



Ekspertrapport af d. 23. oktober 2020

Incidens og fremskrivning af COVID-19 tilfælde



Indholdsfortegnelse

1. Sammenfatning	3
2. Formål og baggrund	5
3. Modelleringsstilgange og resultater	6
3.1 Estimering af incidensen af COVID-19 i en referencegruppe.....	8
3.2 Estimering af kontakttallet med korrektion for testintensitet.....	11
3.3 Fremskrivning af antal positive tests givet et antal tests	14
Bilag 1: Overblik over testdata.....	17
Bilag 2: Dannelse af referencegruppen.....	20
Bilag 3: Aldersfordelingen i befolkningen versus referencegruppen	21
Bilag 4: Estimering af kontakttallet justeret for testintensitet.....	23
Bilag 5: Sammenhængen mellem antal test og antal positive COVID-19 tilfælde	24
Bilag 6: Lokal poisson regression med korrektion for antal test	25
Bilag 7: Medlemmer af ekspertgruppen	28



1. Sammenfatning

Der er hen over sommeren løbende sket en væsentlig udbygning af testkapaciteten i Danmark. I løbet af september så man således en kraftig stigning i antallet af personer der blev testet og samtidig en stigning i andelen af positive test i alle regioner. I de seneste uger har testkapaciteten fortsat været høj, men antal daglige test er faldet betragteligt, til et aktuelt niveau på 30-35.000 daglige test.

Den løbende stigning samt efterfølgende variation i testintensitet vanskeliggør estimering af hhv. nysmittede per dag, kontakttallet og dermed den faktiske vækstrate i antallet af daglige tilfælde af COVID-19. Dette skyldes, at man grundlæggende vil finde flere tilfælde, jo flere personer, der testes (tilsvarende vil man finde færre tilfælde, når færre personer testes).

I indeværende tekniske rapport har ekspertgruppen for matematisk modellering udviklet en række metoder til at estimere antallet af nysmittede med COVID-19 per dag og kontakttallet, hvor der tages højde for variationer i testintensiteten. Yderligere, præsenteres regionale 14-dages fremskrivninger af antal daglige COVID-19 positive test (med 95% prædiktionsinterval), estimeret ved et konstant antal daglige test i perioden.

Uge 42.

Der er set et betragteligt fald i antal daglige test i løbet af uge 42, såvel i sundheds-, som i samfundssporet, hvilket formentligt relaterer sig til, at danskerne har holdt efterårsferie. Det bemærkes også, at der har været en betragtelig nedgang i sundhedspersonale, som er testet i uge 42, hvilket tolkes som en generel nedgang i aktiviteten på de danske hospitaler i efterårsferien. Denne ændrede test-adfærd i befolkningen, kan have betydning for, hvem der lader sig teste, og dermed også sandsynligheden for at teste positiv for COVID-19. Ændret test-adfærd i befolkningen influerer muligvis også på estimaterne i denne rapport. Man må derfor forvente justerede estimater i en efterfølgende opdateret rapport, såfremt test-adfærden igen normaliseres. Oplysninger om årsagen til at folk lader sig teste har ikke været tilgængelig for modelgruppen, og indgår derfor ikke i den aktuelle rapport.

Incidensen af COVID-19 i en referencegruppe der ikke er følsom for testintensitet

Incidensen i befolkningen og en referencegruppe, der ikke er følsom for ændringer i testintensitet viser, at der i starten af oktober og frem mod nu, er sket en reel stigning i incidensen, som et udtryk for stigning i smitten med COVID-19 i Danmark.

Korrektionsfaktor for antallet af test

Det estimeres, at det forventede antal bekræftede tilfælde af COVID-19 i perioden 1. juni til 19. oktober 2020 vokser med antallet af daglige test opløftet i potensen 0,7 (β).

Det er vigtigt at bemærke at dette estimat af β opsummerer alle de komplicerede forhold, der afgør hvilket udsnit af befolkning, der bliver testet. Det er altså plausibelt at estimatet af β vil ændre sig



i fremtiden, hvis teststrategien ændres. Dette kunne for eksempel være hvis antal prøver øges markant, eller hvis prøver i større eller mindre grad rettes mod specifikke grupper.

Dette betyder, at ved fordobling af antallet af test, uden at sygdomsudbredelsen har ændret sig, vil antal bekræftede tilfælde øges med 60%.

Som et eksempel ser man, at hvis der testes 50.000 personer på nuværende tidspunkt hvoraf 600 tester positivt og der fx testes 60.000 personer senere vil det forventes, at antallet, der tester positivt, vil være ca. 680 personer, hvis der ikke er sket en ændring i sygdomsudbredelsen.

Det aktuelle kontakttal korrigeret for testintensiteten

Kontakttallet baseret på nye tilfælde af COVID-19 estimeres til 1,24 (SI: [1,02;1,45]) pr. d. 20. oktober, når der korrigeres for testintensiteten.

Fremskrivning af incidens under hensyn til antal test

Fremskrivningen viser, at der forventes en stigende incidens i alle regioner, på nær Region Midtjylland i de kommende 14 dage.

Konklusion

Det vurderes ud fra rapporten at udbredelsen af COVID-19 aktuelt er på niveau med midten af april 2020 under hensyntagen til testintensiteten. Det aktuelle kontakttal justeret for testintensitet er 1,24 og der forventes en stigning i incidens i de kommende 14 dage i alle landets regioner, fraset Region Midtjylland. For Region Midtjylland rapporteres der om nye udbrud, og der forventes derfor også en stigende incidens i Region Midtjylland, ved en kommende opdatering af fremskrivningerne.

Statens Serum Institut (SSI) planlægger fremover, at inkludere elementer fra rapporten i den ugentlige risikovurdering fra SSI. Rapporten vil yderligere bidrage til en intern proces vedrørende opdatering af beregningen af kontakttallet for nye tilfælde, justeret for testintensitet.



2. Formål og baggrund

Statens Serum Institut indgår i det operationelle beredskab med smitsomme sygdomme og yder rådgivning og bistand til regeringen i forbindelse med den aktuelle pandemi. Som en del af denne opgave har Statens Serum Institut nedsat og leder en ekspertgruppe, der har til formål at udvikle matematiske modeller til at belyse udviklingen i COVID-19 i Danmark.

Ekspertgruppens modellering blev i foråret 2020 benyttet til at belyse udviklingen i belastningen af kritiske sygehusfunktioner som almene og intensive sengepladser, samt effekten af myndighedstiltag, herunder effekten ved at ophæve disse.

Modelleringen var baseret på en populationsmodel, der har fokus på den gennemsnitlige adfærd i befolkningen. Populationsmodellen er bedst egnet, når antallet af indlæggelser er højt, hvor udviklingen beskrives godt ved gennemsnittet. Populationsmodellen er dog ikke det bedste værktøj, når antallet af indlæggelser er relativt lavt, og kan dermed ikke beskrive de stokastiske hændelser, som aktuelt driver smittespredningen af COVID-19.

Der er hen over sommeren løbende sket en væsentlig udbygning af testkapaciteten i Danmark, og der har været en kraftig stigning i antallet af personer, der testes. I de seneste uger er der dog set et fald i antal daglige testede, såvel i sundheds- og samfundssporet, på trods af en fortsat høj testkapacitet.

Oplægget til denne rapport, er et ønske om at tilvejebringe estimater, der fremskriver antallet af positive test, som et redskab til at belyse udviklingen i epidemien, såvel regionalt som nationalt.

Fremskrivningen af antal positive test afhænger af antallet af test. Det er dog en generel udfordring, at antallet af test, og dermed andelen af de forventelige smittede, der testes, varierer geografisk og over tid. Dertil er det en udfordring, at testningen naturligvis ikke foretages tilfældigt, men primært foretages, hvis der er mistanke om smitte. Dette medfører, at det er vanskeligt at estimere, hvad incidensen i hele befolkningen faktisk er.

I indeværende rapport anskues udviklingen i COVID-19 tilfælde ud fra tre forskellige metoder, der hver især tager højde for udviklingen i testintensitet. De tre metoder skal ses som enkeltstående uafhængige bidrag.

3.1 Estimering af incidensen af COVID-19 i en referencegruppe, der ikke er påvirket af testintensitet.

3.2 Estimering af kontakttallet med korrektion for testintensitet.

3.3 Fremskrivning af antallet af positive test under hensyn til antal tests



En grundlæggende præmis er, at matematiske modeller altid vil være forsimplede repræsentationer af virkeligheden ud fra antagelser baseret på tilgængelige data og forskningsbaseret viden.

3. Modelleringsstilgange og resultater

Udviklingen i testkapacitet og antal daglige test

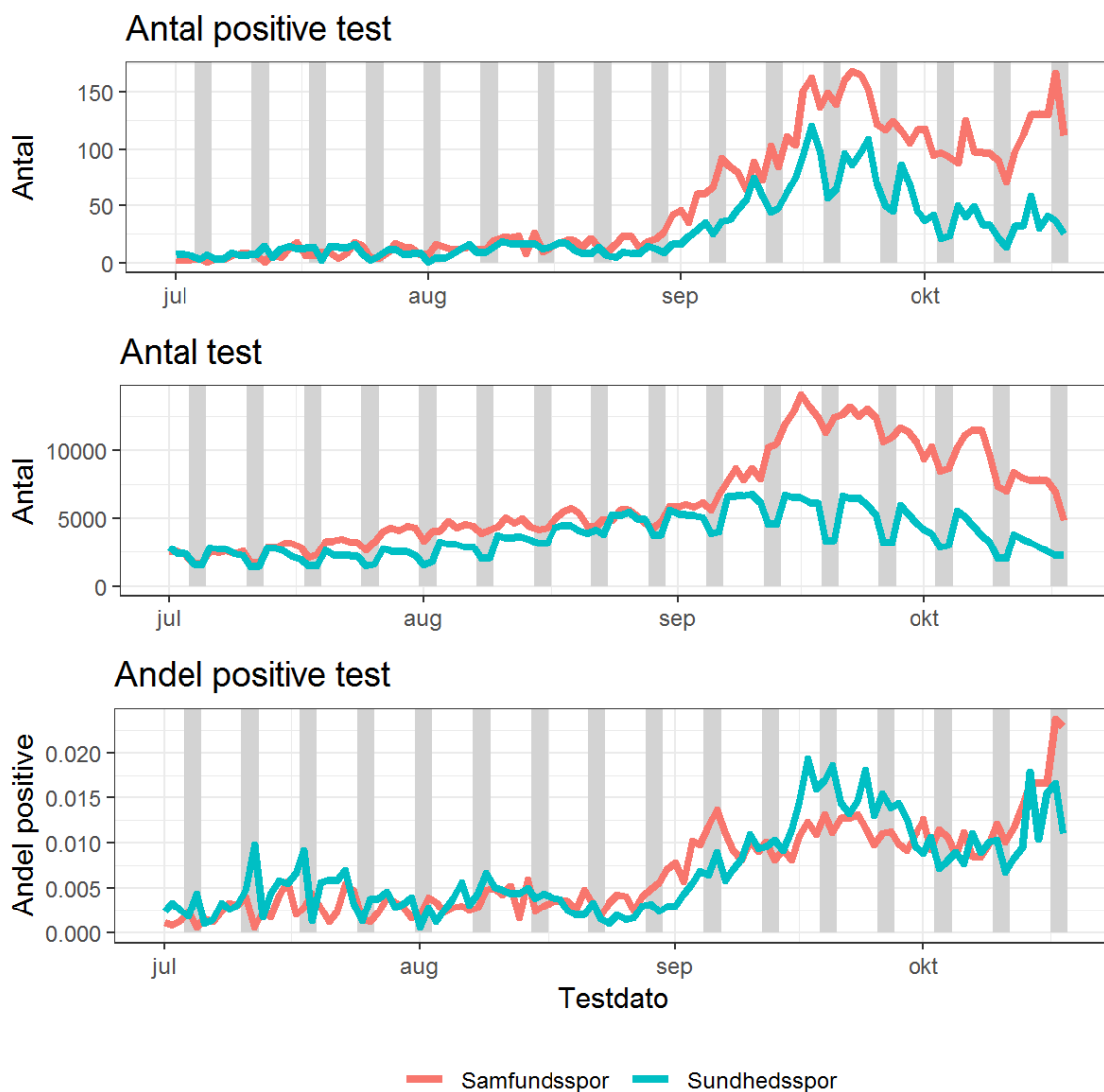
Figur 1 viser udviklingen i antal positive test, antal test og andel positive test per dag i Samfundssporet og Sundhedssporet (kun personer, som hverken er indlagte eller er sundhedspersonale). I figuren er kun medtaget test fra Region Hovedstaden, hvor der er set den største udvikling i antal daglige test.

Det fremgår i figur 1, at der er sket en kraftig stigning i antal testede borgere i samfundssporet i starten af september, hvor der samtidig ses en stigning i antallet af positive test, og en stagnation i andelen af positive test. Dette skal sammenholdes med, at andelen af personer testet positiv for COVID-19, stiger i sundhedssporet i den samme periode, hvor antallet af test ikke er steget i samme grad. Den øgede testintensitet ser ud til at toppe i slutningen af september, hvorefter der ses et fald i antal daglige testede i såvel sundheds- og samfundssporet i perioden frem mod uge 42 2020. På trods af denne opbremsning i antal daglige testede, er der sket en stigning i antal positive test i samfundssporet i uge 42, hvilket har bevirket en fordobling i andelen af positive over de sidste to uger. Det er uvist, hvordan afviklingen af efterårsferien har påvirket hvilke grupper i befolkningen, som har ønsket at blive testet.

Bilag 1 viser tilsvarende grafer for alle regioner.



Figur 1. Udviklingen i antal positive tests, antal test og andel positive per dag for personer, som hverken er indlagte eller sundhedspersonale, i Region Hovedstaden. (De andre regioner er vist i bilag 1)





3.1 Estimering af incidensen af COVID-19 i en referencegruppe

Faktaboks: Definition af incidens

- I det følgende afsnit skal incidens forstås, som andelen af positive test for COVID-19 ud af antal test på en given dag.
- Incidensen af COVID-19 afspejler antallet af daglige nye smittede i befolkningen.

Den observerede incidens af COVID-19 vil til en hver tid være påvirket af den aktuelle testintensitet samt årsagen til at personer lader sig teste. Størrelsen af denne påvirkning kan anskues ved at beregne incidensen for COVID-19 i en referencegruppe, som er et udsnit af befolkningen, der i mindre grad er påvirket af ændringer i testintensiteten samt årsagen til at man bliver testet. Incidensen i referencegruppen bidrager i denne rapport først og fremmest til at illustrere, hvordan testintensitet påvirker incidensen af daglige nye tilfælde. Forskellen i incidensen i referencegruppen i forhold til incidensen iblandt alle testede, illustrerer således både udsving i incidensen, der relaterer sig til ændringer i daglige antal testede, samt til årsagen til at der testes, fx ved forekomst af større udbrud, hvor test-positiv procenten stiger lokalt. En løbende overvågning af forskellen i incidensen, i referencegruppen og befolkningen, vil således kunne bidrage med signaler om, hvorvidt der er en generel stigende incidens i befolkningen (hvilket vil afspejles i en stigende incidens i referencegruppen), eller, om der er tale om en stigning der relaterer sig til øget testintensitet, eller udbrud.

Siden foråret 2020, er alle indlagte patienter, og udvalgte ambulante patienter, rutinemæssigt blevet tilbudt test for COVID-19¹. Referencegruppen i rapporten dannes således af patienter som er testet for COVID-19 i forbindelse med et indlæggelsesforløb. Referencegruppen repræsenterer formentligt et særligt udsnit af befolkningen, som er væsentligt anderledes på en lang række parametre. Det centrale her er dog, at deres sandsynlighed for at blive testet for COVID-19, formodes at være relativt upåvirket af ændringer i testkapacitet, efterspørgsel på test, samt prævalensen af COVID-19 i øvrigt.

Dannelse af referencegruppen

Referencegruppen består af patienter med et indlæggelsesforløb der ikke er relateret til en respiratorisk lidelse, eller graviditet og fødsel. Der er inkluderet indlæggelsesforløb registreret i Landspatientregistret (LPR3), hvor der samtidig er tilknyttet en test for COVID-19 i den Danske Mikrobiologi database (MiBa) i perioden 1. juni til 16. oktober 2020. En detaljeret beskrivelse af referencegruppen er vedlagt i bilag 2.

Alderseffekt

Det er relevant at undersøge, hvorvidt der er en alderseffekt forbundet med testintensitet, som dermed påvirker incidensen af COVID-19. For at undersøge dette, er der lavet et plot, hvor

¹ <https://www.sst.dk/-/media/Udgivelser/2020/Corona/Retningslinjer/Retningslinjer-for-haandtering-af-COVID-19.ashx>

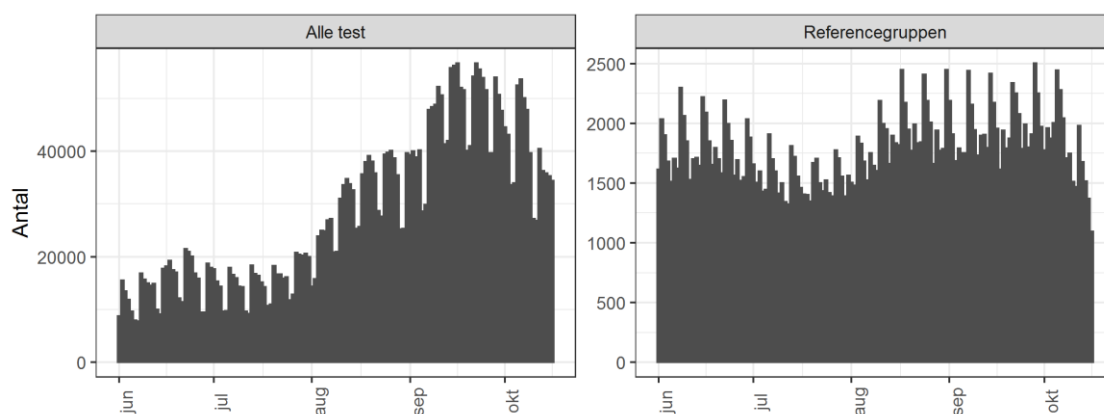


referencegruppen er afgrænset til perioden 1. september til 16. oktober, hvor der kun medtages personer i aldersgruppen 16 til 80 år. Incidensen i den anførte periode beregnes ved, at fitte generaliserede additive modeller til antallet af negative og positive test til en udglattet funktion (thin plate regression spline) af alderen. Baseret på dette er der estimeret en aldersjusteret incidens for COVID-19 i hele befolkningen, samt i referencegruppen. Resultaterne er præsenteret i bilag 3.

Resultater

Figur 2 viser fordelingen af antal daglige test i hele Danmark, samt personer testet i referencegruppen i perioden 1. juni til 16. oktober 2020 (bemærk, at der er anvendt forskellige skalaer i de to figurer). Der ses i hele perioden en stigning i antal daglige test i Danmark, indtil midten af september måned, hvor antal daglige test falder. I samme periode er der i referencegruppen kun set en lille stigning i antal daglige test, hvilket understreger at referencegruppen ikke er følsom for ændringer i testintensitet på samme vis, som den øvrige befolkning. Der bemærkes dog en tilsvarende nedgang i antal daglige test i referencegruppen i den seneste uge, hvilket tolkes som en nedgang i aktiviteten på landets hospitaler i forbindelse med efterårsferien i uge 42. Der bemærkes desuden en tydelig weekendeffekt i begge grafer, dvs. at der generelt testes færre personer i weekender.

Figur 2. Antal personer testet for COVID-19 i Danmark, samt personer testet i en referencegruppe der er mindre følsom for testintensitet, i perioden 1. juni til 16. oktober 2020.



I figur 3 beskriver den røde kurve, den simple incidens af COVID-19 i befolkningen, uden hensyntagen til den stigende testintensitet i august og september. Den blå kurve viser incidensen af COVID-19 for personer, der er testet i referencegruppen, i perioden 1. juni til 16. oktober 2020.

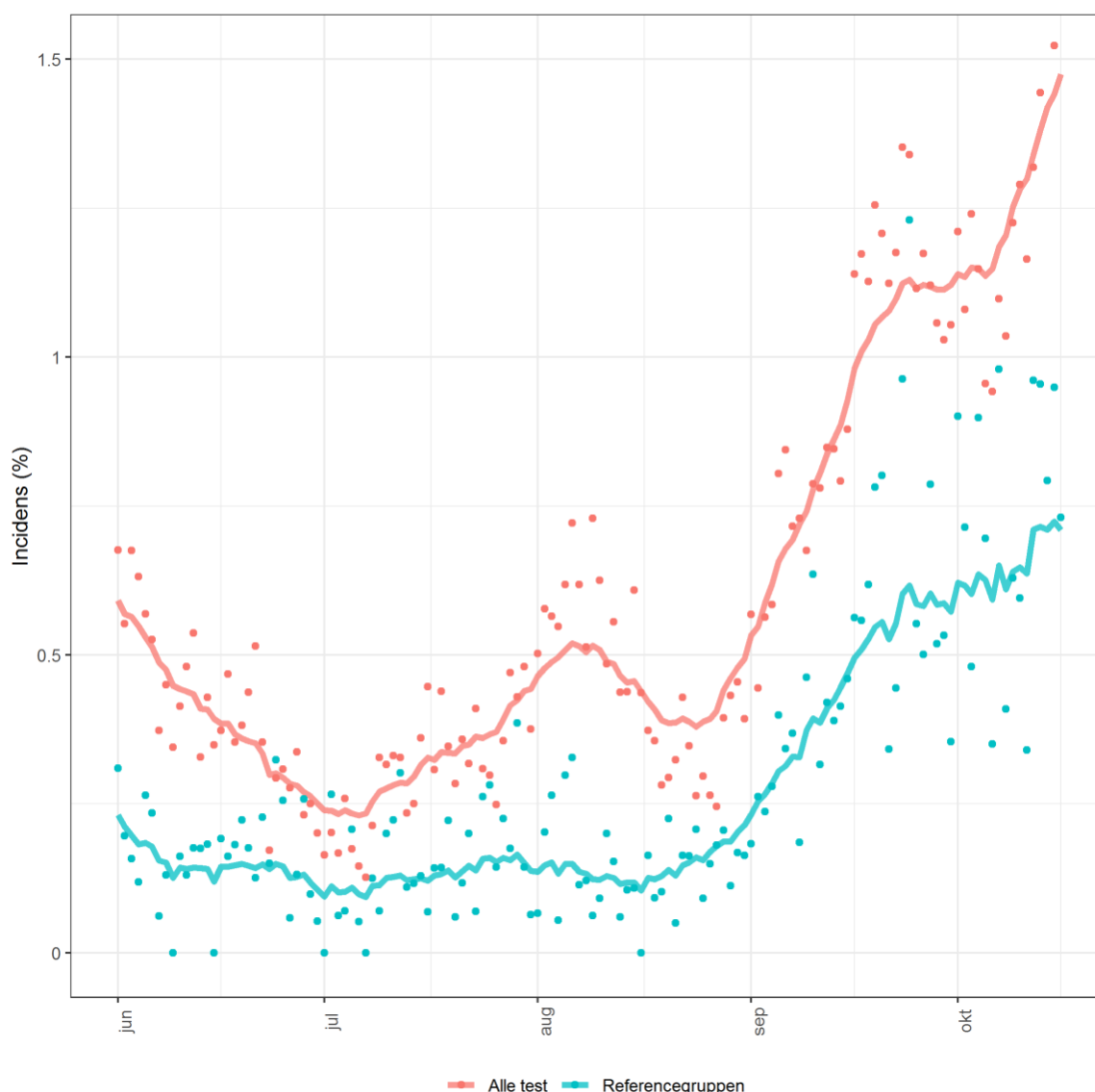
Begge kurver viser et fald i incidensen i løbet af juni måned. Herefter ses incidensen i befolkningen (rød kurve) at stige betragteligt i starten af juli til det højeste niveau tidligt i august, hvorefter der igen ses et fald i slutningen af august. I modsætning hertil, ses incidensen i referencegruppen at være mere stabil i samme periode. Det er vigtigt at bemærke, at sådanne stigninger i incidensen i befolkningen, som ikke umiddelbart kan genfindes i referencegruppen, kan skyldes en øget



testintensitet i forbindelse med lokale udbrud. Som et eksempel, er den ovenfor beskrevne pukkel i incidensen i befolkningen i august måned, tidsmæssigt korreleret til udbruddet som var i Århus på daværende tidspunkt.

Stigninger i incidensen som genfindes i begge grafer, er mere sandsynligt et udtryk for en reel stigning i forekomsten af nye tilfælde af COVID-19. I starten af oktober og frem mod nu, ses således en kraftig stigning i incidensen i hele befolkningen, såvel som i referencegruppen, som et udtryk for stigning i smitten med COVID-19.

Figur 3. Incidens af COVID-19 i befolkningen baseret på alle test (rød), sammenlignet med incidens i en referencegruppe, der er mindre følsom for testintensitet (blå), i perioden 1. juni til 16. oktober 2020 (Prikkerne er de rå incidensestimater. Linjerne er et eksponentielt udglattet gennemsnit.)



Særlige forbehold

Arbejdet med at klassificere hospitalsforløb fra Landpatientregisteret er igangværende og kræver en række subjektive fortolkninger af koder anvendt fra registret.



Der forestår således fortsat et arbejde med at se nærmere på de forskellige typer af hospitalsforløb inden for referencegruppen. Ligeledes forventes referencegruppen justeret, så den afspejler den generelle befolkningen (især unge under 35 år) bedre.

3.2 Estimering af kontakttallet med korrektion for testintensitet

Kontakttallet fortæller hvor mange kontakter en smittet i gennemsnit vil komme til at smitte, og dermed om potentialet for den videre smittespredning.

Kontakttal kan estimeres på forskellige måder. Statens Serum Institut offentliggør beregnede kontakttal baseret på hhv. antal nyindlagte patienter og bekræftede COVID-19-tilfælde med hensyntagen til om de var smittet i udlandet eller i Danmark ².

Når der er en stigning i antal bekræftede tilfælde, kan det både skyldes, at sygdomsudbredelsen stiger, at testaktiviteten stiger eller at strategien ændres for hvem, der testes. Det er a priori ikke klart, hvor stor en del af stigningen, der kan tilskrives hver af disse faktorer. For at undersøge dette analyseres en model, der estimerer kontakttallet ud fra antal bekræftede tilfælde per dag, men korrigeret for det bagvedliggende antal tests.

Metode

En detaljeret gennemgang af metoden i dette afsnit er vedlagt i bilag 4.

Det antages, at det forventede antal positive test på en given dag t er

$$I_t T_t^\beta$$

hvor I_t er et indeks, der beskriver sygdomsudbredelsen i samfundet på den pågældende dato og som estimeres T_t er antal tests på den pågældende dato, og β beskriver, hvordan det forventede antal bekræftede tilfælde afhænger af antallet af tests: Størrelsen af β kendes umiddelbart ikke. Med $\beta=1$ vil man få dobbelt så mange tilfælde, hvis man fordobler antallet af tests, hvilket er hvad man ville forvente hvis man tog tilfældige stikprøver. Med $\beta=0$ vil man ikke forvente et øget antal tilfælde, når man øger antallet af tests; svarende til at de ekstra tests bliver brugt på personer, der ikke har sygdommen. I virkeligheden ligger sandheden et sted mellem disse to yderpunkter, og derfor estimeres β ved hjælp af maksimum likelihood.

Når I_t er estimeret, kan det estimeres hvor mange der ville teste positivt, hvis antallet af daglige tests havde været konstant og lig 50.000. Dette estimat beskriver sygdommens udbredelse.

Estimatet af I_t fluktuerer fra dag til dag, og fra disse fluktuationer beregnes en daglig specifik vækstrate r i I_t . Vækstraten omsættes til et dagligt kontakttal R ud fra formlen $R = 1 + r t$ hvor t er generationstiden som sættes til 4,7 dage.

² <https://files.ssi.dk/COVID19-epi-trendogfokus-11062020>



Specifikke forbehold

Modellen anvender ikke data om dato for centrale tiltag eller deres effekt, ligesom den ikke tager hensyn til tiden fra smitteoverførsel til en positiv test. Derfor vil modellen give et noget forsinket tidsbillede af smittespredningen. Der inddrages kun personer i analysen, der ikke tidligere er testet positiv.

Resultater

Med denne analyse estimeres parameteren β til ca. 0,7 med et konfidensinterval fra 0,6 til 0,8. Dette betyder, at hvis antallet af tests fordobles vil antallet af bekræftede tilfælde øges med 60 % uden at sygdomsudbredelsen ændrer sig. Dette kan bl.a. forklares ved, at når antallet af test øges, vil der i stigende grad testes personer i befolkningen, der har en lavere risiko for at være smittede. Det bemærkes at dette estimat af β opsummerer alle de komplicerede forhold, der afgør hvilket udsnit af befolkning, der bliver testet. Det er altså plausibelt at estimatet af β vil ændre sig i fremtiden, hvis teststrategien ændres. Dette kunne for eksempel være hvis antal prøver øges markant, eller hvis prøver i større eller mindre grad rettes mod specifikke grupper.

Analysen estimerer som eksempel, at hvis der testes 50.000 på nuværende tidspunkt med 500 positive tilfælde og der fx testes 60.000 senere vil det forventes, at antallet af nye positive tilfælde vil være ca. 570, givet at der ikke er sket en ændring i sygdomsudbredelsen. Generelt er antallet af test lavere i weekenderne og dermed ses også ofte et dyk i antal nye daglige tilfælde. Tabel 1 illustrerer effekten af antal test, på antal nye positive tilfælde med den nuværende udbredelse af COVID-19 Figur B4 i bilag 5 viser den samme udvikling relativt til 50.000 daglige test.

Tabel 1. Illustration af betydningen af antal test på nuværende tidspunkt. Den grå række viser situationen, som den nogenlunde er på nuværende tidspunkt.

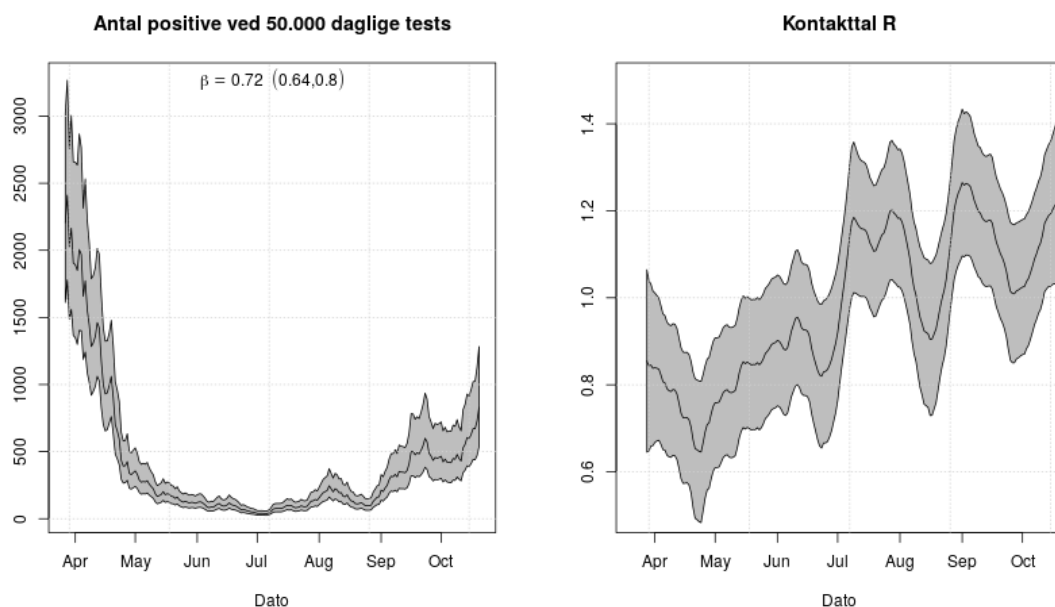
Antal tests	Antal forventede positive tilfælde under forudsætning af samme faktiske sygdomsudbredelse på nuværende tidspunkt
20.000	260
30.000	350
40.000	430
50.000	500
60.000	570
70.000	630
80.000	690
90.000	750
100.000	810

Hvis der regnes med 50.000 daglige test viser modellen, at udbredelsen af COVID-19 på nuværende tidspunkt ca. er på samme niveau som omkring ultimo april korrigeret for testintensitet. Dette er vist i figur 4.

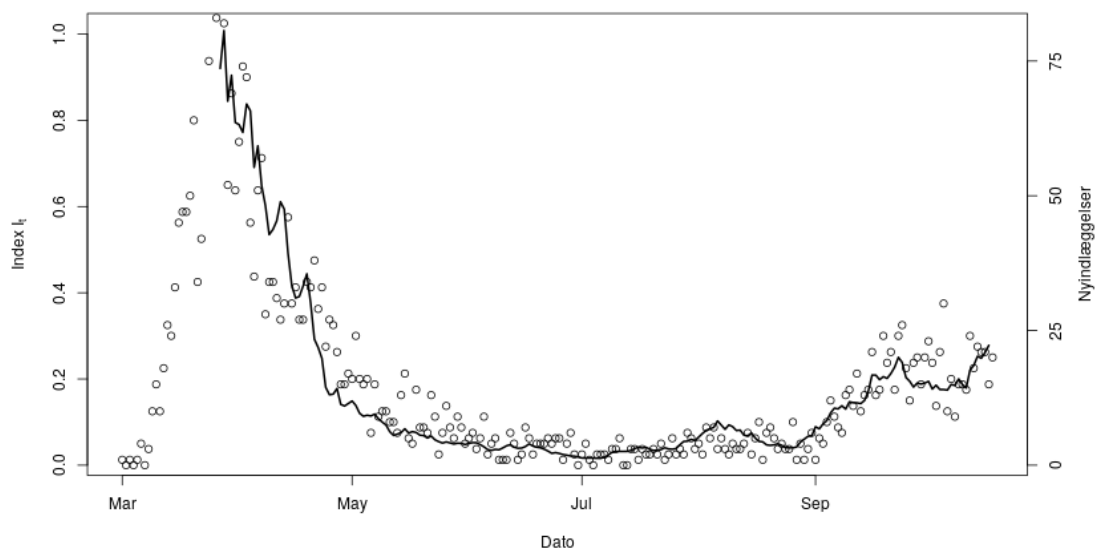
Kontaktallet estimeres således på baggrund af ovennævnte til 1,24 (SI: [1,02;1,45]) pr. d. 20. oktober, når der korrigeres for testintensiteten, figur 4. I denne estimerede kurve for kontaktallet er brugt en tidsskala på 7 dage til udglatning; se bilag 4 for yderligere detaljer.



Figur 4: Antal positive test ved 50.000 daglige test og det estimerede kontakttal baseret på bekræftede tilfælde af COVID-19 korrigeret for testintensitet. Opgjort d. 20. oktober 2020.



Figur 5. Det estimerede indeks I_t for positive prøver og de daglige nyindlæggelser. Opgjort d. 19. oktober 2020.



Figur 5 viser indekset I_t sammen med de daglige nyindlæggelser. Det bemærkes at de to tidsrækker udviser samme mønstre, og dermed samme indikation af sygdommens prævalens, hvilket styrker modellens kredibilitet.



Til at vurdere risikosituationen er nyindlæggelser det mest robuste mål, da det ikke er påvirket af testintensitet, men nyindlæggelserne giver et forsinket billede på smitteudbredelsen, da der går knap 14 dage i gennemsnit fra smitteoverførsel til indlæggelse, hvilket sinker muligheden for at agere med relevante tiltag. Derudover er antallet af nyindlæggelser mindre, hvilket øger den statistiske usikkerhed.

Det er derfor en nyttig observation at antallet af positive prøver, korrigeret for antal af udførte prøver på den måde som her er beskrevet, har samme forløb som nyindlæggelserne. Da tidsforsinkelsen fra smitteoverførsel til positiv prøve er mindre (omend det er ikke kendt præcis hvor meget mindre) og da antallet af prøver er større, kan det korrigerede antal af positive prøver give en tidligere og mere nøjagtig indikation af smitteudbredelsen, til brug for korttidsprognoser.

3.3 Fremskrivning af antal positive tests givet et antal tests

Antallet af positive test samt andelen af positive test afhænger som ovenfor beskrevet bl.a. af antallet af test. Fremskrivning af antal positive test vil dermed også være afhængig af antallet af test. Nedenfor præsenteres regionale 14-dages fremskrivninger af antal daglige COVID-19 positive test (med 95% prædiktionsinterval), estimeret ved et konstant antal daglige test i perioden.

Metode

En detaljeret metodebeskrivelse for dette afsnit, er vedlagt i bilag 6.

I denne model betegnes T_t^β det effektive antal test. I modellen benyttes det samme estimat af parameteren $\beta = 0.7$ som i afsnit 3.2. Smitteraten følger befolkningens adfærd, der bl.a. afhænger af befolkningens opfattelse af risikoen ved at bevæge sig rundt i samfundet. Befolkningens adfærd, og dermed smitteraten, ændrer sig både lokalt og over tid. For at håndtere dette er der i modellen anvendt rullende Poisson regressioner baseret på vinduer på 11 dage. Tiden er brugt som forklarende variabel og den naturlige logaritme af det effektive antal test (T_t^β) er brugt som offset.

For hver region er medianen af antal test siden d. 1. september (afrundet til tusinder) brugt som reference for antal test (Angivet i titlen for hver graf i figur 6). For hver region er der vist et estimat med 95% prædiktionsinterval af det forventede antal positive test på den sidste dag i det benyttede tidsvindue (rød kurve). Dertil er der lavet en 14 dages fremskrivning med konstant vækstrate (blå kurve).

Resultater

I figur 6 beskrives det daglige antal positive test per region samt estimering og fremskrivning af antal positive ved konstant antal daglige test i hele perioden.

Fremskrivningen viser, at der forventes en stigende incidens i alle regioner på nær Region Midtjylland.

Ved at holde antallet af daglige test konstant er det muligt at sammenligne kurverne over tid. Det betyder dog også, at det ikke er muligt at sammenligne direkte med det observerede antal positive test, særligt i perioder, hvor der er større forskel på testreferencen og det faktiske antal test. Dette ses eksempelvis i Region Midtjylland i forbindelse med udbruddet i og omkring Århus i starten af



august, hvor der blev testet relativt få og der var dermed større forskel på testreferencen og det faktiske antal test. Det medfører, at estimeringen (den røde kurve på figur 6) stiger mere end det observerede antal test. Testreferencen varierer på tværs af regionerne og fremgår af figur 6.

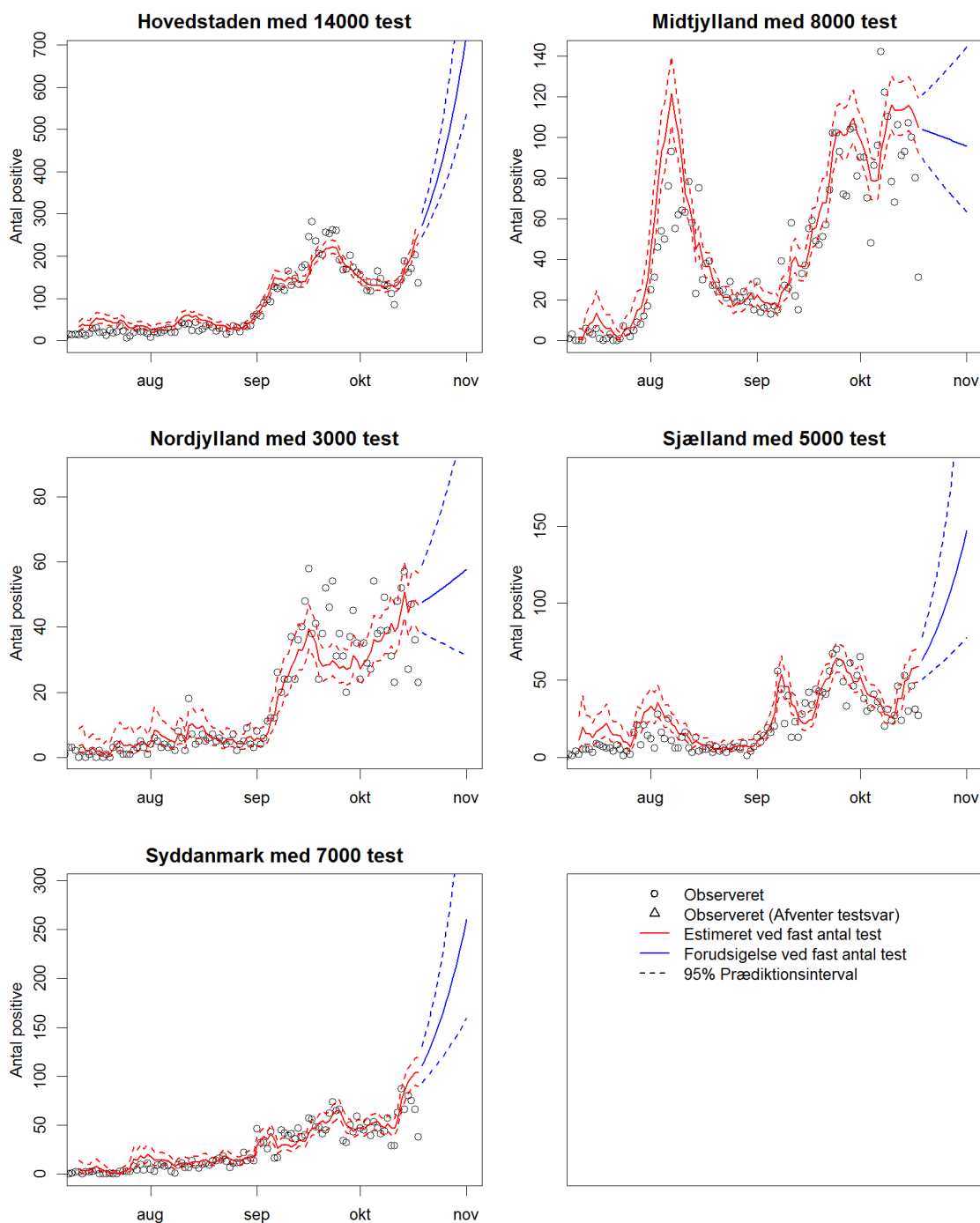
Det bemærkes, at estimererne af incidens faldt i alle regioner i perioden umiddelbart efter, at man indførte en række tiltag (iværksat d. 18. september) for at bremse smittespredningen af COVID-19. Dette er afløst af en stigning i incidens i samtlige regioner, på nær Region Midtjylland.

Særlige forbehold

Der går ca. to uger fra en adfærdsændring til modellen tilpasser sig. Dette skyldes, at der er en forsinkelse fra individer bliver smittet til de bliver testet og dertil yderligere ca. et døgn før der foreligger et prøvesvar. Dette betyder, at der går omkring en uge fra ændring af adfærd til der ses et fald i antal positive og yderligere nogle dage før vækstraten er estimeret til den nye adfærd. Der antages også at der ikke indføres yderligere tiltag i denne periode. Vækstraten i den fremskrevne periode antages derfor at være konstant.



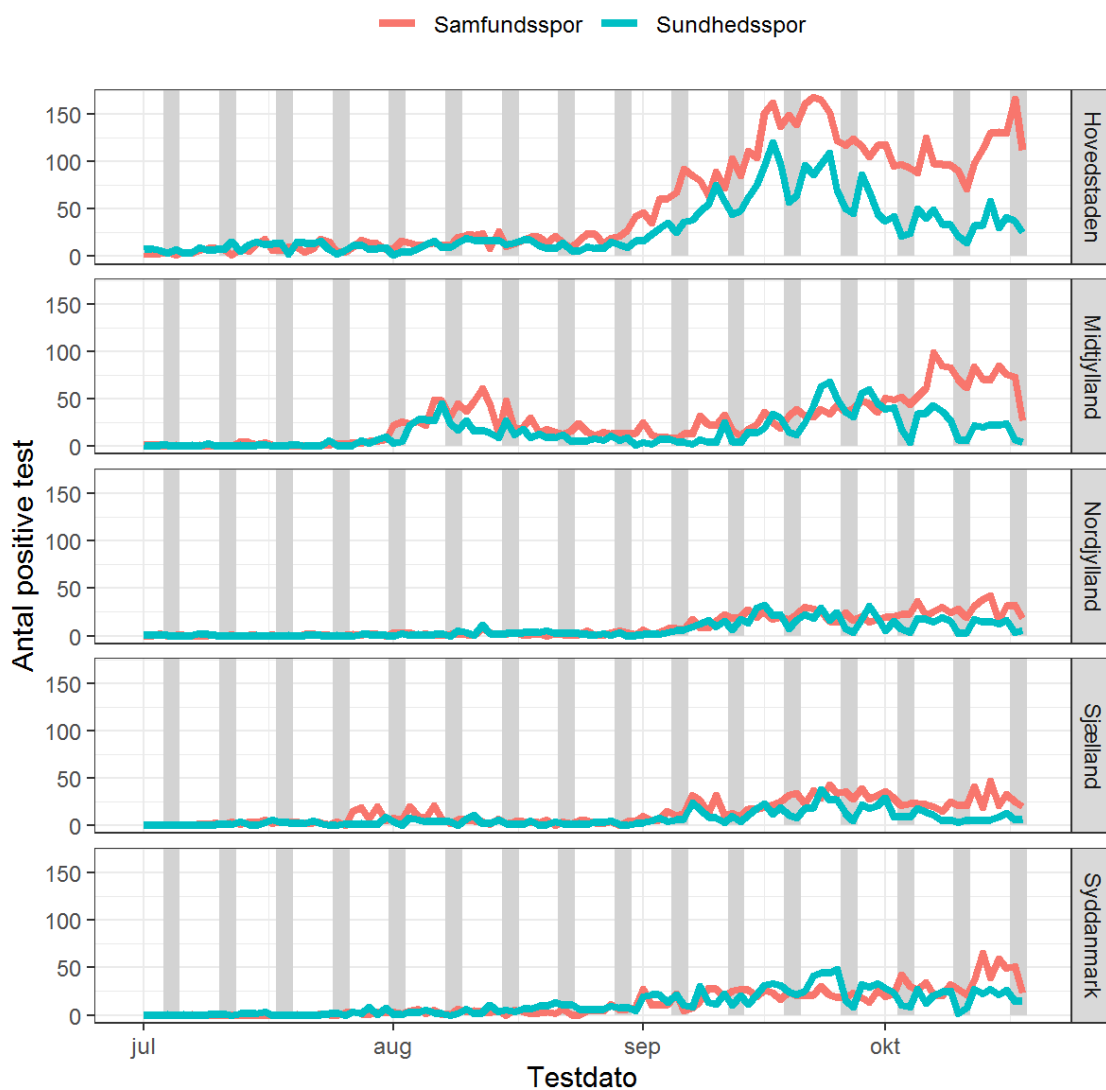
Figur 6. Daglige antal positive test per region, samt estimeret fremskrivning af antal COVID-19 positive (med 95% prædiktionsinterval) ved konstant antal daglige test i hele perioden. Opgjort d. 21. oktober 2020.



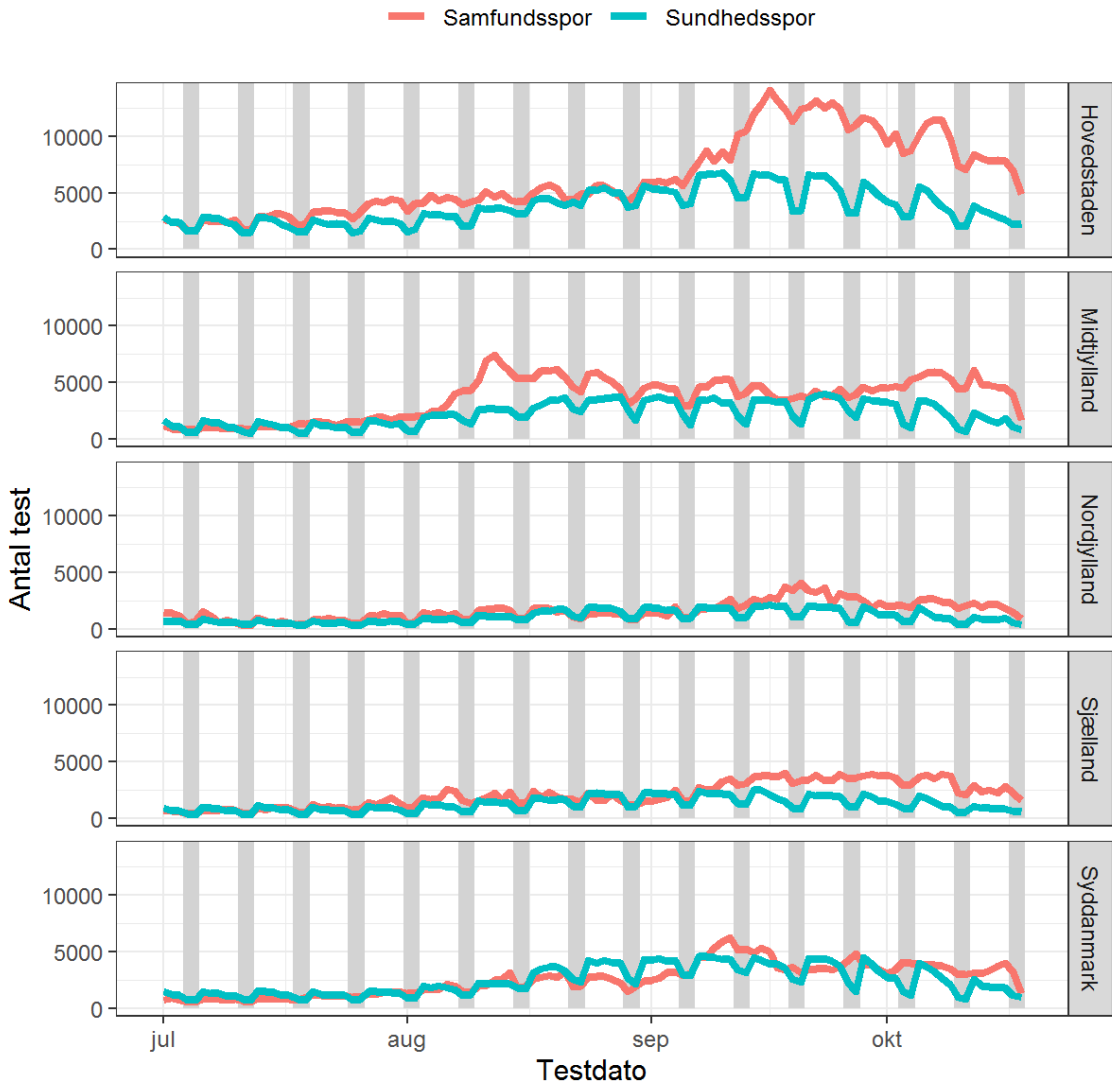


Bilag 1. Overblik over test data

Figur B1. Regional udvikling i antal positive tests per dag for personer, som hverken er indlagte eller sundhedspersonale. Udviklingen er opdelt i de to test spor.

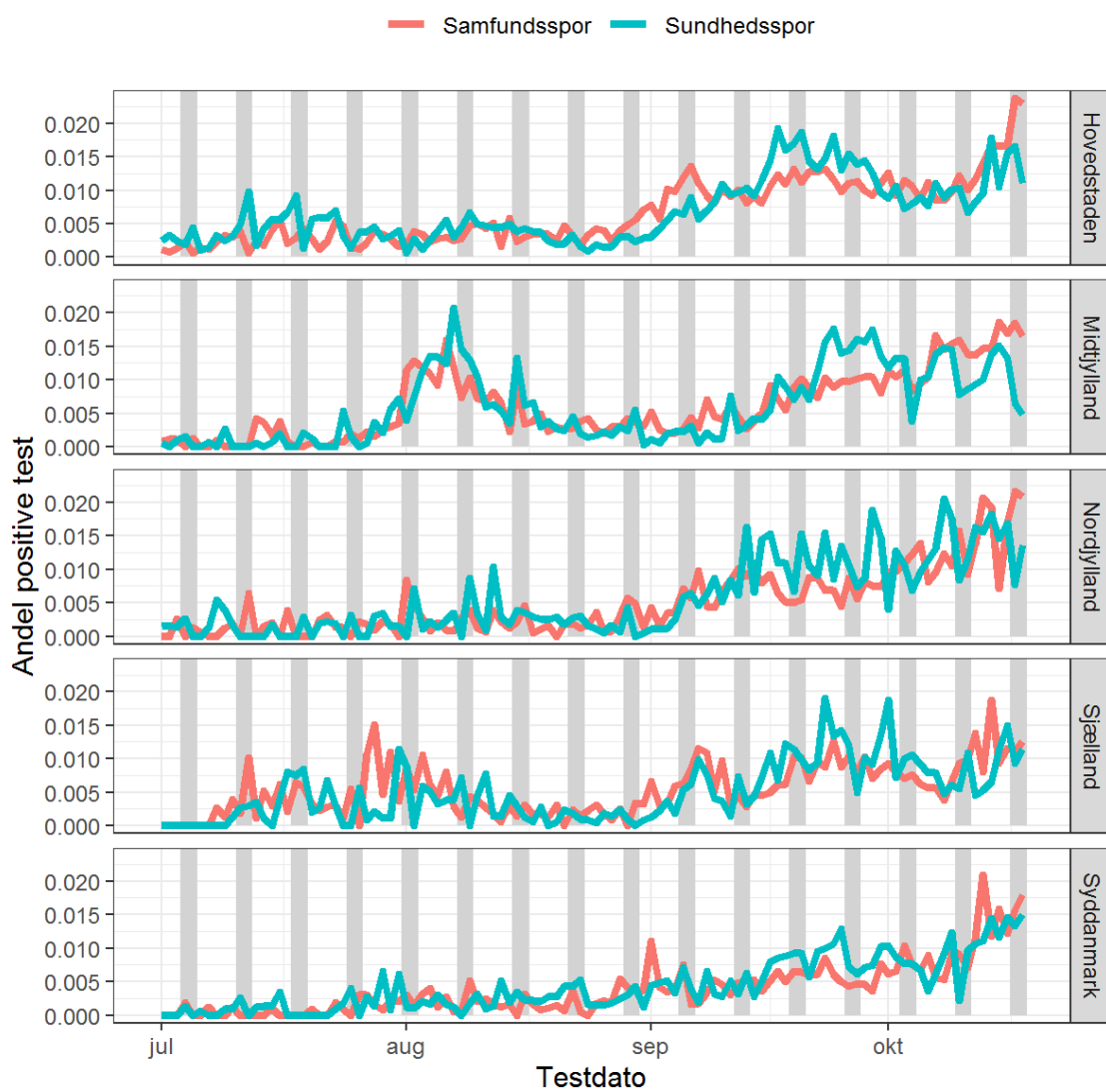


Figur B2. Regional udvikling i antal test per dag for personer, som hverken er indlagte eller sundhedspersonale. Udviklingen er opdelt i de to test spor.





Figur B3. Regional udvikling i andel positive per dag for personer, som hverken er indlagte eller sundhedspersonale. Udviklingen er opdelt i de to test spor.





Bilag 2. Dannelse af referencegruppen

Et indlæggelsesforløb dannes ved at sammenflette overlappende fysiske hospitalskontakter, registreret i LPR3 (inklusive åbne kontakter) i perioden efter 1. maj 2020. Kun indlæggelsesforløb med en varighed på over 8 timer er inkluderet. Referencegruppen afgrænses yderligere således, at hospitalsforløb relateret til respiratorisk sygdom (indlæggelser der ikke er planlagte, som er registreret som en respiratorisk lidelse i aktionsdiagnosekoder, primær/sekundær diagnose koder eller primær/sekundær procedurekoder) ekskluderes, for at undgå at inkludere patienter der er indlagt pga. COVID-19. Indlæggelsesforløb relateret til graviditet og fødsel er ligeledes ekskludert, idet der gælder særlige forholdsregler for gravide i forhold til beskyttelse, og de formodes derfor at have en lavere risiko for at blive smittet med COVID-19. De identificerede indlæggelsesforløb kobles til test for SARS-Cov-2 registreret i den Danske Mikrobiologi database (MiBa), ud fra kriteriet, at en test skal være foretaget 2 dage før til 5 dage efter indlæggelsesforløbet. Såfremt en patient har flere test per dag, vil en positiv test være konklusiv. Såfremt der er flere konklusive test, medregnes den test, som er tættest på indlæggelsesforløbets start, hvor test foretaget før indlæggelsesforløbets start, prioriteres højest.

For patienter hvor flere indlæggelsesforløb matcher et testresultat, er kun medtaget det indlæggelsesforløb der ligger tidsmæssigt tættest på datoen for testresultatet.

Datoen for det koblede testresultat anvendes i beregningen af incidens for referencegruppen.

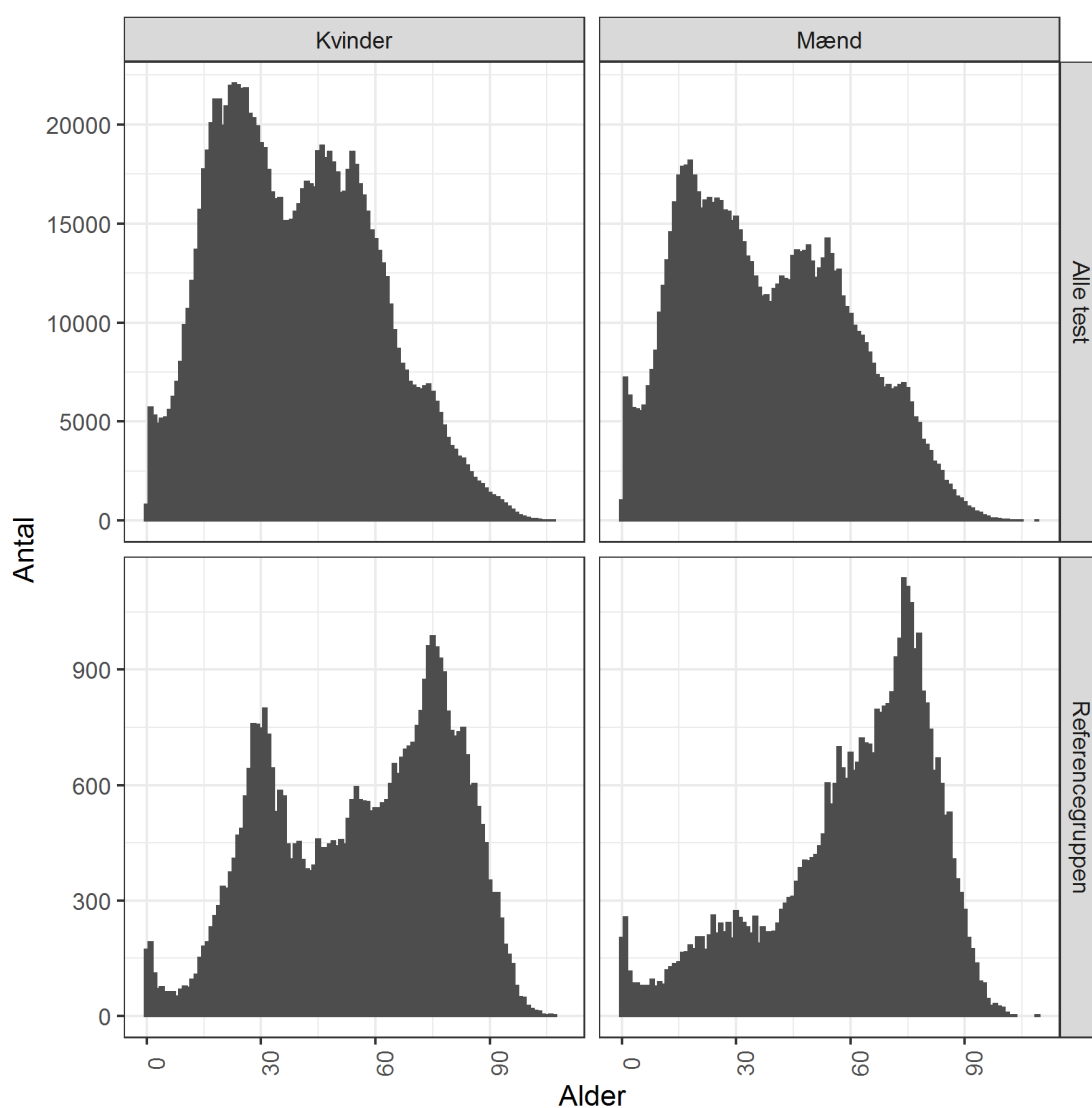
Ud fra de ovenstående kriterier, er referencegruppen afgrænset til 406.163 indlæggelsesforløb. Heraf var der 11.1% (n = 45.219) af patienterne, som aldrig er testet i forbindelse med deres indlæggelsesforløb og 12.7% (n = 51.621) der er testet for COVID-19 uden for tidsperioden på 2 dage før- til 5 dage efter startdatoen for et indlæggelsesforløb. Ydermere var 0.2% (675) af patienterne testet inden for den relevante periode, men hvor testresultatet var inkonklusivt. I alt 5.289 indlæggelsesforløb udgik, da det matchende testresultat allerede var anvendt i et andet indlæggelsesforløb der lå tættere på testdatoen, og 55.070 udgik da testdatoen var uden for perioden 1. juni til 16. oktober. Referencegruppen består således af i alt 248.289 indlæggelsesforløb, hvoraf 736 forløb er tilknyttet en positiv test for COVID-19, og 247.553 er tilknyttet en negativ test for COVID-19.



Bilag 3. Aldersfordelinger i befolkningen versus referencegruppen

I figur 4 beskrives antal test fordelt på alder og køn i perioden 1. september til 16. oktober. Fordelingen af test på alder er relativt ensartet mellem kvinder og mænd i hele befolkningen, dog er der generelt færre mænd der testes. For begge køn ses det, at unge mennesker under 20 år, testes mindre end de ældre generationer. Aldersfordelingen blandt de testede i referencegruppen er præget af, at det primært er ældre mennesker, som indlægges på hospital. Desuden ses kvinder omkring de 30 år at være overrepræsenterede i referencegruppen, hvilket der ikke er nogen egentlig forklaring på. Ved omregning af incidens-estimer til populationsniveau, bør der tages højde for disse forskelle i testintensivitet på tværs af køn og alder.

Figur 4. Antal test for COVID-19 fordelt på alder og køn iblandt testede personer i befolkningen samt i personer testet i referencegruppen i perioden 15. August til 16. oktober 2020.

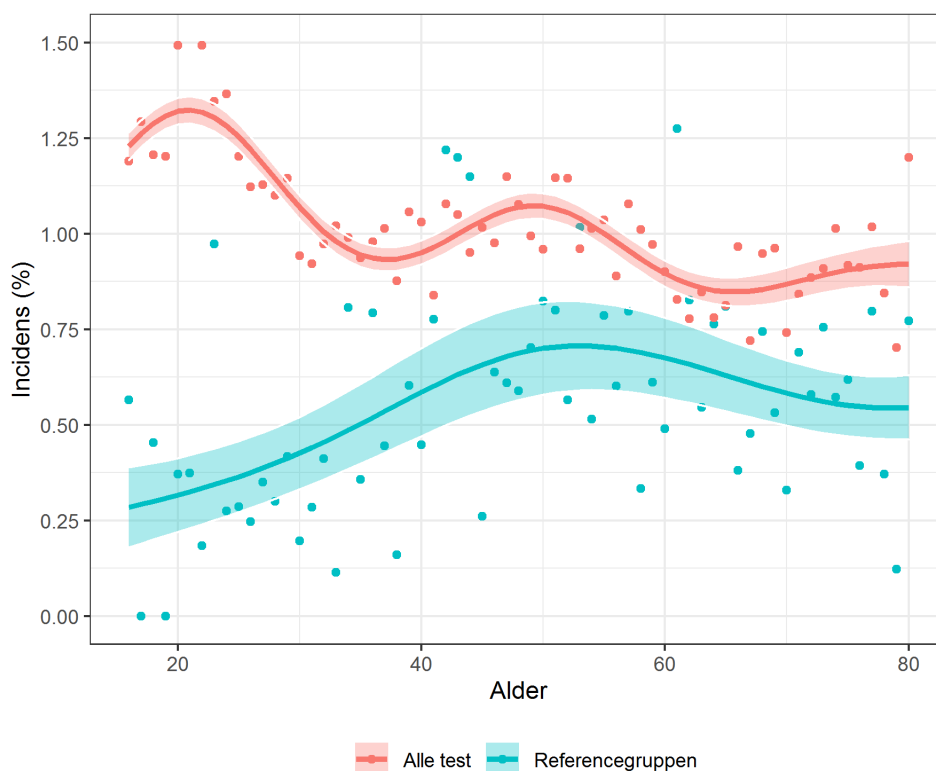




Figur 5 beskriver alderseffekten på COVID-19 incidensen i hele befolkningen og blandt referencegruppen. Incidensen af COVID-19 i referencegruppen er relativt stabil i de 40-67 årige, hvorimod der observeres et faldende incidens for de over 67-årige. Dette fald kan formentligt tilskrives en lavere forekomst af COVID-19 i den aldrende population, som kan have en øget opmærksomhed på fysisk distancering, hvilket er i tråd med resultater fra HOPE projektet. Der ses også en markant lavere incidens i de unge under 35-år. Man må formode at personer i denne aldersgruppe som indgår i referencegruppen, er markant anderledes end den generelle befolkning hvad angår underliggende sygdom. En teori er, at personer med underliggende sygdom, som er i særlig risiko, vil have færre kontakter, og dermed en lavere incidens af COVID-19.

Til sammenligning med referencegruppen, er incidensen af COVID-19 baseret på alle testede, markant anderledes, idet der ses en kraftig stigning i incidens i aldersgruppen 18-28-årige. Forskellen kan til dels forklares ved, at unge mennesker i øjeblikket har en højere sandsynlighed for at blive testet, såfremt de er smittede med COVID-19, pga den intensive smitteopsporing i forbindelse med udbrud iblandt unge mennesker. I samme graf ses også et fald i incidens i aldersgruppen over 55-årige som delvis afspejler incidensen for aldersgruppen i referencegruppen.

Figur 5. Alderseffekten på incidensen af COVID-19, beregnet ved generaliserede additive modeller, for den generelle befolkning (rød kurve) og for referencegruppen (blå kurve), i perioden 1. september til 16. oktober. Prikkerne viser den individuelt beregnede incidens i alle aldre.





Bilag 4. Estimation af kontakttal justeret for testintensitet

Modellen indeholder følgende tidsrækker:

P_t : Antal positive prøver på dagen t

T_t : Antal udførte prøver på dagen t

I_t : Et indeks for sygdommens udbredelse på dagen t , svarende til prævalens r_t : Den daglige tilvækst i indeks

For at specificere disse tidsrækkeres statistiske egenskaber, indeholder modellen en central parameter β , der bestemmer hvor mange positive prøver man forventer at få:

$$P_t \sim \text{Poisson}(I_t T_t \beta)$$

Altså, P_t er (betinget) Poisson-fordelt med den angivne forventningsværdi. Parameteren β estimeres med maximum likelihood og findes typisk til værdier omkring 0.7 med et 95 % konfidensinterval fra 0.6 til 0.8. En værdi på $\beta = 1$ ville svare til at de testede personer var udtaget ved en tilfældig stikprøve, således at en fordobling af testkapaciteten medfører en fordobling af antal positive prøver, og i dette tilfælde ville I_t være prævalensen. Tilsvarende ville en værdi på $\beta = 0$ svare til at en udvidelse af testkapaciteten ikke medfører flere positive prøver, fordi de ekstra prøver bliver brugt på at teste personer der ikke har sygdommen. I dette tilfælde vil I_t være antal personer der tester positivt på den givne dag. Værdier af β imellem 0 og 1 vil optræde når prøverne bliver brugt dér hvor man forventer at finde flest positive, således at en øget testkapacitet medfører en bredere testning og dermed flere positive prøver, men samtidig en lavere incidens. Det skal understreges at denne model skal opfattes som en simplificeret beskrivelse af de komplicerede forhold, der bestemmer hvem der bliver testet.

Indeks I_t antages at udføre en biased random walk i logaritme-domæne, altså: $\log I_{t+1} \sim N(\log I_t + r_t, \sigma^2 I)$

mens den daglige tilvækst r_t udfører en unbiased random walk: $r_{t+1} \sim N(r_t, 2r)$

Modellen indeholder således parametre (β , $\sigma^2 I$, $\sigma^2 r$), og ikke-observerede stokastiske variable ($\log I_t$, r_t), samt data (P_t , T_t). Ikke-observerede stokastiske variable integreres ud (v.h.a. Laplace-approximation) hvorefter den resulterende likelihood maksimeres numerisk.

Den daglige tilvækst r_t konverteres til et kontakttal ud fra formlen $R_t = 1 + T r_t$ hvor T er generationstiden, sat til 4.7 dage.

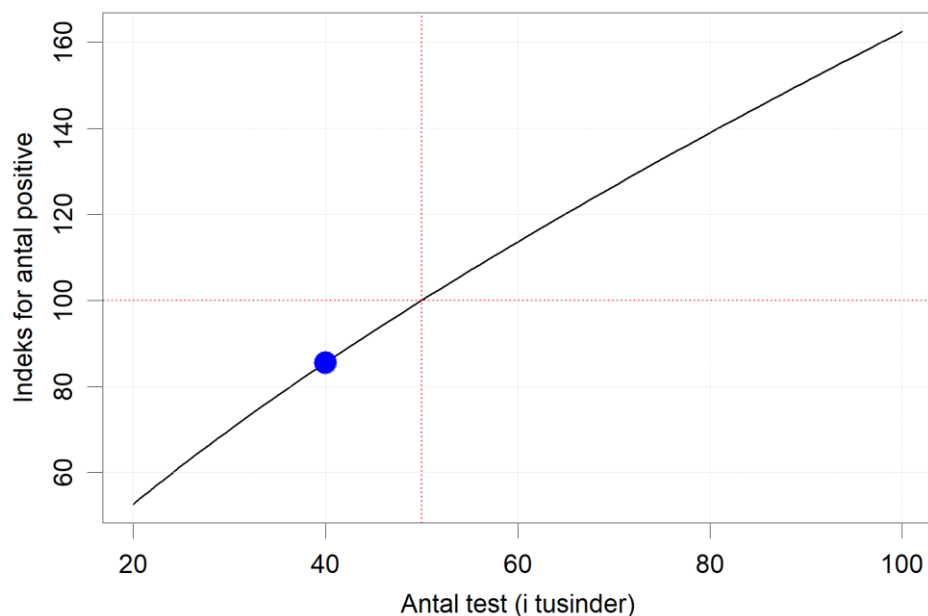
Modellen kan bruges til at standardisere tidsrækken P_t for testintensiteten: Hvis testkapaciteten T_t havde været konstant og lige T_{ref} is hele perioden, havde det forventede daglige antal positive prøver været $I_t T_{ref} \beta$.

Det viser sig at maksimum-likelihood-estimerer af varianserne ($\sigma^2 I$, $\sigma^2 r$) giver anledning til meget glatte estimerede kurver for kontakttallet. Den præcise årsag til dette er ikke klar, men for at omgå dette, fikseres forholdet $\tau = \sigma I / \sigma r$. Dette forhold kan opfattes som en tidsskala, på den måde at det estimerede kontakttal R_t bestemmes ud fra den gennemsnitlige tilvækst i indeks I_t i et eksponentielt af tagende tidsvindue omkring tidspunktet t , hvor bredden af dette vindue skalerer med τ . Denne beskrivelse gør det muligt at sammenligne resultaterne med andre estimationsmetoder, specifikt EpiEstim, hvor vinduet er eksplicit specificeret.



Bilag 5. Sammenhængen mellem antal test og antal positive COVID-19 tilfælde

Figur B4: Indeks for sammenhængen mellem antal test og antal positive. Indeks er sat til 100 ved 50.000 daglige test, som var niveauet ultimo september.



Et eksempel relateret til den blå prik: Hvis man en dag tager 50.000 prøver og finder 500 positive så skulle man forvente at have fundet $500 \cdot 86 / 100 = 428$ positive, hvis man i stedet kun havde taget 40.000 prøver. Dette svarer nogenlunde til det lavere antal test, som tages i weekender.



Bilag 6. Lokal poisson regression med korrektion for antal test

Der sker løbende ændringer af de lokale indsatser for at reducere smittespredningen – både som kommunikation for at få folk til at ændre adfærd og mere direkte med retningslinjer og kontakt-opsporing. Nogle af disse tiltag sker på kommuneniveau mens andre tages på regionalt niveau. På kommunalt niveau er antallet af nye test positive lave i mange kommuner og derfor meget påvirket af tilfældige variationer.

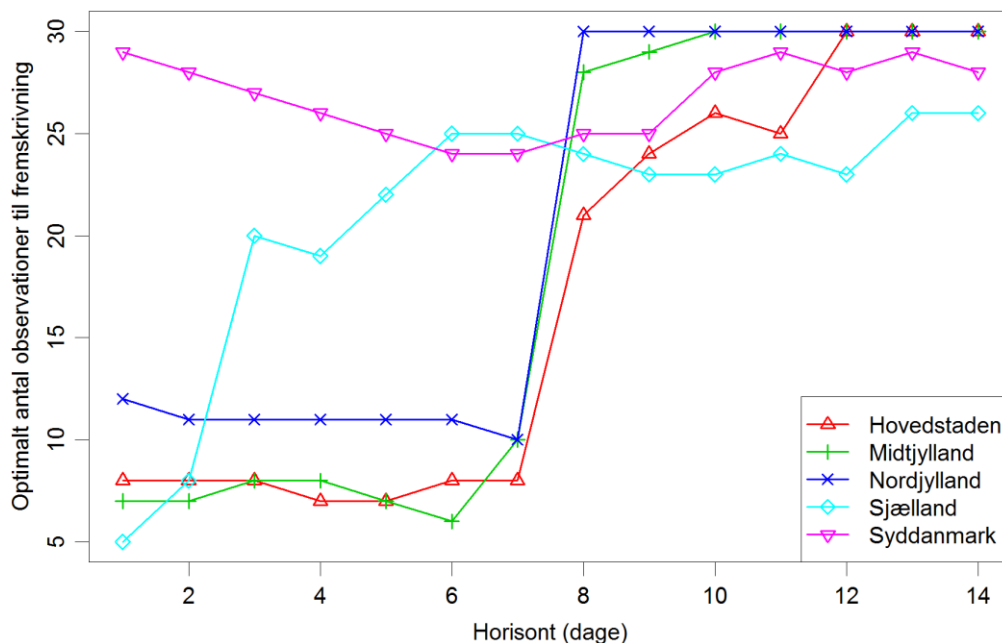
Ved lokale poisson regressioner antages det, at vækstraten for antal test positive per dag (korri-geret for antal test) er konstant i en kortere periode. Der estimeres en generaliseret lineær mo-del med poisson fordelte observationer, den naturlige logaritme som link funktion og tiden som forklarende variabel, samt $\beta \cdot \log(\text{antaltest})$ som et såkaldt offset.

For at lave fremskrivninger med modellen skal der angives et referenceantal test. Her er det valgt at bruge median af antal test siden d. 1. september (afrundet til tusinder). I figur 8 er der benyttet data fra 11 dage til hvert estimat. Dvs. for hver dag er der estimeret en model baseret på dagens observation og de ti forudgående dages observationer af test positive. Referenceantallet for test er dernæst brugt til at lave en fremskrivning for den pågældende dag med et tilhørende 95% prædiktionsinterval (Angivet med rød). Derudover er modellen fra den sidste dag brugt til at lave en fremskrivning 14 dage frem (blå). Det er vigtigt at pointere, at det er en alt andet lige fremskrivning, baseret på væksten i antal positive de seneste 11 dage.

For at undersøge hvilket antal observationer der er bedst at bruge, er der lavet 1 til 14 dages fremskrivninger fra hver dag fra d. 31. august til d. 4. oktober. Det er gjort med modeller baseret på 5 til 30 forudgående observationer. Kvaliteten af fremskrivningerne er bedømt ved at evalu-ere logaritmen af likelihood for at observere et antal positive givet det faktiske antal test på den pågældende dag – disse bidrag er lagt sammen for alle fremskrivninger i perioden. Dette er end-videre gentaget for hver region og figur B5 viser det optimale antal observationer for hver region til hver horisont. Når der skal fremskrives mere end en uge frem er der enighed om at det er op-timalt at bruge et relativt højt antal observationer. Men når der skal laves fremskrivninger få dage frem er der stor forskel på regionerne. Dette skyldes den historiske udvikling i antallet af positive test – et eksempel er Region Sjælland, hvor 1 og 2 dages fremskrivninger laves bedst med få dages hukommelse mens det er optimalt at bruge væsentlig flere observationer, hvis der skal fremskrives mere end et par dage. Dette skyldes, at Region Sjælland har haft en række ret kortvarige opblomstringer af smitte, fx på slagteriet i Ringsted, hvor det fra en optimeringsvin-kel kan være en fordel at fange det generelle lave niveau uden at beskrive opblomstringerne. Omvendt er der et større antal positive i Region Hovedstaden, hvilket gør, at enkelte smittekæ-der ikke fylder så meget og derfor fås et mere jævnt forløb, hvor det er optimalt at bruge relativt få observationer op til en uge frem.



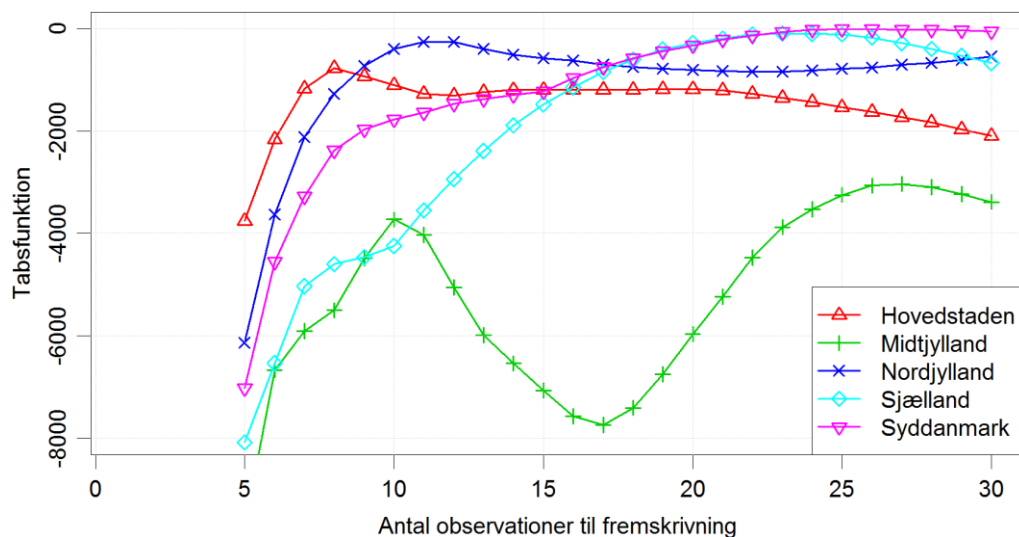
Figur B5: For hver region og hver tidshorisont for fremskrivninger vises det antal observationer, som giver de mest sandsynlige fremskrivninger.



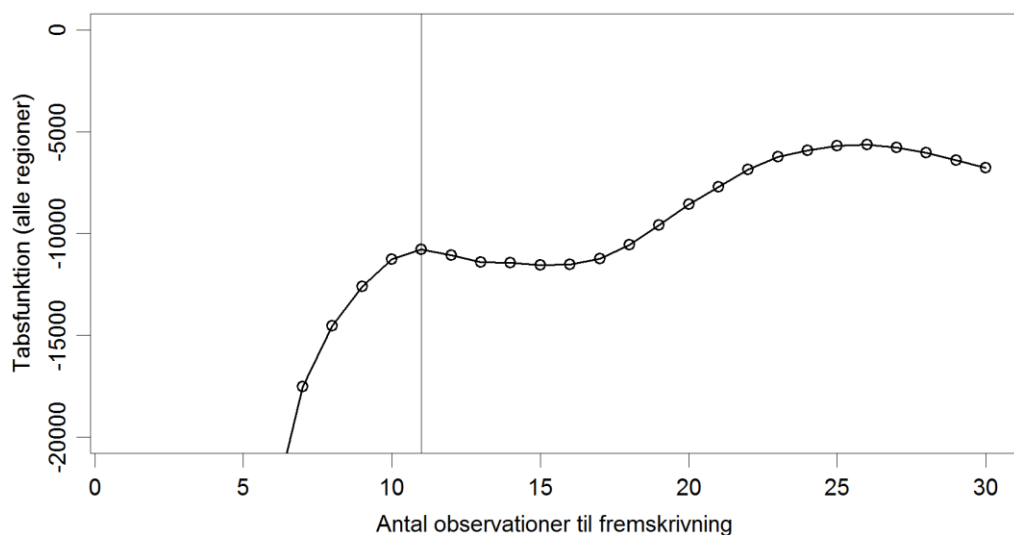
Det ønskes at vælge et fælles antal observation for at forenkle modellen. For hver region og hver horisont er den maksimale log(likelihood) fratrukket, hvorved det er nemt at beregnet tabet ved at vælge et fast antal observationer til alle fremskrivninger. I det efterfølgende er det besluttet at vælge antallet af observationer baseret på fremskrivninger op til 10 dage frem, da antagelsen om homogen vækst i smitte og dermed fravær af adfærdsændringer bliver meget mindre sandsynlig når der kigges længere frem. Figur B6 viser, hvordan denne tabsfunktion afhænger af antal observationer for de enkelte regioner. Tre regioner har et (lokalt) optimum nær 10 observationer, mens Region Syddanmark og særligt Region Sjælland bedst fremskrives ved brug af mere end 20 observationer. I figur B7 er tabsfunktionerne lagt sammen til hele landet. Der er et lokalt optimum ved 11 observationer hvilket er det foretrukne antal observationer, da et valg af 25 observationer vil gøre at stigninger vil ses meget forsinket.



Figur B6: Samlet tabsfunktion (log(likelihood)) per region ved samme antal observationer til fremskrivninger 1 til 10 dage frem relativt til at benytte det optimale antal observationer til de enkelte horisonter.



Figur B7: Samlet tabsfunktion (log(likelihood)) per region ved samme antal observationer til fremskrivninger 1 til 10 dage frem relativt til at benytte det optimale antal observationer til de enkelte horisonter.





Bilag 7. Medlemmer af ekspertgruppen

Ekspertgruppen ledes af overlæge Robert Leo Skov og læge Camilla Holten Møller, Infektionsberedskabet, Statens Serum Institut.

Danmarks Tekniske Universitet, Institut for Matematik og Computer Science

- Kaare Græsbøll, ph.d., MSc, Seniorforsker, Sektion for dynamiske systemer
- Lasse Engbo Christiansen, ph.d., MSc Eng, lektor, Sektion for dynamiske systemer
- Sune Lehmann, Professor, DTU Compute, Institut for Matematik og Computer Science
- Uffe Høgsbro Thygesen, Civilingeniør (ph.d.), lektor, Sektion for dynamiske systemer

Københavns Universitet, Det sundhedsvidenskabelige fakultet, Institut for Veterinær- og Husdyrvidenskab,

- Carsten Thure Kirkeby, Seniorforsker, ph.d., MSc. Sektion for Animal Welfare and Disease Control
- Matt Denwood, BVMS, ph.d., Sektion for Animal Welfare and Disease Control
- Tariq Halasa, BSc, MSc, ph.d., Professor, Sektion for Animal Welfare and Disease Control

Københavns Universitet, Institut for Folkesundhedsvidenskab

- Theis Lange, Vice Institutleder, Lektor i Biostatistik, ph.d., Biostatistisk Afdeling

Københavns Universitet, Niels Bohr Institutet

- Troels Christian Petersen, Lektor, Eksperimentel subatomar fysik

Roskilde Universitets Center, Institut for Naturvidenskab og Miljø

- Viggo Andreasen, Lektor, Matematik og Fysik

Region Hovedstaden

- Anders Perner, Professor, Overlæge, Intensivafdelingen, Rigshospitalet

Danmarks Statistik

- Laust Hvas Mortensen, Chefkonsulent, professor, ph.d., Metode og Analyse

Statens Serum Institut

- Mathias Heltberg, Postdoc ENS Paris samt Statens Serum Institut. Infektionsberedskabet
- Frederik Plesner Lyngse, Postdoc, Økonomisk Institut, Københavns Universitet samt Statens Serum Institut
- Peter Michael Bager, Seniorforsker, ph.d., Infektionsberedskab, Epidemiologisk Forskning, Statens Serum Institut
- Robert Skov, Overlæge, Infektionsberedskab, Statens Serum Institut
- Camilla Holten Møller, Læge, PhD, Infektionsberedskab, Statens Serum Institut