

Ebolas epidemirisk och kontrollstrategier

Emil Hellström

Kandidatuppsats 2016:26
Matematisk statistik
Augusti 2016

www.math.su.se

Matematisk statistik
Matematiska institutionen
Stockholms universitet
106 91 Stockholm

Ebolas epidemirisk och kontrollstrategier

Emil Hellström*

Augusti 2016

Sammanfattning

Vi tittar på hur sjukhusisolering samt hur västafrikansk traditionell begravning påverkar slutstorlekar av epidemiutbrott och risken för att epidemi av ebola kan uppträda. Sjukhusisolering och begränsningar av riskbeteenden som traditionell begravning är typiska kontrollstrategier för att sänka slutstorlekar och eliminera epidemirisk. Slutstorlekarna simuleras fram genom en stokastisk smittmodell. Vi applicerar smittspridningen på en kontaktstruktur av två nivåer där ena nivån är global där kontakt sker mellan hela populationens individer och med likfördelad sannolikhet medans den andra nivån är lokal inom separata hushåll. Vi varierar smittsamhetsantagandena för viruset med avsikt att fånga in en generell bild av riskerna. Vid de högre smittsamhetsantaganden exploderar epidemierna med stora slutstorlekar där det krävs hög grad av sjukhusisolering och låg grad av traditionell begravning för att kontrollera epidemirisken medans det vid de lägre smittsamhetsantaganden krävs mindre del sjukhusisolering för att kontrollera smittspridningen samtidigt som toleransen för den traditionella västafrikanska begravningen ökar.

*Postadress: Matematisk statistik, Stockholms universitet, 106 91, Sverige.
E-post: emil.hellstrom@gmail.com. Handledare: Maria Deijfen, Pieter Trapman.

Innehåll

1	Inledning	4
2	Modellbeskrivning och teori	7
2.1	Kontaktstruktur	7
2.2	Modellformulering	9
2.3	Tillstånd och överföringfördelning	13
2.4	Generering av population och hushåll	18
2.5	Reproduktionstalet R_0^H för hushållstruktur	22
2.6	Durationer för modellens tillstånd	25
3	Simulering av modellen	27
3.1	Simuleringsteknik och beteckningar	27
3.2	Simuleringsresultat	30
4	Slutdiskussion	41
	Appendices	43

Förord och tack

Jag vill tacka Pieter Trapman och Maria Deijfen för vägledning och inspiration under projektets gång.

1 Inledning

Det senaste ebolautbrottet i västafrika är det största som hittills dokumenterats. [4] Utbrottet startade sannolikt i det västafrikanska landet Guinea i december år 2013. Den 23:e mars 2014 blev WHO (World Health Organisation) informerade om att utbrottet eskalerat och den 8:e augusti 2014 utfärdade WHO varningen att det nu blivit en fara av internationella mått. De tre länder som framförallt har drabbats är Sierra Leone, Guinea och Liberia. [8]

När detta skrivs är WHO:s officiella antal rapporterade ebolaoffer totalt 28 616 stycken medans det rapporterade antalet som avlidit till följd av ebolainfektionen är 11 310 stycken. Dessa rapporterade fall motsvarar en lägre dödlighet än många av de historiskt förekommande skattningarna av ebolavirusets dödlighet med dödligheter runt 70%. [4] En del av anledningen till den lägre dödligheten i WHO:s rapporterade ebolafall förmodas bero på underrapportering relaterat till antalet rapporterade ebolaoffer som avlidit. En stor del av de rapporterade fallen gäller ebolaoffer som fortfarande är i ett aktivt sjukdomsstadie och därför ännu inte avlidit eller återhämtat sig. Om denna ebolainfekterade sedan avlider så har statistiken inte justerats vilket ger att antalet avlidna underreporteras. [4] En annan felkälla skulle kunna vara att de fall som är obekräftade och rapporteras in innehåller andra smittor än ebola. Det är nog troligt att Ebola har högre dödlighet än de flesta av dessa obekräftade fall, vilket också tillför en underrapportering till antalet avlidna. En tredje orsak skulle kunna vara att de individer där ebolainfektion misstänks också betingas av att få en högre grad av sjukvård än den totala populationen av ebolainfekterade i genomsnitt får. Sjukvård har visat sig minska dödligheten något om den ebolainfekterade individen inkommer tidigt i sjukdomsstadiet. I den smittmodell vi definierar antar vi dock att dödligheten för en ebolainfekterad på sjukhus är samma som dödligheten för en ebolainfekterad ute i samhället. De två sistnämnda är möjliga orsaker som vi inte lyckats bekräftat med hög säkerhet.

Ebolas infektionscykel inleds med en inkubationstid som betingas av att sjukdomssymptom ännu inte uppvisas samtidigt som risken för smittoöverföring från den infekterade under inkubationsstadiet är liten. [8] Detta leder oss till förenklingen att vi antar smittsamheten som obefintlig. Nästa stadie i infektionscykeln är smittsamhetsstadiet som betingas av tydliga sjukdomssymptom som

kräkningar, feber, diarré och ibland in- och utvärtes blödningar. [4] Sjukdomssymptomen är tydliga och lär utgöra en varningssignal mot de individer som fortfarande är mottagliga för infektion. Detta tillsammans med en snabb smittsamhetsduration där den infekterade antingen avlider eller återhämtar sig med immunitet inom loppet av några dagar motiverar ett resonemang om att ebolaviruset inte har de typiska egenskaper som ger hög spridningsbenägenhet. Storleken på det senaste utbrottet i västafrika visar att resonemanget tål att granskas och att det, åtminstone givet vissa förutsättningar, föreligger en epidemirisk. Det har framkommit att traditionella begravningar, vilket vi fortsättningsvis refererar till som *osäkra begravningar*, har varit betydande för det initialt snabba epidemiutbrottet. WHO listar de traditionella begravningar som en av huvudriskerna för ebolaspridning. [7] På de *osäkra begravningarna* tvättar de anhöriga den avlidna kroppen innan begravningseremonin ägt rum och har även kontakt med den avlidna kroppen vid själva begravningseremonin. [5] En individ som avlidit på grund av ebolainfektion är fortfarande smittsam under en tid efteråt. Eftersom ebola smittar genom kontakt med infekterad kroppsvätska så smittas således många av deltagarna på dessa begravningar. Även långt inne i epidemin har det förekommit fall där de anhöriga inte frångått sina traditionella begravningsritualer trots att ebolainfektion starkt kunde misstänkas. [5]

Bortsett från några få och infrekventa nyrapporterade fall så är det västafrikanska ebolautbrottet, när detta skrivs, över för den här gången men precis som andra virus- och bakteriesmittor finns det alltid en risk för nya utbrott. [8] För att stävja epidemirisker krävs kunskap som ger bättre förutsättningar att minska smittspridningen tillräckligt mycket. En del av denna kunskap kan hämtas från att studera stokastiska modeller som avser beskriva smittspridning. Med stokastiska modeller kan man efterlikna de relevanta egenskaper och förutsättningar från verkligheten och på så sätt ta fram resultat som ger förutsägelser om risker och hur man mest effektivt tillämpar kontrollstrategier.

I detta projekt avser vi därför att beskriva smittspridning, och vilken inverkan kontrollstrategier har på smittspridningen, av ebola i den västafrikanska populationen med hjälp av datorsimuleringar av en stokastisk smittmodell. Modellen kan ses som en expanderad version av den liknande men

förenklade SIR-modellen. [3] [2] Population som studeras delas in i olika tillstånd med avseende på var i infektionscykeln de befinner sig samt huruvida individen är ute i samhället, på sjukhus eller avliden på antingen en säker eller osäker begravning.

De infekterade individerna smittar de mottagliga individerna i populationen enligt antagna smittsannholikheter. Vi antar en kontaktstruktur på två nivåer. Ena nivån är global och den andra är en lokal hushållsnivå där smittsannolikheterna för global respektive lokal nivå skiljer sig åt. [1] En kontaktstruktur likt denna definieras som heterogent mixad kontaktstruktur. [2] Våra hushåll antas bestå av ett varierande antal individer mellan 1 och 10. Fördelningen av andelarna för dessa hushållstorlekar i modellens populationen har valts med avsikt att spegla hushållstrukturen i den verkliga västafrikanska populationen. Motiveringen till att vi implementerar hushåll är att vi tror det kan ha betydelse för smittdynamiken eftersom ebola i hög grad smittar inom hushåll. [7] Hur stor andel av de olika hushållstorlekarna som finns i populationen bestäms genom slumpdragning enligt inhämtad data gällande hushållsdistributionen från det västafrikanska landet Liberia. Vi antar två olika typer av smittbärare, den ena är de infekterade och levande individerna ute i samhället och den andra de avlidna som förts till osäker begravning. Individerna från dessa två smittkällor riktar sin smittsamhet globalt mot samtliga mottagliga individer i populationen. De levande infekterade individerna smittar även lokalt mot de mottagliga individer i det hushåll den tillhör. I samtliga simuleringar antas de avlidna som förts till osäker begravning ha högre global smittsamhet än vad de levande individer har.

Fokus i arbetet ligger på att se hur en varierande andel som förs till *osäker begravning* och varierande andel *sjukhusisolering* påverkar risk för att epidemi uppstår och slutstorlekar på de epidemier som eventuellt uppstår. Att sänka andelen avlidna som förs till *osäkra begravningar* och införa *sjukhusisolering* är typiska *kontrollstrategier* som används för att stävja spridningsbenägenheten för smittor. Implementering av sjukhus i modellen har syftet att isolera ebolainfekterade individer vilket betyder att dessa individer inte längre kan smitta vidare från det att de inkommit till sjukhus. Det är inte orealistiskt att sjukhus skulle kunna vara smittspridare i det läget att medvetenheten om ebolaviruset är låg, men det är ett fall vi bortser från. Följden av implementering av sjukhusisolering blir att den genomsnittliga smittsamhetsdurationen ute i samhället minskar med

den andel som förs till sjukhus. Detta leder till att den aggregerade smittsamhetsdurationen ute i samhället sänks och som i sin tur dämpar det aggregerade "smittspridningstrycket" som leder till att reproduktionen av smittspridningen dämpas.

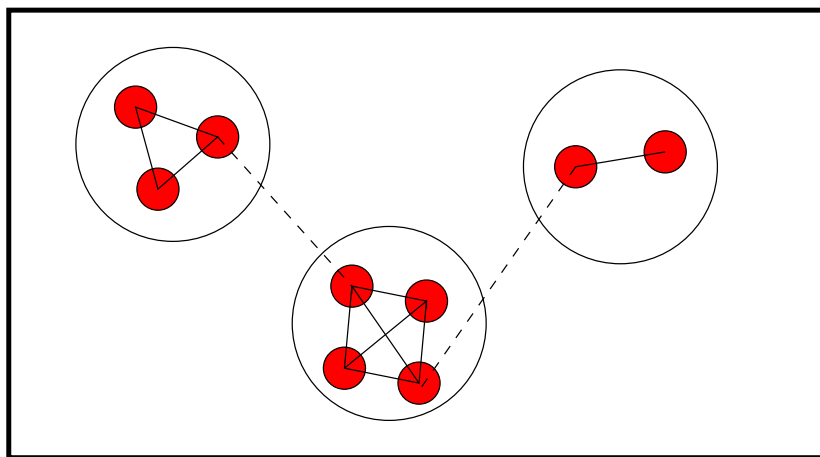
Ovanstående generella modellbeskrivning går vi mer ingående igenom i avsnitt 3. Vi specificerar kontaktstrukturen, formulerar relevant modellteori och visar hur den eboladrabbade populationen genereras. I avsnitt 4 låter vi presentera simuleringsresultat av vår modell och redogör hur simuleringarna utförs. Avslutningsvis förs en övergripande slutdiskussion i avsnitt 5.

2 Modellbeskrivning och teori

2.1 Kontaktstruktur

I den modell vi definierar antas en kontaktstruktur på två nivåer. Det betyder att individerna i populationen antas ha två olika sätt till kontakt som ebola kan överföras via. [1] Den ena nivån betecknas som den *globala nivån*. På globala nivån antas alla individer i populationen ha kontakt med varandra med likafördelad sannolikhet. Vi antar också att antalet globala kontakter, i absoluta mått, en individ i initialskedet har inte får bero på storleken av den genererade populationen. Vi implementerar antagandet genom att skala den globala kontaktsannolikheten med populationsstorleken. Syftet är att generalisera modellen och ger oss möjlighet att generera en godtyckligt stor population utan att det påverkar antalet kontakter en individ förväntas ha. Den andra nivån är den *lokala nivån* som gäller kontakten mellan medlemmar från samma hushåll. Kontaktsannolikheten mellan medlemmarna inom ett hushåll skalas inte med hushållstorleken. Eftersom kontaktsannolikheten inom hushåll inte skalas med hushållstorleken implicerar det att två specifika medlemmar av ett hushåll kommer ha kontakt med varandra lika ofta oberoende av hur många medlemmar som finns i hushållet. Det vill säga, en individ i ett hushåll av storlek $(n + k)$ har fler lokala kontakter än en individ i hushåll storlek k , där $n, k \in \mathbb{N}^+$. Eftersom vi i modellen antar att ebola sprids vid kontakt så betyder detta att en infekterad individ tillhörande ett stort hushåll förväntas smitta fler individer än vad en individ tillhörande ett mindre hushåll förväntas smitta. Utöver det kan ett beroende mellan hushållstorleken och *andelen* av hushållets individer som slutligen infekterats anas.

Detta beroende illustreras till viss del av de kombinationer av smittvägar som beror på hushållets storlek i Figur 1. De helstreckade linjerna representerar kontakterna mellan medlemmar inom hushållet och de streckade representerar globala kontakter mellan individer från olika hushåll. Under avsnitt 3.5 ger vi ett illustrativt exempel på det förväntade antalet och andelen av hushållet som infekteras i respektive hushållstorlek efter att en av medlemmarna infekterats. Resultatet visar att andelen som infekterats, när hushållsutbrottet är över, växer när antalet hushållsmedlemmar växer. Sammanfattningsvis kan man uttrycka det som att varje individ i modellens population antas både ha globala kontakter och lokala kontakter med sina hushållsmedlemmar. I epidemimodeller görs ibland det förenklande antagandet att populationen har en kontaktstruktur med endast globala kontakter där sannolikheterna för kontakt mellan individerna är likafördelade. Detta benämns som en *homogent mixad kontaktstruktur*. [2] Den kontaktstruktur på två nivåer benämns istället som *heterogent mixad kontaktstruktur*. [2]



Figur 1: Hushåll av storlek 2, 3 och 4

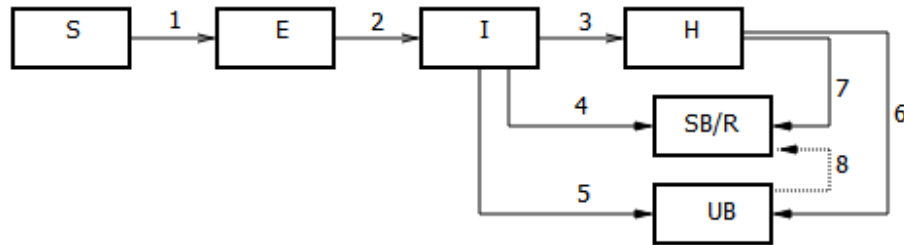
Anledningen till implementering av både en lokal hushållstruktur och en global struktur är att vi tror ebolaviruset har en högre smittsamhet mellan medlemmar av ett hushåll än mellan två individer som kommer från olika hushåll med endast global kontakt. [8] Givet detta antagande så bör följderna också bli att fördelningen av populationens hushållstorlekar har betydelse för smittspridningen. Det kan också ge oss information om hur en del kontrollstrategier mest effektivt tillämpas

med avseende på hushållstorlekar.

I verkligheten är kontaktstrukturen mellan populationens individer mer komplex än vad som fångas av ovan beskriven kontaktstruktur på två nivåer. Det är emellertid orealistiskt att fånga verklighetens komplexitet fullt ut eftersom detta skulle leda till en modell som blir allt för svårhanterlig. Vi nöjer oss med ovan definierade struktur då vi tror att denna fångar det mest väsentliga av ebolas smittdynamik i den västafrikanska populationen.

2.2 Modellformulering

Modellen har tillståndskategorierna; S (Susceptible), E (Exposed), I (Infective), H (Hospital), UB (Unsafe Burial), RSB (Removed/Safe Burial).



Figur 2: Ebolakedjan

Modellen definieras i diskreta tidsteg och varje tillstånd innehåller ett diskret antal individer från den totala populationstorleken N . Populationen är i tidsteg t distribuerad över modellens tillstånd enligt de stokastiska tillståndsvariablerna $S_t, E_t, I_t, H_t, UB_t, RSB_t$. Vid varje tidsteg t sker överföring av individer mellan tillstånd enligt pilriktningarna i Figur 2 och enligt givna sannolikheter för överföring.

De olika tillståndens beteckningar står för:

- S : Tillståndet för de mottagliga individer i populationen som ej blivit infekterade.
- E : Tillståndet för de infekterade individer som befinner sig i inkubationstadiet i ebolas infek-

tionscykel. De smittar inte ännu och uppvisar inga sjukdomssymptom.

- I : Tillståndet för de individer som är infekterade och är i smittsamhetsstadiet i ebolas infektionscykel.
- H : Tillståndet för de infekterade individer som förts till sjukhus och därför inte smittar.
- UB : Tillståndet för de avlidna individer som förts till osäker begravning har hög smittsamhet.
- RSB : Tillståndet för de individer som förts till säker begravning, återhämtat sig eller förts bort från UB efter en tid av smittsamhet. Tillståndet är absorberande och alla individer som lämnat S kommer till slut hit. En återhämtad ebolainfekterad antas bli immun mot att bli återsmittad samtidigt som den inte längre kan smitta någon mottaglig. Detta leder till att de ej återförs till S

I tidsteg $t = 0$ genereras en population av storlek N av fullt mottagliga individer som placeras i tillstånd S . En av dessa individer blir sedan slumpmässigt utvald. Vi låter infektera denna individ från tillstånd S och placerar den sedan i tillstånd I . Smittprocessen startar i tidsteg $t = 0$ med att individen infekterar individerna i S både globalt och i sitt egna hushåll enligt givna smitt sannolikheter. Det vill säga, vid $t = 0$ är populationskompositionen

$$S_0 = N - 1, \quad E_0 = 0, \quad I_0 = 1, \quad H_0 = 0, \quad UB_0 = 0, \quad RSB_0 = 0$$

Om någon eller några individer blir infekterade i ett tidsteg t så är de i inkubationstillståndet E i tidsteg $t + 1$. Dessa individer förs sedan vidare i kedjan enligt samma resonemang. Under varje tidsteg t överförs ett binomialfördelat ($1, 2, 8$ i Figur 2) eller multinomialfördelat ($\{3, 4, 5\}, \{6, 8\}$ i Figur 2) antal individer från tillstånden till det som slumbestämde. Smittprocessen upphör när

$$S_t = N - k, \quad E_t = 0, \quad I_t = 0, \quad H_t = 0, \quad UB_t = 0, \quad RSB_t = k$$

där $I_0 \leq k \leq N$.

Populationens storlek $N = S_t + E_t + I_t + H_t + UB_t + RSB_t$ antas vara fixerad under hela epidemiutbrottet till dess att ebolaspridningen dör ut. Populationen antas sluten där demografiska faktorer så som födselar, dödsfall relaterat till annat än ebolainfektion, immigration samt emigration inte tas hänsyn till. Förenklingen gör modellen mer lättarbetad. Om vi gör antagandet att ett ebolautbrott går relativt fort, vilket ebolas snabba infektionscykeln argumenterar för, så bör inte antagandet om sluten population ha en betydande påverkan på resultaten. För smittmodeller som avspeglar smittspridning som har en mer endemisk karaktär kan det vara viktigare att implementera dessa bakomliggande demografiska faktorer. [2] Då N är fixerad under hela smittprocessen betyder det att vi inte avräknar avlidna till följd av ebola från populationstorleken N . Vi gör det antagandet till följd av att vi inte tror att en levande individ ersätter kontakten med den avlidne med en ny kontakt där smitta kan ske. Om antagandet är rimligt eller inte är diskuterbart men vi tror att avlidna kontakter inte ersätts, åtminstone inte fullt ut, med nya levande kontakter inom det korta tidsperspektivet som ett eventuellt ebolautbrott äger rum. Ett annat alternativ kunde varit att välja någon lämplig vikt som justerar antalet förväntade kontakter en levande individ har när en ebolainficerad kontakt avlider.

I övrigt är sannolikheten för en individ i tidsteg t att hoppa från det nuvarande tillståndet den befinner sig i till nästa tillstånd oberoende av vilka tillstånd individen befunnit sig i innan tidsteg t . Den definierade ebolakedjan uppfyller därmed markovegenskapen. En individs överföring i tidsteg t kan inte ske på annat sätt än vad pilriktningarna i Figur 2 illustrerar.

Modellkommentarer:

- Smittmodellen kan antingen definieras i *diskret* eller *kontinuerlig* tid. Ovan formulerad modell är tidsdiskret och där ett tidsteg definieras som exakt ett dygn. Detta relateras i sin tur till de antaganden vi gör vad gäller durationer en infekterad befinner sig på sjukhus, duration den avlidne befinner sig på osäker begravning samt till den inhämtade datan av durationer för ebolas infektionscykel. I diskret tid sker (eller sker inte) händelser vid ett distinkt tidsteg. En smittprocess är i verkligheten av mer kontinuerlig karaktär, där de händelser så som smittöverföring sker när som helst i tid men med en förväntad intensitet. Dessa intensiteter

kan dock vara tidsberoende vilket ytterligare kan komplicera modellen. Ett konkret exempel är smittspridning av ebola där intensiteten för när smittokontakter sker kan tänkas vara högre dagtid än under nattens timmar. Att definiera modellen som tidsdiskret förenklar simulering och antas vara en tillräckligt god approximation av verkligheten för de resultat vi eftersöker. De tillgängliga epidemiologiska data man utgår från är sällan redovisade i kontinuerlig tid utan snarare i diskret utförande, till exempel på dagsbasis eller till och med vecko- eller månadsbasis. Modellering av sådan data är nödvändigtvis inte bättre med en tidskontinuerlig modell än en tidsdiskret.

- Modellen för smittspridning kan definieras som *deterministisk* eller *stokastisk* och där vi, som framkommit, använder den *stokastiska*. De *deterministiska* modellerna utesluter slump vilket betyder att en epidemi inte heller kan dö ut på grund av slumpmässighet. [2] Smittspridningen definieras ofta i kontinuerlig tid och med differentialekvationer. [2] Exempel på en begränsning med en deterministisk modell är att den inte kan fånga den spridning av smittspridningsutfallen som finns i verkligheten på grund av de slumpmässiga faktorerna. Dessa effekter återspeglas i de stokastiska modellerna där en simulering av smittoprocess kan leda till att smittspridningen dör ut trots att den förväntade smittoreproduktionen mellan smittobärarna är över 1. De slumpmässiga effekterna kan vara särskilt inflytelserika när man studerar smittspridning med initialt få smittbärare och i små populationer.[2]

2.3 Tillstånd och överföringsfördelning

I det här avsnittet redogör vi för fördelningarna av antalet individer som överförs mellan tillstånden per tidssteg. I simuleringarna av modellen använder vi oss av dessa fördelningar.

Följande tabell 1 innehåller beteckningar för överföringssannolikheter och smittintensiteter.

Beteckning	Beskrivning
$p_{SE,I}/N$	Smittsannolikheten per tidssteg från en infekterad i I för mottaglig individ i S
$p_{SE,U}/N$	Smittsannolikheten per tidssteg från en infekterad i U för mottaglig individ i S
p_L	Lokala smittsannolikheten per tidssteg från infekterad i I mot sina hushållsmedlemmar
q_{EI}	sannolikheten för en individ, per tidssteg, att överförs $E \rightarrow I$
q_{IR}	sannolikheten för en individ, per tidssteg, att överförs $I \rightarrow RSB$
q_{IU}	sannolikheten för en individ, per tidssteg, att överförs $I \rightarrow UB$
q_{IH}	sannolikheten för en individ, per tidssteg, att överförs $I \rightarrow H$
q_{HR}	sannolikheten för en individ, per tidssteg, att överförs $H \rightarrow RSB$
q_{HU}	sannolikheten för en individ, per tidssteg, att överförs $H \rightarrow UB$
q_{UR}	sannolikheten för en individ, per tidssteg, att överförs $UB \rightarrow RSB$

Tabell 1: sannolikhetsbeteckningar och smittintensitetsbeteckningar

Susceptible (S):

För att en mottaglig individ i något tidssteg t ska flytta från tillstånd S krävs att den blir globalt infekterad av en individ från tillstånd I eller globalt infekterad av en avliden individ i UB . En reflektion angående de avlidna i UB är att det kan finnas ett relevant beroende mellan den avlidne som förts till UB och dennes hushållsstorlek. Antalet individer som besöker och hanterar en avliden som befinner sig i tillstånd UB kan exempelvis tänkas vara positivt korrelerat med antalet hushållsmedlemmar den avlidne har. Vi har dock valt att bortse från detta eventuella beroende och begränsar modellen till att avlidna individer i UB endast infekterar globalt och likafördelat mot alla mottagliga individerna i S .

I det fall att individen i S har minst en infekterad hushållsmedlem så är den även exponerad för

den lokala smittsamheten. Relaterat till detta har en restriktion införts i modellen. När en mottaglig individ i ett hushåll blir globalt infekterad så tar vi bort möjligheten för resten av hushållet att bli globalt infekterade. Anledningen är främst för att simulering av modellen blir mindre resurskrävande och att man ska kunna relatera beräkningen av virusreproduktionen under avsnitt 3.5 till simuleringsresultaten. Det mer verklighetsspeglade alternativet att inte exkludera ett hushålls medlemmar från möjligheten till att bli globalt smittade bör inte förändra resultatet vad gäller risken för att en epidemi överhuvudtaget uppstår. Detta eftersom populationen av storlek N är stor med stort antal hushåll m vilket gör att sannolikheten för att ett hushåll i initialskedet skall bli återsmittat globalt är liten. En nedsida är att simuleringsresultaten resulterar i lägre slutstorlekar av eventuella epidemier än vad slutstorlekarna hade varit om vi accepterat mer än en global infektion per hushåll. En generell redogörelse över effekten på slutstorlekar och risk för epidemi på grund av restriktion görs senare i avsnitt 4.2. Det exkluderade hushållets hushållsmedlemmarna kan dock globalsmitta andra individer i populationen på samma sätt som tidigare. En individ i tillstånd S kan i ett tidssteg t antingen smittas av individerna i I , av avlidna i UB eller undvika smitta.

Låt oss titta på fördelningen av totala antalet globalt infekterade när den lokala smittspridningen som pågår parallellt bortses från. Vi betecknar smittsannolikheten för en infekterad individ i I som $p_{SE,I}/N$ och smittsannolikheten för en infekterad individ i UB som $p_{SE,U}/N$. Detta ger en total smittsannolikhet, i ett tidssteg, för en mottaglig individ i tillstånd S att undfly smitta från dessa två infekterade individer: $1 - p_{SE} = (1 - p_{SE,I}/N) \cdot (1 - p_{SE,U}/N)$.

Totala sannolikheten för en mottaglig individ att undfly smitta, när vi beaktar totala antalet infekterade individer I_t och UB_t , blir $1 - p_{SE}^{(t)} = (1 - p_{SE,I}/N)^{I_t} \cdot (1 - p_{SE,U}/N)^{UB_t}$.

Detta ger, i ett tidssteg t , den totala sannolikheten för en individ i tillstånd S att övergå till E :

$$p_{SE}^{(t)} = 1 - (1 - p_{SE,I}/N)^{I_t} \cdot (1 - p_{SE,U}/N)^{UB_t}$$

Detta ger att A_S^t är binomialfördelad med parameter S_t och sannolikhet $p_{SE}^{(t)}$:

$$A_{SE}^{(t)} \sim Bin\left(S_t, p_{SE}^{(t)}\right).$$

Sammanfattning:

- De globala smittsannolikheterna skalas med populationsstorleken N . Detta för att det, per tidsteg, förväntade antalet kontakter där smittoöverföring kan ske inte skall få bero på storleken av den genererade populationsstorleken N . En konkret exemplifiering är att vi tror att en individ inte har fler kontakter, där smittoöverföring kan ske, i en population med antalet 5000 individer än i en population med antalet 10 miljoner individer.
- Hur antalet och andelen förväntade infekterade i ett hushåll beror på hushållstorleken, efter att en av hushållsmedlemmarna globalsmittats, studerar vi genom ett exempel i avsnitt 3.5.
- En restriktion införs i modellen där möjligheten för ett hushålls medlemmar att bli globalsmittade tas bort när en av hushållets medlemmar blivit globalsmittad. Restriktionen underlättar simulering, bör inte ha stor påverkan på de gränser där epidemier uppstår men där slutstorlekarna på eventuella epidemier antas bli strikt mindre än utan restriktionen. En redogörelse över restriktionens påverkan finns att läsa i avsnitt 4.2.

Exposed (E):

Inkubationsdurationen är geometriskt fördelad så $T_E \sim \text{Geo}(q_{EI})$ vilket ger $E[T_E] = \frac{1}{q_{EI}}$. Sannolikheten per tidsteg t för överföring från tillstånd E till I blir q_{EI} och sannolikheten $(1 - q_{EI})$ att stanna i E . Sannolikheten är inte beroende av t utan är konstant. Den stokastiska variabeln A_E som betecknar antalet överföringar till I är binomialfördelad:

$$A_E \sim \text{Bin}\left(E_t, q_{EI}\right).$$

Infectious (I):

Från tillstånd I har vi 3 möjliga övergångar per tidsteg plus fallet att individen kan stanna i I . Överföringsannolikheten q_{IH} är för en individ i I att överföras till H , q_{IU} för en individ i I att direkt överföras till UB , q_{IR} för en individ i I att överföras till RSB och q_{II} som sannolikheten att stanna i I . Där $q_{II} + q_{IH} + q_{IU} + q_{IR} = 1$, så $q_{II} = 1 - q_{IH} - q_{IU} - q_{IR}$.

Detta ger smittsamhetsdurationen $T_I \sim \text{Geo}(q_{IH} + q_{IU} + q_{IR})$, där $E[T_I] = \frac{1}{q_{IH} + q_{IU} + q_{IR}}$. Vi låter A_{IH} , A_{IU} , A_{IR} och A_{II} vara stokastiska variablerna för antalet som under tidsteg t överförs till H , UB , RSB alternativt stannar i I . Fördelningen för överföring blir multinomialfördelad enligt tidigare resonemang under Susceptible(S).

Det vill säga:

$$(A_{IH}, A_{IU}, A_{IR}, A_{II}) \sim \text{Multinomial}\left(I_t, q_{IH}, q_{IU}, q_{IR}, q_{II}\right).$$

Smittsamheten från individerna i I redogjorde vi för i Susceptible.

Hospital (H):

Här antas smittsamheten mot S vara mycket låg, så vi betraktar den som obefintlig. Den praktiska implikationen av inkludering av H i vår modell är att den genomsnittliga smittsamhetsdurationen ute i samhälle reduceras. En individ överförs antingen till UB eller RSB eller stannar i H . Detta ger överföringsannolikheterna q_{HU} , q_{HR} , q_{HH} , där $q_{HU} + q_{HR} + q_{HH} = 1$ så $q_{HH} = 1 - q_{HU} - q_{HR}$.

Sjukhusdurationen är $T_H \sim \text{Geo}(q_{HU} + q_{HR})$ vilket ger $E[T_H] = \frac{1}{q_{HU} + q_{HR}}$. Låt A_{HU} vara stokastisk variabel för antalet som överförs till UB , A_{HR} stokastisk variabel för antalet som överförs till RSB och A_{HH} stokastisk variabel för de som stannar i H . Fördelningen blir multinomial:

$$(A_{HR}, A_{HU}, A_{HH},) \sim \text{Multinomial}\left(H_t, q_{HR}, q_{HU}, q_{HH}\right).$$

Unsafe Burial(UB):

Här beräknas de avlidna vara mycket smittsamma. De först till RSB (där de inte längre är smittsamma) med överföringsannolikheten q_{UR} per tidsteg och stannar med sannolikheten $(1 - q_{UR})$. Detta ger att smittsamhetsdurationen $T_U \sim \text{Geo}(q_{UR})$ så $E[T_U] = q_{UR}$. Låt A_{UR} vara den stokastiska variabeln för antalet som övergår till RSB vid varje tidsteg:

$$A_{UR} \sim \text{Bin}\left(UB_t, q_{UR}\right).$$

Removed eller Safe Burial(RSB):

I tillstånd RSB befinner sig de individer från populationen som har avlidit och förts till säker begravning där smittsamheten bedöms vara obefintlig eller återhämtat sig vilket också implicerar obefintlig smittsamhet. Tillståndet är absorberande, när en individ väl kommit hit så kan den inte längre bli smittad eller smittas. Av den anledningen finns ingen överföringsfördelning att redovisa.

2.4 Generering av population och hushåll

De parametrar som används anges i tabell 2.

Beteckning	Beskrivning
$\mu^{(H)}$	genomsnittlig hushållstorlek
m	totalt antal hushåll
m_k	totalt antal hushåll storlek k
h_k	andel av samtliga hushåll som är av storlek k
π_k	sannolikhet slumpmässigt dragen individ är från hushåll storlek k
μ_k	förväntat antal infekterade i hushåll storlek k
μ	förväntat genomsnittligt infekterade räknat över alla hushåll

Tabell 2: parametrar

Populationen av storlek N och dess hushållstruktur genereras på följande vis. Totala antalet hushåll fixeras till m stycken. Vi slumpar ut storlekarna enligt en hushållsdistributionen $h = (h_1, h_2, \dots, h_{10})$. Detta resulterar i m_k stycken hushåll av storlek k där $k \in \{1, 2, \dots, 10\}$.

Totala populationstorleken är $N = \sum_{k=1}^{k=10} k \cdot m_k$ individer. Efter att vi genererat populationen med dess hushållsstruktur så väljs en individ ut som vi låter infektera. Denna individ väljs genom att vi först slumpdrar en hushållstorlek k från fördelningen $\pi_k = \frac{k \cdot h_k}{\sum_{k=1}^{k=10} k \cdot h_k}$. När vi fastställt någon hushållstorlek k väljer vi sedan likformigt något av de genererade hushållen av storlek k . Slutligen låter vi smitta en individ i det hushållet. Individen placeras direkt i tillstånd I varpå smittprocessen startar.

Hushållsdistributionen h är inhämtad från data gällande Liberias fördelning av storleken på hushåll. [6] Även om detta inte är exakt representativt för den genomsnittliga hushållsdistributionen för Liberia, Guinea och Sierra Leone så antas det vara tillräckligt bra eftersom länderna är likartade.

Hushålldistribution	
Antal medlemmar (k)	% av m
1	10.0
2	10.7
3	13.3
4	15.3
5	13.9
6	11.9
7	8.5
8	5.6
9+	10.7

Tabell 3: hushålldistribution

Som kan utläsas i tabell 3 så utelämnar den inhämtade datan de individuella andelarna av totala antalet hushåll för hushållstorlekar 9 och uppåt. Vi väljer något godtyckligt sätta maximal hushållstorlek till 10 samt släta ut 10.7% andelen för hushållstorlek 9+. Vi antar att hushållstorlek 9 är 5.5% av antalet hushåll m samt hushållstorlek 10 till att vara 5.2% av antalet hushåll m . Vi sätter antalet hushåll till $m = 2000$ och slumpar sedan ut dessa enligt denna fördelning. Genomsnittliga antalet individer per hushåll blir $\mu^{(H)} = E[\sum_{k=1}^{k=10} H_k \cdot k] = \sum_{k=1}^{k=10} h_k \cdot k = 4.792$ givet vår h_k . Vi slumpar ut hushållen enligt h vilket betyder att antalet i populationen kommer variera något men där

$$E[N] = E[m \cdot \sum_{k=1}^{k=10} H_k] = m \cdot \mu^{(H)} = 2000 \cdot 4.792 = 9584 \quad (1)$$

För förväntad slutstorleken av en hushållsepidemi, där en medlem blivit globalt infekterad, i ett hushåll av storlek k inför vi nu notationen μ_k där $k \in \{1, 2, 3, \dots, 10\}$.

Som nämndes i avsnitt 3.3 implementerar vi restriktionen att ett visst hushålls individer inte kan bli

globalt infekterade i det fall att någon annan hushållsmedlem redan blivit infekterad globalt. Efter att vi infört restriktionen och ser att vi kan beräkna hushållens slutstorlekar μ_k för $k \in \{1, 2, 3, \dots, 10\}$ så kan vi betrakta vår population som bestående av hushåll som smittar andra hushåll istället för individer som smittar andra individer. [1] Vi kan referera till dessa hushåll som superindivider där dessa superindivider har olika, det vill säga inte likafördelad, smittsamhet mot övriga superindivider i populationen. Kontaktstrukturen där smitta sker kan också ses som att den reduceras till en nivå. Smittsamheten från en superindivid av storlek k , där en av hushållsmedlemmarna blivit globalt infekterad, är smittsamheten från en vanlig individ multiplicerat med slutstorleken μ_k . I det fall att läsaren vill se hur man räknar ut slutstorlek μ_k så hänvisas till appendix där exempelberäkning av specialfallet μ_4 utförs.

Utifrån detta perspektiv kan vi sedan beräkna μ som ses som den genomsnittliga smittsamheten som genereras från en superindivid när en slumpmässigt vald vanlig individ från populationen blivit infekterad.

$$\mu = \sum_{k=1}^{10} \mu_k \cdot \pi_k \quad (2)$$

där π_k definieras som tidigare, det vill säga sannolikheten att den slumpmässigt valda individen kommer från hushåll storlek k .

I nedanstående tabell 4 exemplifieras μ_k för $k \in \{1, 2, \dots, 10\}$ där smittsannolikheten per individ och tidssteg är $p_L = 0.1$ med en smittsamhetsduration på $\frac{1}{1-q_{II}} = 8$ dagar.

Hushållstorlek	Förväntad slutstorlek: μ_k	π_k	Andel infekterade
$k = 1$	1	0.021	0
$k = 2$	1.45	0.044	0.45
$k = 3$	2.08	0.083	0.54
$k = 4$	2.87	0.128	0.623
$k = 5$	3.80	0.145	0.7
$k = 6$	4.79	0.149	0.758
$k = 7$	5.88	0.125	0.813
$k = 8$	6.96	0.110	0.851
$k = 9$	8.05	0.109	0.881
$k = 10$	9.16	0.084	0.907

Tabell 4: Exempel slutstorlek hushåll

Andelen infekterade uttrycks som $\left(\frac{k-1}{\mu_k-1}\right)$, vilket översätts till andelen av de mottagliga i hushållet som smittats när hushållsepidemin är över. I exemplet noteras det intressanta men möjligen väntade resultatet att andelen, det vill säga inte bara det absoluta antalet, infekterade ökar i takt med att storlek k ökar. Ovanstående exempel ger:

$$\mu = \sum_{i=k}^{10} \mu_k \cdot \pi_k \approx 5$$

I ord betyder detta alltså att om en individ i vår population, med vår antagna hushållstruktur och ovan antagna smittsamhet och smittsamhetsduration, blir smittad så kommer denna individ i snitt generera $\mu - 1 = 4$ stycken smittade individer hänförligt den lokala smittdynamiken.

2.5 Reproduktionstalet R_0^H för hushållstruktur

Reproduktionstalet R_0 är en central storhet inom epidemiologi. [2] R_0 är ett tröskelvärde som visar om risk för stor epidemi är möjlig. Man kan uttrycka R_0 som *det förväntade antalet individer i en fullt mottaglig population som smittas av en smittsam individ under hela dennes smittsamhetsduration*. Om $R_0 < 1$ så dör smittspridningen ut innan epidemi uppstår medans om $R_0 > 1$ så föreligger alltid en risk för epidemi. [2] Noteras bör att det i de stokastiska modellerna finns en sannolikhet att smittspridningen dör ut i tidigt stadie även fast $R_0 > 1$ på grund av de slumpmässiga effekterna. [2] Detta återspeglas i simuleringsresultaten av stokastiska modeller men ej i deterministiska modeller. Ett exempel på hur R_0 beräknas i en homogent mixad kontaktstruktur på en nivå, ett smittsamhetstillstånd I , population av storlek N där $N \approx S_0$ och smittsannolikheten per tidsteg som är $\frac{p_1}{N}$ med en sannolikhet p_2 per tidsteg (där infektionsdurationen $T_I \sim Geo(p_2)$, som ger $E[T_I] = \frac{1}{p_2}$) att tillfriskna i tillstånd I är:

$$R_0 = \frac{S_0 \cdot p_1}{p_2 \cdot N} \approx \frac{p_1}{p_2}$$

I detta projekts modell så är beräkningen av reproduktionstalet, R_0^H , något mer komplicerad eftersom att vi dels har kontaktstruktur på två nivåer och dels fler än ett tillstånd av infektivitet. Reproduktionstalet R_0^H kan uttryckas som $R_0^H = \mu \cdot R_0^G$ [1], där R_0^G är reproduktionstalet i vår modell om vi endast hade definierat populationen som på en global nivå av homogent mixad kontaktstruktur. Samt μ enligt (2).

Först beräknar vi R_0^G givet modellens smittsannolikheter, smittsamhetsdurationer och sannolikheterna för en individ att överföras mellan inblandade tillstånd:

$$R_0^G = E \left[S_0 \cdot T_I \cdot \frac{p_{SE,I}}{N} + S_t \cdot T_U \cdot (q_{IU} + q_{IH} \cdot q_{HU}) \cdot \frac{p_{SE,U}}{N} \right]$$

där väntevärdet av summan av två oberoende stokastiska variabler är summan av väntevärdet av samma två stokastiska variabler, det vill säga:

$$R_0^G = E\left[S_0 \cdot T_I \frac{p_{SE,I}}{N}\right] + E\left[S_t \cdot T_U \cdot (q_{IU} + q_{IH} \cdot q_{HU}) \cdot \frac{p_{SE,U}}{N}\right] \quad (3)$$

Den *första termen* i uttrycket (3) motsvarar det förväntade antalet smittade hänförligt *en* ebolainficerad som befinner sig i tillstånd *I* om antalet mottagliga som är exponerade för smittorisk under hela den ebolainficerades smittsamhetsduration är S_0 stycken.

Den *andra termen* i uttrycket (3) motsvarar det förväntade antalet smittade hänförligt *en* avliden ebolainficerad som befinner sig i tillstånd *UB* om antalet mottagliga som är exponerade för smittorisk under hela den ebolainficerades smittsamhetsduration är S_t stycken.

Uttrycket är inte riktigt matematiskt korrekt relaterat till S_0 och S_t . I uttryckets fortsatta härledning kommer S_0 och S_t betraktas som två konstanter där vi gör antagandena att $S_0 \approx N$ och $S_t \approx N$. Detta ger $\frac{S_0}{N} \approx 1$, $\frac{S_t}{N} \approx 1$. Argument till varför vi nöjer oss med dessa förenklingar ges i kommentaren efter uttryckets härledning.

$$\begin{aligned} R_0^G &= \left(S_0 \cdot \frac{p_{SE,I}}{N} \cdot E[T_I] \right) + \left(S_t \cdot (q_{IU} + q_{IH} \cdot q_{HU}) \cdot \frac{p_{SE,U}}{N} \cdot E[T_U] \right) = \\ &= \left(S_0 \cdot \frac{p_{SE,I}}{N} \cdot \frac{1}{1 - q_{II}} \right) + \left(S_t \cdot (q_{IU} + q_{IH} \cdot q_{HU}) \cdot \frac{p_{SE,U}}{N} \cdot \frac{1}{1 - q_{UR}} \right) = \\ &= \left(\frac{S_0}{N} \cdot p_{SE,I} \cdot \frac{1}{1 - q_{II}} \right) + \left(\frac{S_t}{N} \cdot (q_{IU} + q_{IH} \cdot q_{HU}) \cdot p_{SE,U} \cdot \frac{1}{1 - q_{UR}} \right) \end{aligned}$$

och eftersom antagandena $\frac{S_0}{N} \approx 1$, $\frac{S_t}{N} \approx 1$ så fås:

$$R_0^G \approx \left(p_{SE,I} \cdot \frac{1}{1 - q_{II}} \right) + \left((q_{IU} + q_{IH} \cdot q_{HU}) \cdot p_{SE,U} \cdot \frac{1}{1 - q_{UR}} \right)$$

Detta ger reproduktionstalet med vår implementerade hushållstruktur:

$$R_0^H = \mu \cdot R_0^G \approx \mu \cdot \left[\left(p_{SE,I} \cdot \frac{1}{1 - q_{II}} \right) + \left((q_{IU} + q_{IH} \cdot q_{HU}) \cdot p_{SE,U} \cdot \frac{1}{1 - q_{UR}} \right) \right]$$

Kommentar angående R_0^G :

I uttryckets (3) första term så kan det förväntade antalet mottagliga inte antas vara S_0 stycken under alla de tidsteg som den ebolainfekterade förväntas vara smittsam, det vill säga alla de tidsteg som individen befinner sig i tillstånd I . Detta eftersom att ett visst antal mottagliga förväntas smittas under varje tidsteg och där det allmänt gäller att $E[S_0] > E[S_{t_i}]$ för $\forall t_i > 0$, givet att minst en infekterad individ befinner sig i tillstånd I vid tidsteg 0. Även i uttryckets (3) andra term gäller samma resonemang. Beteckningen S_t syftar på antalet mottagliga i tidsteget t och där notationen t syftar på det förväntade antalet tidsteg innan den ebolainfekterade i tillstånd I har avlidit och förts till tillstånd UB . För varje tidsteg den avlidne spenderar i tillstånd UB förväntas antalet mottagliga i tillstånd S strikt minska.

Vi nöjer oss med vår grova förenkling för att:

- I modellen definieras $S_0 = N - 1$
- N är ”stort” i jämförelse med värdet 1, där $E[N] = 9584$ se beräkning av $E[N]$ (1), vilket leder till $S_0 \approx N$.
- Det förväntade antalet individer i S som smittas av den ebolainfekterade, under hela dennes smittsamhetsduration i tillstånd I som vi antar är t tidsteg, är litet i jämförelse med S_0 och N . Därför antar vi $S_t \approx S_0 \approx N$.
- Det förväntade antalet individer i S som smittas av den avlidne ebolainfekterade, under hela dennes smittsamhetsduration i tillstånd UB , är även det litet i jämförelse med S_t .

Nedsidan med vår förenkling är att uttrycket inte kan generaliseras till andra fall när smittsamheten är så hög att den förväntade andelen av S_0 respektive S_t som smittas är mycket hög.

Sammanfattningsvis kan vi se att givet en lokal smittsannolikhet $p_L = 0.1$, smittsamhetsdurationen $\frac{1}{1-q_{II}}$ och den västafrikanska hushållsdistribution h 3 så visar vår utförda exempelberäkning att hushållsstrukturen ger 5 gånger högre virusreproduktion än R_0^G . Det betyder att kravet i det exemplet blir $R_0^G < 0.2$ om reproduktionen av viruset, det vill säga $R_0^H < 1$, ska bli så låg att epidemirisk utesluts.

2.6 Durationer för modellens tillstånd

I [4] hittar vi skattade durationer från ebolautbrott vilket visas i följande tabell.

Durationer	Antal dagar
Inkubationsduration	9.4
Från symptompåslag tills död ute i samhället	7.5
Från symptompåslag till sjukhus	5.0
Från intagen på sjukhus tills död	4.2
Från intagen på sjukhus tills återhämtning	11.8

Tabell 5: durationer

Dessa skattningar var utgångspunkterna i den modell vi simulerar och de är baserade på det senaste ebolautbrottet i västafrika. [4] Det ger väsentlig information om tidsdynamiken i ebolavirusets infektionscykel och hur den västafrikanska regionen hanterade viruset med sjukhusisolering. Vi ser att den skattade durationen till återhämtning är något längre än den skattade durationen till död. Det relevanta för vårt problem är emellertid hur länge den ebolainfekterade är smittsam. Vid sjukhusisolering antar vi smittsamheten som obefintlig och de som återhämtar sig förs till tillståndet RSB och där antas de heller inte smitta. Detta betyder att det inte spelar någon roll hur lång tid den ebolainfekterade som återhämtar sig befinner sig på sjukhus eftersom vi inte har antagit en maximal kapacitet av sjukhusplatser för infekterade. Ute i samhället kan denna tidsmässiga skillnad mellan död och återhämtad ha större betydelse om det påverkar smittsamhetsdurationen. Vi gör dock förenklingen att durationen till återhämtning är samma som durationen till död vilket vilket simplifierar modellen. Exempel på hur durationerna relaterar till våra sannolikheter för överföring mellan tillstånden:

- Sannolikheten för övergång, i ett tidsteg, $E \rightarrow I \implies q_{EI} = \frac{1}{9.4}$, denna hålls alltid fixt.
- Sannolikheten för övergång, i ett tidsteg, $I \rightarrow H \implies q_{IH} = \frac{1}{5}$. Eftersom vi kommer testa

simulera med olika antalet dagar från symptompåslag till sjukhus kommer denna sannolikhet varieras dock. Antalet dagar från symptompåslag tills död ute i samhället kommer dock förbli fixerad.

- Sannolikheten för övergång, i ett tidsteg, $I \rightarrow UB$ eller $RSB \implies q_{IU} + q_{IR} = \frac{1}{7.5}$ detta hålls fixt men att de olika termerna varierar omvänt proportionellt mot varandra då vi ändrar på andelen avlidna som vi låter föra till UB .
- Sannolikheten för övergång, i ett tidsteg, $H \rightarrow RSB$ eller $UB \implies q_{HU} + q_{HR} = \frac{1}{4.2}$ detta hålls fixt men respektive sannolikhet ändras enligt samma logik som punkten ovan när vi varierar andelen av de avlidna som förs till UB .

Vid intresse i hur skattningarna utförts, vad konfidensintervallen är samt övrigt relaterat se [4].

3 Simulering av modellen

3.1 Simuleringsteknik och beteckningar

Koden är skriven i språket R och simuleringen sker i programmet R. Vi utgick från en kod skriven för en enkel stokastisk SIR-modell med homogent mixad kontaktstruktur på en nivå. Vi expanderade den sedan med modellens ytterligare tillstånd E , H , RSB , UB . Sedan implementerade vi modellens heterogena kontaktstruktur på två nivåer. Koden är en direkt avbildning av den teoretiska modellformuleringen. Inga programpaket tillämpas.

Simuleringsprocessen:

- 1. Fyra stycken hushållsvektorer skapas. De 4 vektorerna har alla 2000 element. Heltalsvärdet i respektive element representerar antal hushållsmedlemmar och elementnumret är vilket hushåll det är. Elementnumrena representerar samma hushåll i de 4 vektorerna. Låt oss kalla dem vektor A, B, C och D. Vektor A innehåller de mottagliga individerna i tillstånd S där ingen hushållsmedlem blivit globalt infekterad. Vektor B innehåller de mottagliga individerna i tillstånd S där en hushållsmedlem blivit globalt infekterad. Det är denna vektor som relaterar till den restriktion, se avsnitt 3.3, vi infört med endast lokal smittmöjlighet. Vektor C innehåller de smittade individer som fortfarande befinner sig i inkubationstillståndet E . Vektor D innehåller de smittade individer som kan smitta och befinner sig i tillståndet I . Antalet som befinner sig i hushåll x är alltid (vektor A[x] + vektor B[x] + vektor C[x] + vektor D[x])
- 2. Antalet hushållsmedlemmar k i respektive element i vektor A slumpas ut enligt sannolikheterna i hushållsdistributionen h enligt tabell 3.
- 3. En mottaglig individ blir vald genom att vi slumpdrar en hushållstorlek k från fördelning π_k (2.4). Något av vektorns A element som representerar ett hushåll av storlek k slumpdras med likformig fördelning.
- 4. Den infekterade individen överförs till vektor D och, som ovan framgick, till samma hushållselement.

- 5. Vi placerar de övriga mottagliga hushållsmedlemmarna från det hushållet till vektor B för att de endast ska smittas lokalt.
- 6. *a.)* Simuleringen börjar i tidsteg $t = 0$ med att den infekterade individen i vektor D infekterar lokalt i sitt eget hushåll i vektor B och globalt mot alla hushåll i vektor A.
b.) I nästa tidsteg t när, givet att epidemin inte dör ut direkt, så generaliseras detta steg till att alla infekterade individer i vektor D riktar smittsamheten sina hushållsmedlemmar i vektor B och globalt mot alla hushåll i vektor A. Eventuella individer i vektor C förflyttas till vektor D med sannolikheten q_{EI} . De infekterade i vektor D förflyttas till tillstånd H med sannolikheten q_{IH} , till tillstånd UB med sannolikheten q_{IU} , till tillstånd RSB med sannolikheten q_{IR} . Respektive individ i tillstånd H förflyttas till UB med sannolikheten q_{HU} , till RSB med sannolikheten q_{HR} . Respektive individ i tillstånd UB överförs till tillstånd RSB med sannolikheten q_{UR} . kapitel 3.3
- 7. Punkt 6 *b.)* och nu tidsteg $t + 1$.

Allmänt om simuleringarna:

- Varje loop i simuleringen är ett tidssteg och ett tidssteg representerar en dag.
- För varje given kombination av parametervärden utfördes 350 oberoende simuleringar. Några stickprov gjordes där vi undersökte hur många simuleringar som borde vara tillräckligt givet några godtyckligt valda parametervärden. Slutstorlekarna observerades ha en ganska låg avvikelse vilket betyder att 350 oberoende simuleringar antogs vara mer än tillräckligt.
- Simuleringsutfallen av arbetets stokastiska smittmodell observeras resultera i att epidemin antingen dör ut mycket tidigt eller, om den tar fart, slutar väldigt nära genomsnittliga slutstorleken. Detta "binära" fenomen kan observeras i exempelhistogrammet 9.
- Slutstorlekarna representerar slutstorleken av de simuleringar där en epidemi verkligen blir av. För att inte snitta ut slutstorlekarna med epidemier som inte blev av (där exempelvis endast 0.5% av populationen blev ebolainfekterade innan smittspridningen dog ut) så betingade vi dataurvalet på att minst 5% av populationen måste ha blivit drabbade när utbrottet var över. Vi höll koll på slutstorlekarna så att de inte började närmade sig 5% så snittutfallet inte skulle vara påverkat åt ena eller andra hållet.
- Simuleringar görs för olika andel som förs till H , det vill säga olika värden på överföringssannolikheten q_{IH} . Vi varierar även andel som förs till UB från tillståndet I , det vill säga överföringssannolikheten q_{IU} , och olika andel som förs till UB från tillståndet H vilket betyder olika värden på överföringssannolikheten q_{HU} . Detta framgår i graferna i simuleringsresultaten.
- Dödligheten hos en ebolainfekterad har antagits till 60% både för infekterade i tillstånd H och tillstånd I
- De överföringssannolikheter som i alla simuleringar hålls fixerade är sannolikheten för att övergå från inkubationstillståndet E till det smittsamma tillståndet I , det vill säga $q_{EI} = \frac{1}{9.4}$, och sannolikheten för att övergå från det smittsamma tillståndet UB till icke-smittsamma tillståndet RSB , det vill säga $q_{UR} = \frac{1}{2}$.

Betydelse av beteckningar:

- *Slutstorlek*: Andel av populationen som har infekterats av ebola när smittspridningen dött ut. Y-axeln representerar intervallet 0% till 100% slutstorlek.
- *Andel UB*: Andelen av de ebolainfekterade som, per tidsteg, avlider och förs till tillståndet *UB*. X-axelns intervall mellan 0 till 1 representerar intervallet 0% till 100% av de avlidna. Vi antar att andelen, av de som avlider, som förs till *UB* från *I* är samma som andelen som förs till *UB* av de som avlider i *H*. Ett annat alternativ, som möjligen skulle överensstämt bättre med verkligheten, hade varit att sänka andelen som förs till *UB* om de avlidit i *H*. Vi har dock inte funnit några bra skattningar för detta och har därför antagit denna förenkling. Som förtydligande kan påpekas att med den antagna dödligheten av 60% betyder det exempelvis att $q_{HU} = 0.6$ när andel *UB* uppmäter värdet 1 på x-axeln.
- *Andel sjukhus*: Hur stor andel av de infekterade individerna som, per tidsteg, förs från *I* till *H*. Det vill säga q_{IH} där x-axelns intervall mellan 0 till 1 representerar intervallet 0% till 100%.

Vi uppmärksammar läsaren på de inkonsekventa beteckningarna ”Andel UB” samt ”Andel sjukhus” används istället för ”Andel UB” samt ”Andel H” alternativt ”Andel osäker begravning” samt ”Andel sjukhus”. När detta upptäcktes var graferna redan producerade och tiden knapp.

3.2 Simuleringsresultat

Figurerna (3, 4, 5, 6, 7, 8) presenteras i kronologisk ordning med avseende på den totala smittsamheten. Den högsta smittsamheten återfinns i första figuren. I samtliga figurer (3, 4, 5, 6, 7, 8) representerar den vänstra grafen (*a*) de genomsnittliga slutstorlekarna som funktion av andelen som, per tidsteg, förs till osäker begravning *UB*. I den högra grafen (*b*) representeras de genomsnittliga slutstorlekarna som funktion av andelen av de ebolainfekterade som, per tidsteg, förs till sjukhus *H*. I varje figur så baseras delfigurer (*a*) och (*b*) på samma simuleringsdata. Under varje figur framkommer vilka smittsannolikheter vi använder oss av.

Onaturligt beteende på slutstorlekarna på grund av införd restriktion:

Innan vi presenterar resultaten krävs en kommentar relaterat till onaturliga beteenden i slutstorlekarna i nedanstående simuleringsresultat. Som exempel kan ses i Figur 3(a) respektive Figur 4(a) att slutstorlekarna, där 0% förs till sjukhus, inte ökar trots ökad andel som förs till UB . Givet simuleringens antagna smittsannolikheter bör detta resultat inte förväntas. Ett annat exempel är i Figur 4(a), när vi antar att 88% förs till sjukhus, där vi observerar en snabb ökning av slutstorleken innan den onaturligt planar ut någonstans runt 30%. Vi har identifierat den tidigare nämnda restriktionen, se avsnitt 3.3, som anledning till dessa onaturliga resultat. När den globala smittsamheten är tillräckligt hög så smittas en hög andel av populationens 2000 hushåll snabbt. Till följd av restriktionen exkluderas dessa från möjligheten att bli globalt smittade vilket ger att den globala smittdynamiken, hänförligt smittsannolikheterna $p_{SE,I}$ och $p_{SE,U}$, tappar i effekt eftersom mycket få osmittade hushåll återstår. Den lokala smittdynamiken inom hushåll, hänförligt smittsannolikheten p_L , påverkas dock inte av restriktionen. Vi ser, när restriktionen lett till att nästan alla hushåll tagits bort från möjligheten att bli globalt smittade, att p_L har en stor betydelse för slutstorleken genom att observera resultaten i 3(a), 4(a) och 5(a). Vi kan se att slutstorleken i Figurerna 3(a) och 4(a), när 0% förs till sjukhus, i stort sett är identiska med ett värde på 80% trots att $p_{SE,U}$ halverats. Slutstorleken i 5(a) är däremot 60% när p_L halverats i jämförelse med 4(a) medans $p_{SE,U}$ och $p_{SE,I}$ antar samma värden. Antag som ett tankeexperiment att $p_L = 0$, då inses att restriktionen resulterar i $max(slutstorlek) = \frac{\text{Antal hushåll i populationen}}{\text{Totalt antal individer i populationen}}$. Allmänt inses att slutstorlekarna är strikt mindre med restriktionen än om den ej hade implementerats. Det mest relevanta med simuleringarna är emellertid att identifiera vilken grad av nämnda kontrollstrategier som är tillräckliga för att utesluta epidemirisk. Epidemirisk utesluts då $R_0^H < 1$, se avsnitt 3.5. Områdena kring brytpunkterna där $R_0^H > 1$ övergår till $R_0^H < 1$ betingas av att smittsamheten tryckts ned till sådana nivåer att andelen globalt smittade hushåll också är låg. Restriktionens påverkan i närområdet av dessa brytpunkter antas därför vara mycket liten där vilket rimligen leder till en begränsad och insignifikant influens på resultaten.

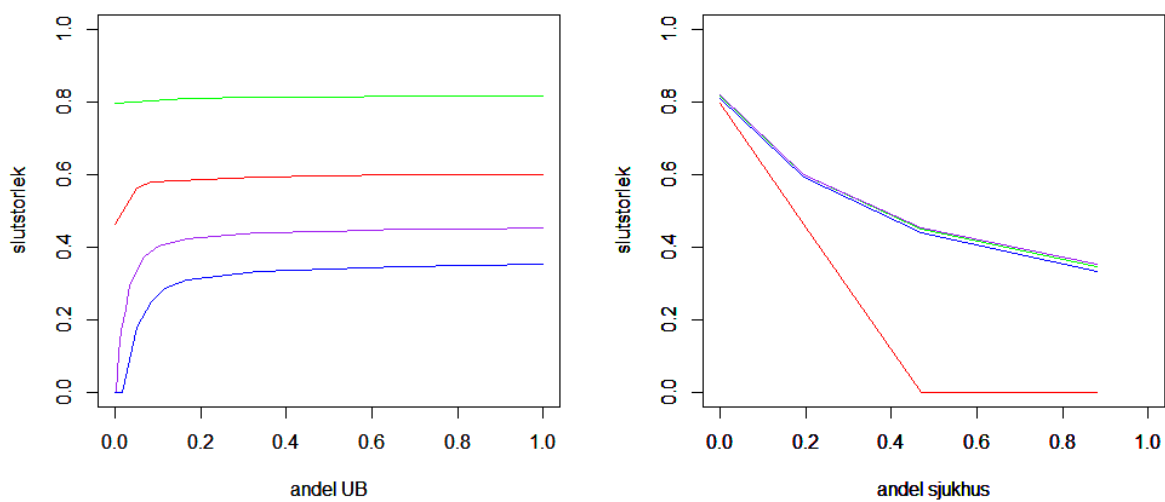
Sammanfattning av ovanstående kommentar:

- Den implementerade restriktionen i avsnitt 3.3 sänker värdet på slutstorleken av epidemiutbrotten på ett onaturligt vis. Särskilt missvisande blir slutstorlekarna om den globala smittsamheten hänförligt $p_{SE,I}$ och $p_{SE,U}$ är så hög att andelen globalt smittade hushåll stiger snabbt.
- Restriktionen har försumbar påverkan på en av den viktigaste frågeställningen vi har i detta projekt, det vill säga vilken grad av våra kontrollstrategier som krävs för att utesluta att det blir ett större ebolautbrott, det vill säga $R_0^H < 1$.

Resultatpresentation:

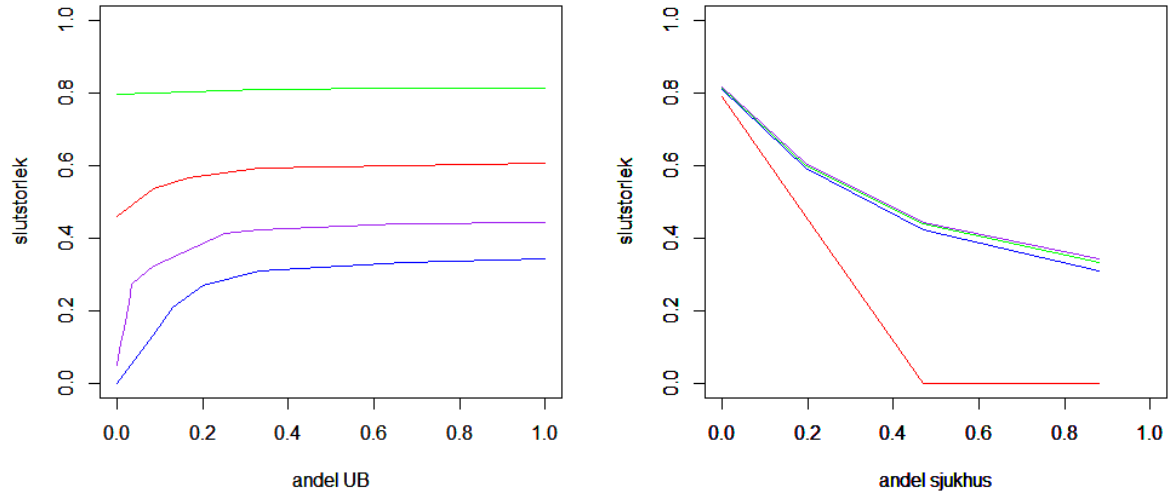
- De antagna smittsannolikheterna i Figurerna 3, 4 och 5 påvisar en liten och orealistisk möjlighet till att kontrollera smittspridningen så mycket att $R_0^H < 1$ och att epidemirisk därför utesluts. Det krävs att man nästan helt utesluter osäker begravning i samtliga av figurenas smittsamhetsantaganden för att $R_0^H < 1$ skall bli möjligt. Detta behöver samtidigt kombineras med en hög andel sjukhusisolering av ebolainfekterade. I Figur 3(a) noteras att slutstorlekarna stiger brant och utan restriktionen, som förklaras under Kommentarer om beteende på slutstorlekarna i detta avsnitt, skulle slutstorlekarna komma närmare 100% även vid relativt låga andelar av de avlidna som förs till UB. Det krävs både hög sjukhusisolering och en närmast obefintlig andel som förs till UB för att hindra en mycket stor epidemi. I Figuren 4(a) ser vi att om 47% och 88% förs till sjukhus så stiger slutstorlekarna något långsammare än i 3(a). I Figur 5(a) noteras att man vid hög andel som förs till sjukhus kan undvika epidemier om andelen som förs till UB är mycket låg. Givet dessa smittsamhetsantaganden kan det ifrågasättas om $R_0^H < 1$ i praktiken kan nås i västafrika.
- De antagna smittsannolikheterna i Figurerna 6, 7 och 8 påvisar en större möjlighet till att kontrollera smittspridningen så mycket att $R_0^H < 1$ och att epidemirisk därför utesluts. I Figur 6(a) observeras att en relativt hög tolerans för osäker begravning finns om 47% ebolainfekterade förs till sjukhus. En tolerans för osäker begravning förmodas även vid något lägre andel som förs till sjukhus. Man kan se på slutstorleken, givet att 47% ebolainfekterade förs till sjukhus, att epidemin skulle nå sin kulmen och dö ut innan majoriteten av befolkningen har blivit ebolainfekterade även om en hög andel avlidna förs till osäker begravning. I Figur 6(b) noterar vi att $R_0^H < 1$ är uppnåbart även vid en fullständig tolerans för osäker begravning, givet att andelen som förs till sjukhus är tillräckligt hög. I Figur 7(a) ser vi att om 19.5% av de ebolainfekterade förs till sjukhus så är $R_0^H < 1$ fram till att strax under 20% avlidna förs till osäker begravning. I Figur 7(b) ses att $R_0^H < 1$ är nåbart även då alla avlidna förs till osäker begravning. Slutstorlekarna faller mycket snabbt i takt när andelen sjukhusisolering ökar. När samtliga avlidna förs till osäker begravning minskar slutstorleken något långsammare. De

lägsta och sista smittsannolikhetsantagandena i Figur 8(a) visar att $R_0^H < 1$ fram till dess att strax under 40% förs till osäker begravning om 19.5% förs till sjukhus. I Figur 8(b) noteras, likt de föregående figurerna, att $R_0^H < 1$ kan uppnås oavsett hur hög andel avlidna som förs till osäker begravning bara den andel som förs till sjukhus är tillräckligt hög. Vi noterar att slutstorlekarna mycket drastiskt minskar när vi ökar andelen som förs till sjukhus.



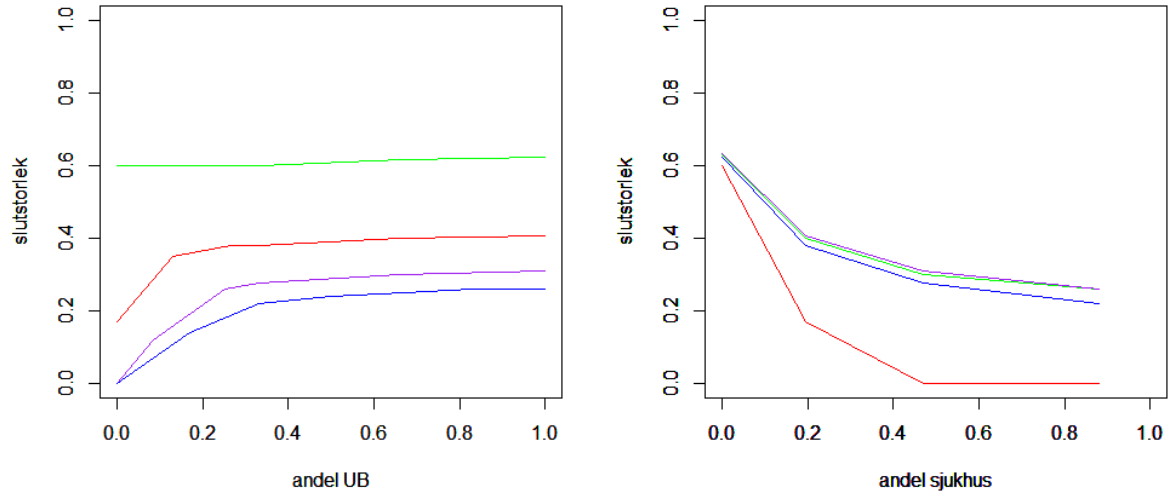
(a) Andel sjukhus: Grön: 0%, Röd: 19.5%, Lila: 47%, Blå: 88% (b) Andel osäker begravning: Röd: 0%, Blå: 33%, Grön: 66%, Lila: 100%

Figur 3: $p_{SE,I}/N = 0.2/N$, $p_L = 0.1$, $p_{SE,U}/N = 10/N$



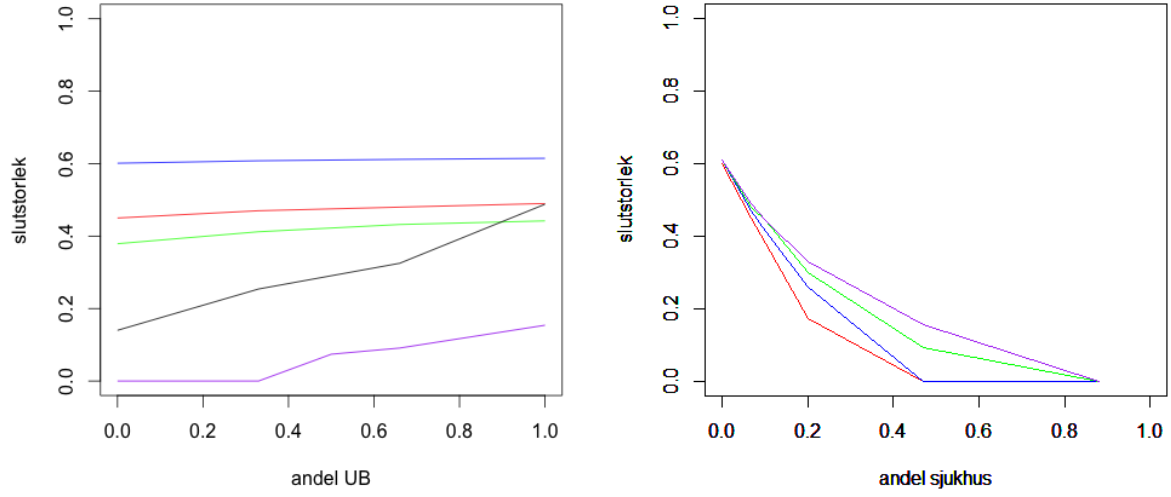
(a) Andel sjukhus: Grön: 0%, Röd: 19.5%, Lila: 47%, Blå: 88% (b) Andel osäker begravning: Röd: 0%, Blå: 33%, Grön: 66%, Lila: 100%

Figur 4: $p_{SE,I}/N = 0.2/N$, $p_L = 0.1$, $p_{SE,U}/N = 5/N$



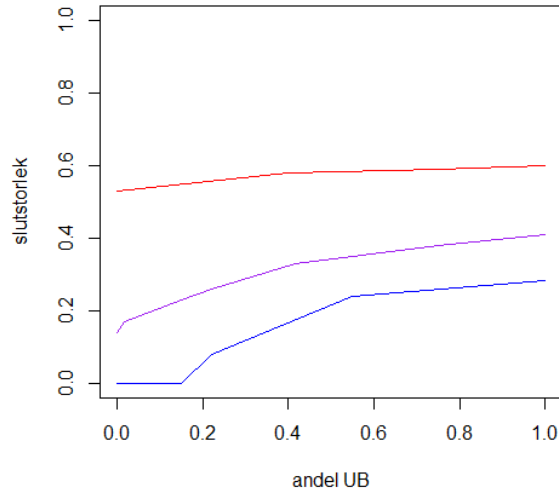
(a) Andel sjukhus:Grön: 0%,Röd: 19.5%,Lila:47%,Blå:88% (b) Andel osäker begravning:Röd:0%,Blå:33%,Grön:66%,
Lila:100%

Figur 5: $p_{SE,I}/N = 0.2/N$, $p_L = 0.05$, $p_{SE,U}/N = 5/N$

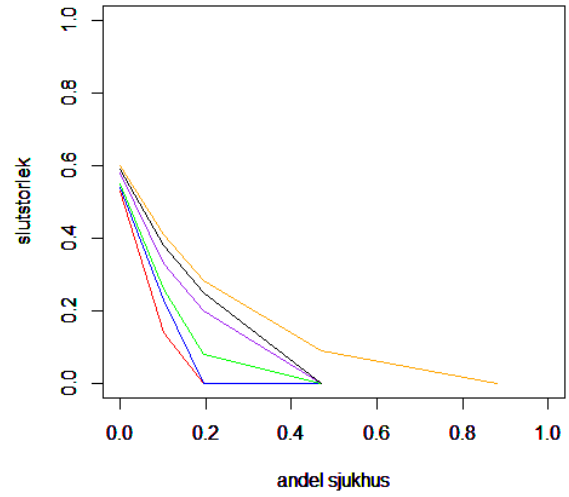


(a) Andel sjukhus: Blå:0%, Röd:6.6%,Grön:10%, Svart:19.5%,(b) Andel osäker begravning:Röd:0%,Blå:33%,Grön:66%, Lila:47%
Lila:100%

Figur 6: $p_{SE,I}/N = 0.2/N$, $p_L = 0.05$, $p_{SE,U}/N = 0.5/N$

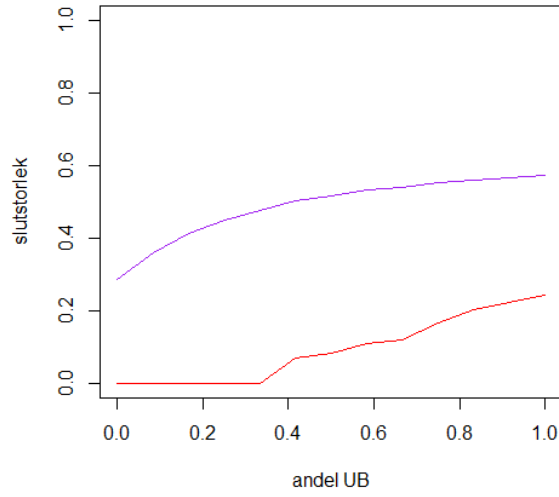


(a) Andel sjukhus: Röd:0%, Lila:10%, Blå:19.5%

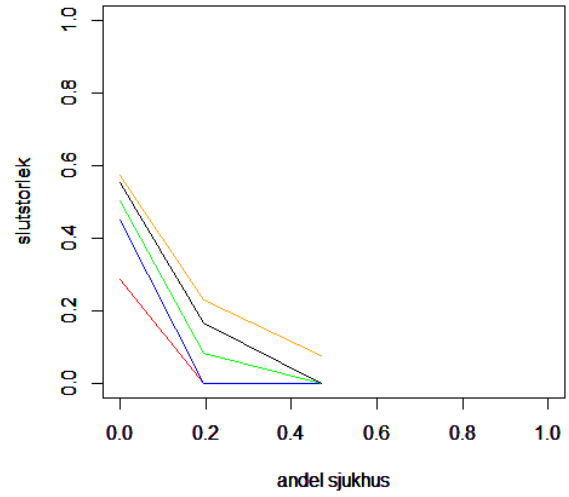


(b) Andel osäker begravning: Röd:0%, Blå:15%, Grön:22%, Lila:40%, Svart:75%, Orange: 100%

Figur 7: $p_{SE,I}/N = 0.1/N$, $p_L = 0.05$, $p_{SE,U}/N = 0.5/N$

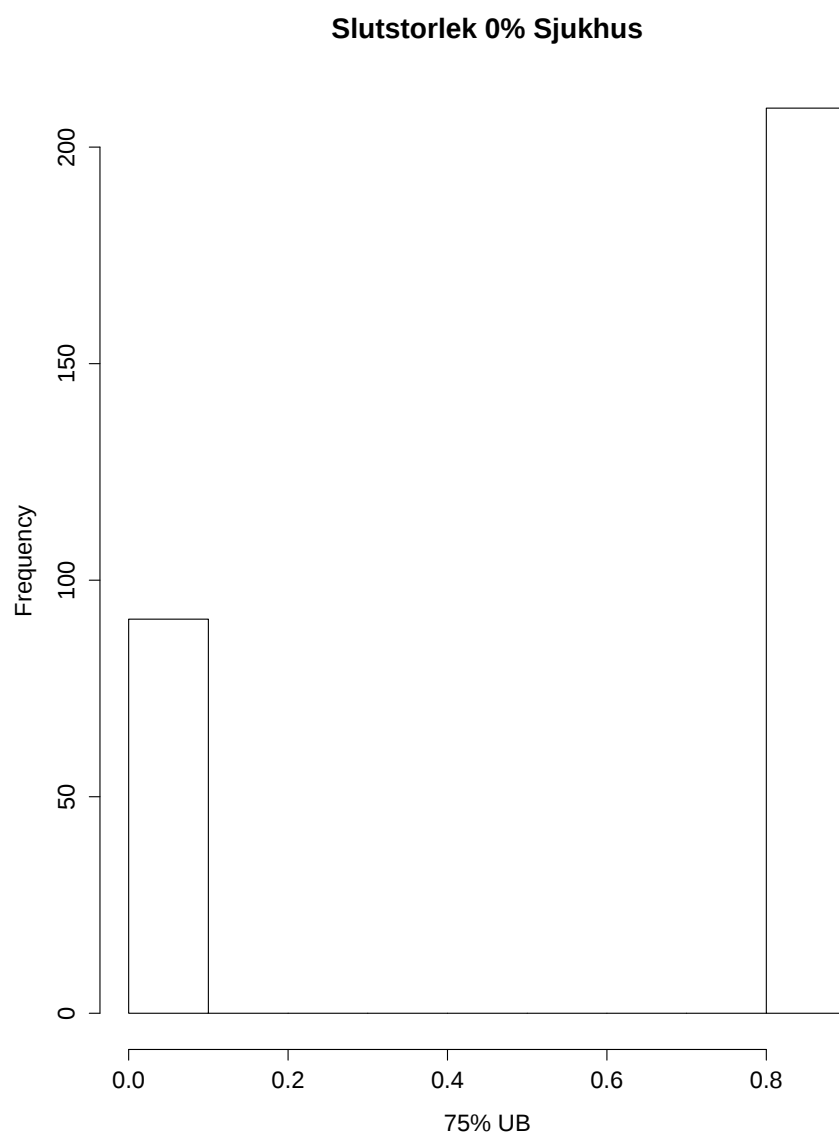


(a) Andel sjukhus: Lila:0%, Röd:19.5%



(b) Andel osäker begravning: Röd:0%, Blå:33%, Grön:50%, Svart:75%, Orange:100%

Figur 8: $p_{SE,I}/N = 0.05/N$, $p_L = 0.05$, $p_{SE,U}/N = 0.5/N$



Figur 9: Exempel på slutstorleksutfall

4 Slutdiskussion

Målet med projektet var att studera hur kontrollstrategierna *sjukhusisolering* och reduktion av *osäker begravning* påverkar epidemirisk och slutstorlekar. Vi vill besvara vilken grad av kontrollstrategierna som krävs för att helt eliminera epidemirisk givet antagna smittsannoliketskombinationer. Som förväntat, givet den modell vi formulerat, ses att implementering av kontrollstrategierna tydligt sänker epidemirisk och slutstorlekar. Simuleringsresultaten påvisar att det finns brytpunkter, det vill säga punkter med avseende på andel avlidna som förs till osäker begravning samt andel ebolainfekterade som förs till sjukhus, där $R_0^H > 1$ blir $R_0^H < 1$ och att epidemirisk därför utesluts.

- *Vilken kontrollstrategi är mest effektiv för att avvärja epidemirisk och sänka slutstorlekar?*

I Figurerna 6(b), 7(b) och 8(b) noteras att tillräckligt hög sjukhusisolering kan $R_0^H < 1$ åstadkommas oavsett vilken andel avlidna som förs till *UB*. I Figurerna 6(a), 7(a) och 8(a) noteras att vid tillräckligt låg andel som sjukhusisolerats så räcker det inte ens med att helt exkludera möjligheten för en avliden att föras till *UB*. Låg grad av andel ebolainfekterade som förs till *UB* räcker alltså inte i flera av fallen för att nå $R_0^H < 1$ när andelen sjukhusisolering är låg. Slutstorlekar faller snabbare när andelen sjukhusisolering ökar än när andelen som förs till *UB* minskar. Konklusionen är, givet våra smittsannolikhetsantaganden och smittdurationer, att sjukhusisolering av ebolainfekterade är det som mest effektivt kan undanröja en risk för epidemi samt sänka slutstorlekar.

- *Har de antagna smittsannolikheterna någon betydelse för kontrollstrategiernas effektivitet?*

Ja. En förändring av smittsannolikheten $p_{SE,U}$ påverkar smittsamheten hänförligt tillståndet *UB*. En förändring av $p_{SE,I}$ samt p_L påverkar smittsamheten hänförligt tillstånd *I*. Detta ger att kontrollstrategin där vi sänker andelen som förs till *UB* endast resulterar i en sänkt smittsamhet relaterat till smittsannolikheten $p_{SE,U}$ medans kontrollstrategin sjukhusisolering endast resulterar i en sänkt smittsamhet relaterat till smittsannolikheterna $p_{SE,I}$ och p_L . Detta betyder att det inte går att uttala sig om vilken kontrollstrategi som är mest effektiv om man inte förutsätter givna smittsannolikheter.

I allmänhet visar resultaten i de högst antagna smittsannolikheterna (Figurerna 3, 4 och 5) att det

är mycket svårt att förhindra epidemier och en mycket hög grad av kontrollstrategierna krävs. Det är troligen inte realistiskt att beredskapen i västafrika är så hög att den skulle klara av att kontrollera en ebolaspridning med dessa smittsamhetsantaganden innan en epidemi uppstår. Implementering av andra kontrollstrategier, så som karantän av alla individer som varit i kontakt med en bekräftat ebolainficerad, skulle kunna göra en kontroll av smittspridningen, så att $R_0^H < 1$ kan uppnås, mer realistisk.

Vid de lägre smittsamhetsantagandena (Figurerna 6, 7 och 8) verkar det däremot mer realistiskt att västafrika kan kontrollera en ebolaspridning innan en stor epidemi uppstår. Även i ett tidigt stadi under ebolautbrottet i västafrika, då vetskapen var låg angående riskerna relaterat till osäker begravning, så verkar det vara realistisk att kunna avvärja en förödande epidemi. Toleransen för osäker begravning blir högre även vid rimliga nivåer av sjukhusisolering.

Vi har också tittat på hur implementering av den västafrikanska hushållsstrukturen, se avsnitt 3.5, multiplicerar virusreproduktionen och därför höjer epidemirisken. I simuleringarna har våra smittsamhetsantaganden varierats kraftigt och resultaten skall inte ses som en god prediktion av verkligheten. Vi observerar att en halvering eller dubbling av smittsannolikhet har betydande påverkan på slutstorlek och huruvida epidemi uppstår. Detta visar att smittsamhetsantagandena kräver god precision för att modellen skall ge en god prediktiv kvalitet. Vi gör förenklingar i modellen som kan ge missvisande resultat så som att vi utelämnar eventuella beroenden mellan en avliden ebolainficerad och dess hushållstorlek den är hemmahörande i. Utöver det kan andra viktiga nivåer av kontaktstrukturer förekomma som vi inte tagit hänsyn till. För att sänka resursåtgången vid simulering införde vi även en restriktion, se avsnitt 3.3, som exkluderar hushåll från att bli globalt infekterade mer än en gång. Denna förenkling sänker slutstorlekarna på ett onaturligt vis men har däremot ingen signifikant påverkan på var brytpunkterna ligger vad gäller om det föreligger epidemirisk eller inte.

Referenser

- [1] F. Ball, D. Mollison, G. Scalia-Tomba. *Epidemics with two levels of mixing*.1997. The Annals of Applied Probability 7 (1), pp. 46-89
- [2] Tom Britton. *Stochastic epidemic models: a survey*.2009. Math Biosciense, 225, pp.24-35
- [3] Linda J. S. Allen. *An Introduction to Stochastic Epidemic Models*.Mathematical Epidemiology.2006. pp. 81-130
- [4] WHO Ebola Response Team. *The First 9 Months of the Epidemic and Forward Projections*.2014. New England Journal of Medicine 371(16), pp.1481-1495
- [5] Jamie J. Shah. *The Dead Bodies of the West African Ebola Epidemic: Understanding the Importance of Traditional Burial Practices*.2015. Inquiries Journal/Studentpulse volume 7 number 11, pp. 1-4
- [6] Liberia Institute of Statistics and Geo-Information Services (LISGIS), Ministry of Health and Social Welfare [Liberia], National AIDS Control Program [Liberia], and ICF International. *Liberia Demographic and Health Survey 2013*.2014. <https://dhsprogram.com/pubs/pdf/FR291/FR291.pdf> p. 20
- [7] WHO team. *Frequently asked questions on Ebola virus disease, World Health Organization*. 2016. <http://www.who.int/csr/disease/ebola/faq-ebola/en/>
- [8] WHO team. *Ebola virus disease outbreak, World Health Organization*. 2016. <http://www.who.int/csr/disease/ebola/en>

Appendices

Beräkning hushållsepidemi för exemplifierade specialfallet μ_4 :

För härledning se F. Ball. (1997)[1].

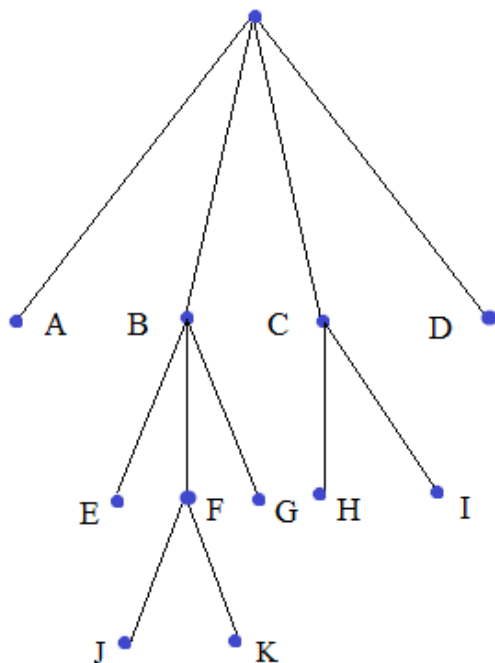
Låt förutsättning vara $S = 3$ och $I = 1$ i $t = 0$ och M_k vara stokastisk variabel för antalet infekterade i hushållstorlek k , där $\mu_4 = E[M_4]$

Först det förväntade antalet infekterade uttrycks enligt väntevärdesdefinitionen:

$$E[M_4] = \sum_{k=1}^4 P(M = k) \cdot k$$

Där $P(M = k)$ är sannolikheten för k stycken infekterade. Vi ska beräkna $P(M = 1)$, $P(M = 2)$, $P(M = 3)$, $P(M = 4)$. Eftersom $\sum_{k=1}^4 P(M = k) = 1$ så kan vi skriva sista sannolikheten som $P(M = 4) = 1 - P(M = 3) - P(M = 2) - P(M = 1)$. Först en illustrationsbild över smittförgreningen för detta hushåll bestående av $S = 3$, $I = 1$:

Figur 10: trädfördelning



A	M = 1, X ₀ = 0	B	M = 2, X ₀ = 1
C	M = 3, X ₀ = 2	D	M = 4, X ₀ = 3
E	M=2, (X ₁ X ₀ = 1) = 0	F	M=3, (X ₁ X ₀ = 1) = 1
G	M=4, (X ₁ X ₀ = 1) = 2	H	M=3, (X ₁ X ₀ = 2) = 0
I	M=4, (X ₁ X ₀ = 2) = 1	J	M=3, (X ₂ X ₁ = 1, X ₀ = 1) = 0
K	M=4, (X ₂ X ₁ = 1, X ₀ = 1) = 1		

Tabell 6: utfall

Låt den lokala smittsannolikheten vara p_L , då är sannolikheten för att en infekterad skall infektera k av S_t möjliga under ett tidsteg:

$$P(X = k) = \binom{S_t}{k} \cdot p_L^k \cdot (1 - p_L)^{S_t - k}$$

Eftersom varje generation löper över flera tidsteg t så skriver vi om $p_L = 1 - (1 - p_L)^t$, där $1 \leq t$.

Vilket resulterar i

$$P(X = k) = \binom{S_t}{k} \cdot (1 - (1 - p_L)^t)^k \cdot (1 - p_L)^{t(S_t - k)}$$

Och låt oss uttrycka $(1 - p_L) = q$, då får vi

$$P(X = k) = \binom{S_t}{k} \cdot (1 - q^t)^k \cdot q^{t \cdot (S_t - k)}$$

Och sannolikhetsfunktionen för smittsamhetsdurationen är

$$P(T = t) = (1 - p_t)^{t-1} \cdot p_t$$

Vilket ger att totala sannolikheten för att k av n skall bli infekterade är

$$\sum_{t=1}^{t=\infty} (1 - p_t)^{t-1} \cdot p_t \cdot \binom{n}{k} \cdot (1 - q^t)^k \cdot q^{t(n-k)}$$

Detta ger oss:

$$P(M = 1) = P(X_0 = 0)$$

$$P(M = 2) = P(X_0 = 1) \cdot P(X_1 = 0|X_0 = 1)$$

$$P(M = 3) = P(X_0 = 1) \cdot P(X_1 = 1|X_0 = 1) \cdot P(X_2 = 0|X_1 = 1, X_0 = 1)$$

$$+ P(X_0 = 2) \cdot P(X_1 = 0|X_0 = 2)^2$$

$$P(M = 4) = 1 - P(M = 3) - P(M = 2) - P(M = 1)$$

Och beräkning:

$$\begin{aligned} P(M = 1) &= \left(\sum_t^\infty (1 - p_t)^{t-1} \cdot p_t \cdot \binom{3}{0} \cdot (1 - q^t)^0 \cdot q^{t \cdot 3} \right) \\ P(M = 2) &= \left(\sum_t^\infty (1 - p_t)^{t-1} \cdot p_t \cdot \binom{3}{1} \cdot (1 - q^t)^1 \cdot q^{t \cdot 2} \right) \cdot \left(\sum_t^\infty (1 - p_t)^{t-1} \cdot p_t \cdot \binom{2}{0} \cdot (1 - q^t)^0 \cdot q^{t \cdot 2} \right) \\ P(M = 3) &= \left(\sum_t^\infty (1 - p_t)^{t-1} \cdot p_t \cdot \binom{3}{1} \cdot (1 - q^t)^1 \cdot q^{t \cdot 2} \right) \cdot \left(\sum_t^\infty (1 - p_t)^{t-1} \cdot p_t \cdot \binom{2}{1} \cdot (1 - q^t)^1 \cdot q^{t \cdot 1} \right) \\ &\cdot \left(\sum_t^\infty (1 - p_t)^{t-1} \cdot p_t \cdot \binom{1}{0} \cdot (1 - q^t)^0 \cdot q^{t \cdot 1} \right) + \left(\sum_t^\infty (1 - p_t)^{t-1} \cdot p_t \cdot \binom{3}{2} \cdot (1 - q^t)^2 \cdot q^{t \cdot 1} \right) \\ &\cdot \left(\sum_t^\infty (1 - p_t)^{t-1} \cdot p_t \cdot \binom{1}{0} \cdot (1 - q^t)^0 \cdot q^{t \cdot 1} \right)^2 \end{aligned}$$

$$P(M = 4) = 1 - P(M = 3) - P(M = 2) - P(M = 1)$$

Nu kan vi räkna ut $\mu_4 = \sum_{k=1}^4 P(M = k) \cdot k$

Se nu detta hushåll som en superindivider vars globala smittsamhet blir smittsamheten från en vanlig individ multiplicerat med μ_4 . Observera att den initialt infekterade inkluderas i denna beräkning.

När vi räknat ut μ_k för $\forall k \in \{1, 2, \dots, 10\}$ så kan vi sedan räkna ut det genomsnittliga antalet $\mu = \sum_{i=1}^n \mu_i \cdot \pi_i$ som smittas när vi slumpmässigt globalsmittar en vanlig individ.

När vi betraktar population bestående av dessa superindivider så transformeras kontaktstrukturen från två nivåer till endast en nivå, nu med superindivider som har en smittsamhet som beror på genomsnittligt antalet infekterade i ett hushåll. Det vill säga superindividers smittsannolikhet är inte likafördelad och kontaktstrukturen benämns som heterogent mixad. Givet detta resonemang kan vi uttrycka ett reproduktionstal relaterat till vår hushållsmodell som $R_0^H = R_0^G \cdot \mu$

