

Enkle modellberegninger tyder på at reproduksjonstallet for COVID-19 i Norge har vært økende siden før påske, og kan nå være større enn 1

Martin Rypdal og Kristoffer Rypdal, Institutt for matematikk og statistikk, UiT Norges arktiske universitet

Sammendrag

Folkehelseinstituttet (FHI) presenterer daglige oppdateringer på reproduksjonstallet $R(t)$ i for COVID-19 i Norge basert på tidsrekker over antall innlagte i sykehus. Instituttet er sparsomme med detaljer om estimeringsmetode, men vi har inntrykk av at den er basert på en antagelse om en konstant verdi R_0 før 15. mars og en annen verdi R_E etter denne datoen. Dette gjør at en eventuell trend i $R(t)$ framstår som en mindre korreksjon av det konstante nivået etter 15. mars, og dermed ikke avspeiler den reelle utviklingen. I dette notatet legger vi fram resultater av estimerer som gjøres med statistisk tilpasning av en mer fleksibel form av funksjonen $R(t)$. Ved å forutsette at verdien av $R(t)$ holder et konstant nivå fra slutten av mars, får vi resultater svært like det som FHI presenterer for COVID-19 i Norge i dag, med tilsvarende små konfidensintervaller. Hvis vi derimot tillater $R(t)$ å ha en lineær trend med uspesifisert stigningsforhold fra slutten av mars, finner vi en signifikant positiv trend som gir mest sannsynlig verdi $R = 1.08$ den 29. april, med 95% konfidensintervall mellom 0.75 og 1.47. Metoden gir god tilpasning data for antall innlagte, og også resultater konsistente med tidsrekkene for positive testresultat. Den gir også prediksjoner for utviklingen framover. Her er det verd å merke seg at til tross for denne positive trenden i $R(t)$, så vil den mest sannsynlige banen for antall innlagte være en fortsatt nedgang i antallet fram til midten av mai før det vil begynne å stige. Man kan derfor ikke tolke en nedgang i dette antallet i dag som et tegn på at reproduksjonstallet i dag er mindre enn 1. Vi har også sammenliknet tidsutviklingen av $R(t)$ som vi har estimert med utviklingen av mobiliteten i Norge målt gjennom Apple Maps. Det er en slående likhet i formen. Dette gir en pekepinn om at disse mobilitetsdataene kan brukes som en prokxy for reproduksjonstallet, og således gi oss en mulighet for å bruke slike data til å detektere $R(t)$ i sann tid.

Oppdatering for 3. mai basert på siste tall: Median 0.86 og 95% konfidensintervall mellom 0.60 og 1.13. Oppdatert figur ligger i slutten av dokumentet.

Modell og metode

Vi anvender en standard SEIR modell:

$$\frac{dE}{dt} = \alpha R(t) SI/N - a E$$

$$\frac{dI}{dt} = a E - \gamma I$$

$$\frac{dS}{dt} = -\alpha R(t) SI/N.$$

der $a = 1/(5 \text{ dager})$, og måler de respektive varighetene av eksponert og smitteførende periode og N er Norges folketall. Videre er $S(t)$ antall individer mottagelige for smitte, $E(t)$ antall individer eksponert, $I(t)$ er antall smitteførende og $R(t)$ er det basale reproduksjons-tallet, som måler gjennomsnittlig antall sekundære infeksjoner som forårsakes av et gjennomsnittlig smitteførende individ i en populasjon der alle er mottagelige. Funksjonen $R(t)$ har følgende form

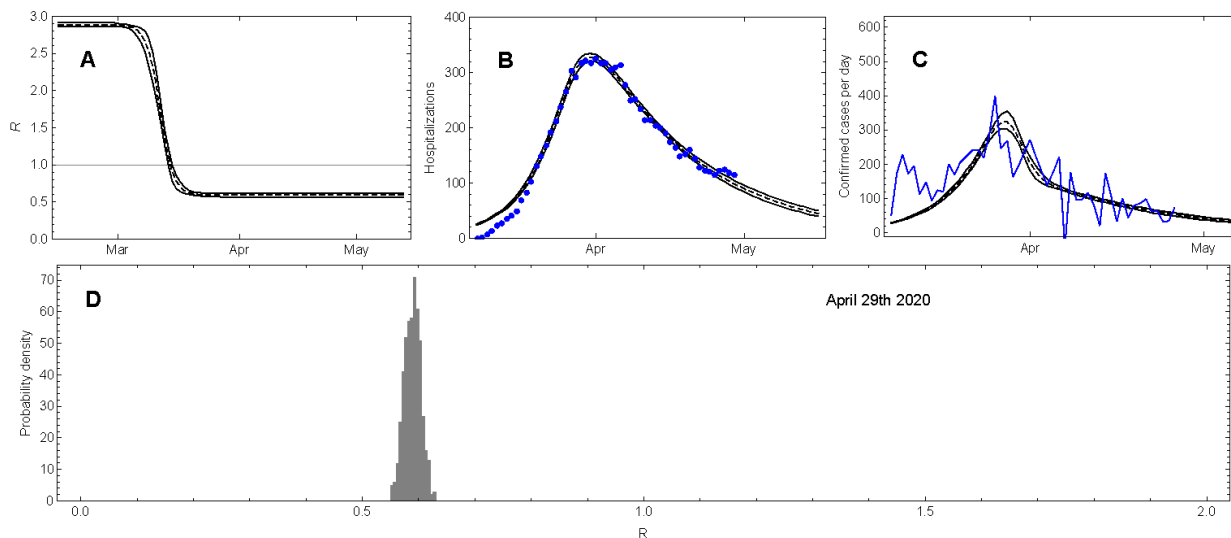
$$R(t) = \frac{R_0 + R_1}{2} - \frac{R_0 - R_1}{2} \tanh \tanh \left(\frac{t - t_0}{d} \right) + x (t - t_0 - 14 \text{ dager}) \Theta(t - t_0 - 14 \text{ dager}),$$

der $\Theta(t)$ er en enhets trappefunksjon. Startbetingelse er $E(0)$ trekkes tilfeldig mellom 1 og 500 personer den 15. februar. Parametere R_0 , R_1 , t_0 , d blir estimert ved miste kvadraters tilpasning til data for innleggelser med en 14 dagers forsinkelse og en skaleringsfaktor $\rho = 0.03$. De ti første data-punktene i tidsrekken for innleggelser er utelatt fra parameterestimeringen. Verdien R_0 er sensitiv for valg av startbetingelsen, men ikke for utviklingen etter 13. mars. Sammenlikning med tidsrekka for positive tester er gjort ved en nedskalering til 10% av $I(t)$ og regne en forsinkelse på fire dager fra man er smitteførende til testresultat foreligger.

Usikkerhet i estimatene er beregnet ved å legge til en normalfordelt støy til data for innleggelser og påfølgende re-estimering av parametere. Standardavviket for støyen er estimert fra differansen mellom en tilpasset (basert på et estimat som beskrevet over) kurve for innleggelser og de faktiske innleggelsestallene. Gjentatte estimater gir et ensemble, og kurvene som er plottet viser medianer og 95% konfidensintervaller. Sammenlikning med mobilitetsdata er gjort ved å skalere kurven $R(t)$ slik at median R_0 tilsvarer 100%.

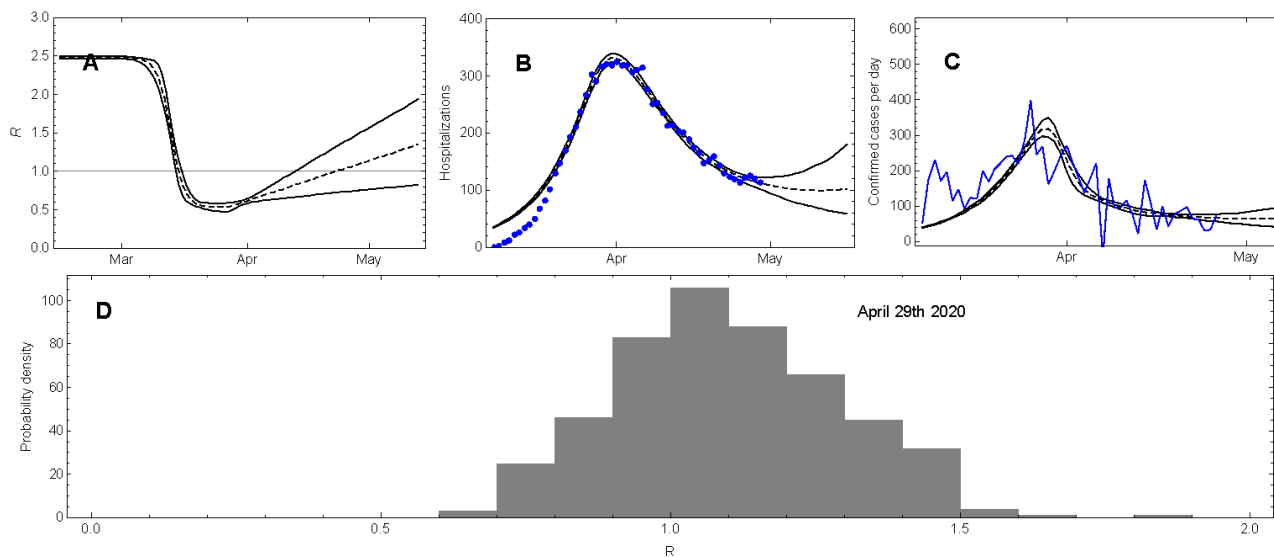
Resultater

Vi gjør to analyser: en der $x = 0$ er fiksert, og en annen der x kan variere fritt (over både positive og negative verdier). Resultatet av å fikserte $x = 0$ er vist i Figur 1.



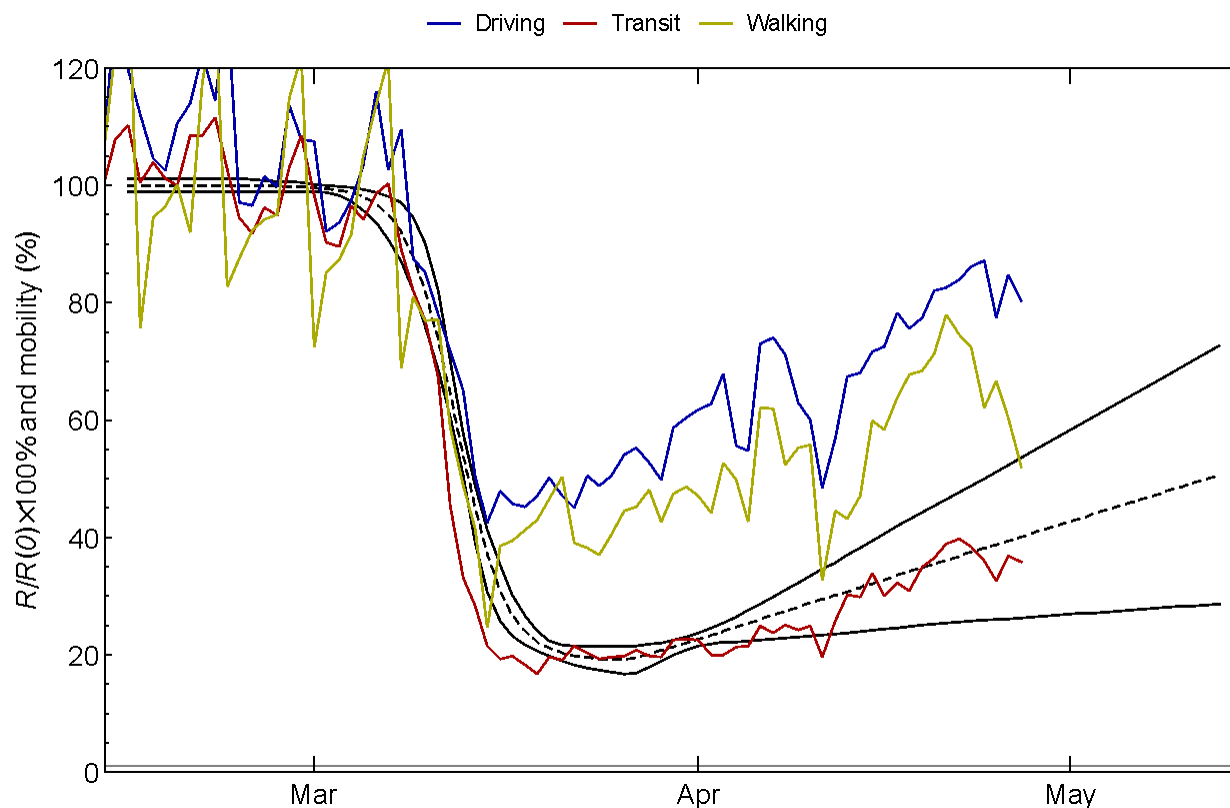
Figur 1. A viser estimert form på $R(t)$ for fiksert $x = 0$. Stiplet kurve er median verdi og det grå feltet representerer 95% konfidensintervall. B viser data for innlagte i sykehus (blå punkter) og predikert forløp med de estimerte $R(t)$. C angir antall daglige positive tester (blå kurve) og predikert forløp. D viser sannsynlighetstetthet for R den 29. april.

Resultater når x får variere fritt, slik at vi tillater en trend er gitt i Figur 2.



Figur 2. A viser estimert form på $R(t)$ når x får variere fritt og estimeres fra data for innleggelser. Ellers er alt som i Figur 1. Legg merke til i B at median prediksjon for sykehusinnleggelser avtar fram til midten av mai, selv om $R(t)$ vokser større enn 1 i denne perioden.

I figur 3 viser vi mobilitetsdata fra Apple Maps¹ og en reskalert versjon av Figur 2A for $R(t)$. Figuren for mobilitet i transitt sammenfaller forbløffende godt med estimert $R(t)$ med trend. Dette bør være en ansporing til å undersøke nærmere om denne typen data kan brukes som en prokxy for $R(t)$ i sann tid.

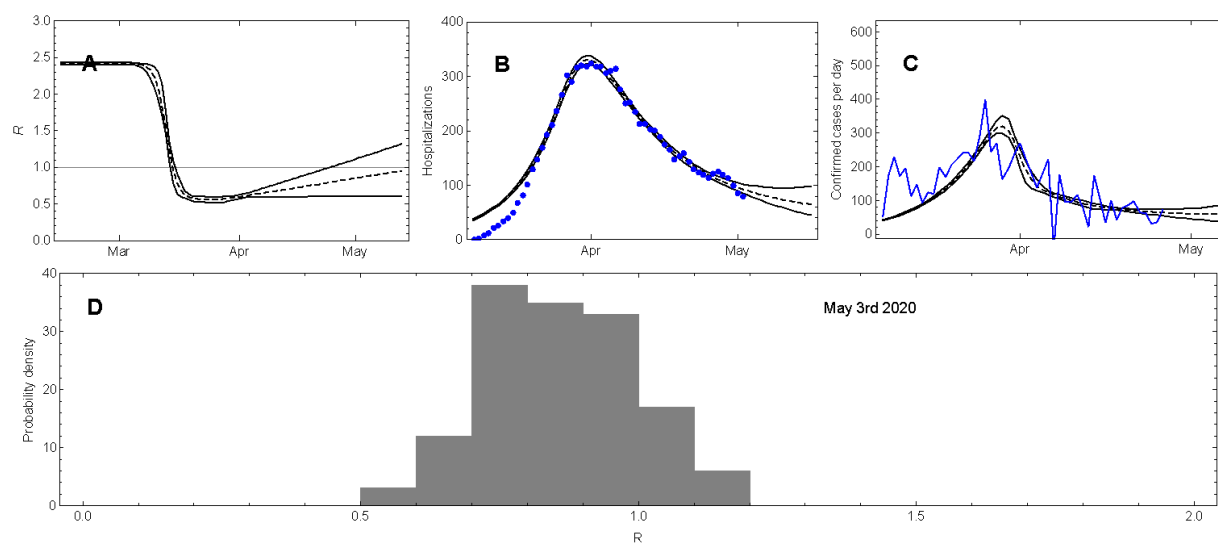


Figur 3. De fargete kurvene er mobilitetsdata fra Apple Maps (se fotnote), og den stiplede kurven og det grå feltet er en reskalert versjon av Figur 2A.

Konklusjoner

Våre estimater tyder på at FHIs daglige estimater av $R(t)$ er satt opp slik at de ikke kan fange opp ferske trender i reproduksjonstallet. Dette bør avklares snarlig, da viktigheten av oppdatert kunnskap om $R(t)$ er udiskutabel. Estimatenes sannsynliggjør at det i dag er en signifikant positiv trend, og at det allerede i dag er like sannsynlig som ikke at R er større enn 1. Analysen viser også at det er en reell mulighet for at mobilitetsdata fra for eksempel Apple Maps kan gi presis informasjon om reproduksjonstallet i sann tid.

¹ <https://www.apple.com/covid19/mobility>



Figur 4. Oppdatert versjon av Figur 2 (3. mai, 2020).