



Teknisk gennemgang af modellerne

Version 1.0 - pr. 10. juni 2020

Statens Serum Institut (SSI) har nedsat og leder en ekspertgruppe, der har til formål at udvikle matematiske modeller til at belyse udviklingen i COVID-19 i Danmark.

Modellerne skal bl.a. belyse udviklingen i smittespredningen i Danmark, belastningen af kritiske sygehusfunktioner som almene og intensive sengepladser samt effekten af nuværende og mulige myndighedstiltag, herunder effekten ved at ophæve disse.

Konkret har ekspertgruppen til opgave at udvikle matematiske modeller til at belyse:

1. Udviklingen i COVID-19 i Danmark
2. Behovet for kritiske sygehusfunktioner, herunder intensivpladser, respirator kapacitet og ECMO-behandling
3. Effekten af nuværende og mulige myndighedstiltag, herunder effekten ved at ophæve disse

I nærværende notat er der lavet en teknisk gennemgang med henblik på at give et indblik i, hvilke modeller samt parametre ekspertgruppen har benyttet sig af. Gennemgangen opdateres løbende i takt med arbejdets fremgang, og er tilgængelig på SSI's hjemmeside.

Det bemærkes, at det er en grundlæggende præmis for matematiske modeller og simuleringer, at de altid vil være forsimplede repræsentationer af virkeligheden ud fra antagelser, som er baseret på tilgængelige data og forskningsbaseret viden. Der er ingen modeller, der vil vise fremtiden præcist, som den udvikler sig, men gode modeller kan give en struktureret ramme og dermed støtte for beslutningstagere.

Ekspertgruppens afrapporteringer kan findes på SSI's hjemmeside¹.

Oversigt over modeller

Kontaktmatricer	2
Analyse af indlæggelsesforløb ved hjælp af multi-state model	7
Populationsmodel	11
Regionsopdelt populationsmodel	15

¹ <https://www.ssi.dk/aktuelt/sygdomsudbrud/coronavirus>



Kontaktmatricer

Formål

Mange sygdomme, inkl. COVID-19, spredes ved kontakt mellem mennesker, og derfor betyder antal og fordeling af kontakter mellem aldersgrupper meget for spredning af disse sygdomme. Der er foretaget flere studier, der beskriver personers kontaktmønstre. Kontaktmønstrene beskrives ofte i kontaktmatricer som viser kontaktmønstre mellem forskellige befolkningsgrupper fx i forhold til alder. I forbindelse med modellering af COVID-19 er kontaktmatricer brugt til at beskrive, hvordan ændringer i befolkningens adfærd og dermed kontaktmønstre kan påvirke smittespredningen af COVID-19 i forbindelse med forskellige myndighedstiltag i den faseopdelte genåbning af Danmark.

Kontaktmatricerne bliver brugt som input i andre modeller. Endvidere kan man gennem undersøgelse af relative ændringer i den maksimale egenverdi for hver matrice få et simpelt mål for effekten af forskellige tiltag. Da mange aktiviteter i samfundet afhænger af de involveredes alder er kontaktmatricerne aldersopdelte.

Teori

Udgangspunktet for de aldersopdelte kontaktmatricer er, at personer i forskellige aldersgrupper møder personer fra egen samt andre aldersgrupper med forskellig hyppighed. Endvidere afhænger hyppighederne af, hvor man mødes fx i hjemmet, på arbejde, i skole eller andetsteds.

Kontaktmatricerne baseres på studier, hvor personer i forskellige aldersgrupper har registreret deres kontakter over en periode, typisk en dag. Kontakterne er grupperet efter egen alder, alder for den man mødes med samt i hvilken kontekst man mødes. Smittetrykket som følge af kontakt kan da beregnes mellem aldersgrupper ved at gange med en risiko per kontakt. I disse modelleringer er kontakter beregnet i 5 års aldersgrupper.

Effekten af myndighedstiltag og adfærdsændringer er implementeret ved at skalere kontaktmatricerne elementvis (dvs. for hvert myndighedstiltag og for hver aldersgruppe). Skaleringen er lavet så en fuld lukning svarer til 0 og en fuld åbning svarer til 1. For hvert myndighedstiltag er det estimeret, hvor meget de enkelte typer af kontakter påvirkes. Fx svarer åbning af 3. klasse til, at én årgang blandt 5-9-årige begynder i skole, svarende til at 20% af skolekontakterne, aktiveres for aldersgruppen. Dertil kommer lærernes arbejdskontakter. Se kode og ekspertgruppens rapporter for flere detaljer.

Data

Til brug for at danne kontaktmatricerne er primært benyttet to studier: BBC Pandemic-studiet med selvrapporterede kontakter i løbet af en dag fra ca. 40.000 englændere over 13 år i 2018 (Klepac 2020), og POLYMOD-studiet med ca. 1.000 englændere tilbage i 2008 (Mossong 2008). BBC Pandemic-studiet er brugt for kontakter for 10-årige og opefter (kontakter for 13-14-årige er overført til 10-12-årige) og POLYMOD-studiet er brugt for kontakter for 0-9-årige. Forskellige studier kan have forskellige rapporteringsbias og derfor er POLYMOD data blevet skaleret til



samme gennemsnitlige antal kontakter blandt 10-14 og 15-19-årige under ét før de er indsat for de 0-9-årige.

Parametre

De enkelte myndighedstiltag har parametre for, hvor mange procentpoint de sænker/øger antallet af kontakter i hjemmet, skoler, arbejde og øvrige steder. Disse parametre er i videst muligt omfang baseret på oplysninger fra de respektive ressortministerier, Danmarks Statistik og diverse brancheorganisationer. De enkelte parameterværdier fremgår af de respektive ekspertrapporter.

I BBC Pandemic-studiet har deltagerne af rapporteret hhv. samtalekontakter og fysiske kontakter. I modellen er antaget, at normal samtaleafstand har fem gange lavere risiko for at overføre smitte end ved fysisk kontakt (en samtale er i studiet defineret som værende ansigt til ansigt med tre eller flere ord) (Klepac 2018).

Fysisk afstand

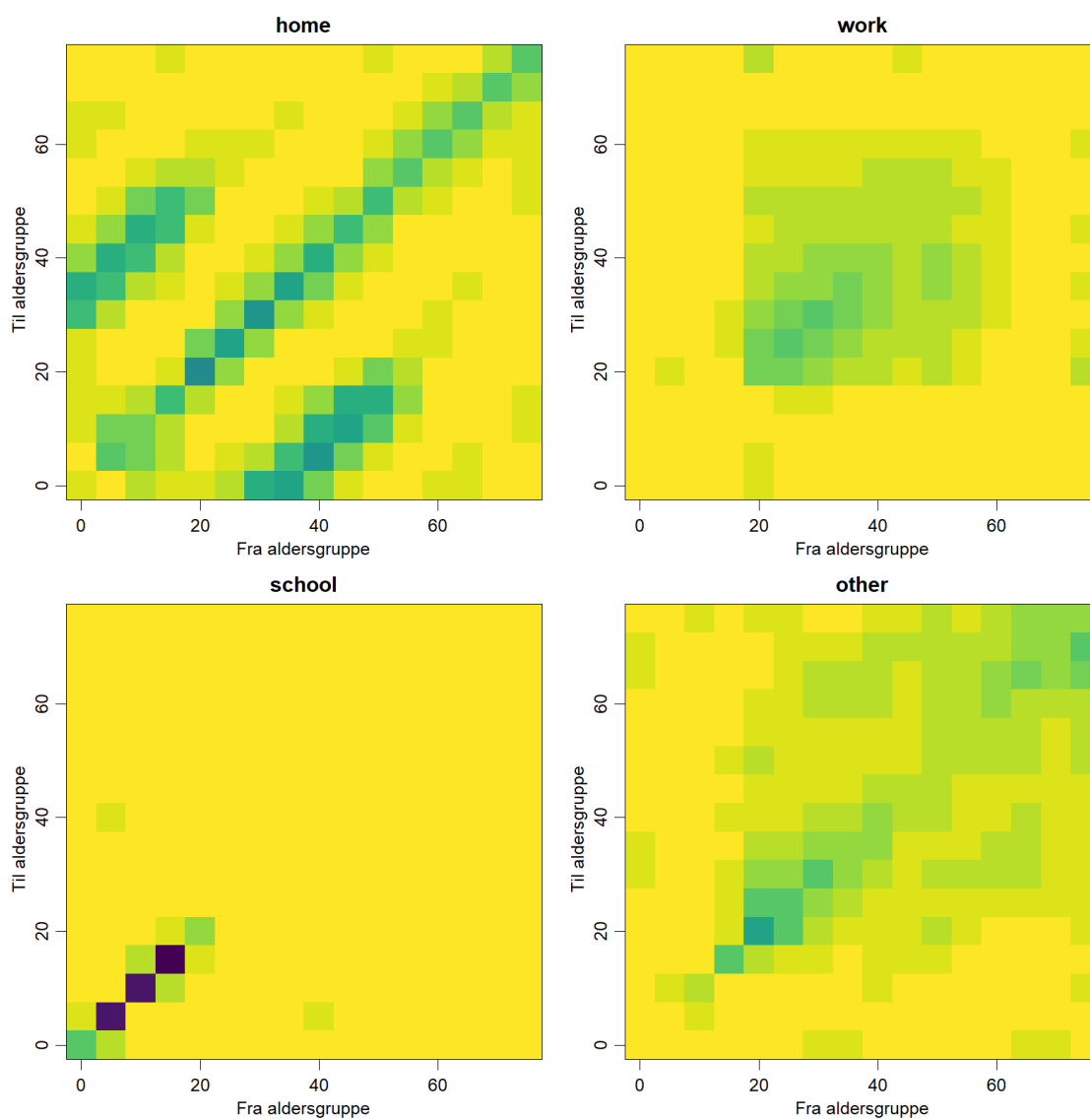
Efter den første nedlukning i fase nul beregnede modelgruppen nedgangen i aktivitet i samfundet baseret på trafikmålinger og nedlukninger af daginstitutioner, skoler, dele af arbejdsmarkedet samt fritidsaktiviteter. Fysisk afstand giver en yderligere reduktion i antal kontakter, men denne er ukendt. Der var international konsensus om daglige vækstrater på 20-24% i den epidemiske fase svarende til R_0 på omkring 2,5 (i en simpel SIR-model). De første estimater af effekten af nedlukningen svarerede til et R_e på 1,3 og effekten af fysisk afstand blev justeret til at opnå dette.

Eksempler

På figur 1 ses fordelingen af kontakter på de fire kontakttypen: hjem, skole, arbejde og øvrige steder. Hver kolonne er normeret med andelen af befolkningen i den pågældende aldersgruppe således, at hvert element angiver antal kontakter til personer i den tilsvarende aldersgruppe. Det ses, at skolebørn har mange indbyrdes kontakter samt, at der i hjemmet er mange kontakter både indenfor egen aldersgruppe og mellem børn og forældre. Figur 2 viser den estimerede effekt af nedlukningen og de første to faser af genåbningen sammenlagt for hjem, skoler, arbejde og øvrige kontakter.

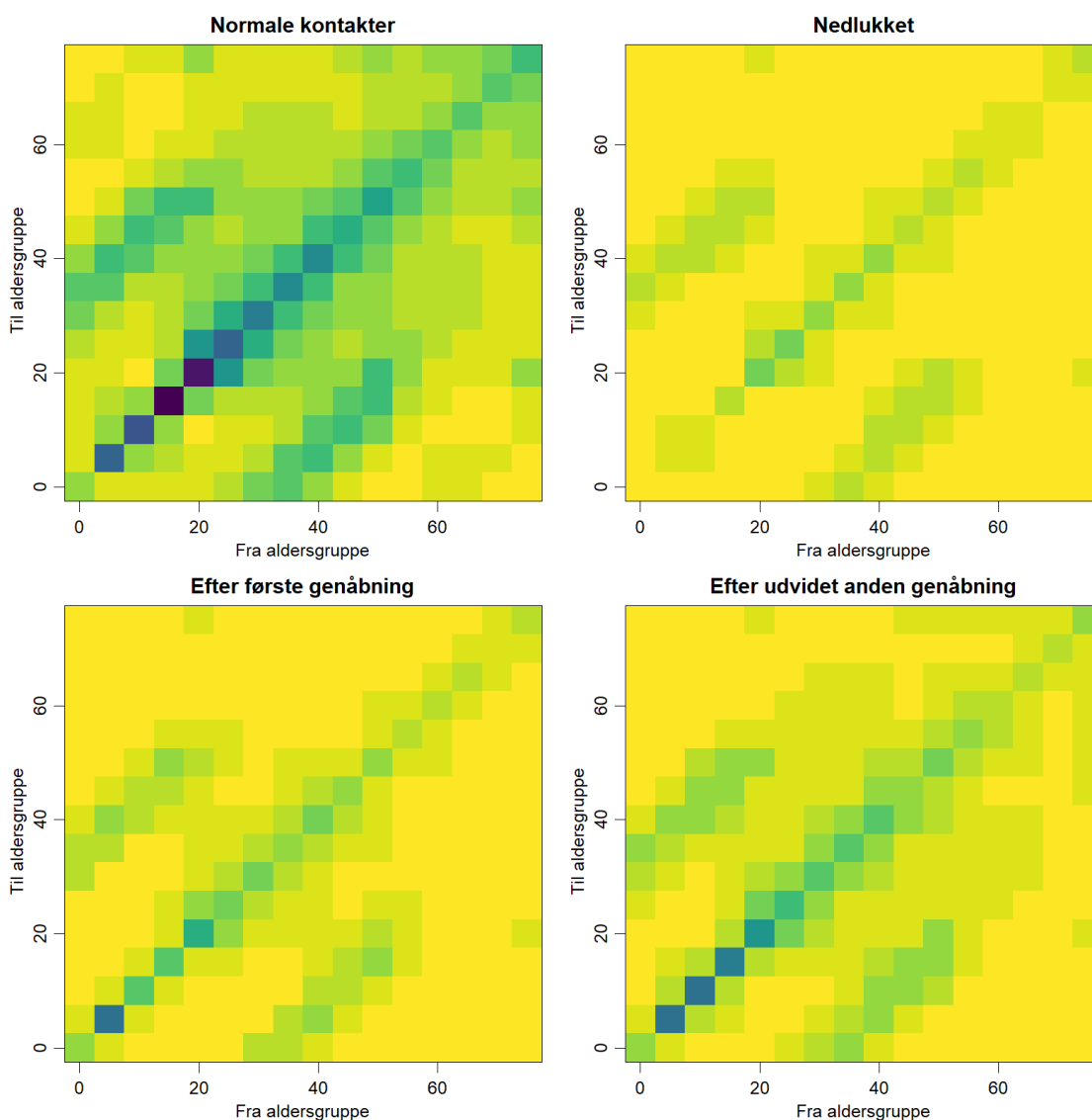


Figur 1: Kontaktmatricer: Fordelingen af kontakter på de fire kontakttyper ved normalt kontaktniveau. Lys betyder relativt få kontakter og mørk betyder relativt mange kontakter.





Figur 2: Kontaktmatricer: Antal kontakter (sammenlagt for både hjem, arbejde, skole og øvrige kontakter) i forskellige faser: Øverst til venstre ses ved fuldt åbent samfund, øverst til højre ses antal kontakter under nedlukningen. Nederst ses antal kontakter efter første fase og den udvidede anden fase af genåbningen. Lys farve betyder relativt få kontakter, mørk farve betyder relativt mange kontakter.



Tekniske oplysninger

Kontaktmatricerne er implementeret i det statistiske software R. Scriptet bag parametre til modelkørsler i ekspertgruppens tillægsrapport d. 20. maj er delt på <https://github.com/laedtu/C19DK> under "Contacts".

Referencer

Klepac, P., Kucharski, A. J., Conlan, A. J., Kissler, S., Tang, M., Fry, H., & Gog, J. R. (2020). Contacts in context: large-scale setting-specific social mixing matrices from the BBC Pandemic project. MedRxiv. <https://doi.org/10.1101/2020.02.16.20023754>



Klepac, P., Kissler, S., & Gog, J. (2018) Contagion! The BBC Four Pandemic – The model behind the documentary. *Epidemics*, Volume 24, September 2018, Pages 49-59. <https://doi.org/10.1016/j.epidem.2018.03.003>

Mossong, J., Hens, N., Jit, M., Beutels, P., Auranen, K., Mikolajczyk, R., ... & Heijne, J. (2008). Social contacts and mixing patterns relevant to the spread of infectious diseases. *PLoS medicine*, 5(3). doi:10.1371/journal.pmed.0050074



Analyse af indlæggelsesforløb ved hjælp af multi-state model

Formål

For at vurdere ressourcetræk på sygehusene er det afgørende at forstå, hvor mange dage en patient er indlagt på hhv. almene sengepladser og evt. på en intensiv sengeplads under hele indlæggelsesforløbet. Da COVID-19-patienter er en ny patienttype for sygehusene, var dette ikke kendt ved epidemiens start.

Ved brug af multi-state model kan den overordnede fordeling af patienters forløb beskrives.

Teori

Patientforløb er forskellige, og under en indlæggelse kan patienter skifte mellem forskellige dele af hospitalet, inden de i sidste ende forlader hospitalet, enten i live eller døde. Specifikt analyseres derfor skift mellem følgende fire tilstande:

1. Indlagt på en almen sengeplads
2. Indlagt på en intensiv sengeplads (der skelnes her ikke mellem intensiv med eller uden respirator).
3. Udskrevet i live
4. Død på hospitalet

Disse tilstande er naturligvis en meget grovkornet beskrivelse af et faktisk patientforløb. Formålet er ikke at forudsige de enkelte patientforløb, men alene at kunne forudsige gennemsnitlige ressourcetræk.

Patienter kan skifte flere gange mellem de to første tilstande, hvis patienten først går ind i tilstand 3 eller 4, vil de derimod forblive der (en genindlæggelse behandles som en ny indlæggelse). Ressourcetrækket på hospitalsvæsnet er bestemt af den samlede varighed af ophold i tilstandene 1 og 2 samt hvor mange patienter, der bliver indlagt til at starte med.

Metoden gennemgås med udgangspunkt i data opdateret frem til d. 27. april 2020. Indtil den dato var, eller havde der været, 1.297 patienter indlagt i Region Hovedstaden og Region Sjælland. Fordelingen af patienterne i de fire tilstande ses i tabel 1:

Tabel 1: Status for de 1.297 patienter i Region Hovedstaden og Region Sjælland indlagt frem til d. 27. april 2020.

	Indlagt (tilstand 1 og 2 sammen)	Udskrevet i live (tilstand 3)	Død på hospitalet (tilstand 4)
Antal patienter	212	839	246

Da mange patienter stadig var indlagt på daværende tidspunkt, vil den gennemsnitlige indlæggelsestid blive undervurderet eftersom patienterne ikke havde afsluttet deres forløb. For at undgå dette benyttes såkaldte multi-state modeller (state=tilstande), hvor varigheden af de forventede ophold i hver tilstand kan bestemmes på trods af, at en betydelig andel patienter stadig er indlagt. Mere specifikt benyttes den såkaldte Nelson-Aalen estimator (Aalen, 1978). Analyserne



er opdelt efter alder (over eller under 60 år). De statistiske modeller giver frekventistiske konfidensintervaller, men når resultaterne fra multi-state modellerne integreres med øvrige modelkomponenter (fx populationsmodellen for smittespredning) benyttes alene punktestimater. Usikkerhed for de samlede output fra modellerne adresseres under de enkelte andre modelkomponenter.

I praksis estimeres såkaldte hazards for transitioner (overgange) mellem de betragtede tilstande som funktion af tid siden indlæggelse. Metoden er ikke-parametrisk, og der laves dermed ikke yderligere antagelser eller benyttes yderligere parametre.

Multi-state modellen sikrer optimal udnyttelse af de tilgængelige data på daværende tidspunkt. Modellen kan dog ikke reparere, at indlæggelser længere end omtrent 21 dage (pr. ultimo april) ikke kunne analyseres ud fra danske COVID-19-data på daværende tidspunkt. Dette skyldes, at der var registreret meget få med så lange indlæggelser, specielt på intensive sengepladser. Denne 'hale' af meget lange indlæggelser kan have en vis betydning for det samlede ressourcetræk og skal derfor inkluderes. For stadig at kunne vurdere det samlede træk på hospitals ressourcer blev det antaget (Anders Perner, Rigshospitalets Intensiv afdeling) - at for indlæggelser længere end 21 dage vil 6% forlade intensiv og almene sengepladser per dag. Det vurderes, at denne antagelse ikke påvirker estimerne væsentligt, da de fleste patientforløb er afsluttet inden.

Derudover blev der for patienter over 60 år antaget en ligelig fordeling mellem tilstand 3 og 4, mens det for patienter under 60 år blev antaget en 3:1 fordeling mellem tilstand 3 og 4. Det skal dog bemærkes, at antagelser om fordeling mellem udskrivning i live og død ikke påvirker modellens forudsigelse for det samlede ressourcetræk på hospitalerne.

Data

Indlagte COVID-19 patienter i Region Sjælland og Region Hovedstaden. Data er baseret på en sammenstilling af oplysninger fra MiBa, CPR-registret, Dødsårsagsregistret, Landspatientregistret og regionernes journalsystemer.

Parametre:

I modellen blev benyttet en antagelse om en daglig udskrivningsrate på 6% efter dag 21 som beskrevet ovenfor, og er derudover drevet af data.

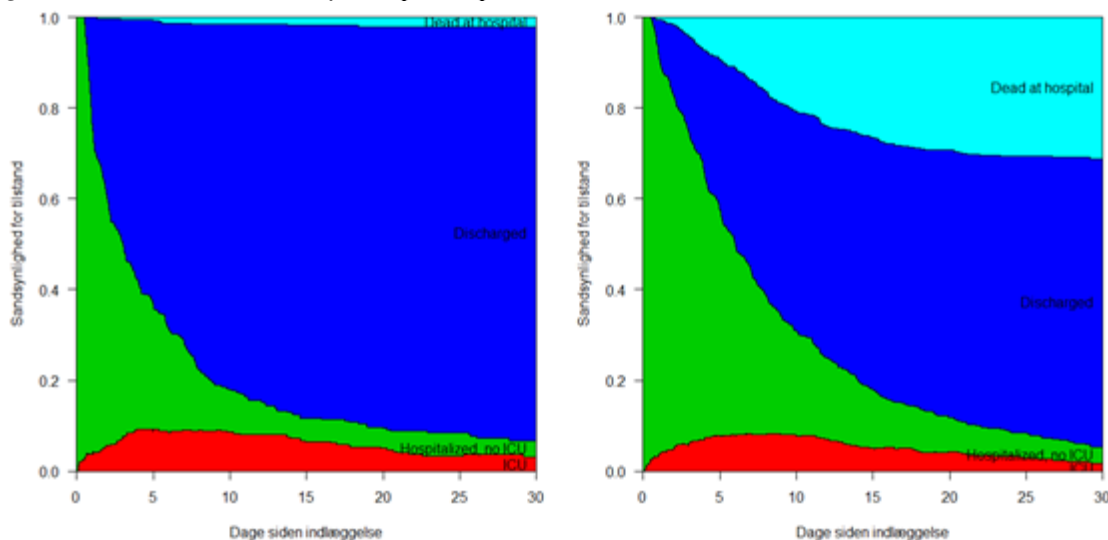
For at tage højde for mulige ændringer i hospitalspraksis eller typen af patienter blev analyserne også gennemført, hvor patienter indlagt før d. 1. april-2020 blev vægtet fire gange lavere end patienter indlagt senere. Resultater i denne analyse præsenterer dog de uvægtede analyser.

Eksempler

Fra modellen fås umiddelbart forudsigelser for fordeling på de 4 tilstande fra indlæggelse og frem. Baseret alene på data fås fordeling på de fire tilstande i figur 3.



Figur 3: De prædikerede fordelinger på de fire tilstande for hhv. personer under 60 år (venstre) og over 60 år (højre). Figuren er baseret på data opdateret frem til d. 27. april 2020. Grøn: 1. Indlagt på en almen sengeplads. Rød: 2. indlagt på en intensiv sengeplads. Mørkeblå: 3. udskrevet i live. Turkis: 4. død på hospitalet.



Ovenstående beskriver fordelingen, når indlæggelse påbegyndes på en almen sengeplads. Et mindretal af patienterne indlægges dog direkte på en intensiv sengeplads, og der er derfor foretaget en sammenvægtning af de få patienter, der indlægges direkte på en intensiv sengeplads og de mange patienter, hvor indlæggelsen påbegyndes på en almen sengeplads.

De forventede fordelinger som funktion af tid siden indlæggelse på hhv. indlagte på almene og indlagte på intensive sengepladser benyttes fx i populationsmodellen.

De estimerede fordelinger kan direkte oversættes til forventede indlæggelsestider. Dette er præsenteret i tabel 2.

Tabel 2: Forventede gennemsnitlige indlæggelsestider og sandsynlighed for at komme på en intensiv sengeplads baseret på data frem til d. 27. april 2020.

	Under 60 år	Over 60 år
Forventet antal dage i almen sengeplads	5,58	8,43
Forventet antal dage i almen sengeplads (givet kommer på intensiv)	7,34	7,64
Forventet antal dage i almen sengeplads (givet aldrig på intensiv)	5,1	8,6
Sandsynlighed for mindst 1 dag i intensiv sengeplads	0,22	0,17
Forventet antal dage i intensiv sengeplads (givet kommer på intensiv)	17,69	14,74



Yderligere/sekundære analyser

Hazards for de forskellige transitioner er desuden analyseret for afhængighed af køn, alder, region og indlæggelsestidspunkt. Køn og alder blev krydschecket for at sikre, at analysen viste de effekter, der forventes på baggrund af eksisterende studier (stigende risiko for indlæggelse på intensiv for mænd og med øget alder). Region og indlæggelsestidspunkt benyttes for at monitorere om analyserne skal opdeles regionalt eller om der er sket skift i patientsammensætning/-behandling over tid.

Tekniske oplysninger

Metoderne er implementeret i det statistiske software R (pakkerne mstate og survival). Data er direkte personhenførbart (de inkluderer præcise tidspunkter for indlæggelser og død) og kan derfor ikke deles offentligt.

Mere specifikt organiseres data først til brug for multi-state modeller, hvilket betyder at der er en række for hver transition en person kan lave og variable for

- tilstanden personen er i
- tidspunktet for start af ophold i denne tilstand
- tidspunktet for transition ud af denne tilstand
- mulige tilstande man kan flytte til (f.eks. fra tilstand 1 kan hoppes til alle de tre andre tilstande)
- variable for hvilken transition, der er tale om (f.eks. fra tilstand 1 til 2)
- indikator for om denne transition faktisk skete
- yderligere basis variable (alder etc).

Først benyttes 'coxph'-funktionen til at fitte en ikke-parametrisk model stratificeret på hver transitionstype. Dernæst konverteres disse fits til kumulative hazards for hver transitionstype har funktionen 'msfit'. Endelige konverteres disse til transitions-sandsynligheder evalueret ved hvert event tidspunkt ved hjælp af funktionen 'probtrans'. Disse indeholder de sandsynligheder, som ligger til grund for ovenstående figur, men kan også bruges til at lave daglige transitions-sandsynligheder, hvilket via simulation kan oversættes til tabel 2.

Referencer

Aalen, Odd (1978). "Nonparametric inference for a family of counting processes". *Annals of Statistics*. 6: 701–726. doi:10.1214/aos/1176344247.



Populationsmodel

Formål

Formålet med populationsmodellen (den initiale SEIR) var at udvikle en relativt simpel model til forudsigelse af den nødvendige hospitalskapacitet. Det var nødvendigt, at modellen kunne håndtere, at ældre har en større risiko for svær sygdom som følge af COVID-19 end yngre, samt køre på de meget få datapunkter, der var tilgængelig i starten af epidemien. Samtidig skulle modellen nemt kunne tilpasses udmeldte nedlukninger og genåbninger, som parameteriseres i kontaktmatricemodellen. I første omgang var der få data for de enkelte regioner, så det primære formål var at lave en enkelt model for hele Danmark.

Teori

Modellen er baseret på en udvidet aldersstruktureret SEIR model, hvor smitterater kan reduceres i givne tidsrum. Den oprindelige model blev udviklet i simpel form i starten af forrige århundrede (Se eks. Kermack 1997). En let forståelig gennemgang kan ses på Wikipedia².

Modellen er udvidet med en dynamisk komponent, hvor parametrene ændrer sig over tid, hvilket gør det muligt at ændre på effektiviteten af myndighedstiltag i et givent tidsrum. Desuden udvides modellen med antal indlagte på hospitaler.

Endvidere er modellen udvidet med en aldersstrukturering, hvilket betyder, at alle parametre, der beskriver sygdommen, er afhængig af alder. Bemærk at disse udvidelser medfører, at en del teoretiske størrelser fra simple modeller ikke bare kan overføres. I modellen opdeles befolkningen i to aldersgrupper; hhv. over og under 60 år. Det var fra epidemiens start klart, at personer over 60 år var mere udsatte for svær sygdom end dem under 60 år. Befolkningen opdeles kun i to aldersgrupper, da der var få datapunkter tilgængelige i epidemiens start, og modellen skulle kunne fitte til disse data.

Figur 4 viser et flowdiagram af den underliggende model: S svarer til den modtagelige del af populationen, som kan blive smittet gennem kontakt med inficerede individer. Inficerede individer (I) bliver inddelt i to kategorier 1) inficerede med milde symptomer, som ikke kræver hospitalsindlæggelse og derfor efterfølgende bliver raske (IR) og 2) inficerede som kræver hospitalsindlæggelse efterfølgende (IH). De hospitalsindlagte er også delt ind i to 1) dem som bliver udskrevet uden kritisk tilstand (HR) og 2) dem, hvor symptomerne udvikler sig til en kritisk tilstand (HC). Individer i kritisk tilstand (C) vil enten få forbedret tilstanden og gå fra kritisk tilstand (CR) til hospitalsindlæggelse med udsigt til at blive rask (HR), eller være i kritisk tilstand (CD) med udgang død (D). Hvis et individ er rask og immun betegnes dette med R og D betegner individer afgået ved døden. Da myndighedsbetjeningsopgaven har fokus på indlagte patienter, er der ikke skelnet mellem om smittede, som ikke indlægges, bliver raske eller dør når de ikke længere smitter. Modellen kan udvides, så det er muligt at gå fra IR til D , men det vil blot tilføje

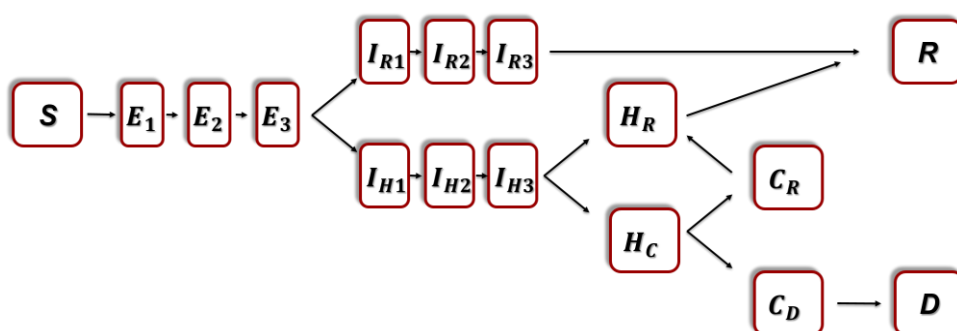
² https://en.wikipedia.org/wiki/Compartmental_models_in_epidemiology



endnu en parameter, som ikke er kendt og som ikke påvirker resultatet - så længe den infektiøse periode i de enkelte aldersgrupper er den samme for alle i IR.

Ovenstående beskriver modellen, som blev brugt til simulationer bag ekspertrapporterne, som blev offentliggjort fra d. 2. til 6. april. Efterfølgende er tilstandene E, IR og IH blevet opdelt i tre undertilstande, fx E1, E2 og E3. Dette er gjort for at lave en ventetidsfordeling, som har mindre varians og for at reducere andelen, som begynder at smitte umiddelbart efter, at de er blevet smittet. Dette svarer til at gå fra en eksponentiefordelt ventetid til en Erlang(3)-fordelt ventetid.

Figur 4: Flowdiagram af populationsmodellen (SSEIR – struktureret SEIR) model.



Note: *S* beskriver modtagelige individer og *E* er eksponerede individer. *IR* er inficerede, som vil blive raske, mens *IH* er inficerede, som vil blive indlagt på hospitalet. *H_R* og *H_C* dækker over individer på hospitalet, hvor *R* og *C* angiver, om de vil blive raske eller udvikle en kritisk tilstand, på samme måde beskriver *C_R* og *C_D* om individer i kritisk tilstand vil hhv. overleve (recover) eller dø. *R* er raske (immune) og *D* er døde.

Modellen med tre undertilstande for hver af E, IH og IR er defineret ved nedenstående differentiaalligningssystem, hvor indeks 'i' angiver aldersgruppen. Blandt de benyttede modeller er nedenstående den, som har flest tilstande per aldersgruppe.

$$\begin{aligned} \frac{dS_i}{dt} &= -S_i \beta_{i.}(IR1 + IH1 + IR2 + IH2 + IR3 + IH3)RR \\ \frac{dE1_i}{dt} &= S_i \beta_{i.}(IR1 + IH1 + IR2 + IH2 + IR3 + IH3)RR - 3\gamma_{E1}E1_i \\ \frac{dE2_i}{dt} &= 3\gamma_{E1}E1_i - 3\gamma_{E1}E2_i \\ \frac{dE3_i}{dt} &= 3\gamma_{E1}E2_i - 3\gamma_{E1}E3_i \\ \frac{dIR1_i}{dt} &= 3\gamma_{E1}E3_i p_{IH} - 3\gamma_{IR}IR1_i \\ \frac{dIR2_i}{dt} &= 3\gamma_{IR}IR1_i - 3\gamma_{IR}IR2_i \\ \frac{dIR3_i}{dt} &= 3\gamma_{IR}IR2_i - 3\gamma_{IR}IR3_i \\ \frac{dIH1_i}{dt} &= 3\gamma_{E1}E3_i (1 - p_{IH}) - 3\gamma_{IH}IH1_i \\ \frac{dIH2_i}{dt} &= 3\gamma_{IR}IH1_i - 3\gamma_{IH}IH2_i \\ \frac{dIH3_i}{dt} &= 3\gamma_{IR}IH2_i - 3\gamma_{IH}IH3_i \end{aligned}$$



$$\begin{aligned} \frac{dHR_i}{dt} &= 3\gamma_{IH}IH3_i p_{HR} + \gamma_{CR}CR_i - \gamma_{HR}HR_i \\ \frac{dHC_i}{dt} &= 3\gamma_{IH}IH3_i(1 - p_{HR}) - \gamma_{HC}HC_i \\ \frac{dCR_i}{dt} &= \gamma_{HC}HC_i p_{CR} - \gamma_{CR}CR_i \\ \frac{dCD_i}{dt} &= \gamma_{HC}HC_i(p_{CR} - 1) - \mu_C CD_i \\ \frac{dR_i}{dt} &= 3\gamma_{IR}IR3_i + \gamma_{HR}HR_i \\ \frac{dD_i}{dt} &= \mu_C CD_i \end{aligned}$$

Data

Nye indlæggelser af patienter med COVID-19. Data er baseret på en sammenstilling af oplysninger fra MiBa, CPR-registret, Dødsårsagsregistret, Landspatientregistret og regionernes journalsystemer.

Parametre

Tablet 3: Oversigt over parametre og værdier. Parametrene bootstrappes i de angivne intervaller.

Parameter	Værdi	
	Personer <60 år	Personer ≥60 år
Latenstid (γ_{EI}) (Perioden fra man bliver smittet til man kan smitte videre)	4-6 dage	4-6 dage
Tid man smitter før indlæggelse (γ_{IH})	6-10 dage	5-9 dage
Andel der indlægges ^a (p_{IH})	0,05-0,50%	5,0-6,2%
Andel der bliver rask efter indlæggelse ^a (p_{HR})	77-97%	70-90%
Tid fra indlæggelse til intensiv (γ_{HC})	0,5-2,5 dage	0,5-1,5 dage
Andel i live efter intensiv ^a (p_{CR})	70-95%	45-55%
Tid tilbragt på intensiv ^b (γ_{CR})	14-28 dage	14-28 dage
Tid tilbragt på hospital (γ_{HR})	5-9 dage	5-15 dage
Tid man smitter udenfor hospital (γ_{IR})	4-6 dage	4-6 dage
Antal smittede den 11/3 2020	40.000-70.000	5.000-14.000
Smitterate (Beta)		
Scenarie, fra den 11/3 til 14/4 2020	0,12-0,22	0,14-0,24
Smitterate mellem aldersgrupper	0,02-0,12	



Grundscenarie, fra den 15/4 2020	0,16-0,26	0,13-0,23
Smitterate mellem aldersgrupper	0,03-0,13	
Grundscenarie, fra den 15/4 2020, normal afstand	0,25-0,35	0,24-0,34
Smitterate mellem aldersgrupper	0,05-0,15	
Skaleringsfaktor for smitterate (RR)	0,5-1,5	

Note: a) Procent analyseres som sandsynlighed, b) Samme tid for patienter der er bortgået eller udskrevet,

Beta-matricen består af diagonalelementer med smitteraten indenfor hver aldersgruppe ('<60 år' og '>=60 år'). Uden for diagonalen er smitteraten mellem de to grupper '<60 år' og '>=60 år'.

Eksempler

Modelresultaterne viser risikoen for antal indlagte på almene og intensive sengepladser, og er fremkommet ved at lave gentagne simuleringer (bootstraps) med forskellige parametre. Disse simuleringer vægtes efter, hvor godt de passer til det observerede antal indlagte. Det vil sige, at modellen fitter til indlæggelsesdata ved at se på tilfældige kombinationer af parameter-værdier inden for de definerede parameter-rum. Denne vægtning giver i realiteten en automatisk optimering af kombinationen af parametre. Da parametrene ikke kendes, og forskellige kombinationer er mulige og kan give samme resultat, er dette en måde at indbygge usikkerheden på. Eksempelvis vil antal smittede per 11. marts og andel smittede der kommer på hospital skalere inverst, men det præcise antal smittede i samfundet er ikke afgørende for antal indlagte på hospital, så længe at den totale andel smittede er lille.

Tekniske oplysninger

Modellen er implementeret i det statistiske software R og kan vises som en hjemmeside (Shiny dashboard) herunder hhv. et skærmdump af brugerflade, hvor det er muligt at vælge parametre og et skærmdump af tilhørende resultater. Koden er tilgængelig via GitHub: <https://github.com/laedtu/C19DK>

Referencer

William Ogilvy Kermack, A. G. McKendrick and Gilbert Thomas Walker 1997 A contribution to the mathematical theory of epidemics. Proc. R. Soc. Lond. A115700–721.



Regionsopdelt populationsmodel

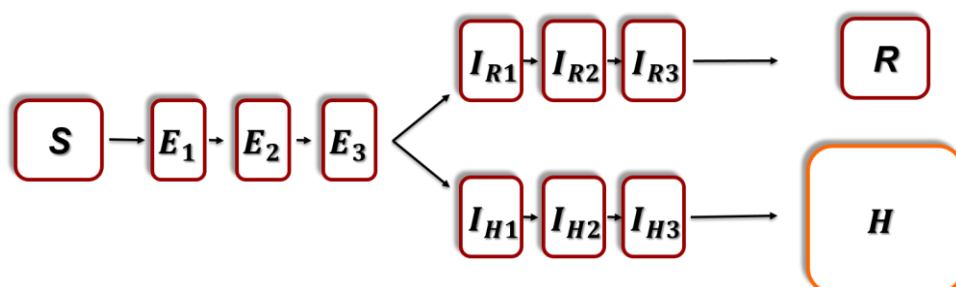
Formål

Efterhånden, som flere data blev tilgængelige per region blev det tydeligt, at der var væsentlige forskelle mellem de fem regioner i Danmark. For at tage højde for denne forskel blev populationsmodellen anvendt på hver region for sig. Desuden indførtes en del mindre tekniske ændringer. Modellen blev ændret i forbindelse med ekspertrapporten d. 6. maj 2020.

Teori

Modeltypen er den samme som nævnt under populationsmodellen, dvs. en SEIR model med erlang(3)-fordelte E og I tilstande. Denne model har dog fuldt ud implementeret multi-state modellen over patienters forløb på hospital og har dermed nedbragt antallet af tilstande og parametre i selve modellen. I visse sammenhænge benævnes denne model som en SEIRH, hvor H står for hospital. Det er HR, HC, CR, CD og D, som er samlet i H, som så fungerer som input til multi-state modellen for forløbet på hospitalet (se figur 5).

Figur 5: Flowdiagram af den regionsopdelte populationsmodel (SSEIR – struktureret SEIR).



I de første populationsmodeller blev alle parametre fundet ved bootstrapping. Da den regionsopdelte model blev opbygget, var der flere dage med oplysninger om antal indlagte. På den baggrund blev det valgt at gå over til at optimere antal smittede i hver aldersgruppe d. 11. marts samt skaleringsparameteren. Optimeringen er lavet som maksimum likelihood af antal dagligt nyindlagte patienter med COVID-19 som antages Poisson-fordelt. Disse likelihood-værdier benyttes også til at kaste de dårligste kombinationer af parametre væk inden der laves percentilplot af fordelingerne af fremskrivningerne.

Data

Nye indlæggelser af patienter med COVID-19 fordelt på regionerne. Data er baseret på en sammenstilling af oplysninger fra MiBa, CPR-registret, Dødsårsagsregistret, Landspatientregistret og regionernes journalsystemer. Data offentliggøres på SSI's hjemmeside og deles efterfølgende på github.



Parametre

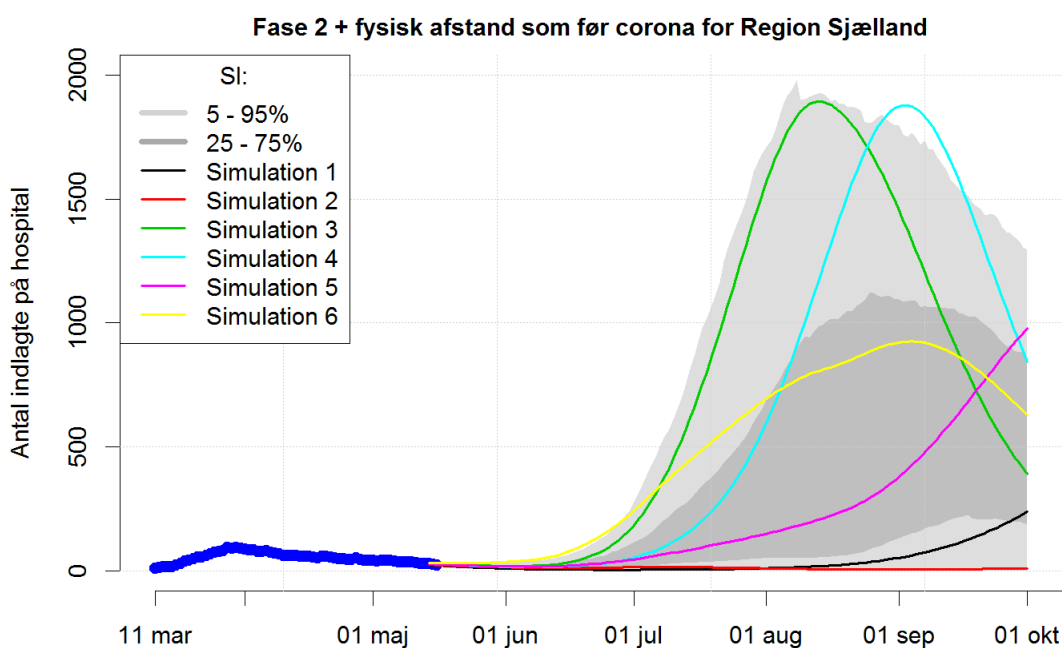
Parametrene er de samme for populationsmodellen dog uden parametre, der beskriver det interne forløb på hospitalet. Derudover fittes et antal smittede per 11. marts til at starte modellen, samt en skalering på smittesandsynligheden indenfor hver region. Det bemærkes, at startbetingelser og sandsynlighed for indlæggelse på hospital er stærkt korrelerede og, at det for nuværende ikke er meningsfuldt at fortolke specifikke estimater for disse tal fra den optimering der foregår i koden. Dette vil give mere mening, hvis der på et senere tidspunkt opbygges viden om hvor mange, der har haft sygdommen i samfundet.

Eksempler

Som eksempel viser figur 6 udvalgte enkelte modelsimulationer for antal indlagte. Disse simuleringer er valgt blandt de gennemførte 500 simuleringer, som ligger til grund for de simulationsintervaller, som er angivet i de lyse- og mørkegrå områder. Det er vigtigt at skelne mellem statistiske forudsigelser angivet i simulationsintervallerne og de enkelte simuleringers forløb.

Den øvre grænse af det mørkegrå område angiver det antal indlagte, som 75% af simuleringerne ligger under på den givne dag. Det skal bemærkes, at de kurver, der definerer 75%-percentilen på en given dag, kan have forskellige forløb, således at nogle kan være opadgående og andre nedadgående. Variationen skyldes usikkerheder i de centrale parametre for modelsimuleringerne.

Figur 6: Eksempel på visning af enkelte simulationer for den simulerede sygehusbelastning for alle indlagte i Region Sjælland.





Tekniske oplysninger

Modellen er implementeret i det statistiske software R. Koden er tilgængelig via GitHub: <https://github.com/laecdtu/C19DK>. Det bemærkes, at grundet fit tager denne model ca. 20 gange så lang tid at køre som populationsmodellen.