



Folkhälsomyndigheten

Skattning av peakdag och antal infekterade i covid-19-utbrottet i Stockholms län februari-april 2020



Denna titel kan laddas ner från: www.folkhalsomyndigheten.se/publicerat-material/. En del av våra titlar går även att beställa som ett tryckt exemplar från Folkhälsomyndighetens publikationsservice, publikationsservice@folkhalsomyndigheten.se.

Citera gärna Folkhälsomyndighetens texter, men glöm inte att uppge källan. Bilder, fotografier och illustrationer är skyddade av upphovsrätten. Det innebär att du måste ha upphovsmannens tillstånd att använda dem.

© Folkhälsomyndigheten, 2020.

Artikelnummer: 20054

Om publikationen

Folkhälsomyndigheten har gjort en matematisk modellering av covid-19-utbrottet i Stockholms län för att skatta tidpunkten när högst antal fall inträffar och det totala antalet infekterade personer vid olika tidpunkter. I modellen har vi delat upp infekterade individer i ”rapporterade fall” och ”obekräftade fall”. De rapporterade fallen har av vården bekräftats vara infekterade med SARS-CoV-2 och utgör antal fall som rapporterats in till Folkhälsomyndigheten mellan 17 februari och 10 april 2020. De obekräftade fallen ingår inte i statistiken och utgör det så kallade mörkertalet och har olika grad av symptom, från mycket milda till mer allvarliga. För skattning av modellen använder vi även resultat från den undersökning som genomförts i Stockholms län för att mäta aktuell förekomst av SARS-CoV-2 i samhället. Undersökning visade att 2.5% av befolkningen i Stockholms län var infekterade mellan 27 mars och 3 april 2020.

Modelleringen har utförts av medarbetare på enheten för analys under april månad 2020.

Folkhälsomyndigheten

Lisa Browsers

Enhetschef, enheten för analys

Innehåll

Om publikationen	3
Sammanfattning	5
Introduktion	5
Syfte	5
Modell för smittspridning	5
SEIR-model	5
Analys	8
Anpassning till observerad data	8
Hälsorapport Stockholms län	8
Metod	9
Resultat	10
Peakdagen och antal smittade	10
Osäkerhet i parametrar	15
Prediktion 3 och 7 dagar fram	17
Begränsningar	19
Referenser	20
Känslighetsanalys	21

Sammanfattning

Vi har modellerat spridningen av covid-19 i Stockholms län för att skatta tidpunkten när högst antal fall inträffar och det totala antalet infekterade personer givet olika antaganden. Våra resultat visar att dagen med högst antal fall i Stockholms län, den så kallade peakdagen, inträffade 15 april. Antalet samtidigt infekterade personer i Stockholms län under peakdagen var 86 000 stycken.

Vi har använt en matematisk modell för att skatta dagen med störst antal fall av covid-19. I modellen har vi delat upp infekterade individer i ”rapporterade fall” och ”obekräftade fall”. De rapporterade fallen har av vården bekräftats vara infekterade med SARS-CoV-2 och utgör antal fall som rapporterats in till Folkhälsomyndigheten mellan 17 februari och 10 april 2020. De obekräftade fallen ingår inte i statistiken och utgör det så kallade mörkertalet och har olika grad av symptom, från mycket milda till mer allvarliga. För skattning av modellen använder vi även resultat från den undersökning som genomförts i Stockholms län för att mäta aktuell förekomst av SARS-CoV-2 i samhället. Undersökning visade att 2.5% av befolkningen i Stockholms län var infekterade mellan 27 mars och 3 april.

Eftersom graden av smittsamhet bland de obekräftade fallen just nu är okänd har vi lagt in olika scenarier med olika antaganden om hur smittsamt ett obekräftat fall är i relation till ett rapporterat fall. Det scenario som ger bäst anpassning till antal rapporterade fall mellan 17 februari och 10 april 2020 är det där obekräftade fall är en tiondel så smittsamma som rapporterade fall. Detta scenario indikerar att det går 1 rapporterat fall per 1000 fall totalt, det vill säga 0.1 % av fallen rapporteras och 99.9% förblir obekräftade.

Introduktion

I denna rapport förklaras den modell som använts för att studera smittspridningen i Stockholms län samt dess anpassning till observerade dagliga antalet nya rapporterade fall i Stockholm mellan 17 februari till 10 april 2020. Vi kommer först gå igenom modellen och metoden som använts och sedan presenteras de resultat vi får.

Syfte

Syftet med modelleringen är att få en bild av det totala antalet infekterade personer i Stockholms län, mörkertalet inkluderat, och skatta när toppen inträffar.

Modell för smittspridning

SEIR-model

Vi har använt oss av en så kallad fackmodell vilket på engelska kallas compartmental model. I en sådan modell delas individer in i olika fack beroende på

förbestämda egenskaper. Inom varje fack antas individer ha samma egenskaper och bete sig likadant.

En frisk individ, och därmed mottaglig för smitta, tillhör facket S , från engelskans *susceptible*. När en frisk individ smittas så blir den inte smittsam på en gång, utan går in i den symptomfria latensfasen E , från engelskans *exposed*, och stannar där i snitt $\frac{1}{\rho} = 5.1$ dagar (1,2).¹

Vi modellerar två smittsamma grupper, de rapporterade fallen och de obekräftade fallen. De obekräftade fallen representerar det som brukar kallas mörkertalet. Vi inkluderar i modellen att de obekräftade fallen kan smitta mindre än de rapporterade fallen, detta bygger på att de obekräftade fallen inkluderar fall med väldigt milda symptom.

En infekterad individ blir efter inkubationstiden, det vill säga när symptom uppstår, antingen provtagen i vården med positivt svar eller förblir ett obekräftat fall som inte rapporteras. Om covid-19 bekräftas och rapporteras i vården flyttas individen till facket $I_r = I_{\text{rapporterad}}$. Om individen inte blir provtagen i vården utan förblir ett oupptäckt fall, flyttas individen till facket $I_o = I_{\text{obekräftat}}$. Andelen av fallen som förblir obekräftade betecknas p_o och andelen som rapporteras blir $p_r = 1 - p_o$. Smittsamheten antas variera över tid mellan två punkter i denna modellering. Mer bestämt varierar smittsamheten mellan värdet θ och värdet $\delta\theta$, där mittpunkten för vändningen mellan θ och $\delta\theta$ sker vid tidpunkten t_b . Vi kommer anta att denna vändnings mittpunkt sker $t_b = 16$ mars 2020 (dag 76 på året), dagen då Stockholmare ombads att arbeta hemifrån. Hastigheten på denna förändring på smittsamheten bestäms av parametern ε . Huruvida det är en minskning eller ökning som sker bestäms av kombinationen av ε och δ . Smittsamheten vid en tidpunkt t beskrivs enligt följande:

$$\text{Tidsberoende infektivitet } b_t = b(t, t_b, \theta, \delta, \varepsilon) = \theta \left(\delta + \frac{1-\delta}{1+e^{-\varepsilon(t-t_b)}} \right).$$

Specialfallet $\varepsilon = 0$ ger en konstant smittsamhet.

Vi antar att en individ som blir ett rapporterat fall har en smittsamhet som följer den tidsberoende infektiviteten och att smittsamheten hos obekräftade fall är en faktor $q_o \in [0,1]$ av smittsamheten hos de rapporterade fallen.

En individ antas vara smittsam i snitt $\frac{1}{\gamma} = 5$ dagar (3) och blir sedan immun och flyttas till facket R från engelskans *recovered*.²

Vi antar att populationen är stängd, ingen ny tillkommer och ingen lämnar populationen. Vi låter antalet individer i varje fack betecknas med S, E, I_r, I_o och

¹ Li et al. (1) får latensfasen/inkubationstiden till 5.2 dagar (95% CI [4.1, 7]), (2) Linton et al. uppskattar inkubationstiden till 5 dagar (95% CI [4, 5.8]).

² . I Wölfel et al. (3) studerades ett litet stickprov där de fann aktivt virus från dag 3 till dag 8 hos patienterna.

R . Det betyder att populationsstorleken är $N = S + E + I_r + I_o + R$.
Smittspridningsdynamiken beskrivs av följande ekvationssystem

$$\begin{aligned}\frac{dS}{dt} &= -S \frac{b_t I_r}{N} - S q_o \frac{b_t I_o}{N} \\ \frac{dE}{dt} &= S \frac{b_t I_r}{N} + S q_o \frac{b_t I_o}{N} - \rho E \\ \frac{dI_o}{dt} &= p_o \rho E - \gamma I_o \\ \frac{dI_r}{dt} &= (1 - p_o) \rho E - \gamma I_r \\ \frac{dR}{dt} &= \gamma (I_o + I_r).\end{aligned}$$

Från detta ekvationssystem kan man beräkna antalet S, E, I_r, I_o och R vid varje tidpunkt t , givet startvärden på antal individer i de olika facken vid önskad begynnelse tidpunkt t_0 .

Notera att det är modellens dagliga incidens av nya rapporterade fall vid tiden t , $b_t S I_r / N$, som anpassats efter dagligt antal faktiskt inrapporterade fall i Stockholms län

Det basala reproduktionstalet R_0 anger det förväntade antal personer som en första infekterad individ smittar i en helt mottaglig befolkning. R_0 varierar i denna modell över tiden t , detta då smittsamheten antas bero på tiden t , och ges av smittintensiteten, det vill säga hur många personer en smittsam individ smittar per tidsenhet, multiplicerat med den förväntade tiden den initialt infekterade individen är smittsam. I vår modell kan en infekterad individ antingen vara av den typen som kommer bli ett rapporterat fall med en smittsamhet på b_t eller ett obekräftat fall med en lägre smittsamhet på $q_o b_t$. Då andelen obekräftade betecknades p_o har vi således

$$R_0(t) = (1 - p_o) b_t / \gamma + p_o q_o b_t / \gamma.$$

För att få fram det effektiva reproduktionstalet $R_e(t)$ så multiplicerar man $R_0(t)$ med andelen mottagliga individer vid tidpunkt t , dvs

$$R_e(t) = R_0(t) S(t) / N.$$

Analys

Anpassning till observerad data

Vi anpassar parametrarna i SEIR-modellen vid tidpunkten t efter de rapporterade inhemska fallen per dag i Stockholms län³, fram till 10 april 2020. Mer bestämt, vi anpassar modellparametrarna så att den skattade incidensen vid varje tidpunkt ($= b_t SI_r/N$) liknar den observerade dagliga incidensen av rapporterade fall. Det sker en viss fördröjning av inrapportering och därmed är de allra senaste dagarna osäkra, varför vi inte tar med de fyra senaste dagarna. Den senaste dagen vi använde för inrapporterade fall var den 10 april 2020. Vi fixerar längden på smittsamheten och längden på latensperioden men skattar alla parametrar till den varierande infektiviteten, θ , δ och ε . Fall som smittats utomlands exkluderas då de inte har smittats inom den modellerade populationen. Vi använder dagen för inrapporterad symptomdebut som dagen då fallen de blev smittsamma.⁴ Då smittspårning skedde för de tidiga importfallen i Stockholms län antar vi att det inte fanns något egentligt mörkertal vid $t_0 = 17$ februari. Vi antar därför att det vid tidpunkten $t_0 = 17$ februari fanns 1 smittsam rapporterad individ men inga obekräftade fall. Ett alternativ är att sätta antalet rapporterade fall vid t_0 till 1 och antalet obekräftade fall till $p_o/p_r = p_o/(1 - p_o)$.

För att anpassa parametrarna efter data minimeras felkvadratsumman⁵ mellan det modellskattade antalet nya fall per dag och det observerade antalet nya fall per dag.

Hälsorapport Stockholms län

Folkhälsomyndigheten har genomfört en studie⁶ där 707 deltagare från en webbpanel av slumpmässigt rekryterade personer i Stockholms län genomförde självprovtagning för covid-19 mellan den 27 mars och 3 april. Vi använder resultat från denna studie för att kalibrera vår modell och få en uppskattning av mörkertalets storlek.

Totalt var 18 av de 707 testade personerna positiva i studien. Viktade andelen positiva var 2.5% (95% KI [1.4% - 4.2%]). Om vi antar att 2.5 % av Stockholmarna var covid-19-positiva mellan dessa datum betyder det att totalt 60 455 (95% KI [33 244, 99 731]) personer i Stockholms län var positiva under denna period. Vi kalibrerar parametrarna p_o och q_o (andelen av smittade som är obekräftade fall och den faktor som beskriver hur mycket obekräftade fall smittar i förhållande till rapporterade fall) på så sätt att vi får en skattad medel-prevalens på ca 2.5% under 27 mars och 3 april.

³ Stockholms län har en population på 2 374 550 individer (SCB, 2019)

⁴ Om *symptomdebut* inte angivits används *rapporteringsdag*

⁵ *residual square error*

⁶ Se: <https://www.folkhalsomyndigheten.se/nyheter-och-press/nyhetsarkiv/2020/april/resultat-fran-undersokning-av-forekomsten-av-covid-19-i-region-stockholm/>

Metod

Vi anpassar den matematiska SEIR-modellen efter olika antaganden om värden för parametrarna p_o , andelen som de obekräftade fallen utgör, och q_o , faktorn som beskriver hur mycket obekräftade fall smittar i förhållande till rapporterade fall.

Vi studerar tre olika scenarier där de tre scenarierna är kalibrerade så att den skattade prevalensen av smittsamma individer blir ca 2.5% i medel mellan 27 mars och 3 april. Det leder till de tre kombinationerna $(p_o, q_o) = \{(0.99, 1), (0.995, 0.55), (0.999, 0.11)\}$. För dessa tre scenarier skattar vi modellparametrarna och får fram deras standardfel. För att sedan få ett parametriskt bootstrap-konfidensintervall (KI) på t.ex. kurvan av antalet nya rapporterade fall per dag så gör vi 1000 dragningar av modellparameterar med deras skattade värde som medelvärde och deras standardfel som standardavvikelse ur en normalfördelning och beräknar kurvan via differentialekvationerna. Metoden används för att beräkna konfidensintervallen till samtliga resultat som presenteras i våra tabeller.

I nästa steg tar vi hänsyn till osäkerheten i p_o och q_o , samt till osäkerheter i de parametrar som beskriver inkubationstiden (ρ) och smittsamhetens längd (γ). Vi utför denna analys för två av våra scenarier, de två scenarierna där $(p_o, q_o) = \{(0.995, 0.55), (0.999, 0.11)\}$. Vi använder 100 olika parameteruppsättningar och skattar det dagliga antalet nya rapporterade fall, där de fyra parametrarna (p_o, q_o, ρ och γ) i varje parameterkombination dras från en fördelning som symboliserar osäkerheten i parametern. På detta sätt kan vi studera hur mycket skattningarna påverkas av osäkerheten i dessa parametrar. För andelen smittade som är obekräftade, p_o , gör vi enligt följande: vi drar 100 observationer från en normalfördelning med väntevärde 0.995 respektive 0.999 för de två olika scenarierna och med en standardavvikelse på 0.01. För q_o , den faktor som beskriver hur mycket obekräftade fall smittar i förhållande till rapporterade fall, antar vi också en normalfördelning med väntevärde 0.55 respektive 0.11 och standardavvikelse 0.01. Vi antar att $1/\gamma$ är gammafördelad med väntevärde 5 dagar och standardavvikelse 0.85 dagar (motsvarar ett 95% KI på [3.4, 7]) och även att $1/\rho$ är gammafördelad men med förväntat värde 5.1 dagar och standardavvikelse 0.37 dagar (vilket motsvarar ett 95% KI på [4, 6.4] (1,2)).

Vi avslutar med att studera hur väl prediktionen lyckas skatta antal nya fall de nästkommande tre och sju dagarna. Detta görs genom att använda data på antal fall fram till 7 april för att skatta kurvan av antal nya fall, och därefter studera hur väl den skattade kurvan passar till de 3 nästkommande dagarnas antal nya fall. Sedan använder vi data fram till 3 april, skattar kurvan, och studerar hur väl kurvan passar till de 7 nästkommande dagarna.

I en känslighetsanalys som finns som bilaga använder vi $p_o \in \{0.1, 0.5, 0.95\}$ och $q_o \in \{0.1, 0.5, 1.00\}$.

Resultat

Utifrån data över rapporterade fall i Stockholms län fram till den 10 april har vi tagit fram tre scenarier som alla ger en utveckling som stämmer överens med antagandet om en prevalens på ca 2.5% mellan 27 mars och 3 april, den prevalens som Stockholmsstudien visade. De tre scenarierna skiljer sig åt avseende andel obekräftade fall av totalt antal fall (p_o) och avseende hur smittsamma de obekräftade fallen är i förhållande till de rapporterade fallen (q_o). Gemensamt för scenarierna är att de obekräftade fallen måste utgöra minst 99% av samtliga fall för att ge den observerade prevalensen.

	Andel obekräftade fall i procent	Smittsamhet hos obekräftade fall i procent
Scenario 1	99	100
Scenario 2	99,5	55
Scenario 3	99,9	11

Scenario 1 ger en medelprevalens på 2.3% mellan 27 mars och 3 april medan scenario 2 och 3 ger 2.5%. Vi har valt att behålla scenario 1 då det ger oss ett minimum på hur stort mörkertalet måste vara för att få en utveckling av antalet fall för att likna antalet smittade i Stockholmsstudien.

Peakdagen och antal smittade

Resultat för ackumulerat antal och andel infekterade, samt peakdag för de tre olika scenarierna presenteras i tabell 1. Vi ser att enligt de tre olika modellerna skattas peakdagen för antal fall av covid-19 i Stockholms län i perioden 15-18 april.

I bilder 1, 2 och 3, visas till vänster den skattade utvecklingen av nya rapporterade fall, *daglig incidens*, medan den högra grafen visar det totala antalet samtidigt infekterade och smittsamma individer, *prevalens*. Perioden som visas är 17 februari till 25 april, bilderna innehåller alltså även prognos två veckor framåt i tiden.

Utifrån inrapporterade fall under perioden 17 februari till 10 april beräknar vi värdet för reproduktionstalet vid en viss tid, $R_e(t)$. I samtliga scenarier ligger $R_e(t)$ den 10 april på ett värde strax över 1. Se tabell 2 för reproduktionstalen med 95% konfidensintervall.

För de tre scenarierna presenterar vi parameterskattningarna samt RSS⁷ i tabell 3, där framgår att för de observerade fallen ger scenario 3 den bästa anpassningen till data, med antaganden om 99.9% obekräftade fall med en smittsamhet på 11% av de bekräftades smittsamhet. Enligt scenario 3 inträffade peakdagen 15 april (95%

⁷ Felkvadratsumman, *residual square error*

KI 9 april – 18 april) och att antalet samtidigt infekterade personer i Stockholms län under peakdagen var 86 000 stycken (95% KI 50 100 – 140 000).

Tabell 1: Ackumulerat antal och andel smittade vid två tidpunkter, samt skattad peakdag och antal smittade och smittsamma den dagen (prevalens). Detta visas för de tre scenarierna med olika andel obekräftade fall (p_o) och olika smittsamhet hos ett obekräftat fall i relation till smittsamheten hos ett rapporterat fall (q_o). Alla skattningar visas med ett 95% konfidensintervall (95% KI).

	p_o	q_o Smittsamhet	Ackumulerat antal smittade och andel av Stockholms läns befolkning som varit smittade fram till				Peakdag (95% KI)	Prevalens på peakdag (95% KI)
			2020-04-11		2020-05-01			
			Ackumulerat antal (95% KI)	Ackumulerad andel (95% KI)	Ackumulerat antal (95% KI)	Ackumulerad andel (95% KI)		
Scenario 1	99.0 %	100%	402 001	0.17	693 744	0.29	2020-04-18	78 699
			[304 766, 530 207]	[0.128, 0.223]	[472 647, 979 996]	[0.199, 0.413]	[2020-04-09, 2020-04-22]	[53 906, 119 121]
Scenario 2	99.5 %	55%	445 459	0.19	757 880	0.32	2020-04-17	86 450
			[319 438, 605 577]	[0.135, 0.255]	[488 033, 1 081 008]	[0.206, 0.455]	[2020-04-07, 2020-04-19]	[55 578, 132 509]
Scenario 3	99.9 %	11%	453 262	0.19	749 782	0.32	2020-04-15	86 162
			[295 311, 651 087]	[0.124, 0.274]	[435 970, 1 132 156]	[0.184, 0.477]	[2020-04-09, 2020-04-18]	[50 873, 139 381]

Tabell 2: Skattat effektivt reproduktionstal (R_e) för sista dagen med rapporterade fall i denna analys, 2020-04-10, med 95% konfidensintervall. Detta visas för de tre scenarierna med olika andel obekräftade fall (p_o) och olika smittsamhet hos ett obekräftat fall i relation till smittsamheten hos ett rapporterat fall (q_o).

	p_o	q_o smittsamhet	$R_e(2020-04-10)$ (95% KI)
Scenario 1	99.0%	100%	1.08
			[0.904, 1.219]
Scenario 2	99.5%	55%	1.07
			[0.919, 1.216]
Scenario 3	99.9%	11%	1.05
			[0.866, 1.25]

Tabell 3: Skattningar av prevalens 27 mars – 3 april, RSS (modell med lägst värde har en bättre anpassning till data) och parameterskattningar. I parentes bredvid skattningen ges dess standardfel (se) och under dess konfidensintervall (95% KI). Detta visas för de tre scenarierna med olika andel obekräftade fall (p_o) och olika smittsamhet hos ett obekräftat fall i relation till smittsamheten hos ett rapporterat fall (q_o).

	p_o	q_o Smittsamhet	Prevalens 27 mars – 3 april	RSS (se) 95% KI	$\hat{\delta}$ (se) 95% KI	$\hat{\varepsilon}$ (se) 95% KI	$\hat{\theta}$ (se) 95% KI
Scenario 1	99.0%	100%	2.3%	19092	0.15 (0.014)	-0.19 (0.024)	1.64 (0.016)
					[0.122, 0.175]	[-0.239, -0.145]	[1.613, 1.676]
Scenario 2	99.5%	55%	2.5%	18621	0.16 (0.015)	-0.19 (0.024)	2.82 (0.031)
					[0.129, 0.189]	[-0.232, -0.138]	[2.757, 2.878]
Scenario 3	99.9%	11%	2.5%	17544	0.18 (0.023)	-0.16 (0.023)	11.68 (0.197)
					[0.137, 0.228]	[-0.203, -0.111]	[11.291, 12.064]

Bild 1: Scenario 1: vi antar att 1% av alla infekterade är rapporterade fall och resterande 99% är obekräftade fall. De obekräftade fallens smittsamhet är 100% av smittsamheten hos de rapporterade fallens. Vänster: skattad daglig incidens av rapporterade fall (röda linjen) med 95% konfidensintervall (streckade linjerna) och observerad data (cirklar). Höger: prevalens av infekterade och smittsamma fall (röda linjen) med 95% konfidensintervall (streckade linjer).

Fitted SEIR model covid-19 Stockholm
Unreported cases 99% of infected and their infectivity 100% compared to reported cases

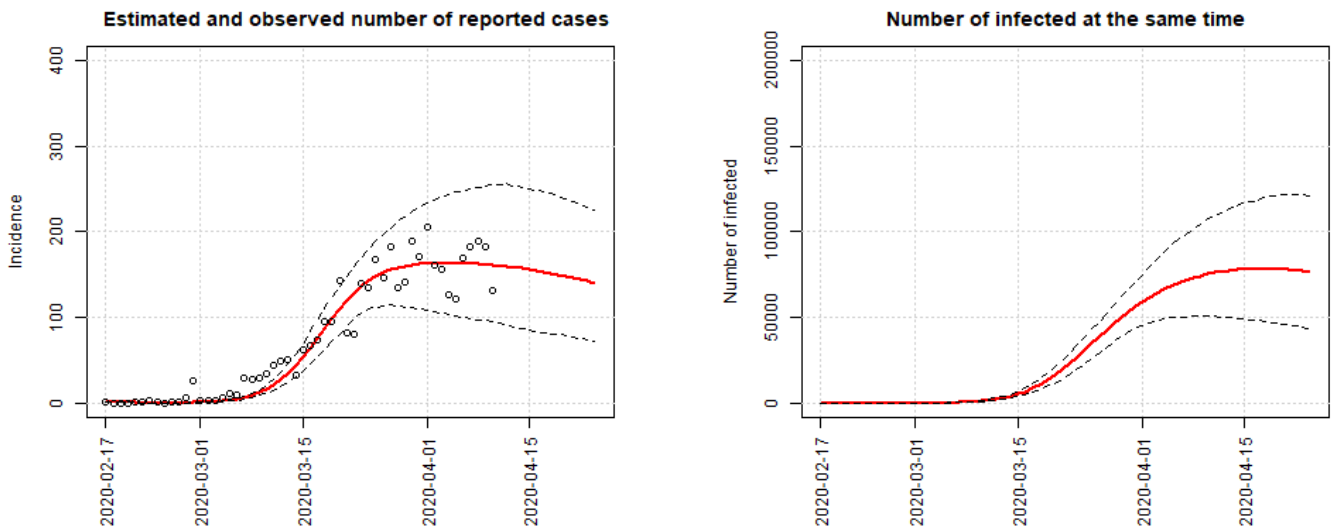


Bild 2: Scenario 2: vi antar att 0.5% av alla infekterade är rapporterade fall och resterande 99.5% är obekräftade fall. De obekräftade fallens smittsamhet är 55% av smittsamheten hos de rapporterade fallens. Vänster: skattad daglig incidens av rapporterade fall (röda linjen) med 95% konfidensintervall (streckade linjerna) och observerad data (cirklar). Höger: prevalens av infekterade och smittsamma fall (röda linjen) med 95% konfidensintervall (streckade linjer).

Fitted SEIR model covid-19 Stockholm
Unreported cases 99.5% of infected and their infectivity 55% compared to reported cases

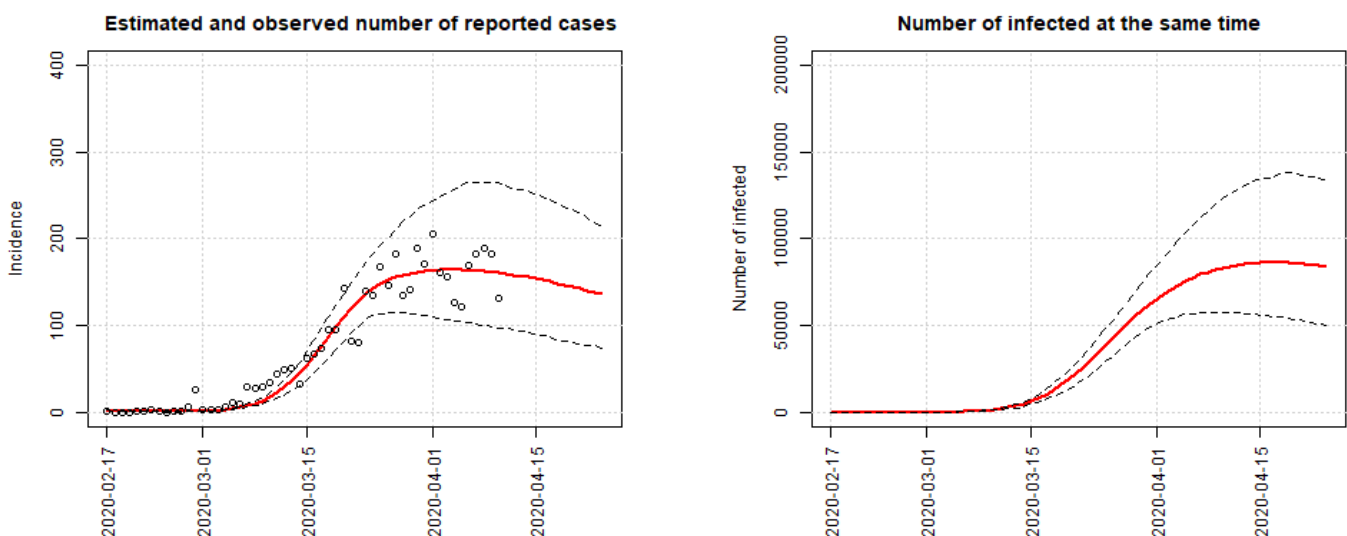
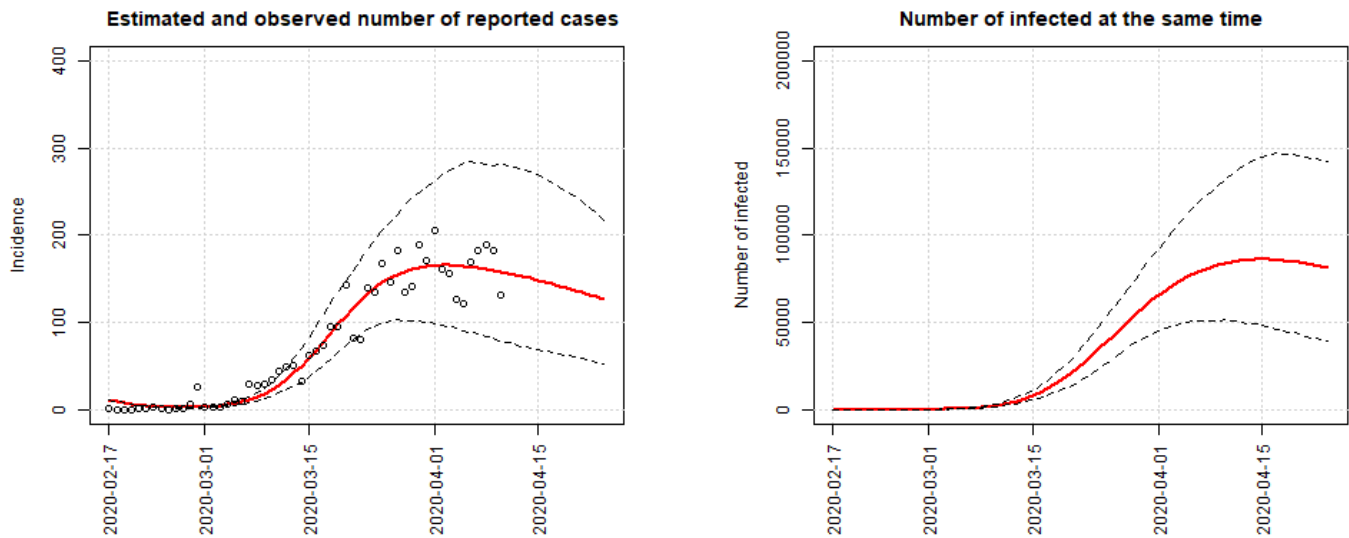


Bild 3: Scenario 3: vi antar att 0.1% av alla infekterade är rapporterade fall och resterande 99.9% är obekräftade fall. De obekräftade fallens smittsamhet är 11% av smittsamheten hos de rapporterade fallens. Vänster: skattad daglig incidens av rapporterade fall (röda linjen) med 95% konfidensintervall (streckade linjerna) och observerad data (cirklar). Höger: prevalens av infekterade och smittsamma fall (röda linjen) med 95% konfidensintervall (streckade linjer).

Fitted SEIR model covid-19 Stockholm
Unreported cases 99.9% of infected and their infectivity 11% compared to reported cases

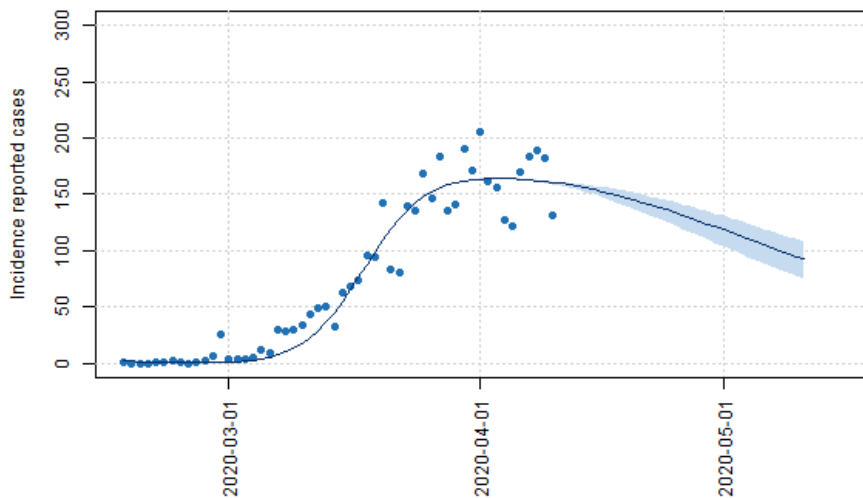


Osäkerhet i parametrar

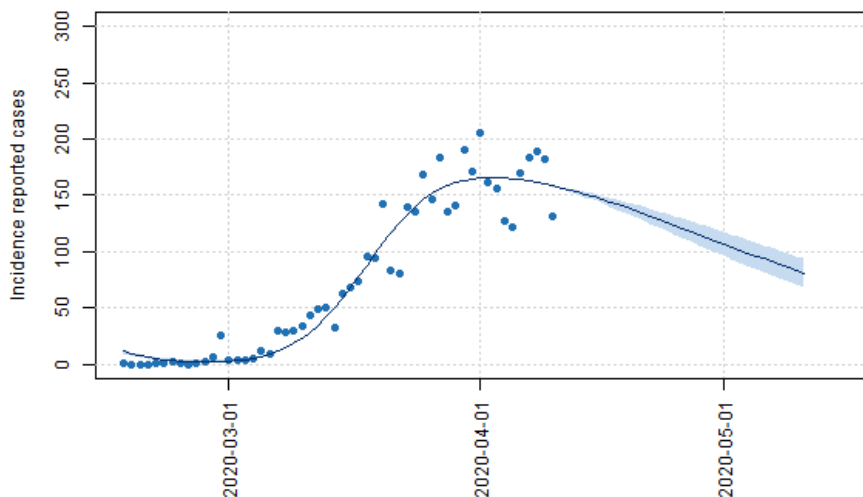
I analysen nedan tar vi hänsyn till osäkerheterna i parametrarna p_o , q_o , ρ och γ mer ingående. Vi väljer att göra det för de två scenarier, scenario 2 och 3, som bäst lyckades återskapa en prevalens på ca 2.5% mellan 27 mars och 3 april. Resultat av denna analys visas i bild 4. I den övre bilden antas de obekräftade fallens smittsamhet i medel vara 55% av smittsamheten hos de rapporterade fallen. I den undre bilden antas smittsamheten hos de obekräftade fallen i medel vara 11% av smittsamheten hos de bekräftade rapporterade fallen. Vi ser att när vi tar hänsyn till osäkerheten i smittsamhetens längd (95% KI [3.4, 7]) och i inkubationstiden (95% KI [4, 6.4]) så varierar inte resultaten avsevärt.

Bild 4: Effekten på skattad daglig rapporterad incidens av att ta hänsyn till osäkerheten i andel smittade som förblir obekräftade fall, smittsamhet hos obekräftade fall, inkubationstiden och hur länge en person är smittsam. I den övre bilden är medelandelan av alla smittade som antas obekräftade 99.5% och medelsmittsamheten hos dessa är 55% av smittsamheten hos rapporterade fallen. I den undre bilden är medelandelan av alla smittade som antas obekräftade 99.9% och medelsmittsamheten hos dessa är 11%. Kurvan visar medianvärdet av 100 parameteruppsättningar där värdet för de 4 olika parametrarna som undersöks dragits från de fördelningar som nämns i metodstycket. Arealen symboliserar ett 95% konfidensintervall.

Fitted SEIR model covid-19 Stockholm
Unreported cases 99.5% of infected and their infectivity 55% compared to reported cases



Fitted SEIR model covid-19 Stockholm
Unreported cases 99.9% of infected and their infectivity 11% compared to reported cases



Prediktion 3 och 7 dagar fram

Slutligen studerar vi hur väl våra skattade modeller baserat på antal fall fram till 7 april kan prediktera nya fall 3 dagar fram, samt hur väl våra skattade modeller baserat på antal fall fram till 3 april kan prediktera nya fall 7 dagar fram. I bild 5 ser vi att de tre modellerna lyckas lika bra att prediktera antalet nya bekräftade fall 3 dagar fram i tiden. I bild 6 ser vi att det scenario där smittsamheten hos obekräftade fall antas vara 100% av smittsamheten hos rapporterade fall sämre predikterar de 7 dagarna efter 3 april än scenarierna med 55% eller 11% smittsamhet. Vi ser att antagandet om 100% smittsamhet överskattar antalet nya fall, medan antagandet om framförallt 11% smittsamhet ger en bättre prediktion 7 dagar fram i tiden.

Bild 5: I de tre scenarierna ser vi hur väl den skattade kurvan utifrån data fram till 7 april predikterar antalet dagliga nya rapporterade fall 3 dagar därefter. Den översta bilden är för scenario 1, den mittersta scenario 2 och den nedersta scenario 3. De vertikala streckade linjerna visar var brytpunkten är för vilken data som användes i skattningarna. De punkter till vänster och fram till denna linje användes för att skatta kurvan och svarta ifyllda punkterna till höger om linjen används för att uppskatta hur väl kurvan kan prediktera nya antal fall.

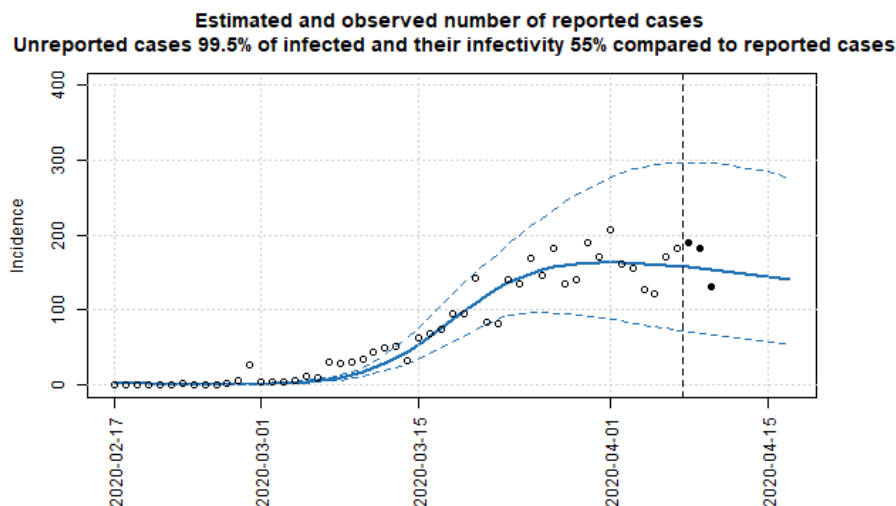
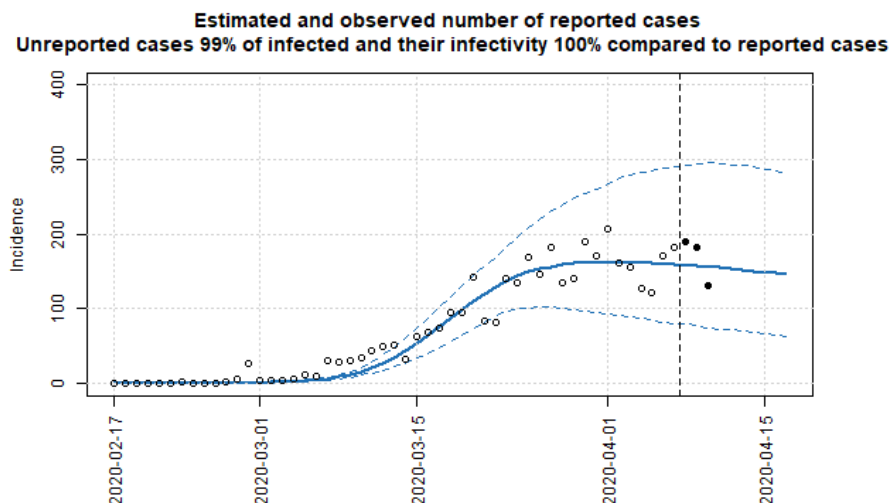
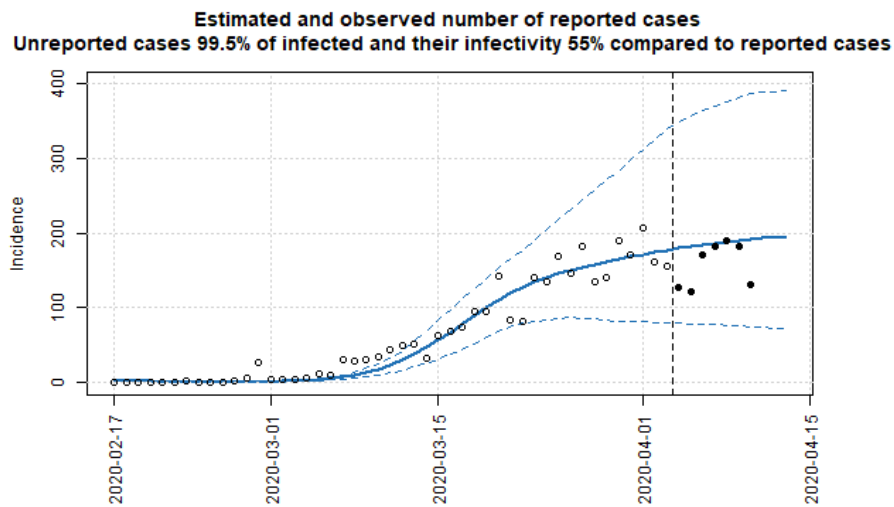
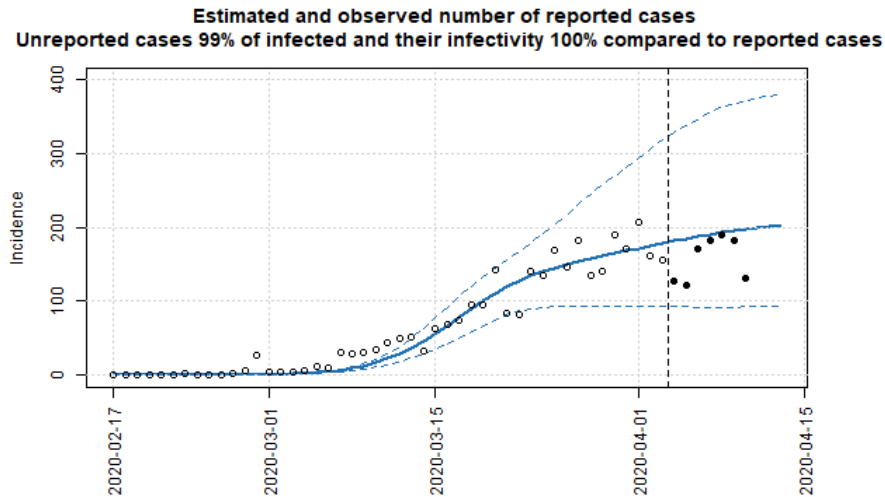
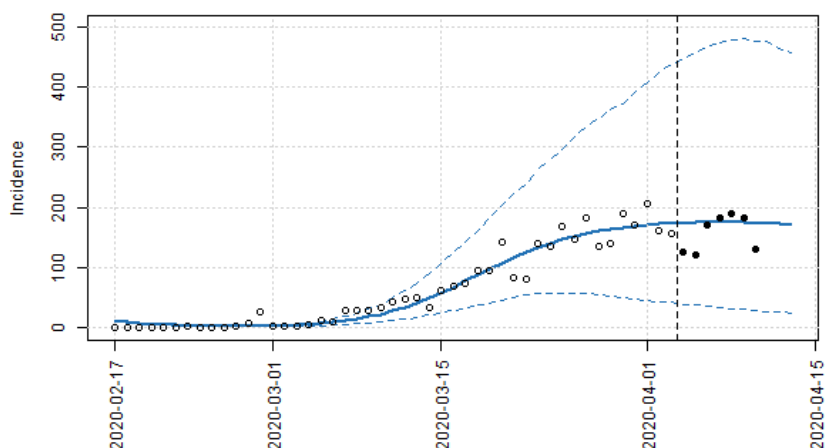


Bild 6: I de tre scenarierna ser vi hur väl den skattade kurvan utifrån data fram till 3 april predikterar antalet dagliga nya rapporterade fall 7 dagar därefter. Den översta bilden är för scenario 1, den mittersta scenario 2 och den nedersta scenario 3. De vertikala streckade linjerna visar var brytpunkten är för vilken data som användes i skattningarna. De punkter till vänster och fram till denna linje användes för att skatta kurvan och svarta ifyllda punkterna till höger om linjen används för att uppskatta hur väl kurvan kan prediktera nya antal fall.



Estimated and observed number of reported cases
Unreported cases 99.9% of infected and their infectivity 11% compared to reported cases



Begränsningar

Covid-19 smittar framförallt via kontakt mellan människor genom droppsmitta, vilket indikerar att den sociala kontaktstrukturen i befolkningen har betydelse för dynamiken. Fackmodellen tar inte hänsyn till den variation i kontakter mellan människor som finns i samhället i verkligheten, där ett fåtal individer kan ha väldigt många kontakter medan flertalet har färre kontakter, man antar istället att alla individer har lika många sociala kontakter. Denna förenkling, det vill säga antaganden om en homogen kontaktstruktur, brukar ge en något snabbare tillväxt av en epidemi än om heterogenitet inkluderas i modelleringen. Modellen riskerar därför att överskatta hastigheten på utbrottet i Stockholms län något. Detta är inte inkluderat i de angivna konfidensintervallen, då ett konfidensintervall inte kan rapportera om sådana osäkerheter. Därför är de än viktigare att nämna.

I vår modell antar vi att smittintensiteten (smittsamheten) varierar över tid men att parametrarna som styr längden på latensfasen (ρ) och hur länge man är smittsam (γ) är konstanta över tid, att de därmed styr exponentialfördelade slumpvariabler. Detta är en (vanligt förekommande) förenkling för att underlätta beräkningar men är även den väl värd att belysas. En exponentialfördelad tid är inte alltid den bästa beskrivningen av vad vi observerar, till exempel är det möjligt att variationen i latensperioden är lägre än vad en exponentialfördelad tid med väntevärde 5.1 ger.

Vi antar vidare att förhållandet mellan rapporterade fall och obekräftade fall är konstant över tid. Detta är en begränsning i modellen då rutinerna för provtagning ändrades 12 mars 2020, fokus flyttades då från att provta individer med symptom som vistats i kända riskområden till personer med behov av inläggande sjukhusvård. Vi tror inte att denna förändring påverkade rapportering av fall från inhemsk smitta avsevärt utan snarare att den påverkade inrapporteringen av de tidigare importfallen.

Referenser

1. Li Q, Guan X, Wu P, et al. "Early Transmission Dynamics in Wuhan, China, of Novel Coronavirus-Infected Pneumonia". *N Engl J Med* (2020). DOI: 10.1056/NEJMoa2001316
2. Linton NM, Kobayashi T, Yang Y, et al. "Incubation Period and Other Epidemiological Characteristics of 2019 Novel Coronavirus Infections with Right Truncation: A Statistical Analysis of Publicly Available Case Data". *Journal of Clinical Medicine* (2020). DOI: 10.3390/jcm9020538
3. Wölfel, R., Corman, V.M., Guggemos, W. et al. "Virological assessment of hospitalized patients with COVID-2019". *Nature* (2020). <https://doi.org/10.1038/s41586-020-2196-x>
4. S. Flaxman, S. Mishra, A. Gandy, et al., "Estimating the number of infections and the impact of non-pharmaceutical interventions on COVID-19 in 11 European countries," tech. rep., Imperial College London, 2020.

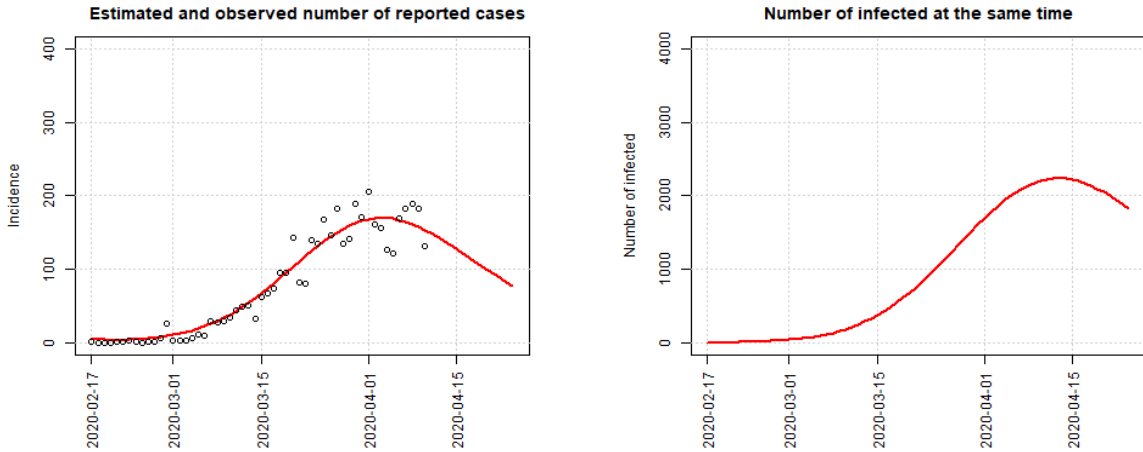
Känslighetsanalys

Vi börjar med att visa anpassning till rapporterade fall för de olika värdena på p_o och q_o , $p_o \in \{0.1, 0.5, 0.95\}$ och $q_o \in \{0.1, 0.5, 1.00\}$. De rapporterade fallen per dag syns som punkter och vår anpassning som kurvan. Därefter visar vi vad detta skulle innebära i totala antalet samtidigt infekterade över tid fram till 25 april 2020. I Bild 7 till Bild 10 ges resultaten av de scenarier som finns återgivna i Tabell 4.

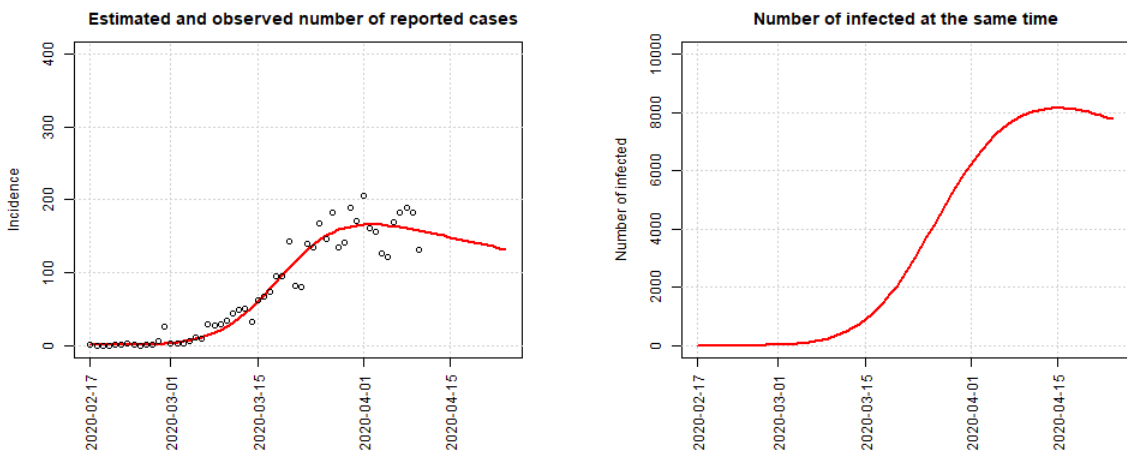
För de olika scenarierna ser vi att en första topp kommer runt 11-19 april. I Figur 7 kan man se att om andel obekräftade fall, mörkertalet, är hög (95%) samt om deras smittsamhet är hög (100% av smittsamheten hos rapporterade fall) så infaller toppen inte med en snabb minskning utan en längre plåtå. I fallen med lägre smittsamhet och lägre andel obekräftade fall sker minskningen efter toppen snabbare.

Bild 7: Scenario där 5% av alla smittade får sådana symptom att de upptäcks av vården (rapporterat fall) och 95% förblir obekräftade fall. De obekräftade fallens smittsamhet är 10% av smittsamheten hos ett rapporterat fall i den översta bilden, 50% i den mittersta och 100% i den nedersta. Den röda linjen i den vänstra bilden visar skattad daglig incidens av rapporterade fall och cirkelarna faktiska data. Den högra bilden visar prevalens av infekterade och smittsamma fall.

Fitted SEIR model covid-19 Stockholm
Unreported cases 95% of infected and their infectivity 10% compared to reported cases



Fitted SEIR model covid-19 Stockholm
Unreported cases 95% of infected and their infectivity 50% compared to reported cases



Fitted SEIR model covid-19 Stockholm
Unreported cases 95% of infected and their infectivity 100% compared to reported cases

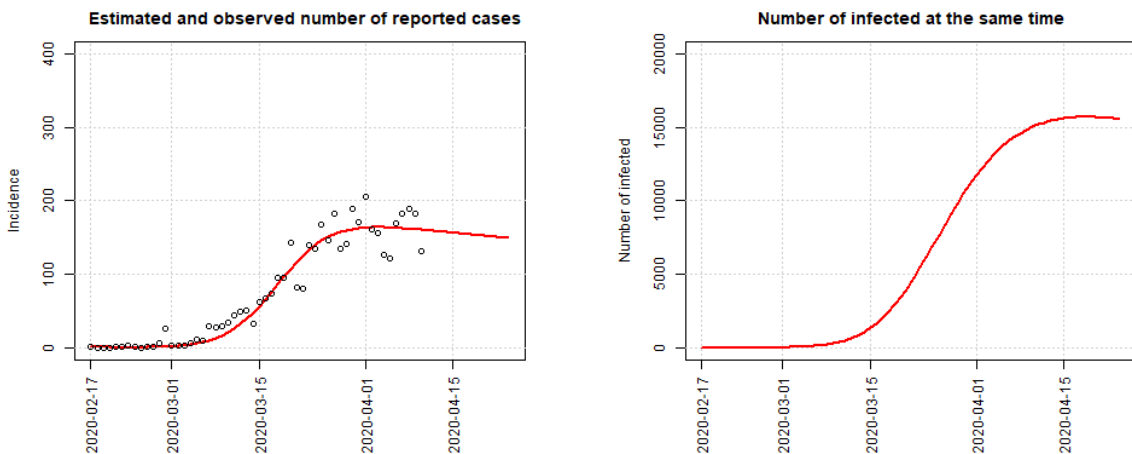


Bild 8: Scenario där 50% av alla smittade får sådana symptom att de upptäcks av vården (rapporterat fall) och 50% förblir obekräftade fall. De obekräftade fallens smittsamhet är 10% av smittsamheten hos ett rapporterat fall i den översta bilden, 50% i den mittersta och 100% i den nedersta. Den röda linjen i den vänstra bilden visar skattad daglig incidens av rapporterade fall och cirkelarna faktisk data. Den högra bilden visar prevalens av infekterade och smittsamma fall.

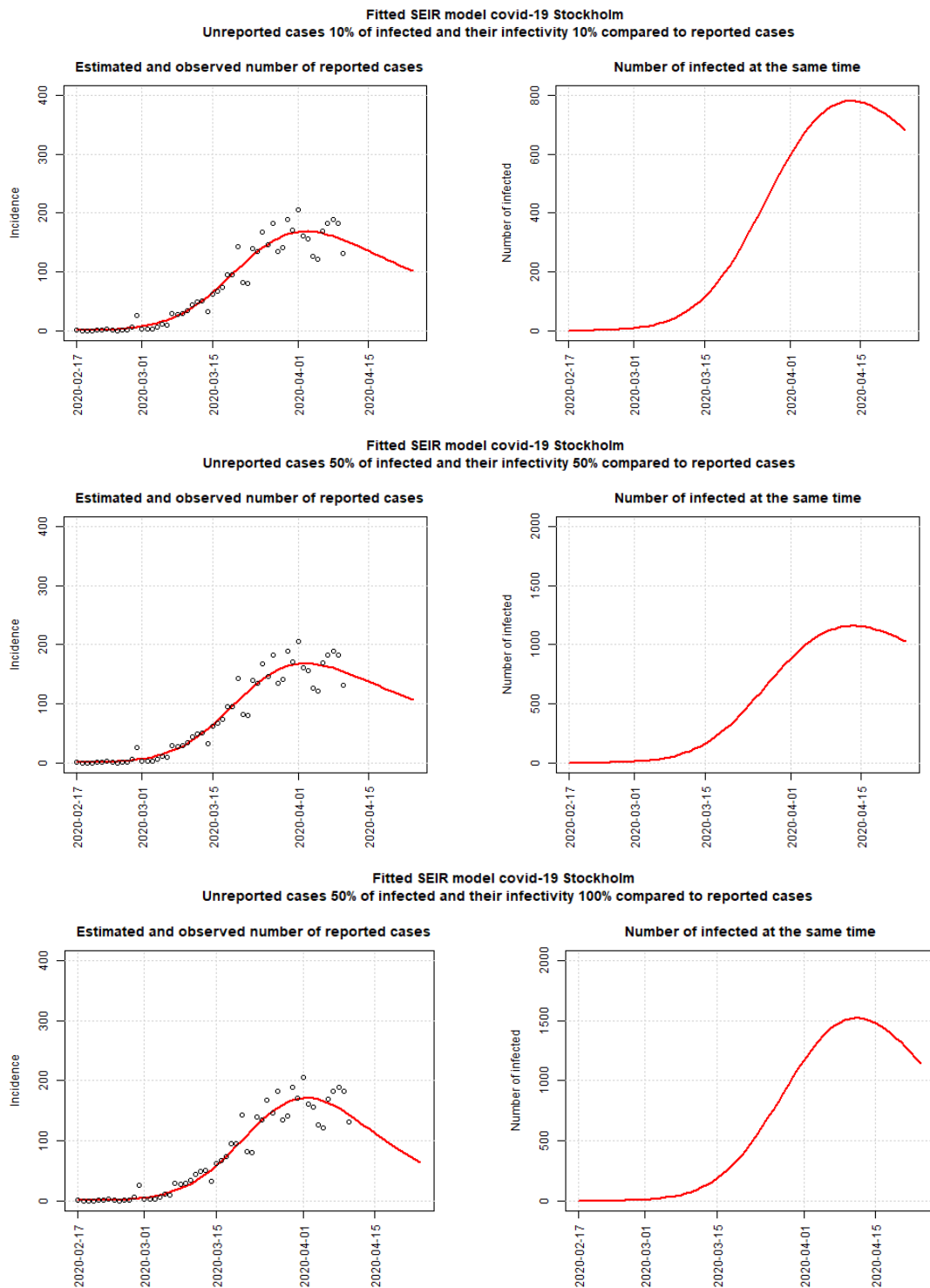
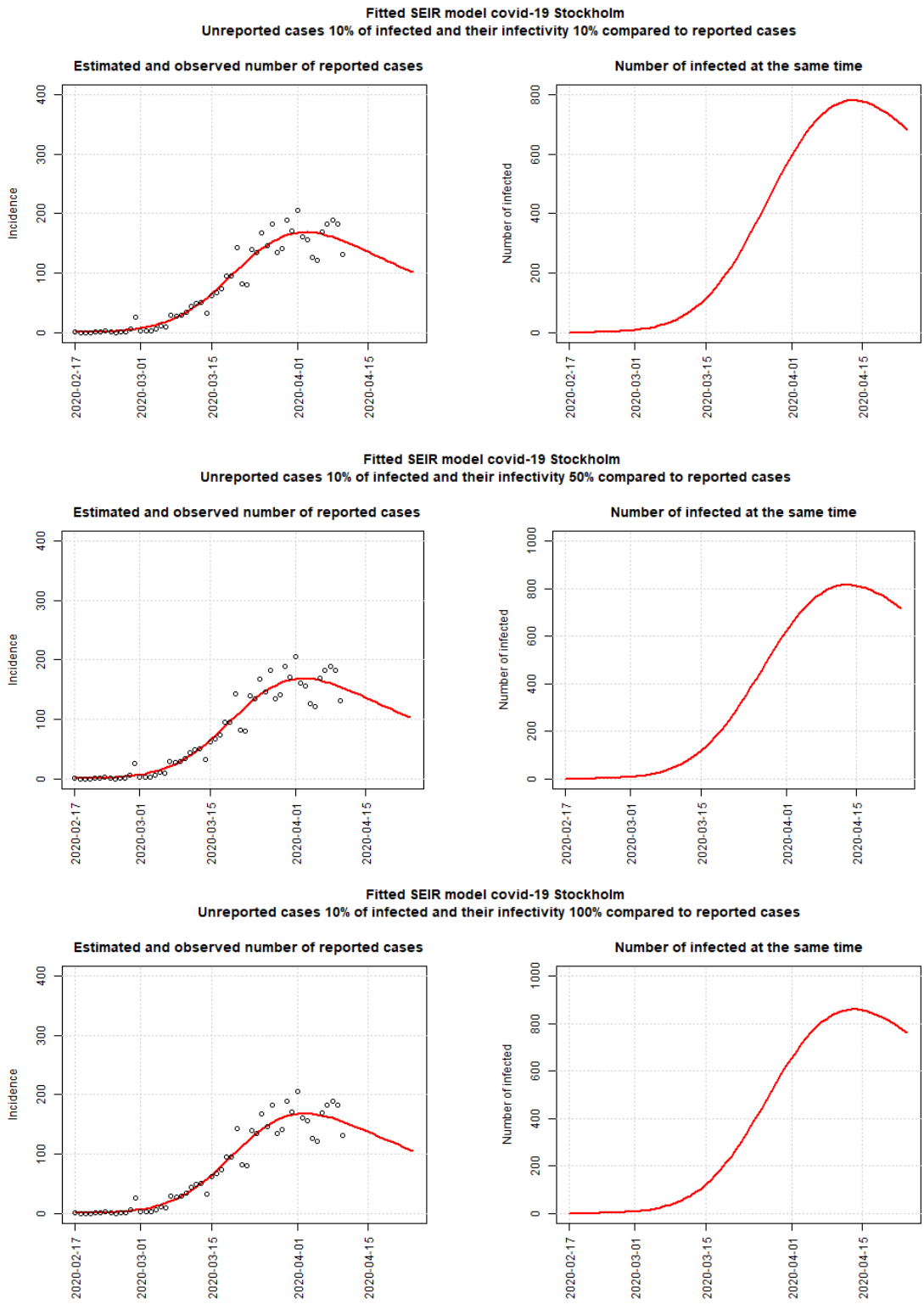


Bild 9: Scenario där 90% av alla smittade får sådana symptom att de upptäcks av vården (rapporterat fall) och 10% förblir obekräftade fall. De obekräftade fallens smittsamhet är 10% av smittsamheten hos ett rapporterat fall i den översta bilden, 50% i den mittersta och 100% i den nedersta. Den röda linjen i den vänstra bilden visar skattad daglig incidens av rapporterade fall och cirkelarna faktisk data. Den högra bilden visar prevalens av infekterade och smittsamma fall.



Tabell 4: Känslighetsanalys av ackumulerat antal och andel smittade vid två tidpunkter, samt skattad peakdag.

		Ackumulerat antal smittade och andel smittade fram till				Peakdag	Prevalens peakdag
q_o	p_o	2020-04-11		2020-05-01			
Obekräftades infektivitet 100%	obekräftade 95%	80 944	0.034	141 591	0.06	2020-04-19	15 710
		[52 972, 118 899]	[0.022, 0.05]	[76 532, 252 053]	[0.032, 0.106]	[2020-04-07, 2020-05-05]	[9 171, 34 672]
	obekräftade 50%	8 194	0.003	11 489	0.005	2020-04-11	1 522
		[2 899, 31 167]	[0.001, 0.013]	[3 328, 80 504]	[0.001, 0.034]	[2020-04-06, 2020-05-05]	[489, 12 924]
	obekräftade 10%	4 682	0.002	7 279	0.003	2020-04-13	859
		[1 536, 13 662]	[0.001, 0.006]	[1 899, 37 415]	[0.001, 0.016]	[2020-04-06, 2020-05-05]	[248, 62 78]
Obekräftades infektivitet 50%	obekräftade 95%	43 273	0.018	72 218	0.03	2020-04-15	8 149
		[24 361, 719 15]	[0.01, 0.03]	[33 090, 157 617]	[0.014, 0.066]	[2020-04-06, 2020-05-05]	[4 164, 22 157]
	obekräftade 50%	6320	0.003	9838	0.004	2020-04-13	1 159
		[2 183, 18 770]	[0.001, 0.008]	[2 618, 48 906]	[0.001, 0.021]	[2020-04-03, 2020-05-05]	[349, 80 77]
	obekräftade 10%	4 454	0.002	6 883	0.003	2020-04-13	816
		[1457, 14449]	[0.001, 0.006]	[1 663, 39 426]	[0.001, 0.017]	[2020-04-04, 2020-05-05]	[228, 7 390]

		Ackumulerat antal smittade och andel smittade fram till				Peakdag	Prevalens peakdag
q_o	p_o	2020-04-11		2020-05-01			
Obekräftades infektivitet 10%	obekräftade 95%	12 415	0.005	17 967	0.008	2020-04-13	2 239
		[1 306, 211 154]	[0.001, 0.089]	[1464, 888 894]	[0.001, 0.374]	[2020-04-05, 2020-05-05]	[181, 162 230]
	obekräftade 50%	4 680	0.002	6 993	0.003	2020-04-13	850
		[971, 29 053]	[0, 0.012]	[1 147, 97 397]	[0, 0.041]	[2020-04-06, 2020-05-05]	[146, 207 78]
	obekräftade 10%	4272	0.002	6565	0.003	2020-04-13	781
		[1369, 16790]	[0.001, 0.007]	[1549, 55741]	[0.001, 0.023]	[2020-04-03, 2020-05-05]	[213, 11693]

Rapporten beskriver en matematisk modell som kan användas för att modellera spridningen av covid-19 i Stockholms län mellan februari och april 2020. Våra resultat visar att dagen med högst antal fall i Stockholms län inträffade 15 april med 86 000 stycken samtidigt infekterade samt att obekräftade fall minst är 99 procent av samtliga fall.

Modellen tar hänsyn till både rapporterade fall som av vården bekräftats vara infekterade med SARS-CoV-2 och obekräftade fall. För skattning av modellparametrar använder vi de rapporterade fallen och för kalibrering använder vi den undersökning som visade att 2.5 procent av befolkningen i Stockholms län var infekterade mellan 27 mars och 3 april.

Rapporten med tillhörande kod är tänkt som ett verktyg för modellerare.

Folkhälsomyndigheten är en nationell kunskapsmyndighet som arbetar för en bättre folkhälsa. Det gör myndigheten genom att utveckla och stödja samhällets arbete med att främja hälsa, förebygga ohälsa och skydda mot hälsohot. Vår vision är en folkhälsa som stärker samhällets utveckling.



Folkhälsomyndigheten

Solna Nobels väg 18, 171 82 Solna. **Östersund** Forskarens väg 3. Box 505, 831 26 Östersund.

www.folkhalsomyndigheten.se