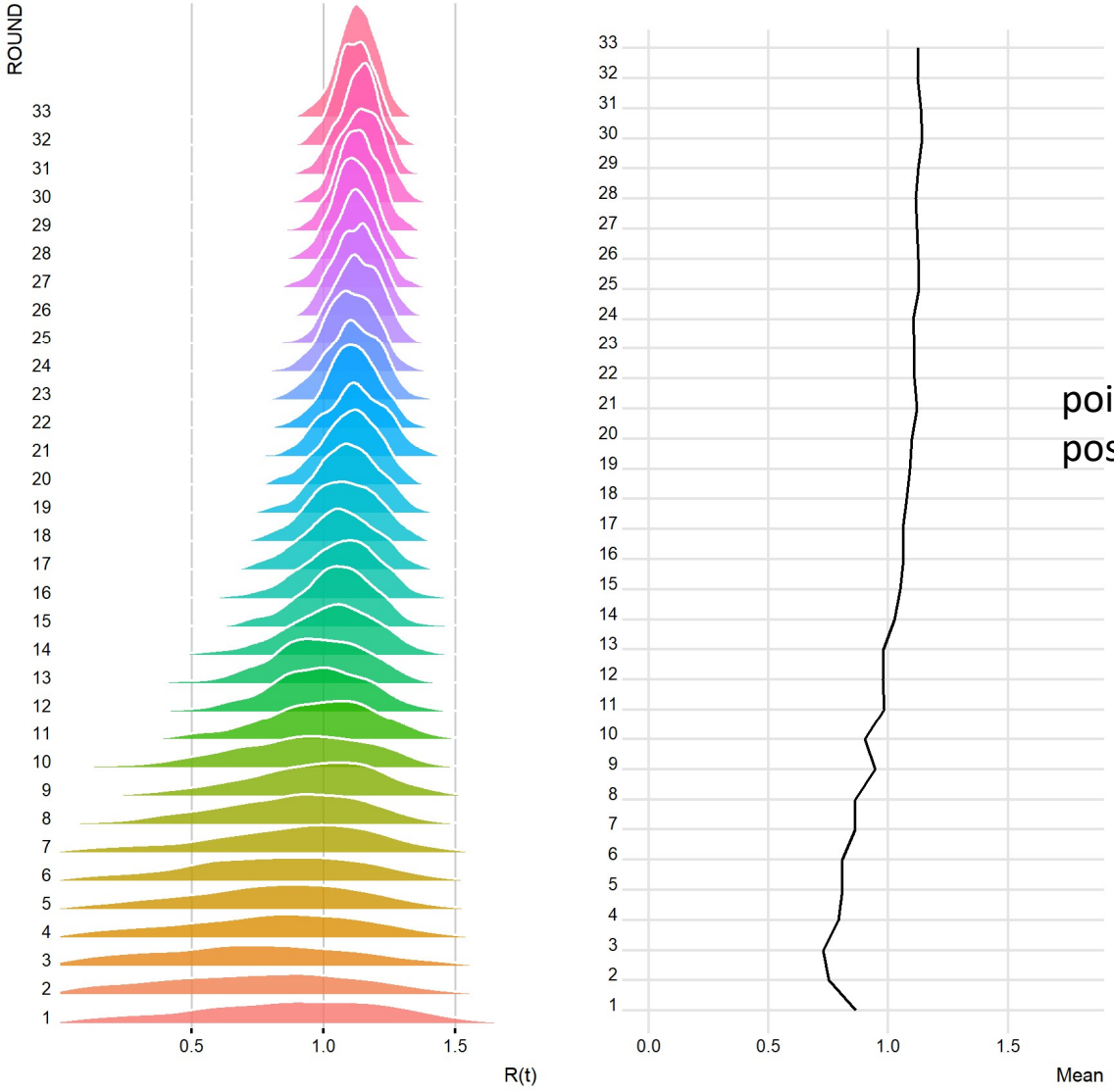
Lack of coherence between test and hospital data



Splitted sequential ABC
to estimate parameters
with uncertainty

Iterative algorithm:
ca. 2 days
on 2000 cores

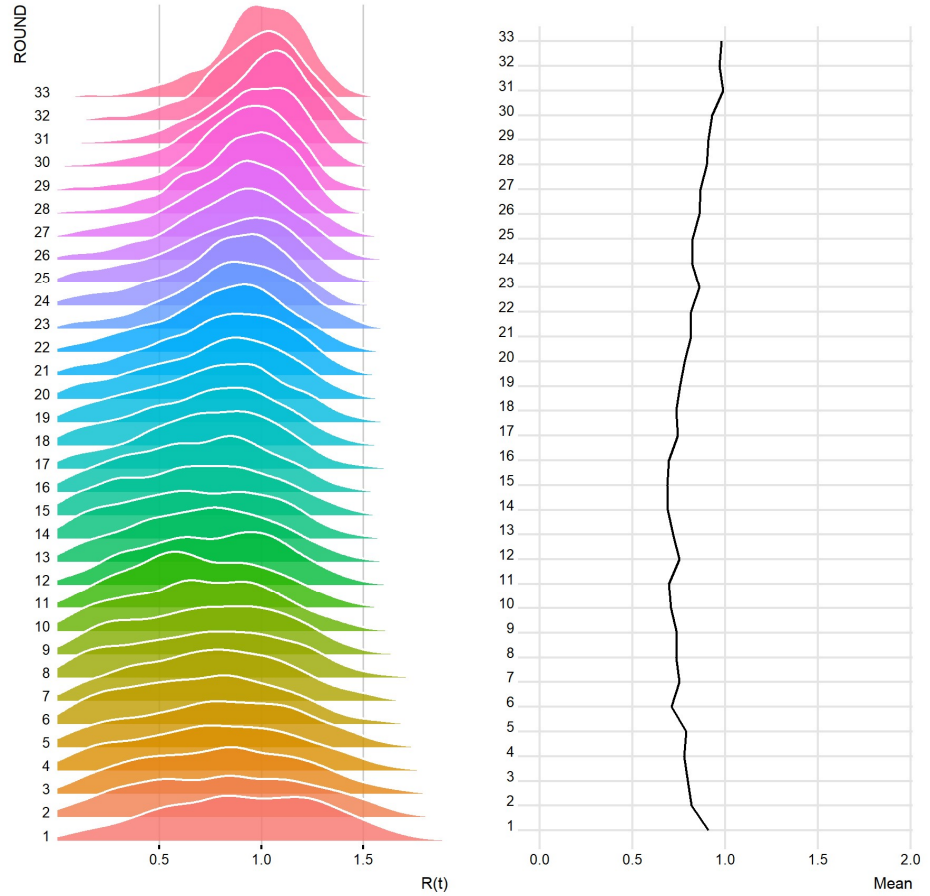
effective reproduction number, now, Oslo



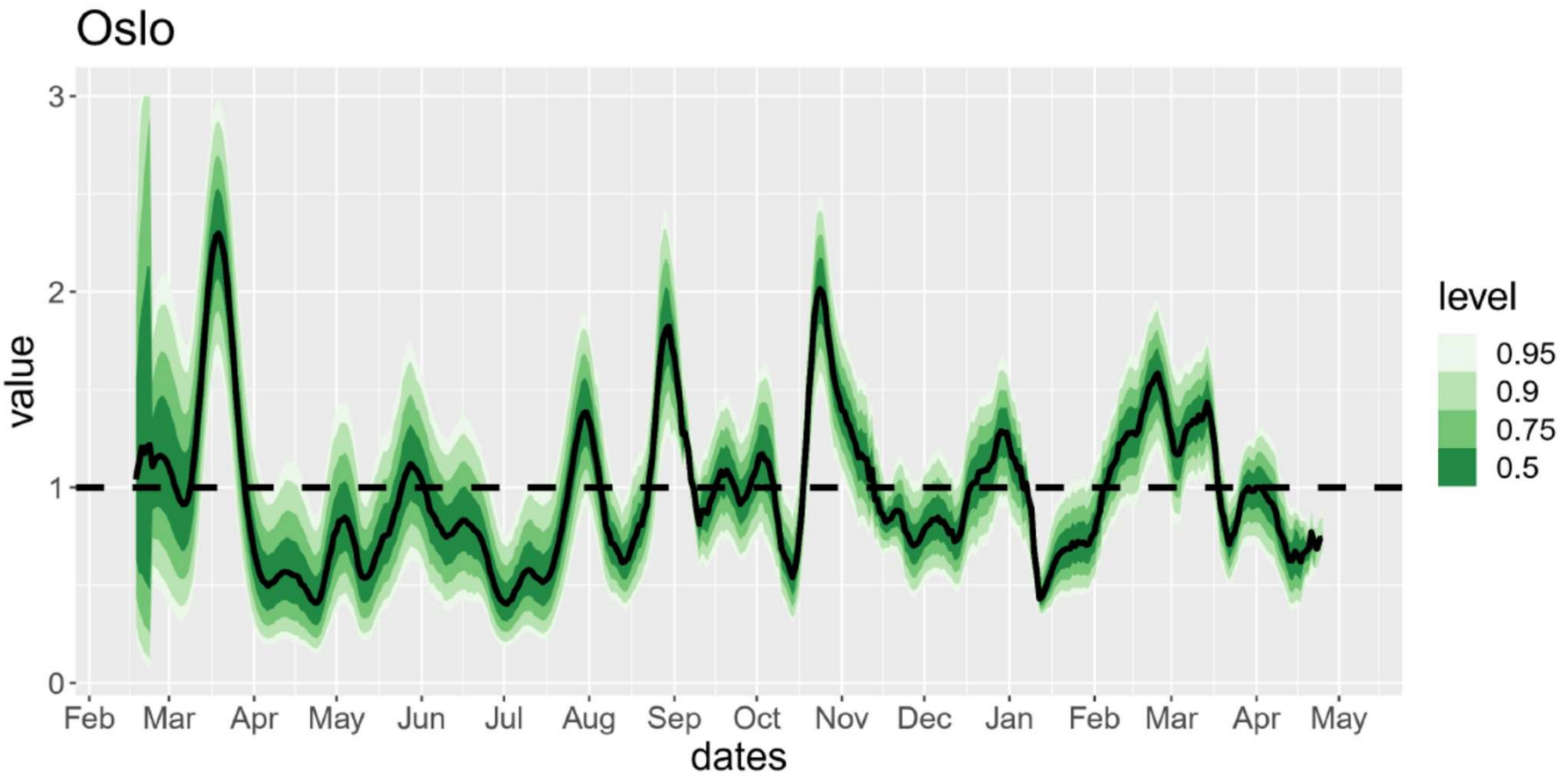
point estimate,
posterior mean

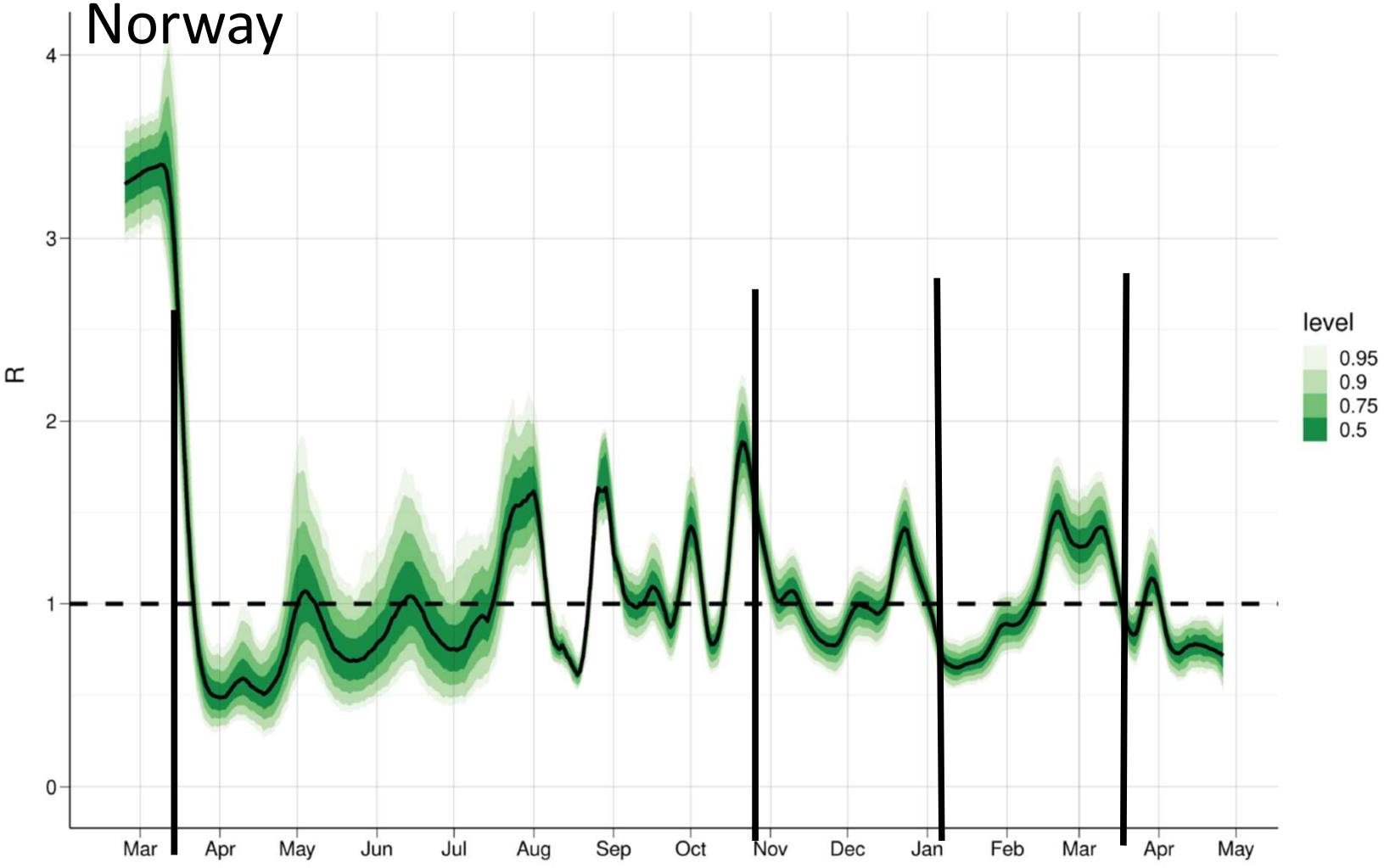
Some counties with
very uncertain estimates.

effective reproduction number, Rogaland



0 1.5


Oslo
value
3
2
1
0
Feb
Mar
Apr
May
Jun
Jul
Aug
Sep
Oct
Nov
Dec
Jan
Feb
Mar
Apr
May
dates
level
0.95
0.9
0.75
0.5



(12 mar – 20 apr)

(28 oct – 2 dec)

National long-term scenarios with vaccination plans and future interventions

1. PMAJ Pfizer, Moderna, AstraZeneca, Janssen
2. PMJ Pfizer, Moderna, Janssen
3. PM Pfizer, Moderna

- Vaccine deliveries follow a realistic scenario.
- The roll-out accounts for regional prioritization,+20% to Oslo, Moss, Sarpsborg Fredrikstad.
- Vaccine uptake: 90% in all age groups and full adherence to the vaccination schedule.
- Vaccine efficacy for Pfizer and Moderna: 1. dose 60%, 2. dose 90%, for others: 60%, 67%.
- With and without seasonal effects.
- No waning immunity after infection or vaccination.

- The initial conditions in each municipalities follows the last estimates of the regional effective reproduction number.
- Number of vaccinations made by today is known.
- The long term scenario results are based on 100 simulations and accounts for stochasticity within the model.

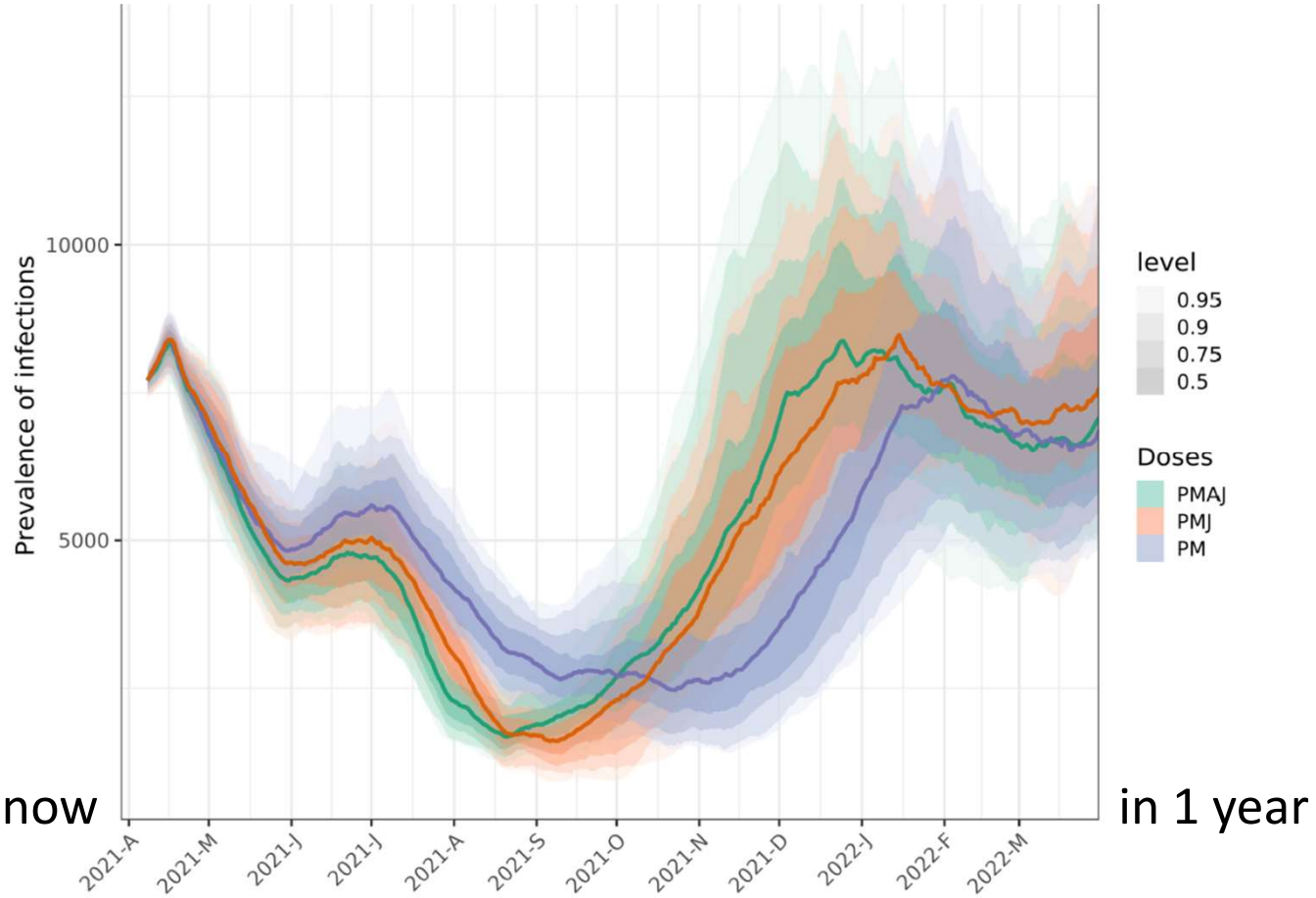
FUTURE interventions:

The government actively controls the reopening of the society:

- When more than 200 hospital beds are occupied, interventions are started that lower the reproduction number to 0.8.
- When less than 50 hospital beds are occupied, interventions are relaxed and the reproduction number increases to 1.2.

The number of hospital admissions is evaluated every three weeks

Scenario
as per end of March

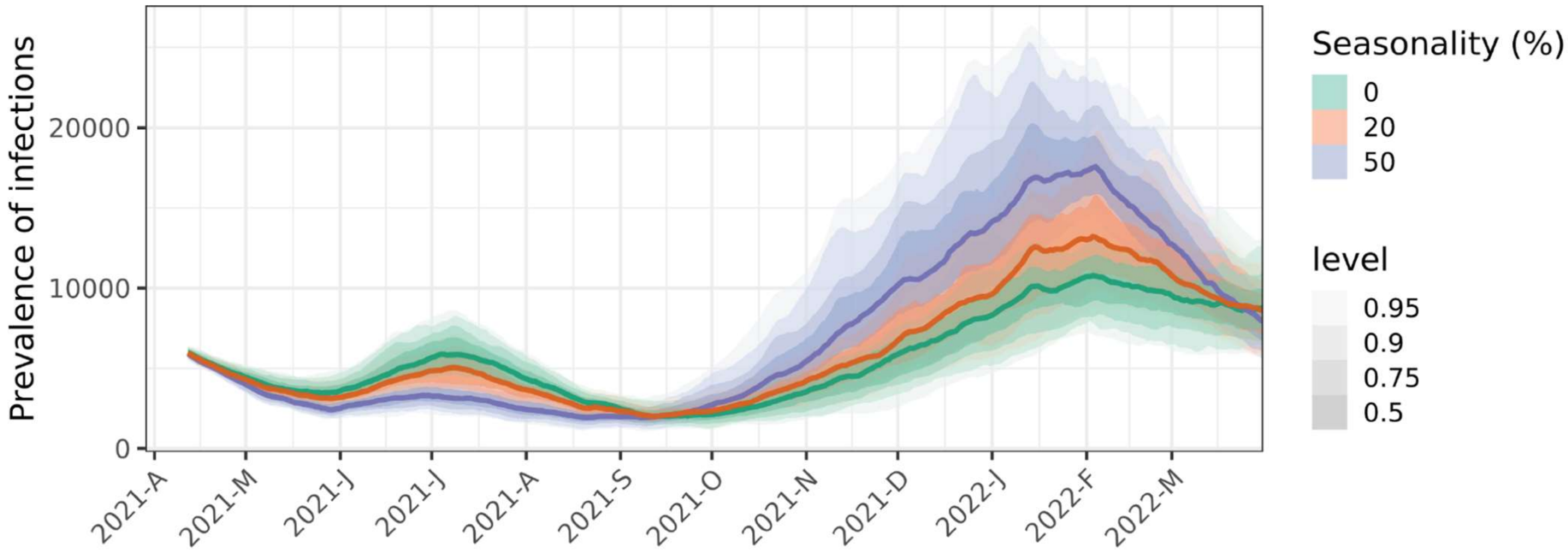


now in 1 year

PMAJ PMJ PM

Infections	374314 (330838-417789)	373283 (334684-411881)	356206 (320438-391974)
Hospitalisations	8237 (7469-9004)	8117 (7450-8785)	7467 (6955-7979)
Ventilator treatments	631 (569-692)	617 (563-671)	561 (520-601)

Scenario only PM, with seasonal effect as per end of April



Science in emergency

Boris Johnson, 12 March press conference

“It is inevitable that most people will get the disease, so we should let the epidemic proceed to allow 60% of the population to become infected and build **herd immunity**.”

https://www.theguardian.com/commentisfree/2020/mar/15/uk-covid-19-strategy-questions-unanswered-coronavirus-outbreak

16 March 2020

16 March 2020 *Imperial College COVID-19 Response Team*

Report 9: Impact of non-pharmaceutical interventions (NPIs) to reduce COVID-19 mortality and healthcare demand

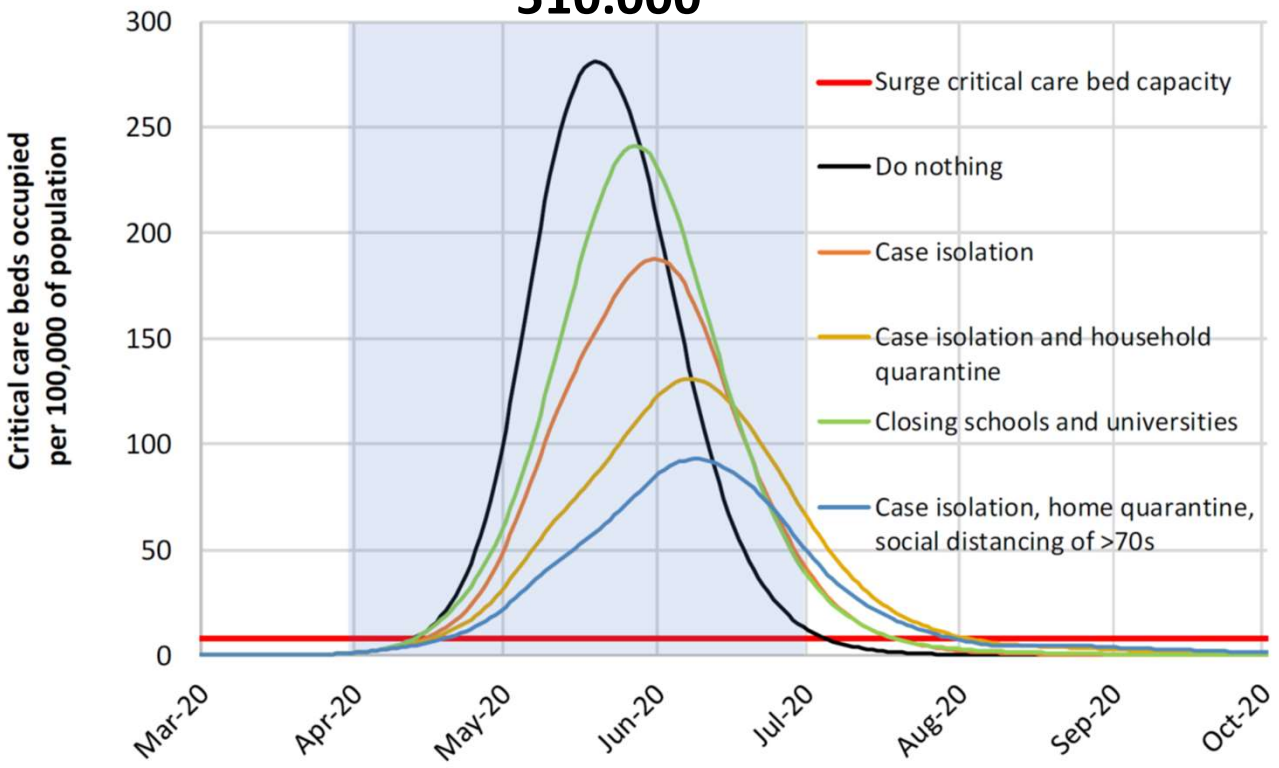
Neil M Ferguson, Daniel Laydon, Gemma Nedjati-Gilani, Natsuko Imai, Kylie Ainslie, Marc Baguelin, Sangeeta Bhatia, Adhiratha Boonyasiri, Zulma Cucunubá, Gina Cuomo-Dannenburg, Amy Dighe, Ilaria Dorigatti, Han Fu, Katy Gaythorpe, Will Green, Arran Hamlet, Wes Hinsley, Lucy C Okell, Sabine van Elsland, Hayley Thompson, Robert Verity, Erik Volz, Haowei Wang, Yuanrong Wang, Patrick GT Walker, Caroline Walters, Peter Winskill, Charles Whittaker, Christl A Donnelly, Steven Riley, Azra C Ghani.

On behalf of the Imperial College COVID-19 Response Team

A stochastic, spatially structured, individual-based discrete time simulation model.

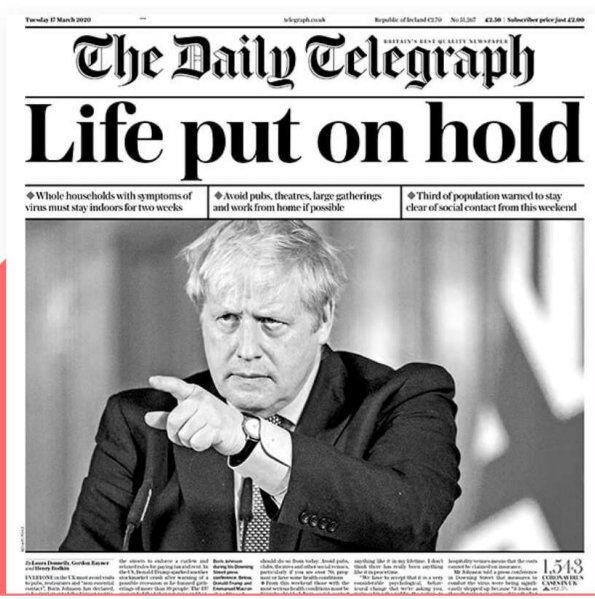
Do nothing

510.000



20.000

Suppression



The Daily Telegraph

Life put on hold

◆ Whole households with symptoms of virus must stay indoors for two weeks

◆ Avoid pubs, theatres, large gatherings and work from home if possible

◆ Third of population warned to stay clear of social contact from this weekend

Mar
17

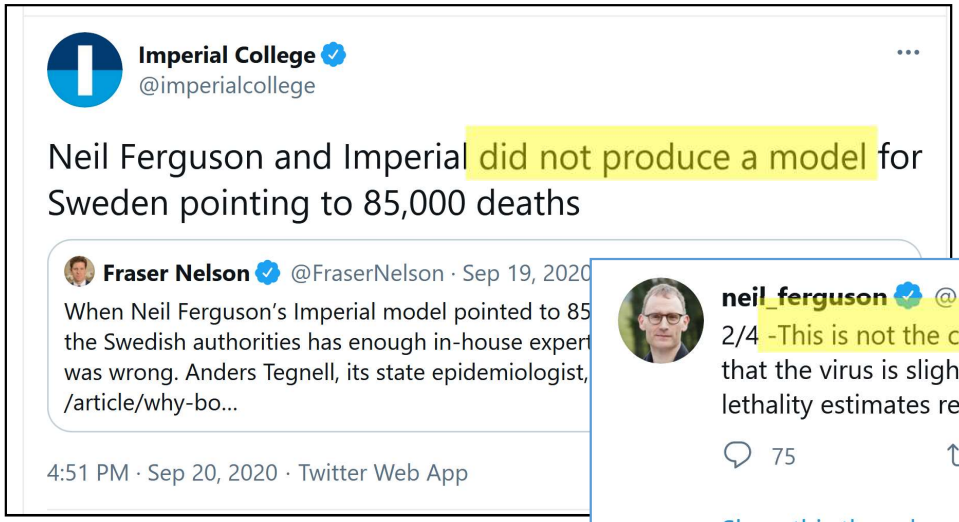
Vallance says 20,000 Covid deaths a "good outcome"

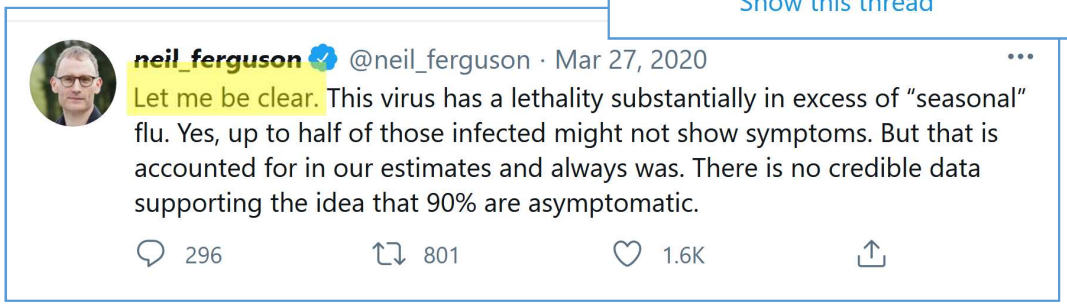
Sir Patrick Vallance tells MPs that **keeping the total number of Covid-19 deaths under 20,000 would be "a**

Mar
23

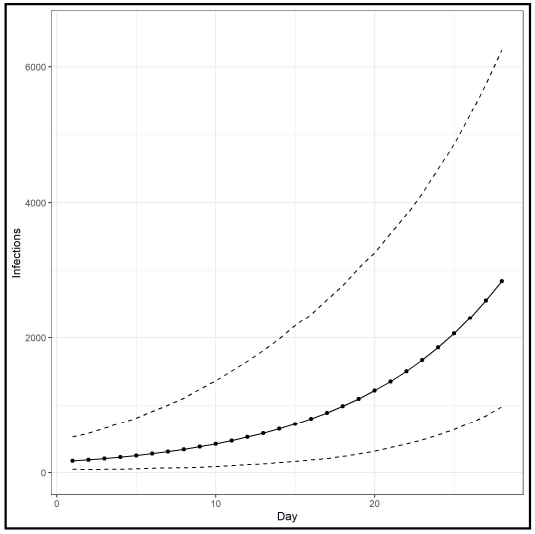
UK lockdown comes into effect

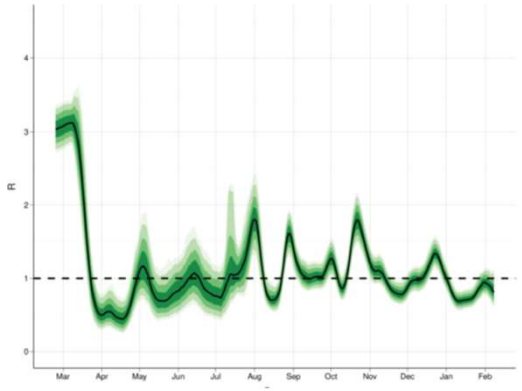
Johnson announces people may only leave their homes for **strictly limited reasons** and gives the police powers to enforce the rules. The furlough scheme introduced by

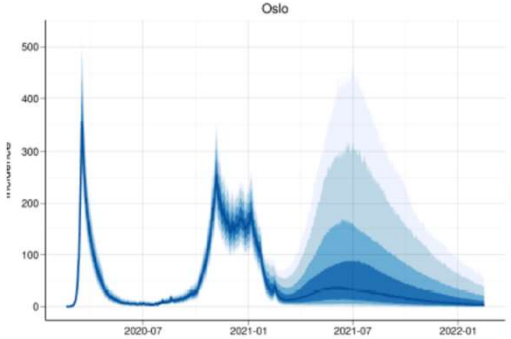


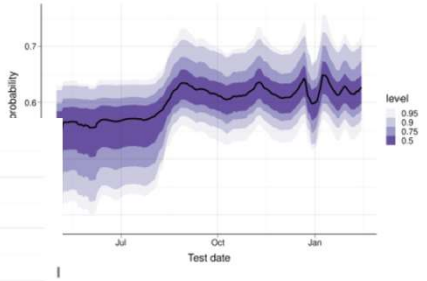


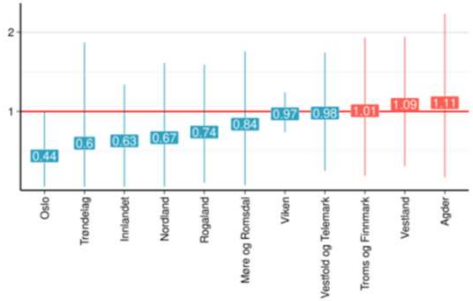
This online publication has been corrected. The first corrected version first appeared at thelancet.com/infection on April 15, 2020 and the second on May 4, 2020


Infections
6000
4000
2000
0
0
10
20
Day

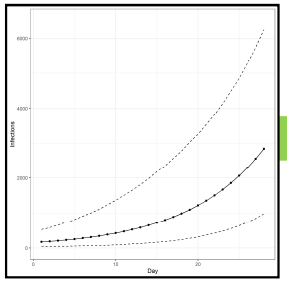

R
4
3
2
1
0
Mar
Apr
May
Jun
Jul
Aug
Sep
Oct
Nov
Dec
Jan
Feb

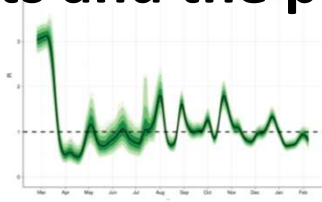

Oslo
500
400
300
200
100
0
2020-07
2021-01
2021-07
2022-01

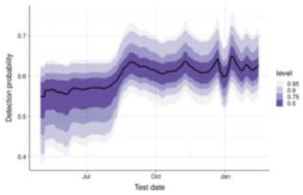

probability
0.7
0.6
level
0.95
0.9
0.75
0.5
Jul
Oct
Jan
Test date

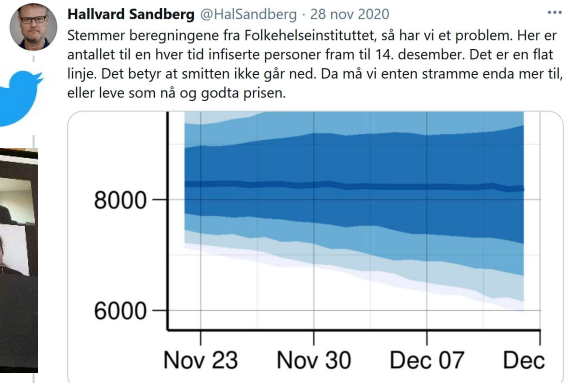

2
1
0.44
0.6
0.63
0.67
0.74
0.84
0.97
0.98
1.01
1.09
1.11
Oslo
Trøndelag
Innlandet
Nordland
Rogaland
Møre og Romsdal
Viken
Vestfold og Telemark
Troms og Finnmark
Vestland
Agder

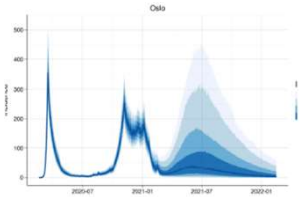
Strongly, repeatedly influenced by the specialists and the public

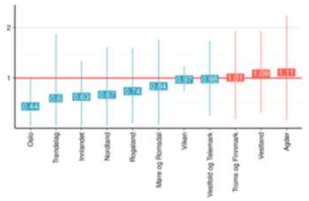


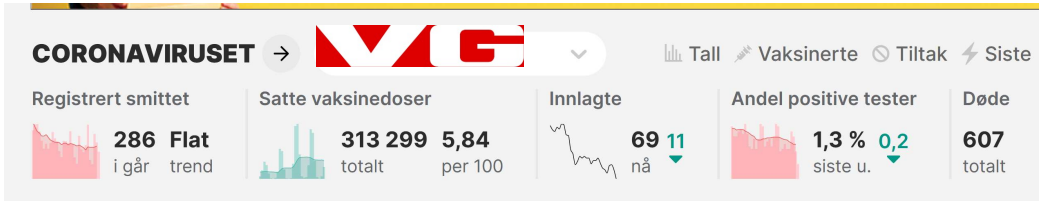


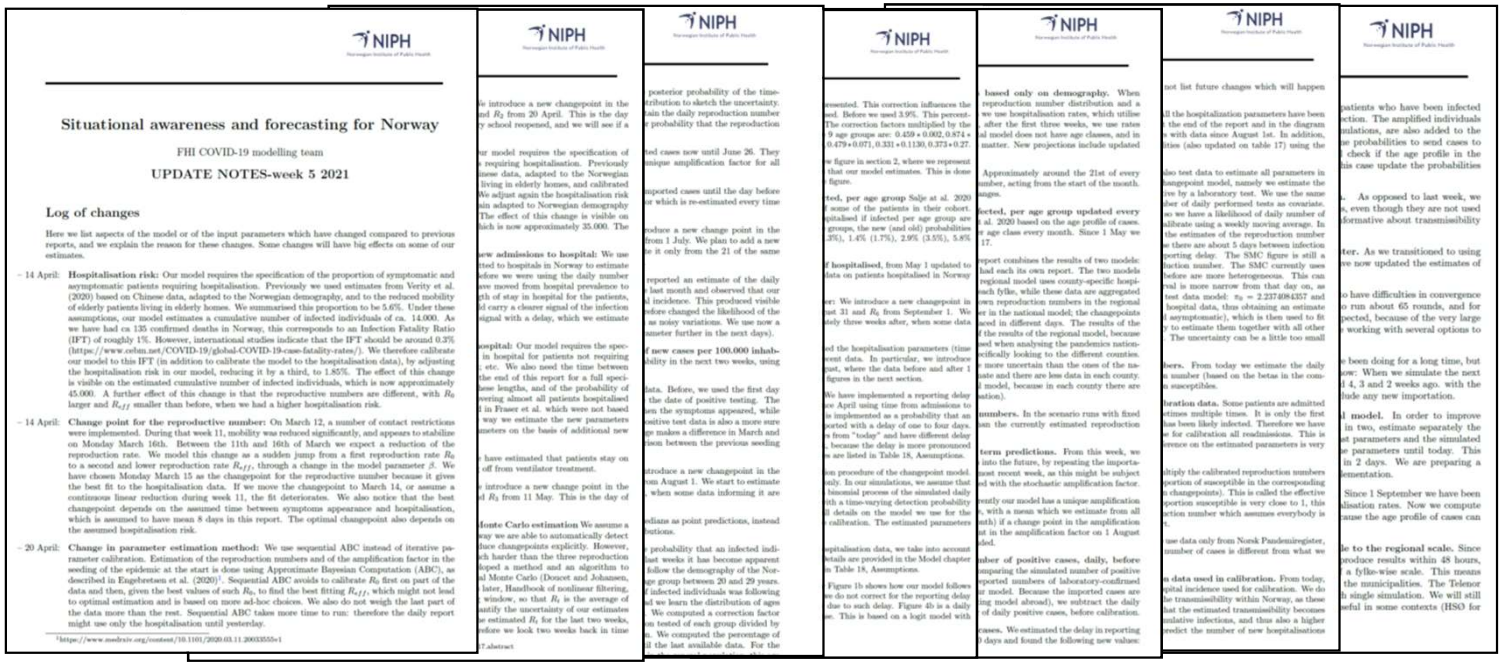












Situational awareness and forecasting for Norway

FHI COVID-19 modelling team

UPDATE NOTES-week 5 2021

Log of changes

Here we list aspects of the model or of the input parameters which have changed compared to previous reports, and we explain the reason for these changes. Some changes will have big effects on some of our estimates.

- 14 April: **Hospitalisation risk:** Our model requires the specification of the proportion of symptomatic and asymptomatic patients requiring hospitalisation. Previously we used estimates from Verity et al. (2020) based on Chinese data, adapted to the Norwegian demography, and to the reduced mobility of elderly patients living in elderly homes. We summarised this proportion to be 5.6%. Under these assumptions, our model estimates a cumulative number of infected individuals of ca. 14.000. As we have had ca 135 confirmed deaths in Norway, this corresponds to an Infection Fatality Ratio (IFT) of roughly 1%. However, international studies indicate that the IFT should be around 0.3% (https://www.cebm.net/COVID-19/global-COVID-19-case-fatality-rates/). We therefore calibrate our model to this IFT (in addition to calibrate the model to the hospitalisation data), by adjusting the hospitalisation risk in our model, reducing it by a third, to 1.85%. The effect of this change is visible on the estimated cumulative number of infected individuals, which is now approximately 45.000. A further effect of this change is that the reproductive numbers are different, with R_0 larger and R_{eff} smaller than before, when we had a higher hospitalisation risk.
- 14 April: **Change point for the reproductive number:** On March 12, a number of contact restrictions were implemented. During that week 11, mobility was reduced significantly, and appears to stabilize on Monday March 16th. Between the 11th and 16th of March we expect a reduction of the reproduction rate. We model this change as a sudden jump from a first reproduction rate R_0 to a second and lower reproduction rate R_{eff}, through a change in the model parameter β. We have chosen Monday March 15 as the changepoint for the reproductive number because it gives the best fit to the hospitalisation data. If we move the changepoint to March 14, or assume a continuous linear reduction during week 11, the fit deteriorates. We also notice that the best changepoint depends on the assumed time between symptoms appearance and hospitalisation, which is assumed to have mean 8 days in this report. The optimal changepoint also depends on the assumed hospitalisation risk.
- 20 April: **Change in parameter estimation method:** We use sequential ABC instead of iterative parameter calibration. Estimation of the reproduction numbers and of the amplification factor in the seeding of the epidemic at the start is done using Approximate Bayesian Computation (ABC), as described in Engbretsen et al. (2020)[1]. Sequential ABC avoids to calibrate R_0 first on part of the data and then, given the best values of such R_0, to find the best fitting R_{eff}, which might not lead to optimal estimation and is based on more ad-hoc choices. We also do not weigh the last part of the data more than the rest. Sequential ABC takes more time to run: therefore the daily report might use only the hospitalisation until yesterday.

[1] https://www.medrxiv.org/content/10.1101/2020.03.11.20033555v1

Christophe Fraser @ChristoPhraser · 8h

I feel like we have a recurring problem on what to in time-periods when evidence is still being collected. Zero effect of vaccines on transmission is not a reasonable prior. Bayesian thinking seems highly appropriate. We can make temporary assumptions and update them.

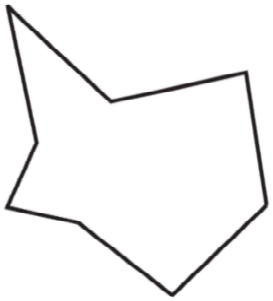
A new way of doing science?

Evidence-based science:

- available scientific knowledge and evidence from data lead to results which are “final given the current situation”.
- Results can be used *as-is*.
- Predictions will be checked in the future, and will be compared to reality.

Evidence-making science:

- models are always **starting points** and not the final truth;
- models, hypothesis, data **change**, also on the basis of needs and questions from the public, media and decision makers.
- Emergency requires **rapid results**. There is no time to curate data, to test code properly.
- How do you communicate to the public that your results might be wrong and are likely to change tomorrow?
- Scientists are not (always) **coherent** with themselves. Scientists must learn to say “I was wrong” and public must appreciate this, instead than getting more sceptical about science.
- **Uncertainty** means that model sets boundaries to describe the possible future. These boundaries create a space, a playground for decisions, where ethics, policitics, economy, culture must be advoked.



https://www.fhi.no/contentassets/e6b5660fc35740c8bb2a32bfe0cc45d1/vedlegg/nasjonale-og-regionale-rapporter/national_regional_model_05may2021.pdf

Lindstrøm, J. C., Engebretsen, S., Kristoffersen, A. B., Rø, G. Ø. I., Palomares, A. D. L., Engø-Monsen, K., ... & de Blasio, B. F. (2021). Increased transmissibility of the B. 1.1. 7 SARS-CoV-2 variant: Evidence from contact tracing data in Oslo, January to February 2021. *medRxiv*.

Engebretsen, S., Engø-Monsen, K., Aleem, M. A., Gurley, E. S., Frigessi, A., & de Blasio, B. F. (2020). Time-aggregated mobile phone mobility data are sufficient for modelling influenza spread: the case of Bangladesh. *Journal of the Royal Society Interface*, *17*(167).

Williams, G. A., Díez, S. M. U., Figueras, J., & Lessof, S. (2020). TRANSLATING EVIDENCE INTO POLICY DURING THE COVID-19 PANDEMIC: BRIDGING SCIENCE AND POLICY (AND POLITICS). *Eurohealth* 2020; 26(2).

Rhodes, Tim, and Kari Lancaster. "Mathematical models as public troubles in COVID-19 infection control: following the numbers." *Health Sociology Review* 29.2 (2020): 177-194.

Mangoni, A. A., & Erre, G. L. (2021). Translating evidence into practice during the COVID-19 pandemic: pitfalls and mileages. *Therapeutic Advances in Drug Safety*, *12*.