

UNIVERSITY OF  
COPENHAGEN



PH.D. THESIS  
by  
*Christian Michelsen*

## Biological Data Science

In ancient genomics, anesthesiology,  
epidemiology, and a bit in between

*Submitted: November 11, 2022*

*This thesis has been submitted to the  
PhD School of The Faculty of Science,  
University of Copenhagen*

Supervisor	Troels C. Petersen	Niels Bohr Institute
Cosupervisor	Thorfinn S. Korneliussen	Globe Institute

Christian Michelsen, *Biological Data Science*, In ancient genomics,  
anesthesiology,  
epidemiology, and a bit in between, November 11, 2022.

*Til kvinderne i mit liv*



# *Contents*

## FRONT

Preface [i](#)

Acknowledgements [iii](#)

Abstract [v](#)

Dansk Abstract [vii](#)

Publications [ix](#)

## MAIN

- 1 Notes on the design [1](#)
  - 1.1 Document layout [1](#)

## APPENDIX

- A SSI Ekstertrapport [7](#)
- B SSI Notat [37](#)



FRONT





# *Preface*

This Ph.D. thesis summarizes my scientific research in collaboration with the Niels Bohr Institute (NBI) and the Globe Institute, University of Copenhagen, and was funded by the Lundbeck Foundation. The research was supervised by Associate Professor Troels C. Petersen (NBI) and Assistant Professor Thorfinn S. Korneliussen (Globe Institute).

Being a multi-disciplinary project, the research presented in this thesis is multi-faceted and covers a wide range of topics with the main scope being the development and integration of a novel statistical methods and machine learning models for the analysis of large-scale biological data. The thesis is organized as follows. First I present a brief introduction to the statistical methods and machine learning models used in the thesis. Then I present the research in the form of four papers, each of which reflects a different aspect of the research.

The first paper presents a novel method I developed for detecting and classifying ancient DNA damage in metagenomic samples taking the full taxonomic information into account. While the first paper focuses on the development of the statistical model in the field of ancient genomics, the second paper focuses on the use of modern machine learning models in medicine and how advanced boosted decision trees can not only improve the accuracy of identifying patients at risk of being readmitted after knee or hip surgery, but doing so in a way that is interpretable as well.

In the beginning of 2020 we all experienced how COVID-19 suddenly changed our lives and impacted our societies in dramatic ways. During this time, I ended up working for Statens Serum Institut, the Danish CDC, on a project to develop an agent based model capable of simulating the spread of COVID-19 in Denmark. This model is presented in the third paper and was used to inform the Danish government on how to best handle the pandemic in the early stages and the effect of contact tracing.

Finally, in the fourth paper I show how advanced Bayesian methods can be utilized to better estimate the diffusion coefficients of molecules in the cell nucleus in XXX experiments.



# *Acknowledgements*

First of all I want to express my sincere gratitude to my long time supervisor, captain, and friend: Troels. You are truly an inspiration to work with. I want to thank you for opening so many doors for me, both academically, professionally, and nautically. I am looking forward to our future adventures together. I also want to thank my co-supervisor, Thorfinn. I want to thank you for introducing me to the field of bioinformatic and helping me to develop my skills in this area. I also want to thank you for your patience and guiding me through the endless amount of (near) identical biological concepts and helping me to understand the minute differences.

I have been fortunate to work with people from a wide range of backgrounds and disciplines during my Ph.D. The author lists on the papers in this thesis include a particle physicist, bioinformaticians, a clinical professor, epidemiologists, a medical doctor, a bio-physicist, a mathematician, a biologist, and the president of the Royal Danish Academy of sciences and letters. Before anything else, I want to thank all of my co-authors for their work and contributions to these papers and for allowing me to be a part of their projects. I have learned a lot from all of you, and I hope that I have been able to contribute something to your work as well.

I am thankful for all the people who have helped me with my work and listened to my complaints when I was stuck, when the code did not compile, or when the small bug was almost impossible to find (which was not a small amount of time). In particular I want to thank the people at Globe who I have spent the most time with; Rasa, Alba, and Rasmus. I also want to thank the Korneliussen Group and the people in my office for helpful advice, suggestions and discussions. This also includes Daniel Nielsen and Rasmus Ørsøe from NBI. Finally, I want to thank Mathias Heltberg for many years of fruitful collaboration and for including me in his projects.

This project would not have been possible had it not been for the Lundbeck Foundation which funded my Ph.D. In addition to the funding itself, I am grateful for the inter-disciplinary aspect of project which has allowed me to meet so many inspiring and talented people and for the freedom to pursue my own interests within the project.

I would also like to express my gratitude to Professor Guido Sanguinetti from the International School for Advanced Studies, SISSA, in Trieste, Italy, for hosting me in his group during the Winter of 2021/2022. My gratitude also goes out to Kosio, Sara, Max, Romina, Noor, Viplove, Anne-Marie, and all the other wonderful people that I met during in Trieste. Thanks for making my stay in Italy so enjoyable and for welcoming me in a way that only non-Danes can do.

I want to thank my friends for always being there for me. A special thanks to my friends from NBI and Borchsen who I know that I can always count on, whether or not that includes a trip in the party bus of the Sea, taking Artemis out for a sail, or board games and beer. Thank you for always being there. I also want to thank my family, especially my parents for their unconditional support and encouragement. I am grateful for the opportunities that they have given me and for the sacrifices that they have made for me.

Lastly, I want to thank my future wife and mother of our child, Anna. I would not have been able to do this without you. Thank you for your patience and support. I am looking forward to our future together. I love you to the moon and back.

# *Abstract*

Basically a thesis (book?) class for Tufte lovers like myself. I am aware that `tufte-latex` already exists but I just wanted to create my own thing.



## *Dansk Abstract*

Her et dansk abstract.





# *Publications*

The work presented in this thesis is based on the following publications:

- Paper 1:** Christian Michelsen, Mikkel W. Pedersen, Antonio Fernandez-Guerra, Lei Zhao, Troels C. Petersen, Thorfinn S. Korneliussen (2022). “*metaDMG: An Ancient DNA Damage Toolkit*”.
- Paper 2:** Christian Michelsen, Christoffer C. Jorgensen, Mathias Heltberg, Mogens H. Jensen, Alessandra Lucchetti, Pelle B. Petersen, Troels C. Petersen, Henrik Kehlet (2022). “*Preoperative prediction of medical morbidity after fast-track hip and knee arthroplasty – a machine learning based approach.*”.
- Paper 3:** Mathias S. Heltberg, Christian Michelsen, Emil S. Martiny, Lasse E. Christensen, Mogens H. Jensen, Tariq Halasa and Troels C. Petersen (2022). “*Spatial Heterogeneity Affects Predictions from Early-Curve Fitting of Pandemic Outbreaks: A Case Study Using Population Data from Denmark*”. In: Royal Society Open Science 9.9. issn: 2054-5703. doi: 10.1098/rsos.220018.
- Paper 4:** Susmita Sridar, Mathias S. Heltberg, Christian Michelsen Judith M. Hattab, Angela Taddei (2022). “*Microscopic single molecule dynamics suggest underlying physical properties of the silencing foci*”.



MAIN



# 1 *Notes on the design*

This class is my personal mix of different book design influences: mainly the works of Edward R. Tufte, (Heltberg et al., 2022; Korneliussen, Albrechtsen, and Nielsen, 2014) known for the big margin and the plentiness of sidenotes and sidecaptions. The margins are however not as prominent as in Tufte’s works, the main text takes a bit more space, more like in Robert Bringhurst’s typographer’s bible (Heltberg et al., 2022).

So it is a bit of a mix of Tufte and Bringhurst, with some of my own choices for other design features, as we will see through this chapter.

## 1.1 *Document layout*

While `tufte-style-thesis` is a class for typesetting theses, the general layout is pretty much the same as in a regular book. A book is traditionally divided into three major sections: the front matter, the main matter and the back matter.



## *Bibliography*

- Heltberg, Mathias Spliid et al. (2022). “Spatial Heterogeneity Affects Predictions from Early-Curve Fitting of Pandemic Outbreaks: A Case Study Using Population Data from Denmark”. In: *Royal Society Open Science* 9.9. ISSN: 2054-5703. DOI: 10.1098/rsos.220018.
- Korneliussen, Thorfinn Sand, Anders Albrechtsen, and Rasmus Nielsen (2014). “ANGSD: Analysis of Next Generation Sequencing Data”. In: *BMC Bioinformatics* 15.1, p. 356. ISSN: 1471-2105. DOI: 10.1186/s12859-014-0356-4. URL: <https://doi.org/10.1186/s12859-014-0356-4> (visited on 2019).





## APPENDIX



## A *SSI Ekstertrapport*

The following pages contain the report from Statens Serum Institut, the Danish CDC: *Ekspertgruppen for matematisk modellering, “Eksperttrapport af den 10. december 2020 – Effekten af kontaktopsporing”* (Statens Serum Institut, 2021).

The report is from December 10 2020 and is a summary on the effect of contact tracing related to COVID-19 in Denmark. The report is in Danish and is based on two agent based models, one from DTU and our model from NBI.



# Eksperttrapport af den 10. december 2020

Effekten af kontaktopsporing



# Indhold

<b>1. Sammenfatning og konklusion .....</b>	<b>3</b>
<b>2. Formål og baggrund .....</b>	<b>4</b>
2.1 Formål og baggrund for modelgruppen .....	4
2.2 Formål med rapporten .....	4
<b>3. Opsporing og håndtering af nære kontakter i Danmark .....</b>	<b>5</b>
3.1 Forudsætninger for en effektiv kontaktopsporing .....	5
3.2 Definition af en nær kontakt .....	5
3.3 Periode for smitteopsporing .....	6
3.4 Opsporing af nære kontakter .....	6
<b>4. Agentbaserede modeller .....</b>	<b>8</b>
4.1 Om agentbaserede modeller .....	8
4.2 Forbehold .....	8
<b>5. Resultater .....</b>	<b>9</b>
5.1 Resultater fra den agentbaserede model udviklet af Niels Bohr Institutet, Københavns Universitet. ....	9
5.2 Resultater fra den agentbaserede model udviklet af DTU Compute, Danmarks Tekniske Universitet .....	10
<b>6. Referencer .....</b>	<b>13</b>
<b>Bilag 1. Beskrivelse af den agentbaserede model fra Niels Bohr Institutet</b>	<b>14</b>
<b>Bilag 2. Beskrivelse af den agentbaserede model fra DTU .....</b>	<b>16</b>
<b>Bilag 3. Regneeksempel .....</b>	<b>22</b>
<b>Bilag 4. Udvikling i antal kontakter fra HOPE projektet .....</b>	<b>24</b>
<b>Bilag 5. Beskrivelse af parametre brugt i rapporten .....</b>	<b>25</b>
<b>Bilag 6. Medlemmer af ekspertgruppen .....</b>	<b>258</b>



# 1. Sammenfatning og konklusion

I indeværende rapport har modelgruppen for matematisk modellering af COVID-19 estimeret hvilke delelementer af kontaktopsporing, som er afgørende for at opnå størst mulig effekt af kontaktopsporing af nære kontakter til COVID-19 smittede personer.

Rapporten præsenterer resultater fra to forskellige agentbaserede modeller, som er udviklet af eksperter fra Danmarks Tekniske Universitet (DTU) og Københavns Universitet, Niels Bohr Institutet (NBI).

En agentbaseret model gør det muligt at modellere enkelte tiltag og deres effekt på smittespredningen af COVID-19. Forudsætningen for en præcis simulation er, at der er tilgængelige data, som kan informere modellen. Der er flere parametre, hvor der i nærværende arbejder er lavet antagelser på basis af de tilgængelige oplysninger. Det forventes, at nogle af disse kan belyses efterhånden som yderligere data frembringes. Hvor der ikke er specifikke eller komplette data, vil en agentbaseret model have unøjagtigheder eller risikere at være baseret på antagelser, som ikke nødvendigvis er retvisende. I modellerne anvendes der endvidere ens ventetidsfordelinger for alle agenter, selvom der i realiteten kan være lokale udsving i ventetider.

Sundhedsstyrelsen udkom d. 23. november 2020 med opdaterede retningslinjer for smitteopsporing af nære kontakter, herunder en udvidet definition af nære kontakter. Indeværende rapport er udviklet i henhold til de tidligere retningslinjer, og tager ikke højde for disse ændringer.

Der er i rapporten heller ikke taget højde for den stigende brug af private antigen test. Coronaopsporingen under STPS foretager også opsporing af nære kontakter, for primærttilfælde som er testet positiv for COVID-19 på sådanne antigen test.

## *Konklusion*

Modellerne peger på, at den største reduktion i kontakttallet kan nås ved effektiv opsporing for flest mulige primærttilfælde. Gevinsten i form af en reduktion i kontakttallet er således større, såfremt der sikres effektiv opsporing for samtlige primærttilfælde, relativt til reduktionen i kontakttallet, som kan opnås ved at nedbringe ventetiden til test og testsvar for primærttilfældet.

Ventetiden til test og testsvar for et primærttilfælde med COVID-19, har stor betydning for den reduktion af kontakttallet, som kan opnås gennem kontaktopsporing. De to uafhængigt udviklede modeller fra hhv. DTU og NBI finder begge, at for hver dag ventetiden til test og testsvar forsinkes for primære tilfælde, stiger kontakttallet med 4%. DTU-modellen finder endvidere, at ventetiden til et primærttilfælde booker en test og samtidig går i isolation har stor betydning for reduktionen i kontakttallet.

Modellerne viser endvidere, at med de anvendte ventetidsfordelinger, vil størstedelen af de nære kontakter som opspores, bliver testet så sent, at det er en mindre del af smitten, som forhindres. Det er derfor vigtigt at opspore nære kontakter hurtigst muligt efter eksponering, så de kan isoleres og blive testet på dag 4 og 6. Dette vil igen afhænge af den samlede ventetid til test og testsvar for primærttilfældet, som er forudsætningen for at opsporingen af nære kontakter kan initieres.

Den agentbaserede model fra NBI finder, at der er yderligere gevinst at hente ved at opspore nære kontakter i de netværk en person indgår i uden for husstand, job og skole. Det skyldes, at relativt få kontakter uden for husstand, job og skole opspores, og at disse kontakter ofte starter nye smittekæder i ikke ellers relaterede netværk. En bredere smitteopsporing har den fordel, at den potentielt finder de nye smittede, som ikke udviser symptomer.



## 2. Formål og baggrund

### 2.1 Formål og baggrund for modelgruppen

Statens Serum Institut indgår i det operationelle beredskab for 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. Medlemmerne af ekspertgruppen fremgår af bilag 5.

Ekspertgruppens modellering var i foråret 2020 baseret på en populationsmodel, der har fokus på den gennemsnitlige adfærd i befolkningen. Populationsmodellen er bedst egnet, når udviklingen beskrives godt ved gennemsnittet. Derimod er populationsmodellen ikke det bedste værktøj til at beskrive de stokastiske hændelser i lokale udbrud, som aktuelt driver smittespredningen af COVID-19 i Danmark.

Siden sommeren 2020 har modelgruppen derfor udviklet to agentbaserede modeller, som er platformen for de analyser, modelgruppen forventes at levere i den kommende periode. De agentbaserede modeller kan, modsat en populationsmodel, estimere effekten ved enkelte tiltag, såsom effekten ved at nedbringe forsamlingsforbuddet, eller effekten af kontaktopsporing.

### 2.2 Formål med rapporten

Opsporingen af nære kontakter, foretaget af Styrelsen for Patientsikkerhed (STPS), er løbende udbygget i Danmark siden foråret 2020. Opgaven er vokset betydeligt i takt med, at det daglige antal nye COVID-19 tilfælde stiger, som følge af både en opblussen af epidemien, men også af, at testkapaciteten i Danmark er væsentligt udbygget hen over sommeren. Der testes aktuelt omkring 70.000 personer dagligt.

Formålet med denne rapport er at belyse, hvilke faktorer der er afgørende for at sikre en effektiv kontaktopsporing. Dette belyses ved at estimere effekten af centrale elementer i kontaktopsporingen, såsom ventetid til test og testresultat hos primærtildfældet, samt ventetid til at nære kontakter bliver opsporet og testet.



## 3. Opsporing og håndtering af nære kontakter i Danmark

### 3.1 Forudsætninger for en effektiv kontaktopsporing

Den vigtigste forudsætning for, at kontaktopsporing er et effektivt redskab til at nedbringe smitten med COVID-19 er, at der til hver en tid identificeres flest mulige smittede personer, som der derved kan udføres smitteopsporing for. Jo lavere mørketallet er, desto flere vil kunne smitteopspores. Det er derfor afgørende, at der sikres nem og hurtig adgang til test, først og fremmest for personer med COVID-19 lignende symptomer, men også for øvrige personer, der kunne have mistanke om at være smittet med COVID-19. Den Nationale Prævalensundersøgelse i Danmark har vist, at op mod 40-50% af dem, som havde antistoffer mod SARS-CoV-2 i blodet, ingen erindring havde om at have haft COVID-19 lignende sygdom<sup>1</sup>. Ved at udbyde adgang til test for flest mulige personer, vil man også finde flere asymptomatiske smittebærere.

### 3.2 Definition af en nær kontakt

*Sundhedsstyrelsen udkom d. 23. november 2020 med opdaterede retningslinjer for smitteopsporing af nære kontakter. Indeværende rapport er udviklet i henhold til de tidligere retningslinjer.*

*Der er således ikke taget højde for den udvidede definition af nære kontakter, eller indførslen af screeningsprøver for personer, som ikke umiddelbart opfylder kriteriet for nære kontakter, men som har været eksponeret i et omfang hvor der tilrådes en screeningstest.*

Kontaktopsporingen af nære kontakter baserer sig på, at personer der testes positiv for COVID-19 isolerer sig, og dernæst at nære kontakter til den smittede opspores, isoleres og testes, for derved at afbryde smittekæder hurtigst muligt.

Definitionen af en nær kontakt er beskrevet i Sundhedsstyrelsens rapport om smitteopsporing af nære kontakter<sup>2</sup>.

En nær kontakt er defineret som en af følgende personer:

- En person der bor sammen med en, der har fået påvist COVID-19
- En person der har haft direkte fysisk kontakt (fx kram) med en, der har fået påvist COVID-19
- En person med ubeskyttet og direkte kontakt til smittefarlige sekreter fra en person der har fået påvist COVID-19
- En person der har haft tæt "ansigt-til-ansigt" kontakt inden for en 1 meter i mere end 15 minutter (fx i samtale med personen) med en, der har fået påvist COVID-19
- Sundhedspersonale og andre som har deltaget i plejen af en patient med COVID-19, og som ikke har anvendt værnemidler på de foreskrevne måder

<sup>1</sup> <https://www.ssi.dk/-/media/arkiv/dk/aktuelt/nyheder/2020/notat---covid-19-prvalensundersogelsen.pdf?la=da>

<sup>2</sup> <https://www.sst.dk/da/Udgivelser/2020/COVID-19-Smitteopsporing-af-naere-kontakter>





### 3.3 Periode for smitteopsporing

Der foretages smitteopsporing for perioden, hvor primærtildfældet vurderes at være smitsom. Smitteperioden er således afgrænset til 48 timer før symptomdebut til 48 timer efter symptomophør. For primære tilfælde der ikke har symptomer på COVID-19, er den smitsomme periode afgrænset til 48 timer før positiv test til 7 dage efter.

### 3.4 Opsporing af nære kontakter

Nære kontakter til en person der er smittet med COVID-19 kan opspores på følgende måder:

- De bliver kontaktet af STPS's Coronaopsporingen
- De bliver kontaktet ifm. kendte udbrud, eksempelvis på skoler
- De bliver kontaktet direkte af primærtildfældet
- De bliver notificeret om, at de har været tæt på en smittet person via appen Smitte|Stop

#### *Nære kontakter opsporet af Coronaopsporingen*

Coronaopsporingen under STPS kontakter smittede mhp. at hjælpe med at identificere og opspore nære kontakter til den smittede. Smittede kan også vælge selv at iværksætte opsporing af nære kontakter, og henvise dem til Coronaopsporingen, hvor de nære kontakter vil modtage rådgivning om, hvornår de bør testes, samt får adgang til at booke test på de pågældende dage.

Aktivitetsrapporter fra STPS viser, at der i hele opsporingsperioden i gennemsnit opspores ca. 5 nære kontakter for hvert primærtildfælde, der foretages kontaktopsporing for. Dette er et samlet gennemsnit for opsporede nære kontakter som STPS opsporer, og som primærtildfældet selv opsporer.

Til sammenligning er det estimeret i HOPE-projektet, at danskere henover sommeren i gennemsnit havde ca. 11 kontakter dagligt. Dette antal er nu faldet til ca. 7 kontakter dagligt, som opfylder kriterierne for en nær kontakt, se bilag 4.

Det skal dog pointeres, at Coronaopsporingen ikke er involveret i opsporing af nære kontakter i relation til udbrud i dagtilbud, skoler, plejehjem og hospitaler. Der vil der være opsporede kontakter fra sådanne udbrud, som kontakter Coronaopsporingen for at få rådgivning om hvilke dage de bør testes, samt for at få rekvisitioner til booking af test på de pågældende dage.

Nære kontakter anbefales at blive testet på dag 4 og dag 6 efter vurderet eksponering. Dette relaterer sig til latenstiden, som er perioden fra, at man bliver smittet, til at man er smitsom, og virus kan påvises. En person som er opsporet som nær kontakt til en smittet skal ifølge Sundhedsstyrelsens retningslinjer gå i selv-isolation, indtil der foreligger testsvar. Såfremt der foreligger et negativt testresultat på dag 4, kan den nære kontakt bryde isolationen, men skal fortsat testes på dag 6. Hvis testresultatet på dag 4 derimod er positivt, skal den nære kontakt ikke testes igen på dag 6, men forblive i isolation indtil 48 timer efter symptomophør.

#### *Nære kontakter der ikke opspores af Coronaopsporingen*

Der vil være nære kontakter, der ikke opspores og rådgives via Coronaopsporingen. Dette kunne fx være nære kontakter, der bliver opsporet af primærtildfældet selv, og som vælger at booke test på coronaprover.dk uden først at have rådført sig med Coronaopsporingen. Det kunne også være personer, som er opsporet via appen Smitte|Stop, eller personer der mener, at de på anden vis



kan være nære kontakter til en smittet – uden nødvendigvis at opfylde kriteriet for at være en nær kontakt.

I oktober måned blev der i alt testet 1.091.966 personer. Heraf havde 62% (n = 675.623) bestilt tid på coronaprover.dk. Blandt disse svarede 58% (n = 391.146) på spørgeskemaet på coronaprover.dk, hvoraf 25% (n = 99.389) anførte, at de blev testet fordi, de var nær kontakt til en smittet (herunder personer som er adviseret af Smitte|Stop app). Kun 13% (n = 12.706) af dem som svarede, at de blev testet fordi de var nær kontakt til en smittet, var testet på én af de rekvisitionskoder, som der anvendes i Coronaopsporingen (Tabel 1). Samlet set blev 45.616 personer testet på én af de rekvisitionskoder som anvendes i Coronaopsporingen i oktober måned, hvor test-positivprocenten var ca. 4%. Til sammenligning var positivprocenten for de personer, der svarede, at de var nær kontakt til en smittet på Coronaprover.dk omkring 2,5 %. Dette indikerer at Coronaopsporingen har større succes med at opspore de korrekte nære kontakter, sammenlignet med hvis befolkningen selv booker test som nær kontakt, uden forudgående rådgivning fra Coronaopsporingen.

*Tabel 1. Oversigt over antal testede i oktober måned 2020.*

	<b>Oktober</b>		
	N	<i>Testpositive (1. test)</i>	
		n	%
Testede personer	1.091.966	14.723	1,35
Total antal tests rekvireret via Coronaopsporingen	45.616	1.941	4,26
Bestilt på coronaprøver.dk	675.623	10.335	1,53
Svaret på spørgeskema	391.146	5.387	1,38
Ja, nær kontakt til smittet (herunder adviseret på Smitte Stop app)	99.389	2.544	2,56
Rekvireret test via Coronaopsporingen	12.706	524	4,12



## 4. Agentbaserede modeller

### 4.1 Om agentbaserede modeller

I indeværende rapport er resultaterne for effekten af kontaktopsporing frembragt fra to forskellige agentbaserede modeller, som er udviklet på henholdsvis Danmarks Tekniske Universitet (DTU) og Niels Bohr Institutet, Københavns Universitet (NBI).

En agentbaseret model simulerer et antal agenter (individer i en population) og deres interaktioner med andre agenter, svarende til de interaktioner som en befolkning normal vis har. Hver agent er således en person, som er knyttet til en lokation i Danmark, svarende til deres bopæl. Agenterne indgår i flere forskellige netværk, f.eks. husstand, job og skole hvor de har kontakt til andre personer. Desuden har de andre kontakter til tilfældige personer i samfundet i den tid, hvor personen ikke er hjemme, på job eller i skole.

Hvis en agent bliver smittet med SARS-CoV-2, er forløbet for den enkelte agent beskrevet således, at agenten først er eksponeret (E) og derefter infektiøs (I), hvorefter agenten ikke længere er smitsom og betragtes som rask (R). De gennemsnitlige tider i hvert sygdomsstadie kan findes i bilag 1 og 2.

Hver kontakt som en agent eksponeres for tildeles en sandsynlighed for at blive smittet af en anden agent, hvis denne er smitsom. Sandsynligheden er sat til et niveau, som afspejler den nuværende situation med et kontakttal omkring 1.

Ud fra de ovenstående generelle antagelser simuleres en epidemi. For en mere detaljeret beskrivelse af de agentbaserede modeller, herunder de inkluderede parametre, henvises til bilag 1 (NBI) og 2 (DTU).

### 4.2 Forbehold

Mens en agentbaseret model kan medtage mere detaljerede dynamikker i en epidemi, så kræver en præcis simulation input fra data, som ofte ikke er tilgængelige eller forefindes, fx hvem en person mødes med i løbet af en dag. Derfor kan en sådan model have unøjagtigheder eller bygge på antagelser, som ikke er retvisende. Det er ikke muligt at kvantificere den nøjagtige størrelse eller effekt af disse potentielle fejlkilder.

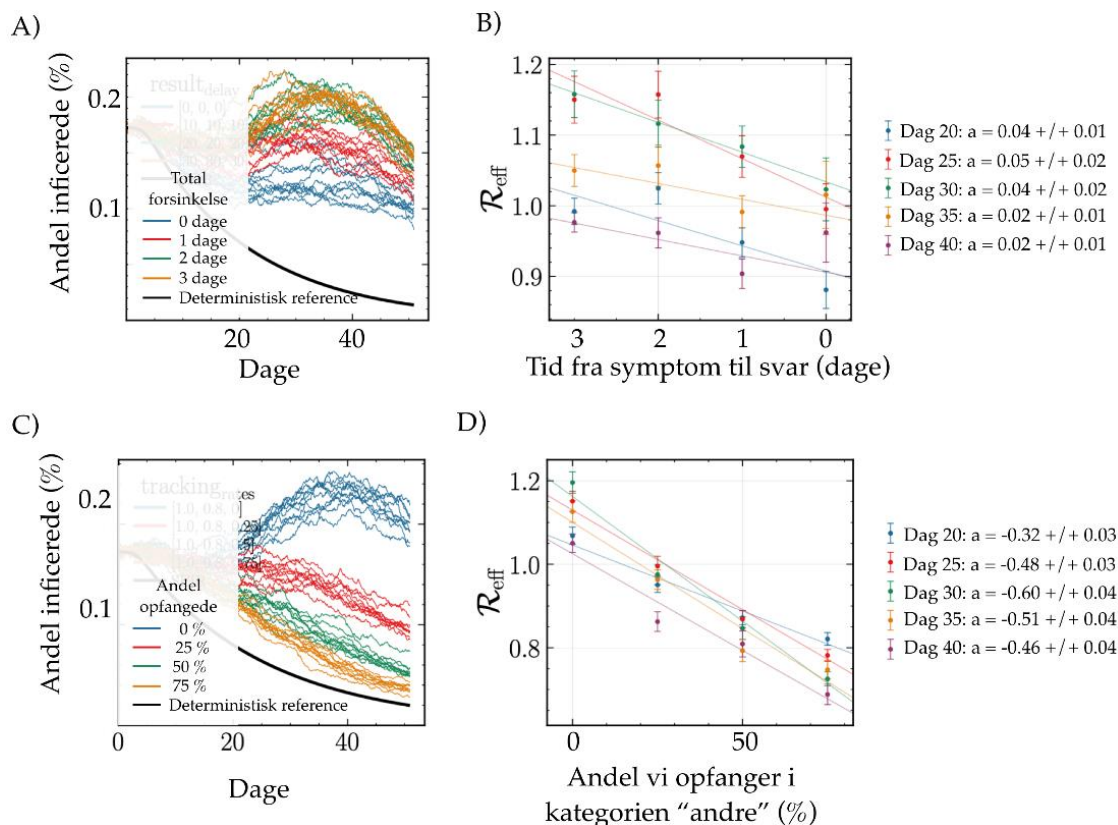
## 5. Resultater

### 5.1 Resultater fra den agentbaserede model udviklet af Niels Bohr Institutet, Københavns Universitet.

Modelkørslerne viser, at når 80% af de sekundære tilfælde i netværkenes husstande, arbejde og skole opspores, vurderes det, at ville nedsætte kontakttallet med omkring 30% sammenlignet med et hypotetisk scenarie uden opsporing af nære kontakter. Dette fremgår af figur 1. Hvis det af logistiske eller kapacitetsmæssige årsager ikke lykkedes at kontakte alle nye COVID-19 tilfælde, vil det betyde en forøgelse af kontakttallet i proportion til dette tal. Dvs. hvis opsporingen ikke kommer i kontakt med 20% af nye COVID-19 tilfælde, vil man potentielt miste 6 procentpoint ( $0.2 \times 0.3 = 0.06$ ) af reduktionen i kontakttallet, som ellers kunne opnås ved kontaktopsporing.

Ventetiden fra at et primært tilfælde ønsker en COVID-19 test (fx hvis man har symptomer), til at vedkommende har modtaget resultatet fra en test har indflydelse på effekten af både selvisolation og kontaktopsporing. Ved en række simulationer med forskellige antagelser finder modellen, at for hver dag man forkorter tiden mellem bestilling af test og testresultat mindskes kontakttallet med omkring 4%. Effekten er lidt større ved højere kontakttal end 1.

Effekten af kontaktopsporing kan øges ved at opspore flere i netværket af øvrige kontakter (ud over husstand og job og skole). Den agentbaserede model viser, at hvis man opsporer 25% af øvrige kontakter, vil kontakttallet falde med omkring 10%. En mere komplet kontaktopsporing (evt. yderligere hjulpet af apps på mobiltelefoner) vil således nedsætte kontakttallet væsentligt. Tilsvarende resultater er fundet i lignende modeller (Plank et al. (september 2020) og Kretzschmar et al. (august 2020)).

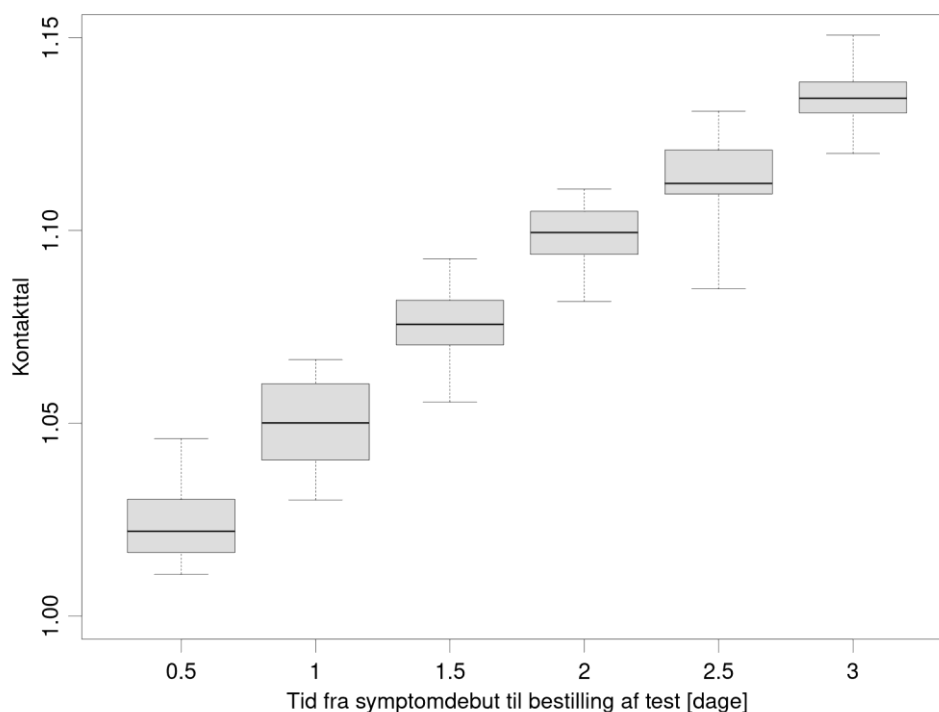


Figur 1: A) Simuleret model, hvor hver kørsel (markeret med samme farve) gentages 10 gange for forskellige værdier af tiden fra symptom til svar. B) Værdien af kontakttallet estimeret på forskellige tidspunkter i simulationen vist i A). Den lineære sammenhæng giver en værdi for hvor mange procent kontakttallet sænkes for hver dag, man gør opsporingen hurtigere. C) Samme som A, men her for forskellige værdier af hvor mange man opsporer blandt øvrige kontakter D) Samme som B) men som funktion af hvor mange man opsporer blandt øvrige kontakter.

## 5.2 Resultater fra den agentbaserede model udviklet af DTU Compute, Danmarks Tekniske Universitet

Denne agentbaserede model er baseret på tilhørsforhold til grupper (hjem, arbejdsplads, m.fl.). Modellen indeholder en række ventetider fra et primærtifælde får symptomer til sekundære tilfælde er opsporet. Modellen er nærmere beskrevet i bilag 2. Modellen er kørt med en række forskellige kombinationer af parametre. For hver kombination er der lavet 40 gentagelser for at illustrere variabiliteten. For hver gentagelse simuleres 30 dage som en transient, hvorefter kontakttallet estimeres baseret på de efterfølgende 30 dage.

De to parametre, som betyder mest for effekten af kontaktopsporingen, er den gennemsnitlige ventetid fra en smittet får (milde) symptomer til at denne går i isolation og samtidig bestiller en test, samt andelen af kontakter som personen reducerer i perioden fra der bestilles en test til der foreligger et testsvar – det antages, at nære kontakter som opspores opretholder samme grad af isolation som andre, der venter på testsvar, hvilket vil sige, at nære kontakter går i isolation fra de bliver notificeret og indtil de får svar på deres første test.

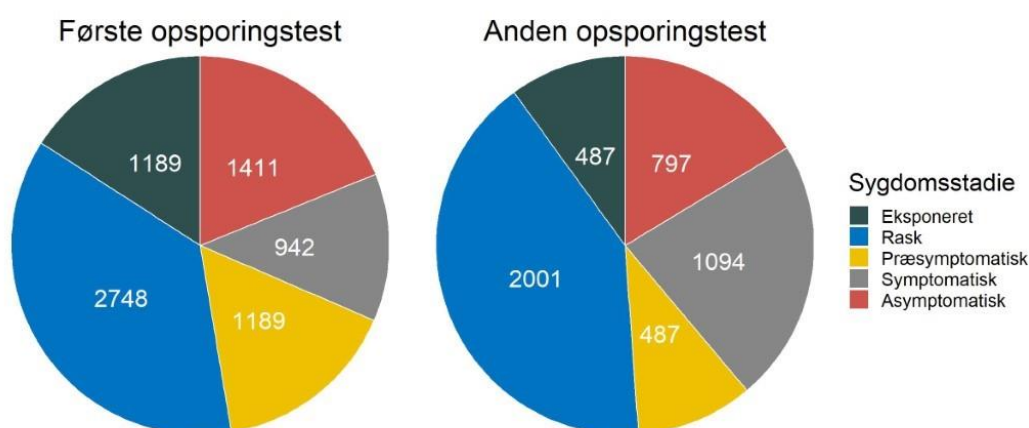


Figur 2: Kontakttallets afhængighed af den gennemsnitlige tid fra at primærtifældet har symptomdebut til der bestilles en test og personen går i en grad af isolation. For hver parameter værdi er der foretaget 40 simulationer, og boxplottet viser median, de indre kvartiler samt minimum og maksimum af disse.



På figur 2 ses en klar effekt af tiden fra symptomdebut til isolation og samtidig bestilling af test. For hver dag den gennemsnitlige person går hurtigere i (delvis) isolation estimeres det, at kontakttallet reduceres med 0,04 (når referencen er et kontakttal omkring 1).

Modellen viser også, at omkring 25% af alle test positive, er fundet gennem kontaktopsporing. Det er her antaget, at der udføres kontaktopsporing for alle tilfælde (Se detaljer i bilag 2), samt at test af nære kontakter bestilles på de foreskrevne tidspunkter. Endvidere viser modellen, at over halvdelen af alle smittede aldrig bliver testet positiv (både falsk negative test og asymptomatiske tilfælde). Disse starter derfor nye smittekæder uden forudgående opsporing. Dette kan være årsagen til, at det kun er 25% som findes gennem kontaktopsporing.



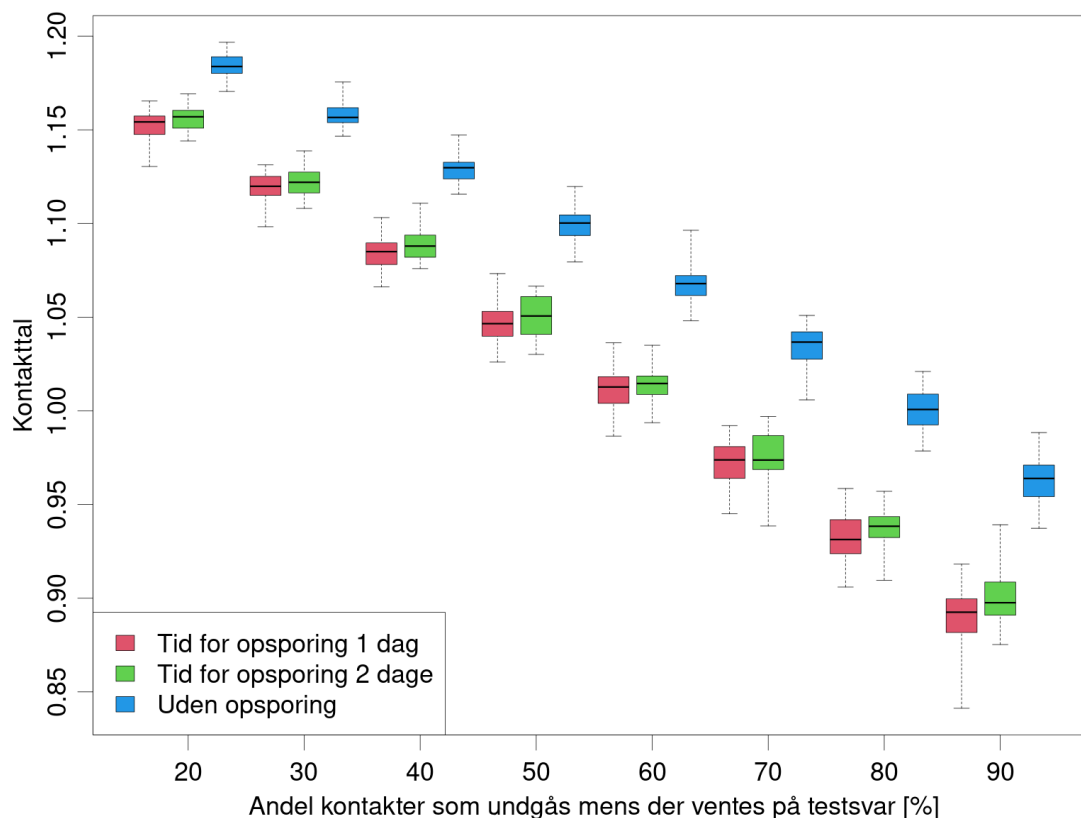
Figur 3. Antal opsporede og smittede i hvert sygdomsstadie, når de får foretaget hhv. første og anden test i opsporingsprocessen. Der er flere, som ikke kommer til anden test, bl.a. fordi de tester positiv i første test eller efter negativt testsvar vælger ikke at få taget den opfølgende test. Derudover vil der være en andel, hvor kontaktopsporingen er initieret sent, således at det kun er foreskrevet at teste personen én gang.

Figur 3 beskriver de forskellige sygdomsstadier for smittede personer, som er opsporet som nære kontakter. Det ses, at en betydelig andel af de opsporede personer, med de i modellen anvendte ventetidsfordelinger, på tidspunktet for opsporingen allerede har overstået deres infektiøse periode, når de testes første gang – en del af disse vil være smittet tidligere og ikke i forbindelse med den nærværende kontaktopsporing. I praksis vil nogle af disse teste positiv, da qPCR kan detektere virus 17 dage efter symptomdebut (Cevik et al., 2020). Desuden ses det, at personer i det præsymptomatiske stadie - hvor ca. halvdelen af smitten sker - kun udgør en lille andel af de opsporede smittede personer ved både første og anden test. Ved begge test er det således under halvdelen af dem, som er smittede, som reelt er infektiøse. Kontaktopsporingen vil derfor kunne optimeres yderligere, hvis man identificerer flere nære kontakter i den præsymptomatiske fase. Dette kan ske ved at nedbringe ventetiden fra symptomdebut til testsvar for primærttilfældet.

Personer, som tidligere er testet positiv er ikke medtaget her og bidrager derfor ikke til antallet af raske. Endvidere vil personer som modtager et positivt testresultat på deres første opsporingstest ikke få foretaget anden opsporingstest. Ovenstående diagrammer er produceret på baggrund af referenceparametrene som beskrevet i bilag 2.



Graden hvorved en smittet person isolerer sig, dvs. hvor stor en andel af ens kontakter man reducerer i perioden fra bestilling af test til testsvar, har stor betydning for kontakttallet. Referenceværdien antages at være 50% reduktion i antallet af kontakter i denne periode. Som det fremgår af figur 4 så opnås der i modellen en reduktion i kontakttallet på knap 0,04 for hver 10 procentpoint graden af isolation øges, hvis der udføres kontaktopsporing (rød og grøn). Mens reduktionen er på 0,03 når der ikke udføres kontaktopsporing (blå). Således har andelen af kontakter, der reduceres hos primærtilfældet og opsporede nære kontakter i ventetiden fra bestilling af test til testsvar, større betydning for en reduktion i kontakttallet, end en reduktion i ventetiden til opsporing af nære kontakter.



Figur 4. Kontakttallets afhængighed af andelen af kontakter et primærtilfælde og opsporede nære kontakter reducerer, i ventetiden fra der bestilles en test til at testsvar foreligger, samt betydningen af ventetiden til at en nær kontakt opspores og går i tilsvarende isolation. For hver parameter værdi er der foretaget 40 simulationer og boxplottet viser median, de indre kvartiler samt minimum og maksimum af disse.



## 6. Referencer

Cevik, M., Kuppalli, K., Kindrachuk, J. & Peiris, M. (2020). Virology, transmission, and pathogenesis of SARS-CoV-2. The BMJ. Lokaliseret: <http://dx.doi.org/10.1136/bmj.m3862>

Kretzschmar, M., Rozhnova, G., Bootsma, M., van boven, M., Wijgert, J & Bonten, M. (2020). Impact of delays on effectiveness of contact tracing strategies for COVID-19: a modelling study. The Lancet Public Health. Lokaliseret: [https://doi.org/10.1016/S2468-2667\(20\)30157-2](https://doi.org/10.1016/S2468-2667(20)30157-2)

Kucirka, Lauren M., Stephen A. Lauer, Oliver Laeyendecker, et al., (2020). Variation in False-Negative Rate of Reverse Transcriptase Polymerase Chain Reaction–Based SARS-CoV-2 Tests by Time Since Exposure. Annals of Internal Medicine. Lokaiseret: <https://doi.org/10.7326/M20-1495>

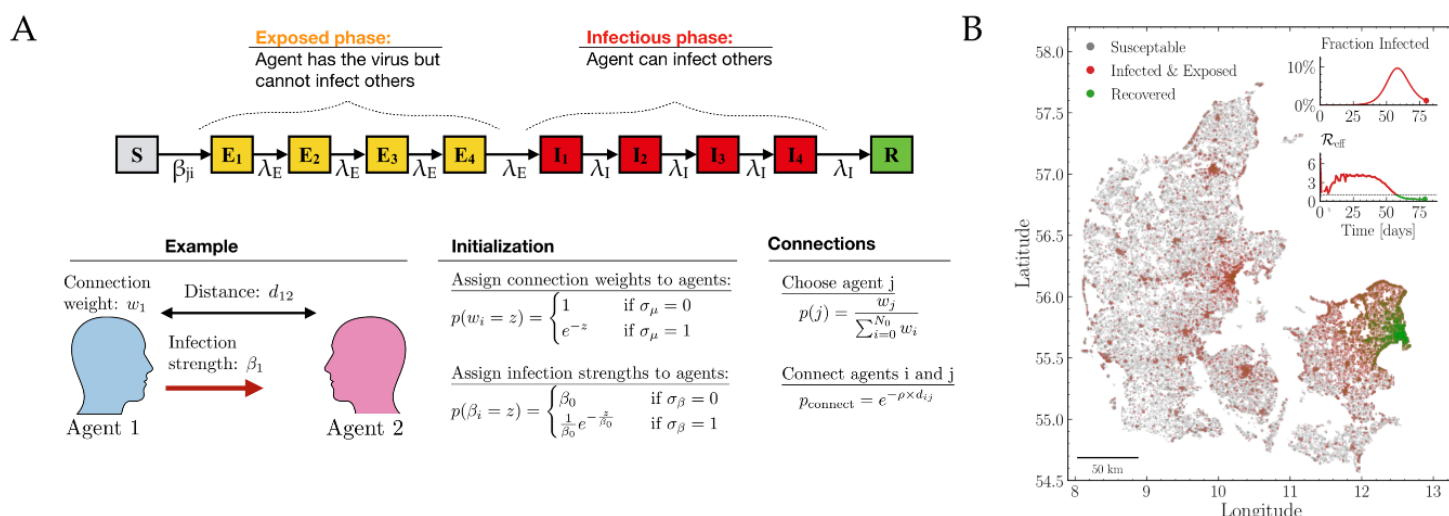
Plank, M., James, A., Lustig, A., Steyn, N., Binny, R. & Hendy, S. (2020). Potential reduction in transmission of COVID-19 by digital contact tracing systems. MedRxiv. Lokaliseret: <https://doi.org/10.1101/2020.08.27.20068346>



# Bilag 1. Beskrivelse af den agentbaserede model fra Niels Bohr Instituttet

Bidrag og udvikling: Christian Michelsen, Emil Martiny, Tariq Halasa, Mogens H. Jensen, Troels C. Petersen og Mathias L. Heltberg

Den agentbaserede model fra NBI baseres på agenter, dvs. individer hvis karakteristika er tildelt ud fra statistiske fordelinger i befolkningen. Dette er f.eks. en aldersfordeling og en fordeling over pendlerafstande. Modellen starter med at fordele Danmarks bopæle ud i landet baseret på det danske hussalg over de sidste 15 år. Herefter placeres agenter i hver husstand baseret på deres alder og geografiske placering.



Figur 5: A) Skematisk oversigt over hvordan interaktionsnetværket i modellen ser ud. B) Eksempel på simulation af smittespredning i Danmark i modellen, gennem et simuleret tilfælde af flokimmunitet i København.

Et afgørende element i modellen er opbygningen af alle personers interaktionsnetværk. Dette genereres ved, at hver agent har et netværk, de interagerer med. Dette opdeles i tre dele: 1) kontakter i hjemmet, 2) kontakter på arbejdet, 3) kontakter i kategorien andre kontakter. Der er ikke nogen geografisk afhængighed af antallet af kontakter på arbejdet, men i den kategori der kaldes “andre”, vil der generelt være flere kontakter for dem der bor i tæt befolkede områder i forhold til dem der bor på landet. Måden hvorpå netværket dannes er vist i Figur 5A.

Ud fra data fra HOPE-projektet har vi estimeret, hvor mange personer hver agent vil interagere med, og i denne model vil alle agenter have mellem 3 og 15 daglige kontakter.

Når modellen simuleres vil alle inficerede agenter gennemgå et forløb, hvor de er i en latent periode, hvor de ikke smitter, hvorefter de vil rykke over i en infektiøs periode, hvor de kan smitte agenter i deres netværk. Denne model simuleres ud fra det der kaldes Gillespie algoritmen, således at netværket opdateres instantant for alle smittebegivenheder. En samling af de væsentligste parametre er vist herunder (Tabel 2).



*Tabel 2: Parametre i modellen.*

<b>Parameter</b>	<b>Værdi interval for middelværdien</b>	<b>Reference</b>
Antal kontakter per dag	3-15	HOPE projektet
Latent tid (dage)	3-5	Litteratur se referenceliste i bilag 5
Infektios tid (dage)	4-8	Litteratur se referenceliste i bilag 5
Andel af kontakter i "andre" (%)	30-80	HOPE projektet
Typisk afstand mellem kontakter (km)	5-20	Trafik data
Andel afstandsuafhængige kontakter (%)	3-5	Trafik data
Tid fra symptom til test (Dage)	0-2	Fordeling fra spørgeskemaundersøgelse i foråret 2020 (ikke offentliggjort)
Sandsynlighed for at få symptomer og blive testet (%)	20-60 %	Prævalensundersøgelsen
Sandsynlighed for at kontakte husstand (%)	100%	Antagelse
Sandsynlighed for at kontakte kollegaer (%)	40-80	Antagelse
Sandsynlighed for at kontakte andre (%)	0-75	Antagelse



## Bilag 2. Beskrivelse af den agentbaserede model fra DTU

*Bidrag og udvikling: Freja Terp Petersen, Jacob Bahnsen Schmidt, Kasper Telkamp Nielsen, Rebekka Quistgaard-Leth, Kaare Græsbøll og Lasse Engbo Christiansen*

Den agentbaserede model fra DTU baseres på en befolkningstabel, hvor hver række i tabellen svarer til en agent - eller et individ – og hver kolonne indeholder data, der beskriver den pågældende agent, herunder aldersgrupper med 5 års-intervaller, bopælskommune, netværks-ID og forskellige smitteparametre.

I sygdomsmodellen bæres smitten fremad ved, at agenter der deler netværks-ID, f.eks. husholdnings-ID, skole/job-ID eller omgangskreds-ID, kan smitte hinanden. Hver dag får alle agenter udregnet deres sandsynlighed for at blive smittet på baggrund af antal infektiose i deres forskellige netværk og på baggrund af deres individuelle antal nære kontakter, som de er blevet tildelt baseret på en fordeling fra totalt antal kontakter inden for 1m i HOPE projektet.

Der er 7 forskellige netværkstyper, som en agent kan være en del af:

- Husholdning (alle agenter har en husholdning)
- Daginstitution (børn mellem 0 og 4 år)
- Grundskole (børn mellem 5 og 14 år)
- Ungdomsuddannelse (unge mellem 15-24 år samt voksne på erhvervsuddannelser)
- Arbejdsplads med kontorinddelinger (voksne op til 65 år)
- Omgangskreds (alle agenter har en omgangskreds)
- Kommune (alle agenter har en kommune)

Agenterne er blevet tildelt netværk baseret på data fra Danmarks Statistik (husholdninger og arbejdspladser), Undervisningsministeriet (grundskoler og ungdomsuddannelser) samt Institution.dk (daginstitutioner).<sup>3</sup> Det antages i modellen, at den gennemsnitlige kontorstørrelse og den gennemsnitlige omgangskreds uden for skole og arbejde er på 8 personer.

---

<sup>3</sup> FAM122N: <https://www.statistikbanken.dk/FAM122N>

FAM133N: <https://statistikbanken.dk/FAM133N>

FAM55N: <https://statistikbanken.dk/FAM55N>

PEND100: <https://www.statistikbanken.dk/PEND100>

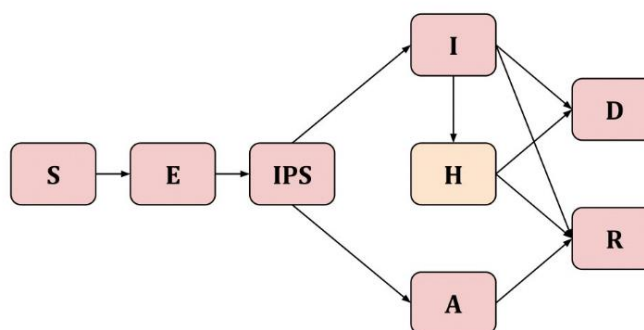
ERHV6: <https://www.statistikbanken.dk/ERHV6>

UVM (Normering grundskoler): <https://uddannelsesstatistik.dk/Pages/Reports/1577.aspx>

UVM (Normering gymnasier): <https://uddannelsesstatistik.dk/Pages/Reports/1851.aspx>

UVM (Normering erhvervsuddannelse): <https://uddannelsesstatistik.dk/Pages/Reports/1850.aspx>

Daginstitutioner: <https://www.institutioner.dk/>



Figur 6. Flowdynamisk diagram af bevægelse gennem sygdomsstadier.

Agenter i modellen kan være i et af følgende sygdomsstadier: Modtagelig (S), Eksponeret (E), Præ-symptomatisk (IPS), Symptomatisk (I), Asymptomatisk (A), Rask (R) eller Død (D). Agenter, som befinder sig i det præ-symptomatiske, symptomatiske eller asymptomatiske stadie, er infektiøse og kan således viderebringe smitte til agenter, som befinder sig i det modtagelige stadie. Bliver en modtagelig agent inficeret, overgår de til at være eksponeret. Dette sygdomsstadie repræsenterer den latente periode, hvor den inficerede agent endnu ikke er infektiøs. Agenterne kan bevæge sig gennem sygdomsstadierne, som vist på det flowdynamiske diagram, figur 6. Modellen antager, at 2/3 af agenterne bliver symptomatiske og at 1/3 forbliver asymptomatiske ved infektiøs tilstand. En andel symptomatiske agenter får et behandlingsbehov i løbet af deres sygdomsforløb og bliver indlagt på et Hospital (H). Sandsynligheden for indlæggelse blandt symptomatiske agenter er opdelt efter regioner og 10-års aldersgrupper baseret på data over indlæggelser i Danmark i september-oktober 2020.

Når en agent skifter til et nyt sygdomsstadie, tildeles de den ventetid, som de skal opholde sig i stadiet. Ventetiden i de forskellige stadier er beskrevet ved gamma-fordelinger med parametre, som vist i tabel 3. Modellen simuleres i diskret tid. Hvert tids-skridt svarer til en halv dag.

Tabel 3. Parametre og kvartiler for varighed af de enkelte stadier.

Stadier	Parametre		Kvartiler			Referencer
	Shape	Periode [Dage]	Nedre kvartil [Dage]	Median [Dage]	Øvre kvartil [Dage]	
Eksponeret (E)	3	3	2	3	4	Litteratur se referenceliste i bilag 5
Præsymptomatisk (IPS)	5	1,25	1	2	2	Litteratur se referenceliste i bilag 5
Symptomatisk (I)	4	7	5	7	9	Litteratur se referenceliste i bilag 5
Asymptomatisk (A)	4	7	5	7	9	Litteratur se referenceliste i bilag 5



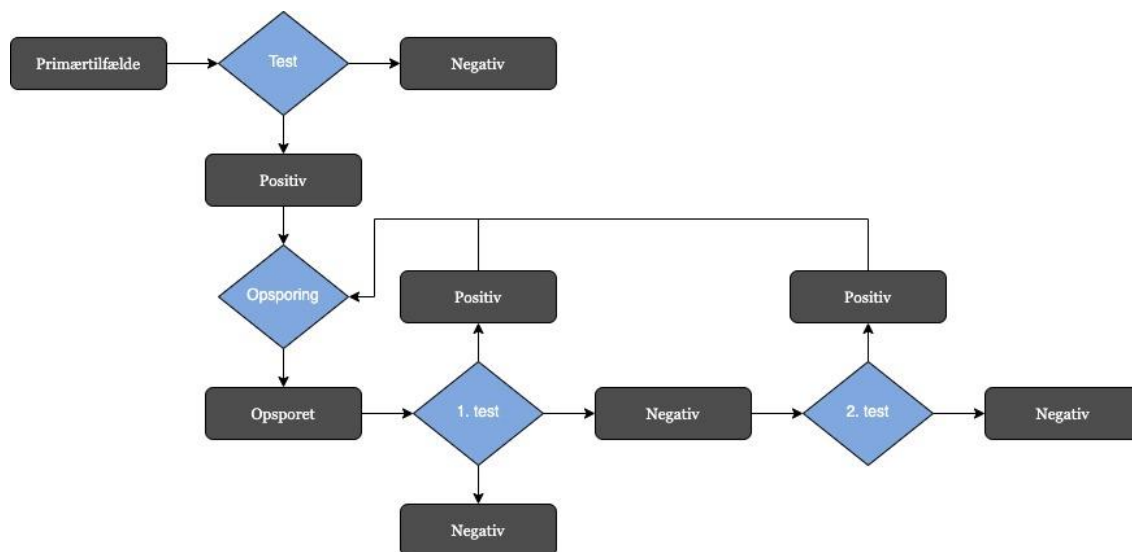
Hospitaliseret (H) Under 60 år	2	3	2	3	5	Linelisten SSI
Hospitaliseret (H) 60 år og der-over	2	5	3	5	7	Linelisten SSI
<b>Ventetider</b>						
timeSymp-ToOrderTest	5	1	1	1	1	Antagelser - af-venter STPS data
timeOrderTo-Test	2	2	1	2	3	Antagelser - af-venter STPS data
timeTestToRe-sult	6	1,5	2	2	2	Ventetider fra samfundssporet
traceDelay	5	1	1	2	3	Antagelser – af-venter STPS data

Sandsynligheden for, at en modtagelig agent bliver inficeret af en infektiøs agent og overgår til at være eksponeret i et givent netværk stiger med antallet af infektiøse agenter i netværket, de infektiøse agenter i netværkets smitsomhed, samt antallet af kontakter som både de modtagelige og infektiøse agenter har i netværket. Raten hvormed en modtagelig agent bliver inficeret er summen af smitterater fra de enkelte netværk, som agenten deltager i. Test og opsporing er indført i modellen ved følgende regler:

- Når en agent får symptomer, er der en sandsynlighed ( $p_{\text{TestGivenSymptoms}} = 80\%$ ) for, at de bestiller en test efter en gammafordelt ventetid ( $\text{timeSympToOrderTest}$ ). Hvis der er bestilt en test, vil personen reducere sine kontakter til 50% (undtagen i husholdninger, hvor kontakter reduceres til 70%).
- Der er en gammafordelt ventetid fra testen bestilles, til testen udføres ( $\text{timeOrderToTest}$ ).
- Der er en gammafordelt ventetid fra testen udføres, til der kommer svar ( $\text{timeTestToResult}$ ).
- Hvis der kommer positivt svar, vil agenten isolere sig yderligere; kontakter reduceres til 10% (husholdning: 50%). Derudover påbegyndes opsporing af netværk under følgende regler:
  - I skoleklasser, ungdomsskoleklasser, institutioner og i husholdninger opspores alle personer (i husholdninger foregår det dobbelt så hurtigt som i de øvrige netværk).
  - På kontorer (arbejdspladser) og i omgangskredse opspores et antal nære kontakter givet ved fordeling af kontakter under 1m i data fra HOPE projektet.
  - Personer, som tidligere er testet positiv, får ikke tildelt yderligere test.
  - Der opspores med en gammafordelt forsinkelse ( $\text{traceDelay}$ ) fra den positive test.
  - Ved opsporing efter en person testes positiv tildeles de opsporede personer test-tider relativt til 48 timer før den positive fik symptomer - eller blev testet i et asymptomatisk tilfælde. Hvis muligt, gives test på dag 4 og dag 6, ellers dag 5 og 7, og ellers én test hurtigt muligt.
  - Personer, som er i et igangværende opsporingsforløb, får kun tildelt test, hvis de venter på mindre end to testsvar.
  - Den opsporede person har samme ventetider på testsvar, som symptomatiske personer.



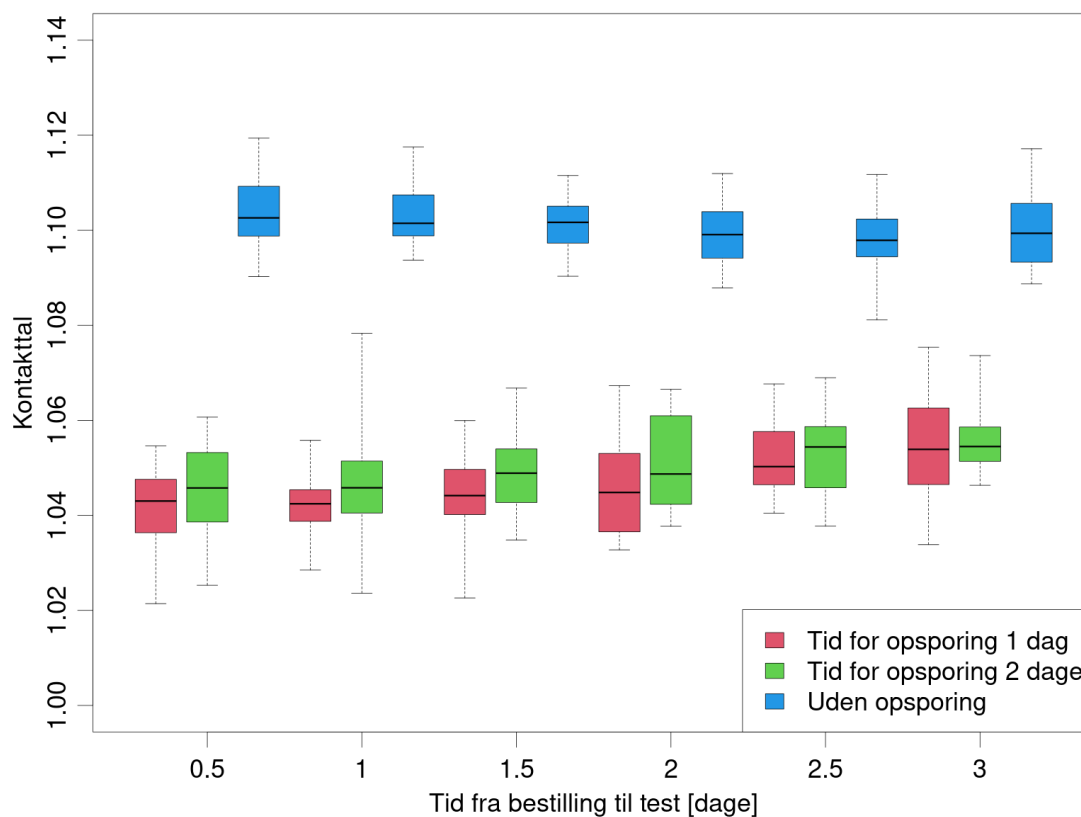
- Mens der ventes på test og testsvar, isoleres den opsporede person på samme måde som en symptomatisk, der venter på svar.
  - Hvis en opsporet person får negativt svar på den første test, vil der være en sandsynlighed for ( $p_{NoShow2ndTest} = 40\%$ ) at de ikke tager test nummer 2.
  - Efter et negativt svar på test nummer 1, vil isolationen brydes. Hvis der fås et positivt svar, inden test nummer 2 er taget, annulleres test nummer 2, og personens egne netværk opspores.
- For alle tests – om det er en opsporet person eller ej – antages der en sandsynlighed på 20% for en falsk negativ test (Kucirka et al., 2020).



Figur 7. Diagram, der viser test og opsporing i den agentbaserede model fra DTU.

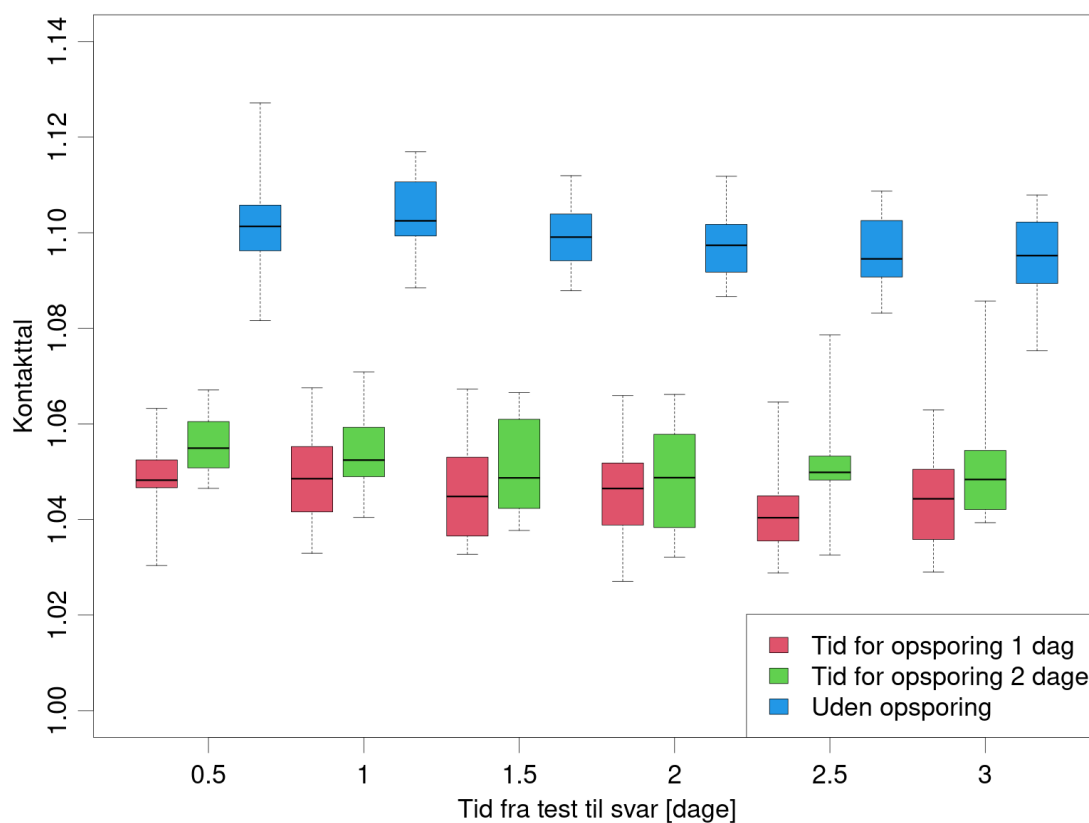


## Yderligere resultater



Figur 8. Kontakttallets afhængighed af ventetiden på at få taget en test hos primærtilfældet, samt betydningen af hvor lang tid der går før nære kontakter opspores og går i tilsvarende isolation. For hver parameter værdi er der foretaget 40 simulationer og boxplottet viser median, de indre kvartiler samt minimum og maksimum af disse.

Af figur 8 fremgår det, at der ikke er nævneværdig forskel på reduktionen i kontakttallet, hvorvidt man reducerer ventetiden til at primærtilfældet testes, i forhold til at reducere ventetiden til opsporing af nære kontakter.



Figur 9. Kontakttallets afhængighed af tiden fra der testes til at der foreligger et testsvar, samt betydningen af hvor lang tid der går indtil nære kontakter opspores og går i tilsvarende isolation. For hver parameter værdi er der foretaget 40 simulationer og boxplottet viser median, de indre kvartiler samt minimum og maksimum af disse.

Af figur 9 fremgår det, at der ikke er nævneværdig forskel på reduktionen i kontakttallet, hvorvidt man reducerer ventetiden fra at primærttilfældet og opsporede kontakter testes til der foreligger et testsvar, i forhold til at reducere ventetiden til opsporing af nære kontakter. En årsag kan være, at ventetiden til testsvar gør, at en masse opsporede og modtagelige kontakter er isoleret i længere tid og derfor ikke bliver smittet. Det er ikke undersøgt om dette kun ses når kontakttallet er nær 1.





## Bilag 3. Regneeksempel

Følgende er et illustrativt regneeksempel på den agentbaserede model fra Niels Bohr Institutet beskrevet i bilag 1. Udregningerne er baseret på modellens underliggende antagelser, nemlig at perioden for eksposition ( $E$  ( $T_E$ ), hvor den latente fase er en gammafordeling med middelværdi på 4.7 dage, og perioden for den smitsomme fase er en gammafordeling med middelværdi på 7 dage, samt en antagelse om, at 40% af cases findes uden kontaktopsporing. Det antages, at for de COVID-19 tilfælde der findes uafhængigt af kontaktopsporingen, er de smittede uniformt fordelt i den smitsomme periode ( $I$ ). Vi udregner nu tiden man er asymptomatisk men smitsom ved at trække tal fra fordelingen af tider for hele perioden, man er smitsom og tester en andel  $p$ , på et uniformt tilfældigt tidspunkt. Det giver en fordeling og en gennemsnitlig eksponeringstid (se figur 10A).

Vi kigger nu på et sekundært tilfælde, der blev smittet på et uniformt tilfældigt tidspunkt i den smitsomme periode for primærttilfældet. Denne person kan enten findes tilfældigt, eller ved at primærttilfældet testes, og at sekundærttilfældet opspores efter en tidsperiode ( $d$  for delay). Denne ventetid, er tiden fra at primærttilfældet testes til at sekundærttilfældet kontaktes, og afspejler således både ventetid til test samt ventetid til opsporing. Igen antages det, at sekundærttilfældet går i isolation øjeblikkeligt. Ved igen at trække tal tilfældigt fra de relevante fordelinger fås en eksponeringsperiode, hvori sekundærttilfældet måske opspores, forhåbentligt inden smitten er ført videre.

### Resultat

I figur 10B vises det gennemsnitlige antal dage en kontakt er eksponeret for smitte, som en funktion af den samlede ventetid til test og opsporing. Herudfra estimeres effekten af kontaktopsporing på det effektive kontakttal,  $R_t$ . Det antages, at en given andel ( $f_c$ ) af alle smittede tilfælde, findes via kontaktopsporing, og derved reduceres smitten, idet eksponeringsperioden for opsporede kontakter reduceres. Herved fås et simpelt estimat af effekten af kontaktopsporing på kontakttallet  $R_t$ . Dette vises i figur 10C. Farverne på graferne viser, hvor stor en andel af smitten der kan reduceres, såfremt eksponeringsperioden reduceres, som følge af kontaktopsporing. Hvis det f.eks. antages, at der er 2000 nye smittede med COVID-19 per dag (ca. 1000 fundne smittede + et mørketal), så svarer 0.05 grafen (orange) til at 100 smittede bliver fundet gennem kontaktopsporing dagligt.

En væsentlig begrænsning er, at disse udregninger ikke medtager effekten af, at flere COVID-19 tilfælde bliver fundet pga. kontaktopsporing, men er udelukkende baseret på effekten ved at for korte eksponeringsperioden for kontakter.

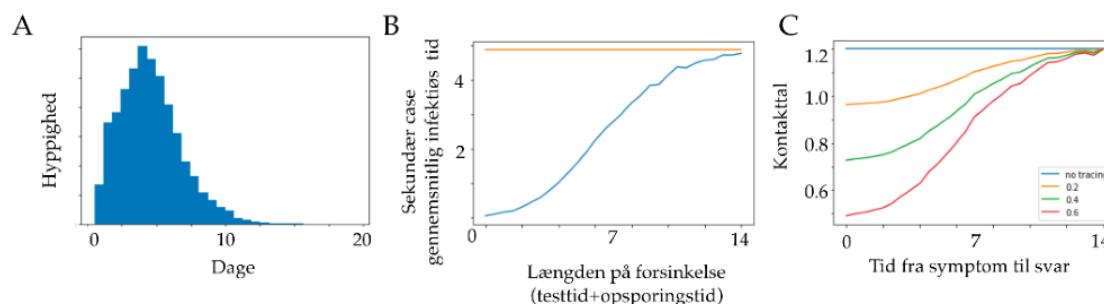
I modellen indgår 4 mulige eksponeringsperioder. 1: Kontakter opspores ikke, hvorved eksponeringsperioden ikke afkortes (blå graf), 2: 20% af kontakter opspores (gul graf), 3: 40% af kontakter opspores (grøn graf) og 4): hvis 80% af kontakter opspores (rød graf).

Af regneeksemplet fremgår det, at givet antagelserne i eksemplet vil kontakttallet kunne reduceres med ca. 50%, såfremt man opsporer 50% af alle kontakter inden for ca. 3 dage.

Bemærk at alle kurverne i figur 10C er meget flade i intervallet mellem dag 0 og 3. Dette betyder, at der kun opnås en lille gevinst ved at afkorte den samlede ventetid fra symptomer til der foreligger et test svar inden for denne periode, men at der til gengæld er en stor gevinst ved at øge andelen af opsporede kontakter.



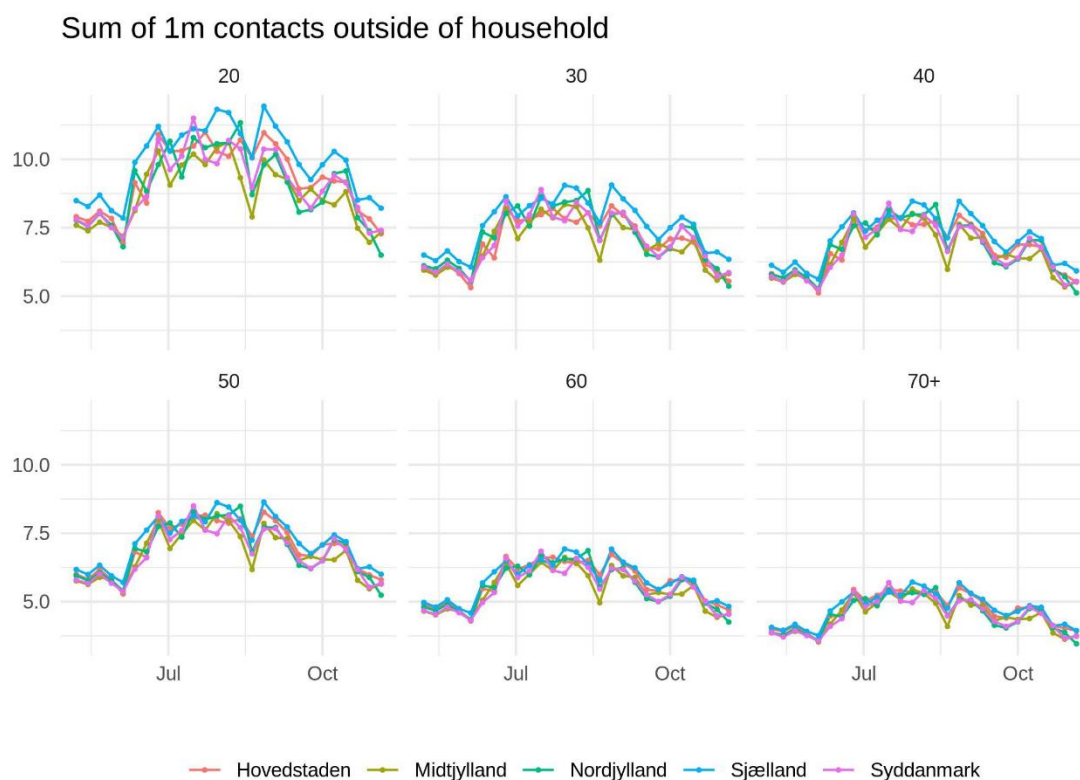
Det skal i øvrigt bemærkes, at det i eksemplet antages, at opsporede kontakter går i isolation, indtil de får svar på deres test.



Figur 10: A) Fordeling af eksponeringstiden, gennemsnit = 4.9 dage. B) Gennemsnitlig eksponeringstid for sekundære tilfælde (blå), som funktion af den samlede ventetid til test og opsporing. Den orange graf viser gennemsnittet i ventetiden til test og opsporing for primært tilfælden som reference. C) Det effektive kontakttal  $R_t$  efter kontaktopsporing som funktion af ventetiden fra symptomer til testsvar hvor udgangspunktet er et kontakttal på 1.2, inden der iværksættes opsporing. Farverne indikerer hvor stor en andel af kontakter der opspores, hvorved eksponeringstiden reduceres.



## Bilag 4. Udvikling i antal kontakter fra HOPE projektet



Figur 11. Kilde: Hope-projektet (12.11.2020). *Estimating Local Protective Behavior in Denmark with dynamic MRP.* [https://github.com/mariefly/HOPE/raw/master/HOPE\\_report\\_2020-11-12.pdf](https://github.com/mariefly/HOPE/raw/master/HOPE_report_2020-11-12.pdf)



## Bilag 5. Beskrivelse af parametre brugt i rapporten

Modellerne i rapporten bygger på en række parametre. Estimerne, som parametre er baseret på er udvalgt af den relevante institution, der har udarbejdet modellerne. Begrundelsen for valg af estimerne er beskrevet nærmere i dette bilag.

Overordnet set er parametre om sygdomsforløb primært baseret på international litteratur på emnet, men også på data fra den danske befolkning. Estimer over befolkningens adfærd i forbindelse med covid-19 bygger på en række danske undersøgelser fra i år, samt på data over danskernes rejsemønstre.

### **Estimer for latensperiode, inkubationsperiode og infektiøs periode fra litteraturen:**

Særligt relevant for simuleringerne over effekten af kontaktopsporing er estimerne bag sygdomsforløbet, herunder hvor lang tid der går fra eksponering til, at vedkommende kan smitte, og derefter til, at vedkommende vises symptomer. Estimerne i modellen er blandt andet baseret på andre forskeres data, som er offentliggjort i international litteratur om covid-19.

For at finde de bedste estimat på *latensperioden* har modelgruppen trianguleret distributioner fra nedenstående kilder. Estimatet er 3,6 dage med et interval på mellem 3-5 dage.

- Read et al. (2020). Novel coronavirus 2019-nCoV: Early estimation of epidemiological parameters and epidemic predictions. *Preprint*.
- Li et al. (2020). Early Transmission Dynamics in Wuhan, China, of Novel Coronavirus–Infected Pneumonia. *N. Engl. J. Med.*
- Li et al. (2020). Substantial undocumented infection facilitates the rapid dissemination of novel coronavirus (SARS-CoV2). *Science*.
- Milne and Xie (2020). The Effectiveness of Social Distancing in Mitigating COVID-19 Spread: a modelling analysis. *Preprint*.

For at finde det bedste estimat af *inkubationsperioden*, har Ekspertgruppen gennemgået nedenstående litteratur. Estimatet er 5 dage med et interval på mellem 3-7 dage.

- Lauer et al. (2020). The Incubation Period of Coronavirus Disease 2019 (COVID-19) From Publicly Reported Confirmed Cases: Estimation and Application. *Ann. Int. Med.*
- Li et al. (2020). Early Transmission Dynamics in Wuhan, China, of Novel Coronavirus–Infected Pneumonia. *N. Engl. J. Med.*
- Anderson et al. (2020). How will country-based mitigation measures influence the course of the COVID-19 epidemic. *The Lancet*.
- Linton et al. (2020). Incubation Period and Other Epidemiological Characteristics of 2019 Novel Coronavirus Infections with Right Truncation: A Statistical Analysis of Publicly Available Case Data. *J. Clin. Med.*
- Liu et al. (2020). Transmission dynamics of 2019 novel coronavirus (2019-nCoV). *bioRxiv*.
- Shen et al. (2020). Modelling the epidemic trend of the 2019 novel coronavirus outbreak in China. *bioRxiv*.



- Backer et al. (2020). Incubation period of 2019 novel coronavirus (2019-nCoV) infections among travellers from Wuhan, China, 20–28 January 2020. *Euro Surveill*.
- Gostic et al. (2020). Estimated effectiveness of symptom and risk screening to prevent the spread of COVID-19. *eLife*
- Hellewell et al. (2020). Feasibility of controlling COVID-19 outbreaks by isolation of cases and contacts. *The Lancet Global Health*.
- Milne and Xie (2020). The Effectiveness of Social Distancing in Mitigating COVID-19 Spread: a modelling analysis. *Preprint*.

For estimatet af *den infektiøse periode*, hvor det bedste estimat er 5 dage, mens det bedste interval er mellem 3-7 dage, har Ekspertgruppen gennemgået følgende artikler:

- Read et al. (2020). Novel coronavirus 2019-nCoV: Early estimation of epidemiological parameters and epidemic predictions. *Preprint*.
- Prem et al (2020). The effect of control strategies that reduce social mixing on outcomes of the COVID-19 epidemic in Wuhan, China. *Centre for the Mathematical Modelling of Infectious Diseases COVID-19 Working and Jit, Mark and Klepac, Petra, The Effect of Control Strategies that Reduce Social Mixing on Outcomes of the COVID-19 Epidemic in Wuhan, China*.
- Milne and Xie (2020). The Effectiveness of Social Distancing in Mitigating COVID-19 Spread: a modelling analysis. *Preprint*.

### HOPE rapporter og data:

En del af estimaterne i modellerne om befolkningens adfærd, herunder kontaktmønstre, bygger på både data og rapporter for Hope-projektet (<https://hope-project.dk/#/>).

HOPE-projektet udsender løbende spørgeskemaer til tilfældigt udvalgte personer i Danmark vedrørende både deres tillid til myndighederne, og til deres adfærdsmønstre, herunder hvor mange de ses med i forskellige kontaktkategorier, hvor meget afstand de holder fra andre mennesker etc. Denne information samles i rapporter, der løbende offentliggøres.

Udover HOPE-rapporten, der henvises til i Bilag 4 ([https://github.com/marie-fly/HOPE/raw/master/HOPE\\_report\\_2020-11-12.pdf](https://github.com/marie-fly/HOPE/raw/master/HOPE_report_2020-11-12.pdf)), oversender HOPE-projektet løbende anonymiserede data om befolkningens adfærd under covid-19 til Ekspertgruppen, der anvender det i deres modeller. Ekspertgruppen har også adgang til HOPE-projektets rapporter, der samenskriver data.

### Trafik data:

Antagelser om befolkningens adfærd bygges ligeledes på trafikdata, hvorudfra man kan bestemme danskernes rejsemønstre. Efter aftale med Trafik-, Bygge- og Boligstyrelsen får Ekspertgruppen løbende adgang anonymiserede data over danskernes bevægelse rundt i landet. Data er bl.a. brugt til at bestemme den typiske afstand mellem kontakter og afstanden mellem afstands-uafhængige kontakter. Data bygger på 5 forskellige kilder:

- Overblik over rejsende, der bruger rejsekort, som kommer fra Rejsekort og Rejseplanen A/S
- Overblik over biltrafik på Øresunds- og Storebæltsbroen fra Sund og Bælt A/S



- Overblik over flytrafik (antal passagerer) til og fra Københavns Lufthavn og Billund Lufthavn
- Overblik over biltrafikken på Statsvejsnettet og cykeltrafikken (samlet ud fra tællestationer) leveret af Vejdirektoratet.
- Overblik og færgetrafik på 5 rederier, der dækker over 17 færgeruter. Data er leveret af Danske Rederier.

### **Estimater for ventetider til test**

Estimater for ventetider til test og svar på test er taget fra TCDKs hjemmeside (<https://tcdk.ssi.dk/vente-og-svartider>).

### **Data fra SSIs Linelisten**

Linelisten på SSI indeholder informationer om de covid-19 podninger, der tages en given dag. Data fra Linelisten er bl.a. brugt til at modellere risikoen for at blive hospitaliseret i løbet af et covid-19-forløb for personer over og under 60 år.

### **Spørgeskemaundersøgelse blandt covid-19 syge lavet af SSI i foråret:**

I foråret 2020 foretog SSI en telefonisk spørgeskemaundersøgelse blandt en række personer, der fik konstateret covid-19. Spørgsmålene undersøgte deltagernes sygdomsforløb, herunder symptomer, hvorvidt nære kontakter i hustanden var smittet og lignende.

Data fra spørgeskemaundersøgelsen blev i modellerne brugt til at estimere tiden fra symptomdebut til tests i dage.

### **Den nationale prævalensundersøgelse for covid-19:**

SSI iværksatte i maj en undersøgelse af, hvor udbredt covid-19 var blandt danskerne. Undersøgelsen bestemmer seroprævalencen blandt et repræsentativt udsnit af danskerne fra maj og til i dag. Informationer fra prævalensundersøgelsen har været anvendt i modellerne til at estimere sandsynligheden for at få symptomer og blive testet.



## Bilag 6. Medlemmer af ekspertgruppen

Ekspertgruppen ledes af læge Camilla Holten Møller og overlæge Robert Leo Skov, 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, Afdelingen for Kognitive Systemer
- 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

### **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, Infektionsberedskabet
- Peter Michael Bager, Seniorforsker, ph.d., Infektionsberedskabet, Epidemiologisk Forskning, Statens Serum Institut
- Robert Skov, Overlæge, Infektionsberedskabet, Statens Serum Institut
- Camilla Holten Møller, Læge, PhD, Infektionsberedskabet, Statens Serum Institut





## **B**    *SSI Notat*

The following pages contain the report from Statens Serum Institut, the Danish CDC: *Ekspertgruppen for matematisk modellering, "Scenarier for udviklingen i den engelske virusvariant af SARS-COV-2 (cluster B.1.1.7)" (Statens Serum Institut, 2021).*

The report is from January 2 2021 and is a summary of the estimated spread of the "alpha" variant of COVID-19 (B.1.1.7) in Denmark. The report is in Danish and is based on two models, one from DTU and our agent based model from NBI.



d. 2. januar 2021

*Notatet er opdateret d. 22. januar 2021 med en præcisering af formuleringer vedrørende udviklingen i forholdet mellem Cluster B.1.1.7 og øvrige virusvarianter.*

### **Scenarier for udviklingen i den engelske virusvariant af SARS-COV-2 (cluster B.1.1.7)**

Ekspertgruppen for matematisk modellering, der ledes fra SSI, bringer i dette notat en række estimater for den forventede udbredelse af cluster B.1.1.7 i den kommende periode, dels ved logistisk regression af udviklingen i forekomsten af varianten, og dels ud fra simuleringer af spredningen af varianten i en agentbaseret model.

### **Sammenfatning**

- Den observerede udvikling i forekomsten af cluster B.1.1.7 i Danmark, svarer til en ugentlig vækstrate for forholdet mellem cluster B.1.1.7 og de øvrige virusvarianter på 72% (95% CI: [37, 115] %).
- Med udgangspunkt i den aktuelle situation hvor 2,3% af virusvarianterne i den rutinemæssige helgenomsekventering tilhører cluster B.1.1.7, estimeres det, at varianten vil udgøre halvdelen af de cirkulerende virusstammer i Danmark om 40-50 dage såfremt ovennævnte stigning fortsætter.
- Det nuværende niveau af restriktioner forventes ikke at være tilstrækkeligt til at få kontakttallet for cluster B.1.1.7 under 1. Derfor vil denne vokse eksponentielt upåagtet at det samlede kontakttal (for alle virusvarianter) kan være under 1 indtil cluster B.1.1.7 overtager om omkring en måned.
- Forekomsten af cluster B.1.1.7 er højest i Region Nordjylland, og udviklingen i forekomsten er ca. fire uger foran Region Hovedstaden.
- Det er på baggrund af engelske data estimeret at kontakttallet er ca. 1,5 gange højere for den nye virusvariant i forhold til andre virusvarianter.
- Den reduktion i smittetal og indlæggelser, der kan opnås i den kommende måned vil give et lavere udgangspunkt for den forøgede smitte og stigende kontakttal, som vi må forvente.

Disse beregninger er behæftet med usikkerheder af forskellige grunde. I perioden op til jul var der stor efterspørgsel på tryktest, og i samme periode er der udført et stigende antal antigen test. Derimod så vi i juledagene, at kun ganske få har ladet sig teste. Disse ændringer i testdynamikker gør det svært at følge udviklingen i covid-19, idet de vanlige indikatorer såsom incidenser, positivprocenter og kontakttallet påvirkes af den ændrede fordeling af covid-19-positive blandt de testede. Et lignende mønster forventes i dagene op til og efter nytår. Desuden har vi endnu ikke set effekten af de sidst indførte tiltag, herunder lukning af detailhandlen og liberale erhverv. Samlet set giver dette usikkerhed omkring det aktuelle kontakttal. Analysen er baseret på 76 isolater med cluster B.1.1.7 fordelt på de fem regioner. Den lille stikprøve giver relativt store statistiske usikkerheder. Der vil derfor være behov for at løbende at opdatere estimaterne og lave nye analyser.



### Logistisk regression for spredningen af cluster B.1.1.7

Som det fremgår af nedenstående tabel, er der stor forskel på, hvornår man har fundet cluster B.1.1.7 i de enkelte regioner.

*Tabel 1. Forekomst af cluster B.1.1.7 i de fem regioner baseret på helgenomsekventering af stikprøver af SARS-CoV-2 positive isolater.*

Uge	Hovedstaden		Midtjylland		Nordjylland		Sjælland		Syddanmark	
	B.1.1.7	Total	B.1.1.7	Total	B.1.1.7	Total	B.1.1.7	Total	B.1.1.7	Total
45	0	656	0	283	0	238	0	181	0	200
46	4	420	0	327	0	305	0	132	0	168
47	0	588	0	297	0	240	0	143	0	241
48	3	679	0	291	0	169	0	165	0	195
49	0	825	0	332	3	64	0	246	0	208
50	2	892	0	360	7	92	0	214	1	431
51	3	753	0	524	9	254	3	310	4	354
52	8	774	5	221	12	169	10	193	1	225

Ud fra udbredelsen af cluster B.1.1.7 i Danmark samt andelen af nye isolater i overvågningen som er relateret til clusteret, anvendes logistisk regression til at estimere den forventede udbredelse af cluster B.1.1.7. Da fokus er på spredningen af virusvarianten, og ikke på introduktioner af denne, er det kun regioner, hvor der er detekteret isolater tilhørende cluster B.1.1.7 i mindst fire uger – dvs. Region Hovedstaden og Region Nordjylland, der er medtaget i denne første analyse.

Der er lavet logistisk regression med uge og region som forklarende variable. Der er også testet en interaktion, men den er ikke signifikant.

*Tabel 2. Estimer for logistisk regression af andelen af cluster B.1.1.7. Referencen repræsenterer Region Hovedstaden.*

	Estimate	Std. Error	z value	Pr(> z )
(Intercept)	-32.812	5.679	-5.778	0.000
Uge	0.540	0.112	4.844	0.000
Region Nordjylland	2.221	0.311	7.133	0.000



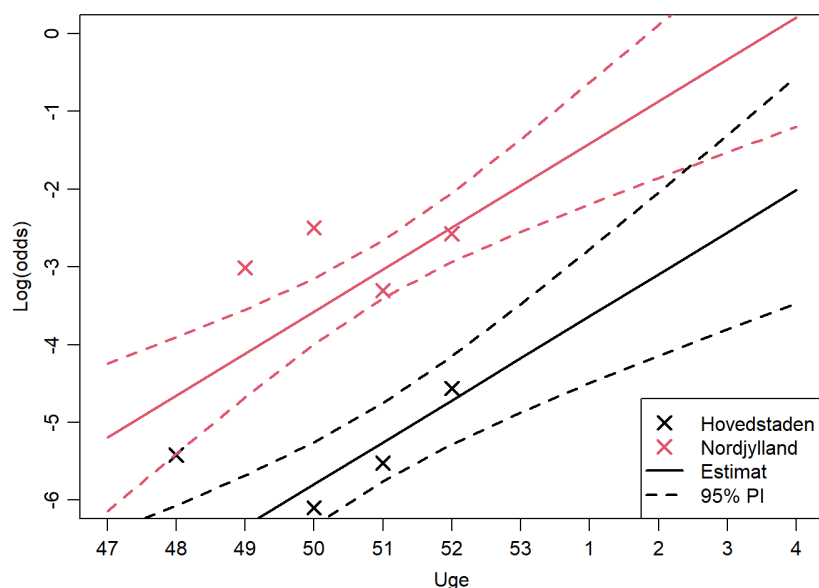
Det ses, at  $\log(\text{odds})$  for at detektere cluster B.1.1.7 er 2.2 højere i Region Nordjylland end i Region Hovedstaden. Det svarer til odds på 9.2. Det mest interessante er den tidlige udvikling, hvoraf det ses at  $\log(\text{odds})$  øges med 0.54 for hver uge. Dette svarer til at cluster B.1.1.7 har en ugentlig vækstrate i odds (forholdet mellem antal cluster B.1.1.7 og øvrige virusvarianter) på 72% (95% CI: [37, 115] %), hvilket med den nuværende lave andel af cluster B.1.1.7 svarer til den samme stigning i andelen af cluster B.1.1.7 blandt alle positive prøver. Usikkerheden på estimatet er endnu ganske stort og estimatet er følsomt over for hvilke uger der medtages. Uanset usikkerheder, svarer det fundne estimat til de der er rapporteret fra England for denne virusvariant og det tyder på, at cluster B.1.1.7 har samme forøgede transmissionsrate i Danmark som i England.

Det ses, at  $\log(\text{odds})$  for at detektere cluster B.1.1.7 er 2.2 højere i Region Nordjylland end i Region Hovedstaden. Det svarer til odds på 9,2, dvs. at sandsynligheden for at detektere cluster B.1.1.7 her er 9,2 gange højere. Det svarer også til at Region Nordjylland er fire uger foran Region Hovedstaden i andelen af cluster B.1.1.7

Det forventes, at usikkerhederne vil blive reduceret væsentligt når der er data for 1-2 uger mere. Men givet at B.1.1.7 er så meget mere smitsom end hidtidige varianter vil det kræve længerevarende restriktioner at sænke smittetallet.

De seneste estimater af kontakttallet er lige under 1,0. Dette er dog påvirket af den ændrede testaktivitet og adfærd hen over jul og nytår, og vi har endnu ikke et overblik over konsekvenserne af sammenkomster i forbindelse med jul og nytår. Endvidere har vi endnu ikke set effekten af nedlukningen af de liberale erhverv og detailhandlen omkring jul. Derfor er det forventningen, at en fastholdelse af de nuværende restriktioner vil give et fald i kontakttallet, hvis man kigger på de virusvarianter som vi har set før introduktionen af cluster B.1.1.7. I England har man estimeret, at deres reference kontakttal var 0,8 for andre virusvarianter og 1,2 for cluster B.1.1.7. Det observerede kontakttal er et vægtet gennemsnit af virusvarianterne i populationen.

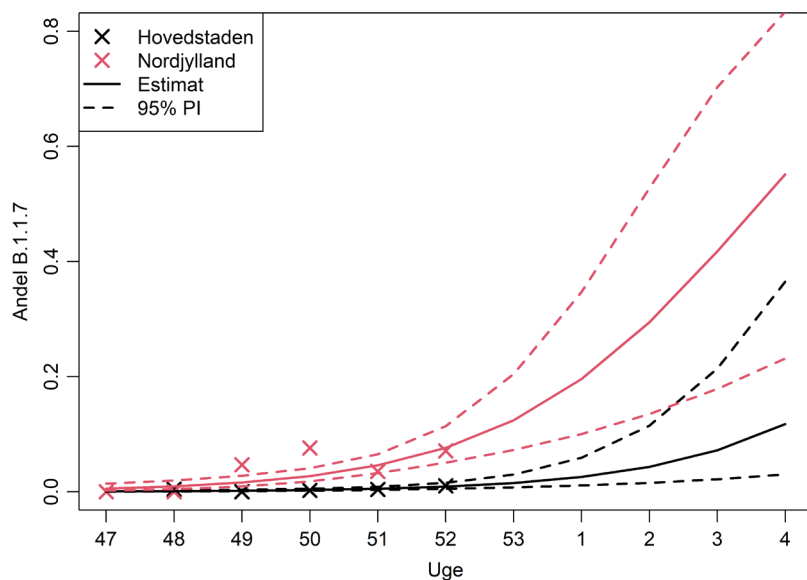
Figur 1 viser en fremskrivning af  $\log(\text{odds})$  for B.1.1.7 mod andre virusvarianter baseret på ovenstående logistiske regression. Estimatet er, at cluster B.1.1.7 allerede i uge 4 vil udgøre halvdelen af alle positive test i Region Nordjylland. Dette er dog behæftet med stor usikkerhed på baggrund af de nuværende data.



Figur 1. Log(odds) for at detektere cluster B.1.1.7 i hhv. Region Hovedstaden og Region Nordjylland

Ved sammenligning med England er vi nu, hvor de var i starten af november, hvor South East havde log(odds) på -2 svarende til Nordjylland og både London og East of England havde log(odds) omkring -4 svarende til Hovedstaden<sup>1</sup>

Figur 2 viser den samme fremskrivning som i figur 1. Blot er der transformeret tilbage til andelen af positive test, som tilhører cluster B.1.1.7.



<sup>1</sup> 2020\_12\_23\_Transmissibility\_and\_severity\_of\_VOC\_202012\_01\_in\_England.pdf (cmmid.github.io)



*Figur 2. Udviklingen i forekomsten af cluster B.1.1.7 i de kommende uger. Fremskrivningen viser, at halvdelen af isolaterne i Region Nordjylland vil være cluster B.1.1.7 omkring uge 4.*

Det skal bemærkes, at udviklingen i Hovedstaden er ca. 4 uger efter udviklingen i Nordjylland. Det er endnu for tidligt at udtale sig om niveauet i de andre tre regioner, men særlig Region Sjælland synes at have oplevet en hurtig stigning, om end det er baseret på meget lidt data. De næste par uger vil forbedre estimatet af niveauet i alle regioner.

Hen over julen har der været et nyt toppunkt i antal indlagte og der er endnu kun set små fald. Det er først i uge 1, at vi kan forvente at se eventuelle indlæggelser som følge af smitte i julen. Alt andet lige må dette forventes at give en yderligere kortvarig pukkel i antal nye indlæggelser.

På nuværende tidspunkt er prognosen, at vi har omkring en måned før det samlede kontakttal for alle virusvarianter hurtigt vil stige på grund af øget udbredelse af cluster B.1.1.7. Hvis restriktionerne skærpes i den kommende tid, vil det give en reduktion i smittetal og indlæggelser og dermed et lavere udgangspunkt for den forøgede smitte og stigende kontakttal, som vi må forvente.

Et første estimat af kontakttallet for cluster B.1.1.7 for perioden uge 47 til 52 og baseret på observationer fra Region Hovedstaden og Region Nordjylland er 1.5 (95% CI [1,2 ; 1,7]) - dette er estimeret vha. Poisson regression med offset lig med  $0.7 \cdot \log(\text{antal sekventerede})$ . Det gennemsnitlige kontakttal (baseret på SSIs publicerede kontakttal 2020-12-29) for perioden er 1,1. Da kontakttallet for cluster B.1.1.7 er så meget højere må det selv med de nuværende restriktioner forventes, at det vedbliver med at være over 1 og dermed forventes cluster B.1.1.7 at vokse eksponentielt, hvis det nuværende niveau af restriktioner fastholdes.

### **Simulering af spredningen af cluster B.1.1.7 i en agentbaseret model**

#### *Agentbaserede modeller*

Spredningen af cluster B.1.1.7 er simuleret i en agentbaseret model, som er udviklet af Niels Bohr Institut, Københavns Universitet (NBI). En agentbaseret model simulerer et antal agenter (individer i en population) og deres interaktioner med andre agenter, svarende til de interaktioner som en befolkning normal vis har. Hver agent repræsenterer således en person, som er knyttet til en lokation i Danmark, svarende til deres bopæl. Agenterne indgår i flere forskellige netværk, f.eks. husstand, job og skole hvor de har kontakt til andre personer. Derudover har de kontakt til tilfældige personer i samfundet i den tid, hvor personen ikke er hjemme, på job eller i skole. Hvis en agent bliver smittet med SARS-CoV-2, er forløbet for den enkelte agent beskrevet således, at agenten først er eksponeret (E) og derefter infektiøs (I), hvorefter agenten ikke længere er smitsom og betragtes som rask (R). De gennemsnitlige tider i hvert sygdomsstadie kan findes i bilag 1. Hver kontakt, som en agent eksponeres for, tildeles en sandsynlighed for at blive smittet af en anden agent, såfremt denne er smitsom. For en detaljeret beskrivelse af den agentbaserede model, herunder de inkluderede parametre, henvises til bilag 1.

#### *Forbehold*



Mens en agentbaseret model kan medtage mere detaljerede dynamikker i en epidemi, så kræver en præcis simulation input fra data, som ofte ikke er tilgængelige eller forefindes, fx hvem en person mødes med i løbet af en dag. Derfor kan en sådan model have unøjagtigheder eller bygge på antagelser, som ikke er retvisende. Det er ikke muligt at kvantificere den nøjagtige størrelse eller effekt af disse potentielle fejlkilder. Da datagrundlaget for disse simuleringer er sparsomt, fordi vi endnu har få datapunkter for cluster B.1.1.7, vil resultatet være behæftet med væsentlig usikkerhed.

### *Resultater*

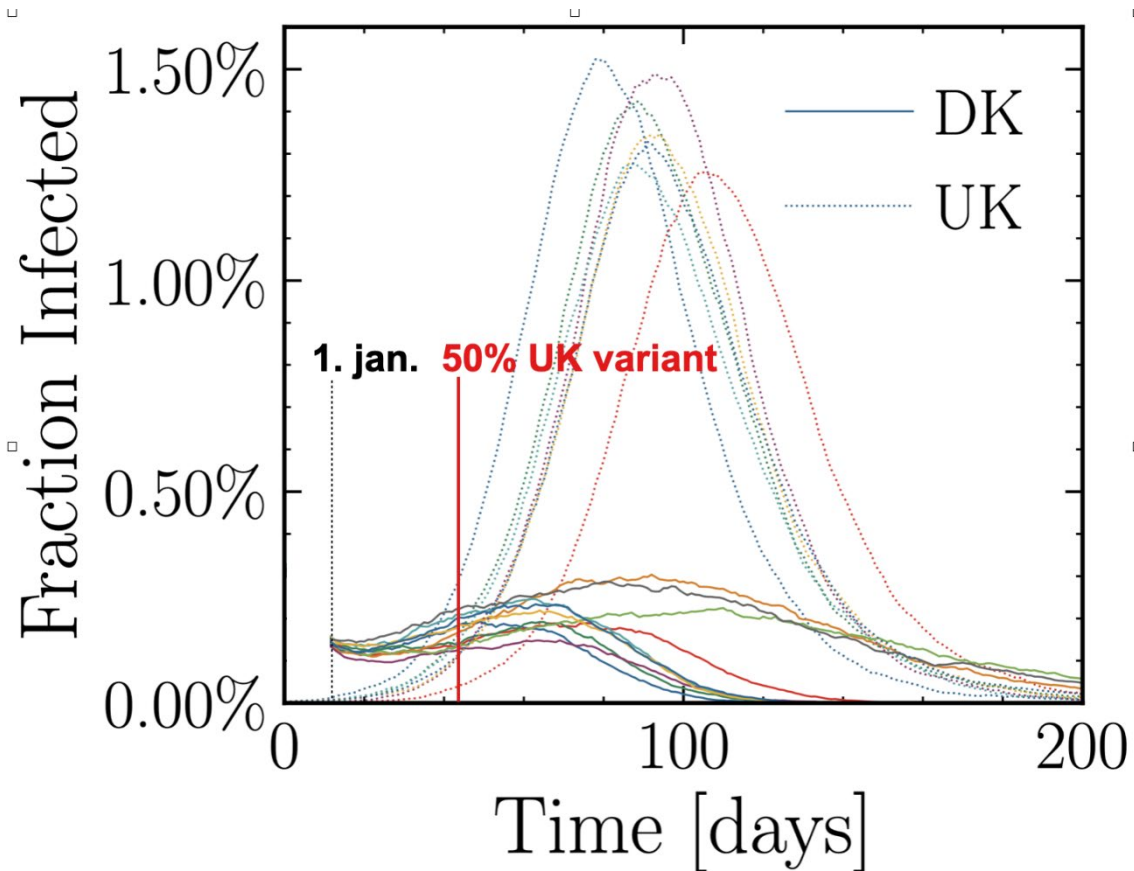
I det følgende er udviklingen simuleret i en model, hvor udgangspunktet er 1/10 af Danmarks befolkning, og hvor cluster B.1.1.7 fra starten udgør omkring 5% af de cirkulerende virusvarianter. Epidemien simuleres ud fra et kontakttal på omkring 1,0, samt en antagelse om, at cluster B.1.1.7 smitter 50% mere, som rapporteret fra England<sup>2</sup>

Figur 3 viser, hvordan en epidemi vil udvikle sig i tid, forudsat at det simulerede scenarie ikke ændres. Der opdeles i hhv. de nuværende virusvarianter (DK, fulde linjer) og det engelske cluster B.1.1.7 (UK, stiplede linjer). Simulationen er gentaget flere gange (forskellige farver) for at se, hvor store variationer der forekommer. Som det kan ses, så udfases DK-versionen af smitten, mens UK-versionen B.1.1.7 giver ophav til en eksponentiel vækst, idet kontakttallet for denne er væsentligt over 1.

Af figuren fremgår det, at cluster B.1.1.7 ca. 35-40 dage fra simulationens start ("1. jan.") udgør omkring 50% af de cirkulerende virusvarianter. Da simulationen er startet med en større andel UK-varianter (5%) end det aktuelle landsgennemsnit (2.3%), så bliver estimeret 40-50 dage til at halvdelen af de sekventerede varianter tilhører cluster B.1.1.7. I de viste simulationer er de første smittede med cluster B.1.1.7 varianten placeret i Hovedstadsområdet. I andre scenarier, hvor cluster B.1.1.7 varianten i starten udvikler sig i et tyndere befolket område tager udviklingen lidt længere tid, op til 60 dage.

---

<sup>2</sup>2020\_12\_23\_Transmissibility\_and\_severity\_of\_VOC\_202012\_01\_in\_England.pdf  
(cmmid.github.io)



Figur 3. Den forventede udvikling i cluster B.1.1.7 sammenholdt med udviklingen i øvrige virusvarianter, simuleret i en agentbaseret model. Ud fra simulationerne estimeres det, at B.1.1.7 varianten vil være dominerende efter 40-50 dage.



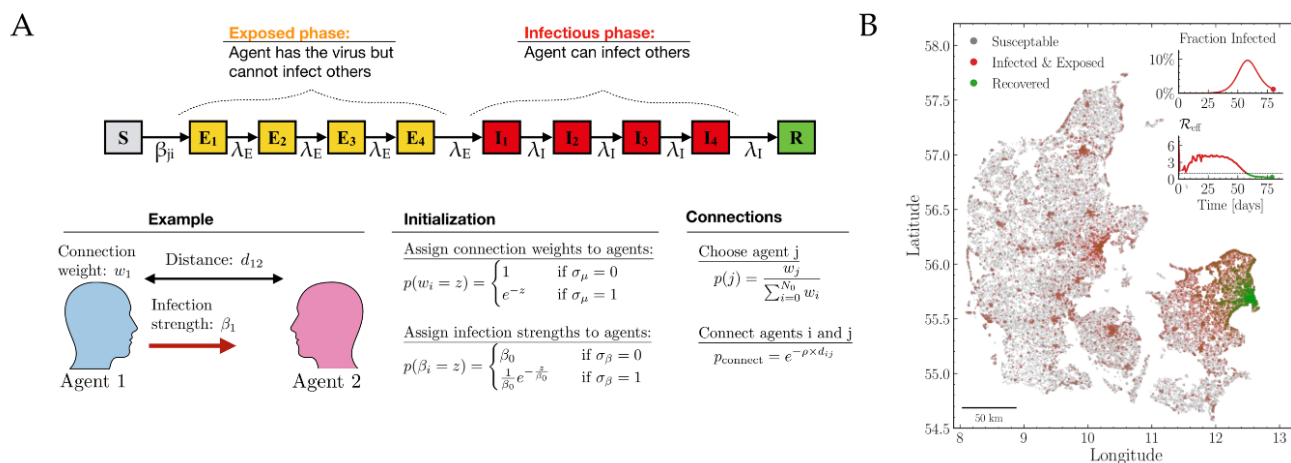


## Bilag 1. Beskrivelse af den agentbaserede model

Den nedenstående modelbeskrivelse er et uddrag fra ekspertrapporten “effekten af kontaktopsporing” der er publiceret d. 16. december 2020

*Bidrag og udvikling: Christian Michelsen, Emil Martiny, Tariq Halasa, Mogens H. Jensen, Troels C. Petersen og Mathias L. Heltberg*

Den agentbaserede model baseres på agenter, dvs. individer hvis karakteristika er tildelt ud fra statistiske fordelinger i befolkningen. Dette er f.eks. en aldersfordeling og en fordeling over pendlerafstande. Modellen starter med at fordele Danmarks bopæle ud i landet baseret på det danske hussalg over de sidste 15 år. Herefter placeres agenter i hver husstand baseret på deres alder og geografiske placering.



Figur 5: A) Skematisk oversigt over hvordan interaktionsnetværket i modellen ser ud. B) Eksempel på simulation af smittespredning i Danmark i modellen, gennem et simuleret tilfælde af flokimmunitet i København.

Et afgørende element i modellen er opbygningen af alle personers interaktionsnetværk. Dette genereres ved, at hver agent har et netværk, de interagerer med. Dette opdeles i tre dele: 1) kontakter i hjemmet, 2) kontakter på arbejdet, 3) kontakter i kategorien andre kontakter. Der er ikke nogen geografisk afhængighed af antallet af kontakter på arbejdet, men i den kategori der kaldes “andre”, vil der generelt være flere kontakter for dem der bor i tæt befolkede områder i forhold til dem der bor på landet. Måden hvorpå netværket dannes er vist i Figur 5A.



Ud fra data fra HOPE-projektet har vi estimeret, hvor mange personer hver agent vil interagere med, og i denne model vil alle agenter have mellem 3 og 15 daglige kontakter.

Når modellen simuleres vil alle inficerede agenter gennemgå et forløb, hvor de er i en latent periode, hvor de ikke smitter, hvorefter de vil rykke over i en infektiøs periode, hvor de kan smitte agenter i deres netværk. Denne model simuleres ud fra det der kaldes Gillespie algoritmen, således at netværket opdateres instantant for alle smittebegivenheder. En samling af de væsentligste parametre er vist herunder (Tabel 2).

*Tabel 2: Parametre i den agentbaserede model*

Parameter	Værdi interval for middelværdien	Reference
Antal kontakter per dag	3-15	HOPE projektet
Latent tid (dage)	3-5	Litteratur se referenceliste i bilag 5
Infektiøs tid (dage)	4-8	Litteratur se referenceliste i bilag 5
Andel af kontakter i "andre" (%)	30-80	HOPE projektet
Typisk afstand mellem kontakter (km)	5-20	Trafik data
Andel afstandsuaafhængige kontakter (%)	3-5	Trafik data
Tid fra symptom til test (Dage)	0-2	Fordeling fra spørgeskemaundersøgelse i foråret 2020 (ikke offentliggjort)
Sandsynlighed for at få symptomer og blive testet (%)	20-60 %	Prævalensundersøgelsen
Sandsynlighed for at kontakte husstand (%)	100%	Antagelse
Sandsynlighed for at kontakte kollegaer (%)	40-80	Antagelse
Sandsynlighed for at kontakte andre (%)	0-75	Antagelse

This document was typeset using  $\LaTeX$  and the `tufte-style-thesis` class.  
The style is heavily inspired by the works of Edward R. Tufte and Robert Bringhurst.  
This is available on here:

<https://github.com/sylvain-kern/tufte-style-thesis/>.

Feel free to contribute!