

# SHORT-TERM PREDICTIONS OF COVID-19 HOSPITALIZATION NUMBERS

---

## Håkon K. Gjessing

Professor/Principal Investigator

Centre for Fertility and Health, Norwegian Institute of Public Health, Oslo

Department of Global Public Health and Primary Care, University of Bergen

### In collaboration with

Geir Bukholm (NIPH)

Jostein Starrfelt (NIPH)

Robert Neil Whittaker (NIPH)

Hilde Kløvstad (NIPH)

Per Magnus (NIPH)

Line Vold (NIPH)

Siri Skodvin (NIPH)

---

Makerere, Friday 9 June 2023

## BACKGROUND & MOTIVATION

- The emergency registry BeredtC19 merges information from a large number of sources
  - MSIS, Norwegian Pandemic Registry, Norwegian intensive care registry, Sysvak etc....
- Speed of update has been remarkable
  - Confirmed positive tests are usually updated within days of testing
  - Hospitalizations with similar speed, though perhaps some more lag in particular cases
- **Amazing job!** by so many people to get the registry up and running

## BACKGROUND & MOTIVATION

The **NIPH modelling team** has done an excellent job from the very start of the pandemic in computing R-numbers, both regionally and nationally, providing long-term scenarios and also short-term predictions.



- Their work has been presented in regular report, as well as in answers to numerous “oppdrag” (requests) from the Ministry of Health and Care Services as well as the Norwegian Directorate of Health.
- Various other groups/teams have provided solid and frequently updated reports on the development, both for the use by NIPH as well as for use by the press.

# BACKGROUND & MOTIVATION

**Standard Reporting:**  
Incidences  
Weekly totals  
Various summary statistics  
Etc.!

**Research papers:**  
Vaccine efficacy  
Vaccine side effects  
Intervention effects  
Etc.!

**Simulation models for spread of infection:**  
Computationally very intensive  
Can include lots of factors  
Can provide future “scenarios”  
Etc.!

- However, in between the “standard reporting”, regular research, and heavy simulation models there is room & need for data-driven models in surveillance.
- Speed is a decisive factor.

# STATISTICAL MODELS FOR SHORT-TERM PREDICTION AND CONTROL

- **Extremely useful to inform day-to-day decisions in pandemic control**

## Rationale

- Pandemic simulation-based long-term modelling
  - Complex, very sensitive to assumptions, but also very flexible
  - Computationally very heavy and time-consuming
  - Very useful to generate long term “scenarios”
  - Needed for short-term modelling?
- Short-term models
  - Need not worry quite that much about feedback, non-linearity etc.
  - Short-term models can be fast and “easy” to update on a regular basis
  - Based on combinations of “regular” statistical models... huge speed gains
  - Can be seen as extensions of standard reporting... or simplifications of simulation models

# BACKGROUND & MOTIVATION, ASTRAZENECA-ANALYSES

Oppdrag fra HOD om koronavirusvaksinasjonsprogrammet nr. 21

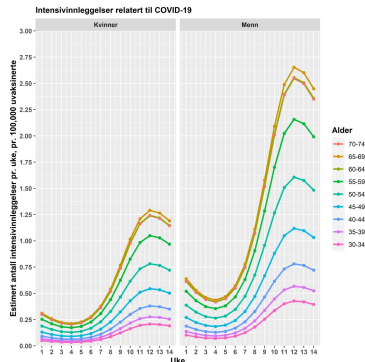


Fig 1: Resultatene fra prediksjonsmodellen er her vist, avgrenset til 2021. Beregningene er vist separat for kjønn og alderskategorier. Estimert Covid-19-relaterte intensivinnleggelses er beregnet som forventet antall intensivinnleggelses pr. uke pr. 100.000 uvaksinerte personer. Vaksinerte er tatt ut av modellen etter vaksinerings.

52

Oppdrag fra HOD om koronavirusvaksinasjonsprogrammet nr. 21

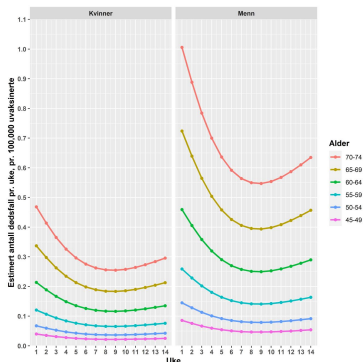
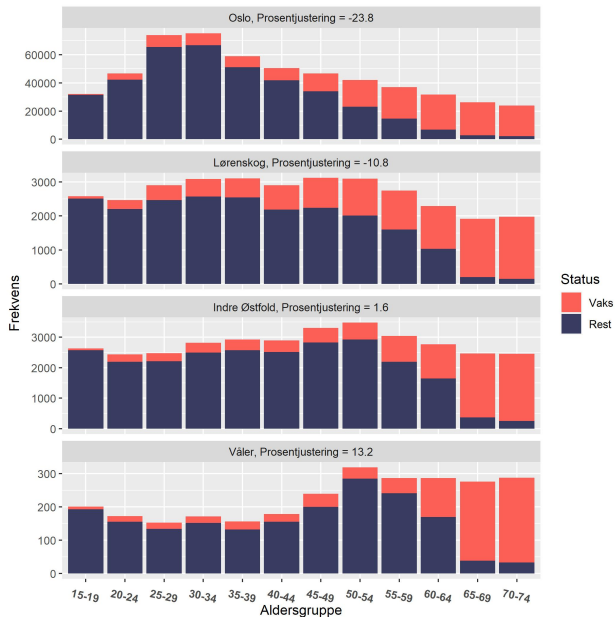


Fig 1: Resultatene fra prediksjonsmodellen er her vist, avgrenset til 2021. Beregningene er vist separat for kjønn og alderskategorier. Estimert Covid-19-relatert dødelighet er beregnet som forventet antall dødsfall pr. uke pr. 100.000 uvaksinerte personer. Vaksinerte er tatt ut av modellen etter vaksinerings.

47

# BACKGROUND & MOTIVATION, VACCINE DISTRIBUTION

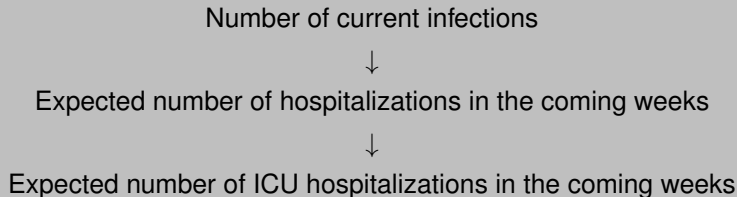


# BACKGROUND & MOTIVATION, SHORT-TERM PREDICTIONS

## Objective

- Predict hospitalizations and ICU admissions from current infection rates
- “Infection” means positive PCR test (MSIS)
- Hospitalizations have covid-19 as main cause

## FOCUS ON SURVEILLANCE



## How valid is this relationship?

- Quite stable, *as long as conditions remain the same*

- The number of infections are high; relatively stable estimates of infections
- An “undisturbed” pandemic is exponential (over modest time intervals)
- The relationship between infections and hospitalizations is quite reliable and can be estimated over a fairly large time interval
- Predictions 2–3 weeks into the future will provide you with some time to implement interventions

## VARIABLES TO BE USED IN THE MODEL

- Age (5-year intervals)
- Vaccination Status: Unvaccinated, Vaccinated (1 or 2 doses), Vaccinated (2+booster)
- Risk group (risk of serious development of infection)
- Sex
- **Calendar Time**
- (Municipality, county)

# NECESSARY MODEL COMPONENTS

- 1 Infection prediction:  
# of infections until today → # of infections in coming weeks
- 2 Estimated probability: infection → hospital admission
- 3 Estimated probability:  
admitted to hospital same days as positive PCR test
- 4 Estimated time distribution:  
positive test → hospital admission later than positive PCR test
- 5 Estimated time distribution: hospital length of stay if admitted
- 6 Estimated probability: hospital admission → ICU admission
- 7 Estimated time distribution: hospital admission → ICU admission
- 8 Estimated time distribution: length of stay in ICU

... and all of these may depend on the listed covariates(!)

... in particular, they may change over time...

## SHORT-TERM PREDICTIONS, STARTUP TIMELINE, 2021

- Tuesday 3 Aug, first meetings about the project (initiated by Geir Bukholm)
- Wednesday 4 Aug, I sent the data request in the morning



## SHORT-TERM PREDICTIONS, STARTUP TIMELINE, 2021

- Tuesday 3 Aug, first meetings about the project (initiated by Geir Bukholm)
- Wednesday 4 Aug, I sent the data request in the morning
- Mid Friday 6 Aug, I got access to (an excellent) data set



# SHORT-TERM PREDICTIONS, STARTUP TIMELINE, 2021

- Tuesday 3 Aug, first meetings about the project (initiated by Geir Bukholm)
- Wednesday 4 Aug, I sent the data request in the morning
- Mid Friday 6 Aug, I got access to (an excellent) data set



- Saturday 7 Aug



# SHORT-TERM PREDICTIONS, STARTUP TIMELINE, 2021

- Tuesday 3 Aug, first meetings about the project (initiated by Geir Bukholm)
- Wednesday 4 Aug, I sent the data request in the morning
- Mid Friday 6 Aug, I got access to (an excellent) data set



- Saturday 7 Aug



- Monday 9 Aug at 09:00, temporary results presented to the Minister of Health and the political heads of the Health Ministry



## SHORT-TERM PREDICTIONS, STARTUP TIMELINE, 2021

- Tuesday 3 Aug, first meetings about the project (initiated by Geir Bukholm)
- Wednesday 4 Aug, I sent the data request in the morning
- Mid Friday 6 Aug, I got access to (an excellent) data set



- Saturday 7 Aug



- Monday 9 Aug at 09:00, temporary results presented to the Minister of Health and the political heads of the Health Ministry
- Wednesday 11 Aug: New presentation of results to the Minister of Health. Most of the results ready
- Thursday 12 Aug: Camilla presented the essence of it at a government conference deciding on whether re-opening the country was possible
- Friday 13 Aug: Government press conference about re-opening plans



# PROBLEM

The roof of my summer cabin wasn't entirely waterproof at the moment...



# NECESSARY MODEL COMPONENTS

- 1 Infection prediction:  
# of infections until today → # of infections in coming weeks
- 2 Estimated probability: infection → hospital admission
- 3 Estimated probability:  
admitted to hospital same days as positive PCR test
- 4 Estimated time distribution:  
positive test → hospital admission later than positive PCR test
- 5 Estimated time distribution: hospital length of stay if admitted
- 6 Estimated probability: hospital admission → ICU admission
- 7 Estimated time distribution: hospital admission → ICU admission
- 8 Estimated time distribution: length of stay in ICU

... and all of these may depend on the listed covariates(!)

... in particular, they may change over time...

# POISSON REGRESSION FOR THE NUMBER OF POSITIVE TESTS

$$\log(\lambda) = \text{sex} + \text{risk group} + (\text{age} * \text{vaccination status}) / \text{testdate}$$

$\lambda$  = expected number of positive tests

- sex: Two levels
- risk group: Three levels
- age: 5-year groups
- vaccination status: Two levels
- testdate: Three previous weeks for fitting, three coming weeks for predictions
- (municipality, county)

Plus: A weight variable, with weights from 0.1 up to 1 from the earliest to the latest date when fitting (or variants thereof)

# POISSON REGRESSION FOR THE NUMBER OF POSITIVE TESTS

$\log(\lambda) = \text{sex} + \text{risk group} + (\text{age} * \text{vaccination status}) / \text{testdate}$

$\lambda = \text{expected number of positive tests}$

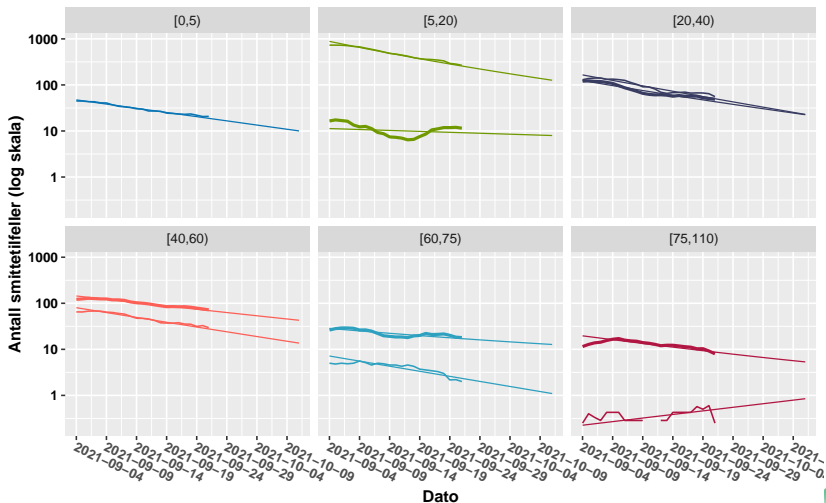
## Notation:

- $(\text{age} * \text{vaccination status}) / \text{testdate}$
- All combinations of `age` and `vaccination status`
- Effect of `testdate` estimated separately in each category

That is, a three-way interaction, but a bit more conveniently coded

# OBSERVED AND MODEL EXTRAPOLATED NUMBER OF POSITIVE TESTS

Vaksinestatus — Uvaksinert — DelFullvaksinert



# POISSON REGRESSION FOR THE NUMBER OF POSITIVE TESTS

$$\log(\lambda) = \text{sex} + \text{risk group} + (\text{age} * \text{vaccination status}) / \text{testdate}$$

$\lambda$  = expected number of positive tests

## Note:

- No  $\log(\text{offset})$ , i.e. “exposure” in the model
- That is, the model is entirely “pragmatic”,...  
Descriptive, for prediction purposes
- Coefficients cannot be directly interpreted as “vaccine effects”, rates, or suchlike
- May be more correct with, e.g. negative binomial if standard errors are to be computed

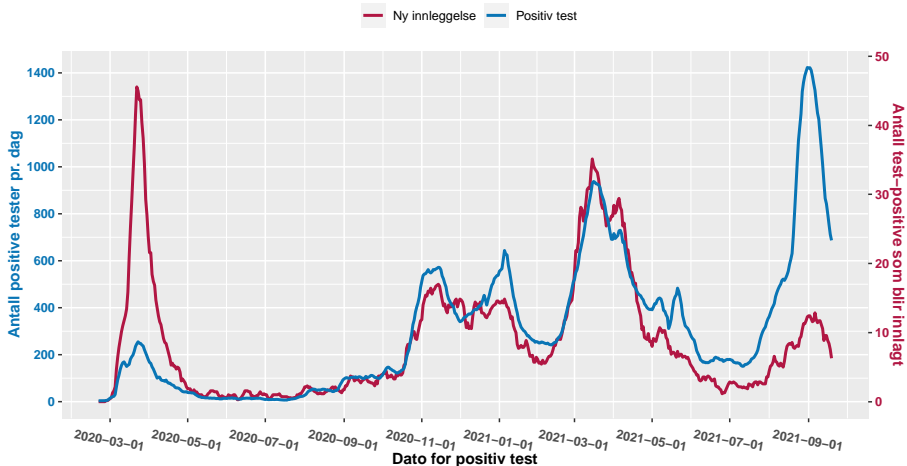
# NECESSARY MODEL COMPONENTS

- 1 Infection prediction:  
# of infections until today → # of infections in coming weeks
- 2 Estimated probability: infection → hospital admission
- 3 Estimated probability:  
admitted to hospital same days as positive PCR test
- 4 Estimated time distribution:  
positive test → hospital admission later than positive PCR test
- 5 Estimated time distribution: hospital length of stay if admitted
- 6 Estimated probability: hospital admission → ICU admission
- 7 Estimated time distribution: hospital admission → ICU admission
- 8 Estimated time distribution: length of stay in ICU

... and all of these may depend on the listed covariates(!)

... in particular, they may change over time...

# POSITIVE TESTS AND HOSPITALIZATIONS



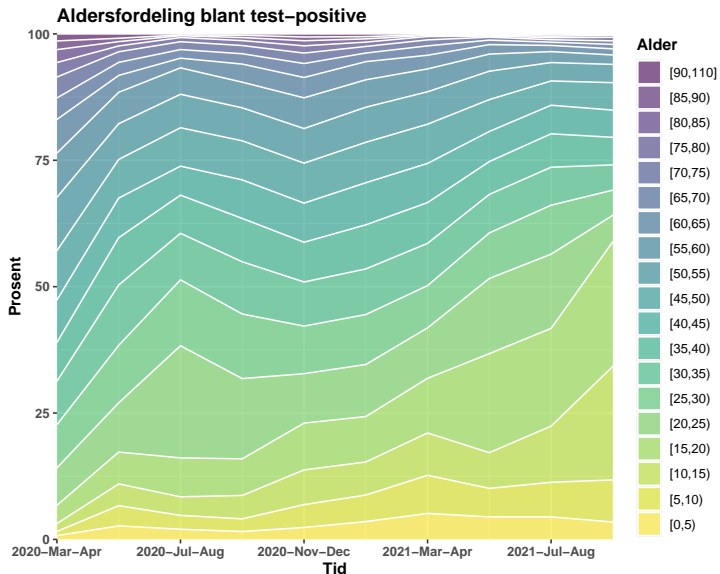
Blue curve: Number of positive tests on a given day

Red curve: Number of those who at some point gets admitted to a hospital

## Probability of hospitalization

- Probability of hospitalization depends on
  - sex
  - age
  - risk group
  - vaccination status
  - calendar date
  - (municipality, county)

# AGE DISTRIBUTION AMONG TEST-POSITIVES



# FROM POSITIVE TESTS TO NEW HOSPITALIZATIONS

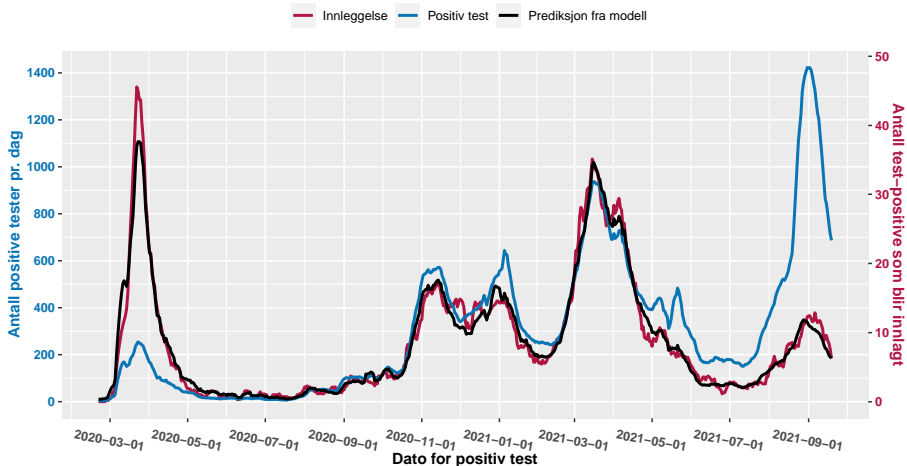
## Probability of hospitalization

Binomial family GAM (Generalized Additive Model) with a smooth term  $s()$

$$\begin{aligned} \text{logit}(P(\text{hospitalized} \mid \text{positive test})) \\ = \text{sex} + \text{age} + \text{risk group} + \text{vaccination status} + s(\text{testdate}) \end{aligned}$$

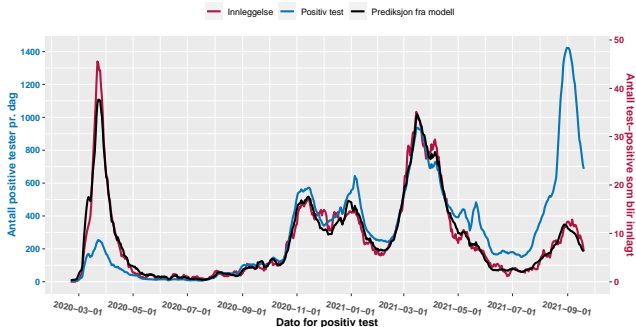
- Hopefully,  $s(\text{testdate})$  does not change quickly
- Variants with interactions also relevant
  - $\text{age} * \text{vaccination status}$  can be relevant, but isn't very strong
  - Changes in testing strategies specific to age groups, e.g. school testing, can introduce interactions between age and  $s(\text{testdate})$

# POSITIVE TESTS AND HOSPITALIZATIONS



- Checking the model fit on the number of test-positives being hospitalized

# POSITIVE TESTS AND HOSPITALIZATIONS

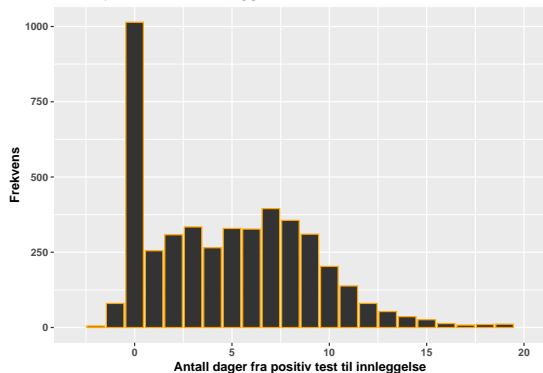


- Checking the model fit on the number of test-positives being hospitalized
- The divergence between positive tests and hospitalizations from April/May 2021 can be explained by increasing vaccination (which also leads to infections being driven by younger individuals)

# TIME FROM POSITIVE TEST TO HOSPITALIZATION

- Conditional on being hospitalized, what is the distribution of time until hospitalization?

Tid fra positiv test til innleggelse med Covid-19 som hovedårsak



Zero inflation: Some people get ill, get hospitalized, and *then* tested

# TIME FROM POSITIVE TEST TO HOSPITALIZATION

Probability of zero time

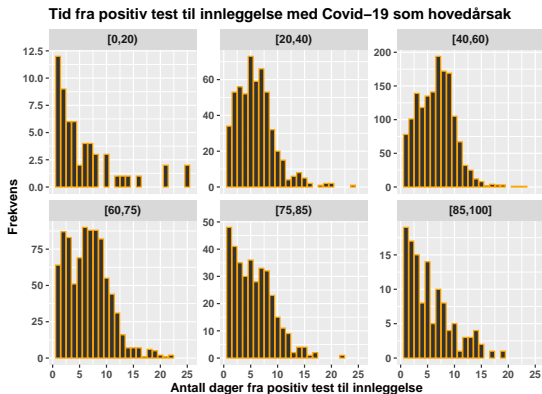
Binomial family GAM (Generalized Additive Model) with a smooth term  $s()$

$$\text{logit}(P(\text{zero time}|\text{hospitalized}))$$
$$= \text{sex} + \text{risk group} + \text{age} + \text{vaccination status} + s(\text{testdate})$$

- Variants with interactions also relevant

# TIME FROM POSITIVE TEST TO HOSPITALIZATION

(Some type of) survival/non-parametric model for the remaining



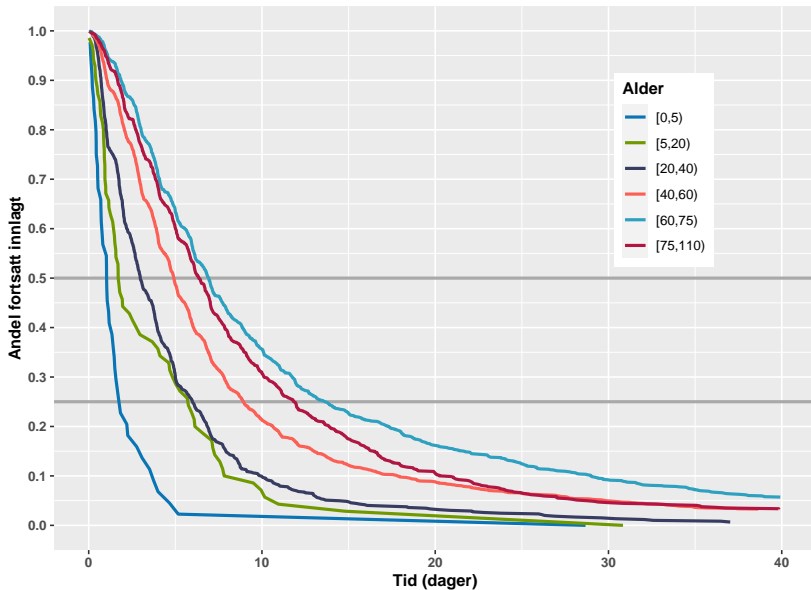
# NECESSARY MODEL COMPONENTS

- 1 Infection prediction:  
# of infections until today → # of infections in coming weeks
- 2 Estimated probability: infection → hospital admission
- 3 Estimated probability:  
admitted to hospital same days as positive PCR test
- 4 Estimated time distribution:  
positive test → hospital admission later than positive PCR test
- 5 Estimated time distribution: hospital length of stay if admitted
- 6 Estimated probability: hospital admission → ICU admission
- 7 Estimated time distribution: hospital admission → ICU admission
- 8 Estimated time distribution: length of stay in ICU

... and all of these may depend on the listed covariates(!)

... in particular, they may change over time...

# “SURVIVAL” DISTRIBUTIONS FOR LENGTH OF HOSPITAL STAY

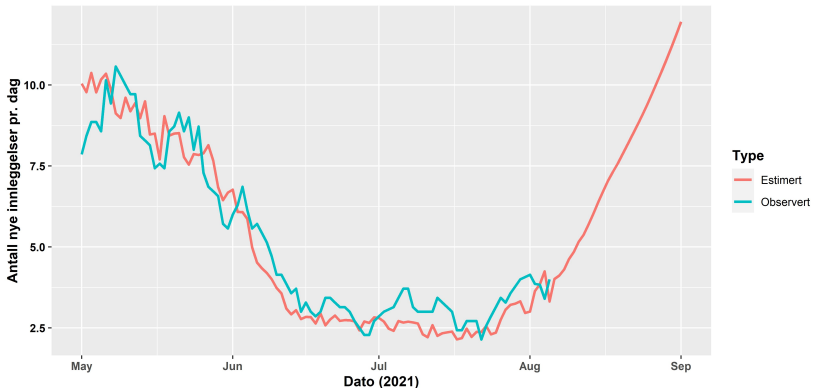


# OBSERVED AND ESTIMATED NUMBER OF NEW HOSPITALIZATIONS

## First predictions

Totalt antall nye innleggelser, observert og modellestimert

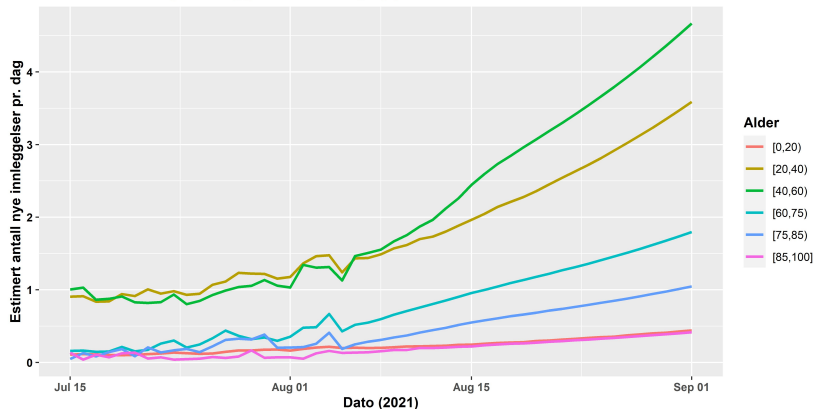
Observert er 7 dagers glidende gjennomsnitt



# OBSERVED AND ESTIMATED NUMBER OF NEW HOSPITALIZATIONS

## First predictions

Estimert antall nye innleggelser i aldersgrupper

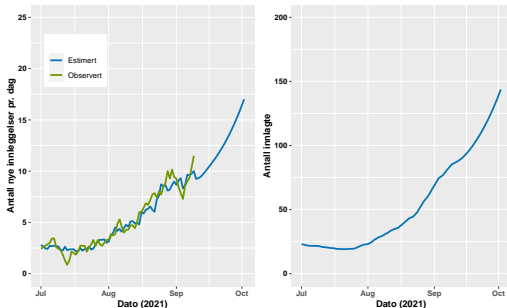


# OBSERVED AND ESTIMATED NUMBER OF NEW HOSPITALIZATIONS

## Predictions 12 Sept, expected hospitalizations



### Estimert antall innleggelser, totalt



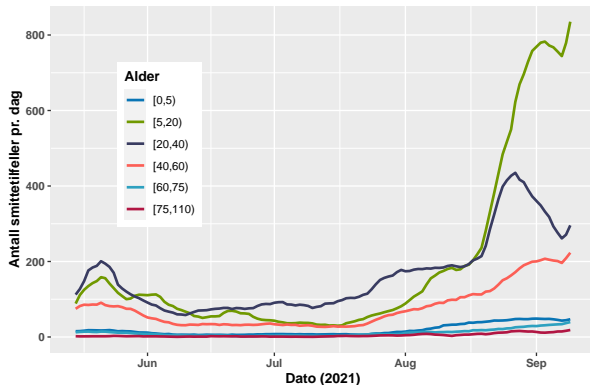
- Merk et betydelig fall i forhold til tidligere framskrivninger.

# OBSERVED NUMBER OF NEW HOSPITALIZATIONS

## 12 Sept, number of positive tests

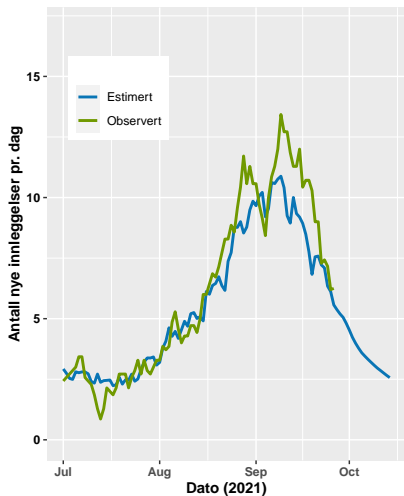
### Registrerte positive tester, i aldersgrupper

Antall smittetilfeller pr. dag i aldersgrupper, siste månedene  
7 dagers glidende gjennomsnitt

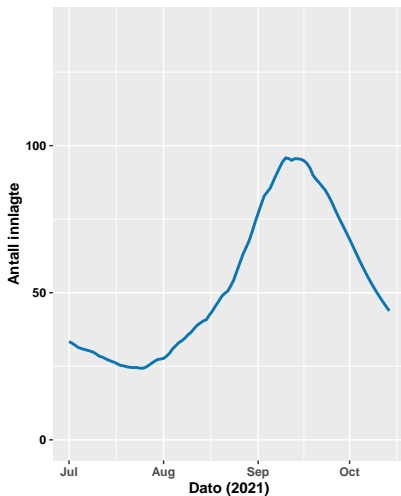


# ESTIMATED HOSPITALIZATIONS, TOTAL

## New hospitalizations



## In hospital

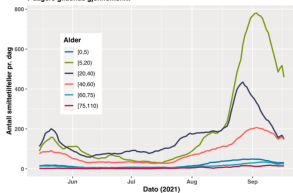




## CORONAVIRUSET

### Antall smittetilfeller pr. dag i aldersgrupper, siste månedene

7 dagers glidende gjennomsnitt



**KLARER** Disse grafene viser at smitten går størst ned i aldersgruppene 0-20 og 20-40. Dette er en klar vaksineeffekt. Gruppene 40-60 går svakt nedover. Dette er en gruppe som har vært vaksinert i en del, som forklarer hvorfor den ikke går så bratt nedover. 75-110 år: dette er den eneste som går svakt oppover, men stigningen er marginal. [Les mer](#) [Sjå illustrasjon](#)

## Dette er grafen som kan åpne Norge

De siste analysene fra FHI viser at tiden snart er inne for å åpne, sier smitteverndirektøren.

Av LINA FAUZEK og MARTHAN C. S. HOLMES  
Oppdatert 21. september



## Dette viser grafen

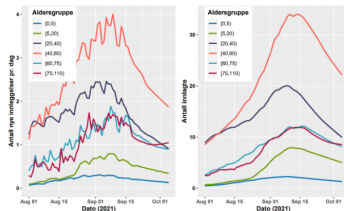
Antall smittetilfeller brutt ned på alder:

I alle aldersgrupper, er det en knekk, bortsett fra 75-110 år.

- Aldersgruppe 0-20: smitten går tydelig nedover – det er en kombinasjon av at ungdommer blir vaksinert og at smitten går ned som følge av befolkningsimmunitet
- 20-40 går sterk nedover, som er en klar vaksineeffekt
- 40-60 svakt nedover. Dette er en gruppe som har vært vaksinert i en tid, som forklarer hvorfor den ikke går så bratt nedover
- 75-110 år: dette er den eneste som går svakt oppover, men stigningen er marginal
- Estimert antall innleggelser, går veldig tydelig ned

**KLARER** [Les mer](#) [Sjå illustrasjon](#)

## Estimert antall innleggelser, etter alder



**SYNKER:** Denne grafen viser for første gang antall forventede innleggelser fordelt på alder. Til venstre ser...



## Korttidsframskrivninger sykehusinnleggelser

Dato: 28. september 2021

### Data:

- Data ekstrahert fra BeredC19 28/9 kl. 20:00.  
Siste inkluderte dag er 26/9.
- Smitte er her definert som registrert positiv test.
- Innleggelser er avgrenset til de med covid-19 som hovedårsak.

### Modellbeskrivelse:

En kort beskrivelse av modellene finnes på [slutten av rapporten](#).

- Separate reports (about 3 times a week)

**Korttidsframskrivninger:** <https://www.fhi.no/sv/smittsomme-sykdommer/corona/koronavirus-modellering/>

- ... with all figures in the official NIPH color palette!

## CAN YOU HAVE TOO MUCH DATA?



(Scarface “sampling” his “data”...)

One of the wisest remarks ever made: “Don’t get high on your own supply”

# CONCLUSIONS I

- This type of modelling indeed seems useful...
- 2-3 weeks warning is probably enough in most cases
- Similarly, important indicators for, say, vaccine distribution can be updated “in real time”
- However, model structure may need frequent updates/modifications over time

# CONCLUSIONS II

And! New roof as well...

