Mathematical modelling for determining COVID-19 incidence from testing data

Rasmus Kristoffer Pedersen

Postdoc, PandemiX Center, IMFUFA
Dept. Science and Environment,
Roskilde University, Denmark
Email: rakrpe@ruc.dk
Joint work with Christian Berrig and Viggo Andreasen

Contributed talk at "Data-driven mechanistic mathematical modelling for life-science applications"
Göteborg, October 23rd, 2023

Introduction to the problem

- Different approaches to COVID-19 mitigation throughout the world

Introduction to the problem

- Different approaches to COVID-19 mitigation throughout the world
- To compare mitigation-strategies, the impact of differences in data-collection must be understood.

Introduction to the problem

- Different approaches to COVID-19 mitigation throughout the world
- To compare mitigation-strategies, the impact of differences in data-collection must be understood.
- The role of testing: Confirmation of symptoms, required for various activities or entirely voluntary?

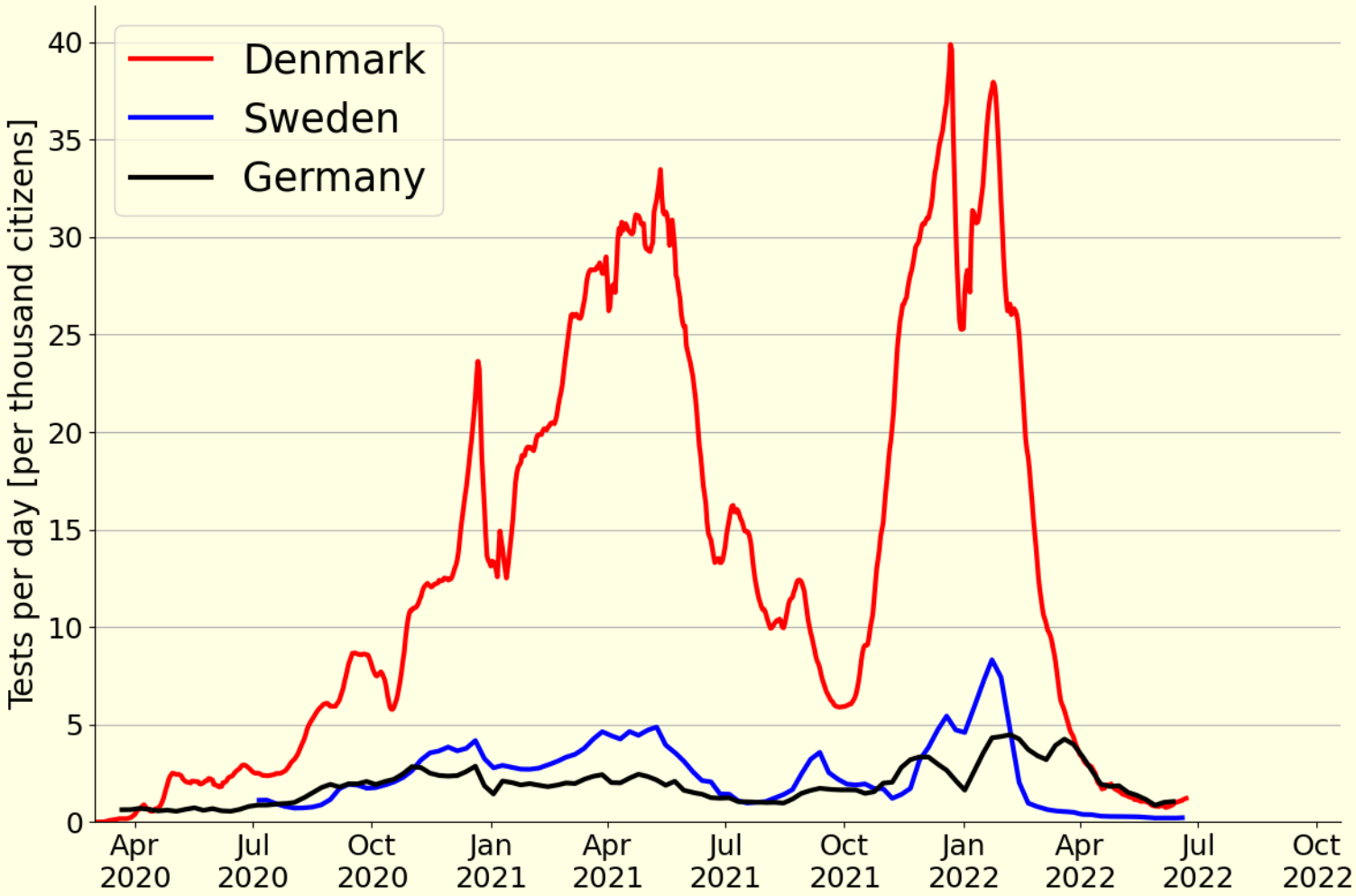
Introduction to the problem

- Different approaches to COVID-19 mitigation throughout the world
- To compare mitigation-strategies, the impact of differences in data-collection must be understood.
- The role of testing: Confirmation of symptoms, required for various activities or entirely voluntary?
- **How do we compare case-counts between periods and places where testing activity was different?**

Introduction to the problem

- Different approaches to COVID-19 mitigation throughout the world
- To compare mitigation-strategies, the impact of differences in data-collection must be understood.
- The role of testing: Confirmation of symptoms, required for various activities or entirely voluntary?
- **How do we compare case-counts between periods and places where testing activity was different?**

Let's look at some data...



Our approach

- Overall question: For each reported case of COVID-19, how many unidentified cases?

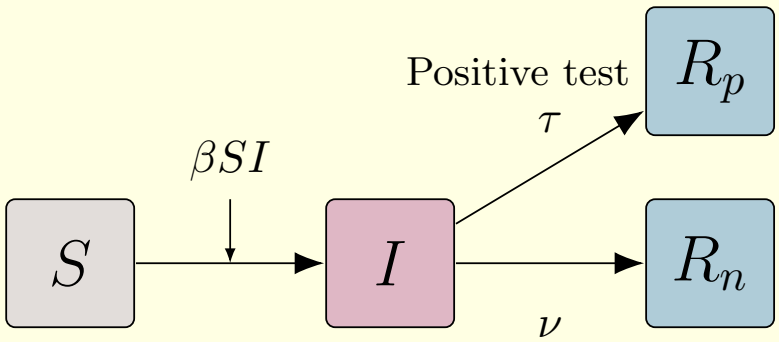
Our approach

- Overall question: For each reported case of COVID-19, how many unidentified cases?
- We aim to determine a correction factor for observed data.

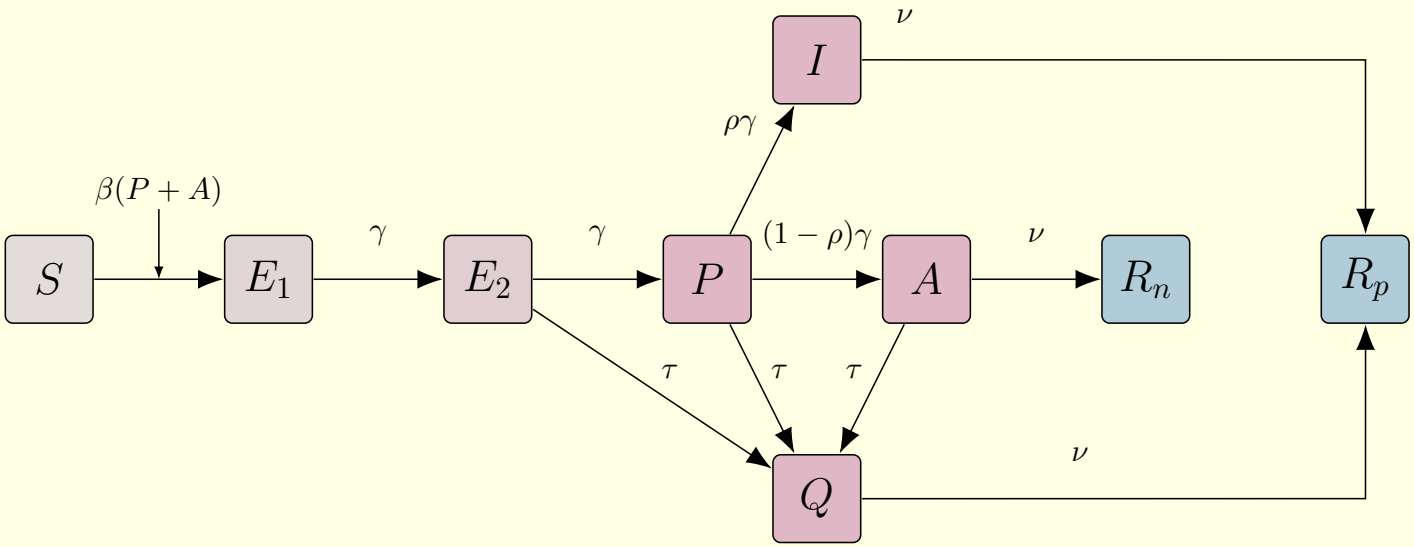
Our approach

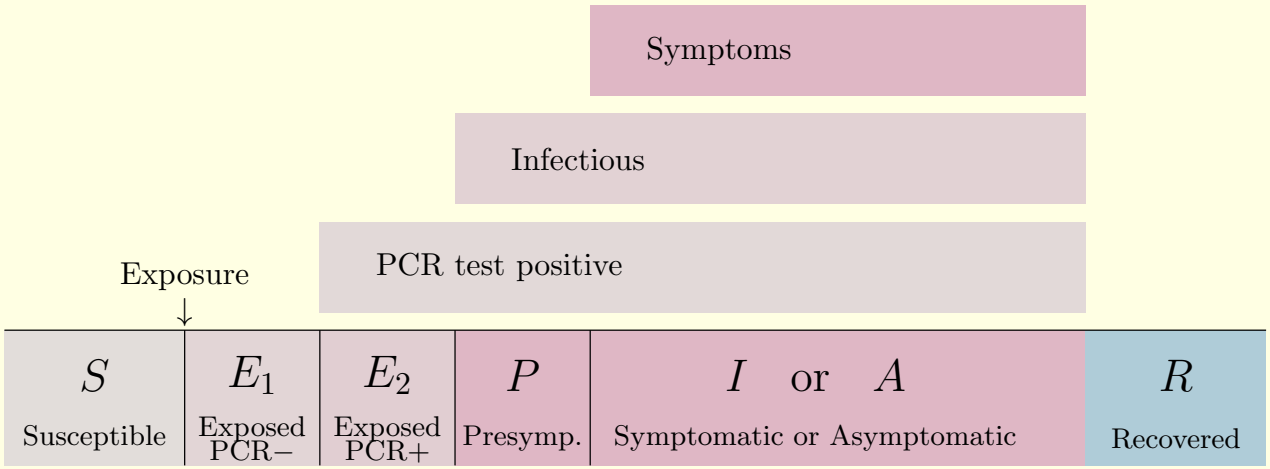
- Overall question: For each reported case of COVID-19, how many unidentified cases?
- We aim to determine a correction factor for observed data.
- Approach: Extend the classic SIR-model to include testing.

The conceptual idea

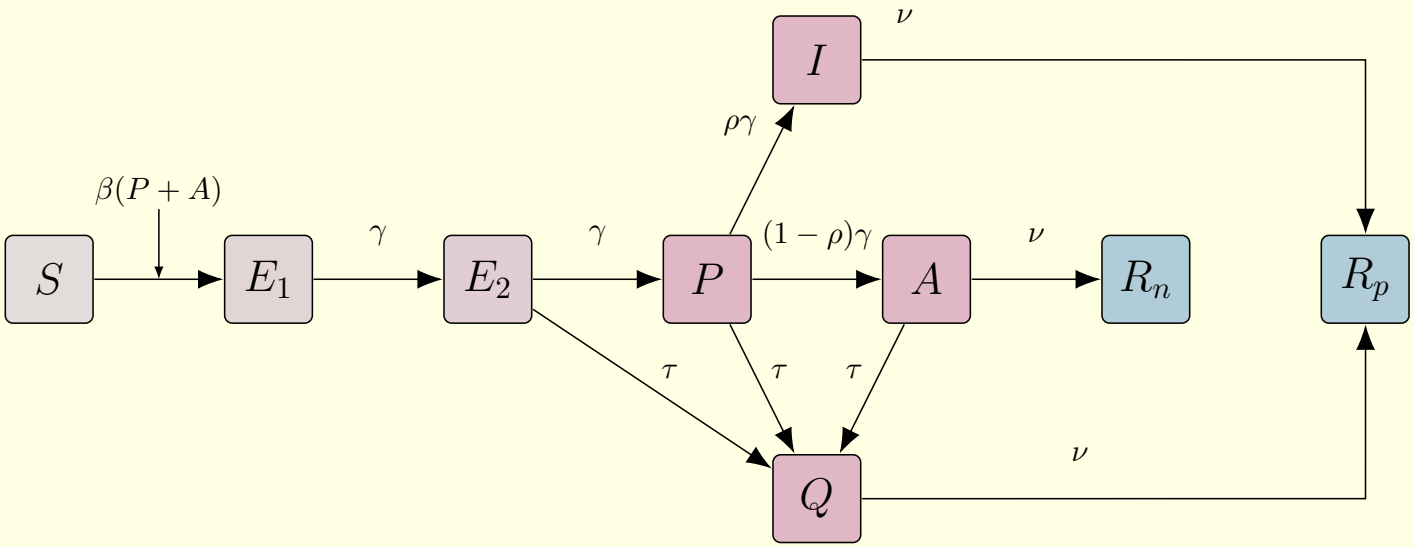


The model


ν
I
$\rho\gamma$
$\beta(P+A)$
S
γ
E_1
γ
E_2
P
$(1-\rho)\gamma$
A
ν
R_n
R_p
τ
τ
τ
ν
Q


Symptoms
Infectious
PCR test positive
Exposure
S
Susceptible
E_1
Exposed PCR−
E_2
Exposed PCR+
P
Presymp.
I or A
Symptomatic or Asymptomatic
R
Recovered

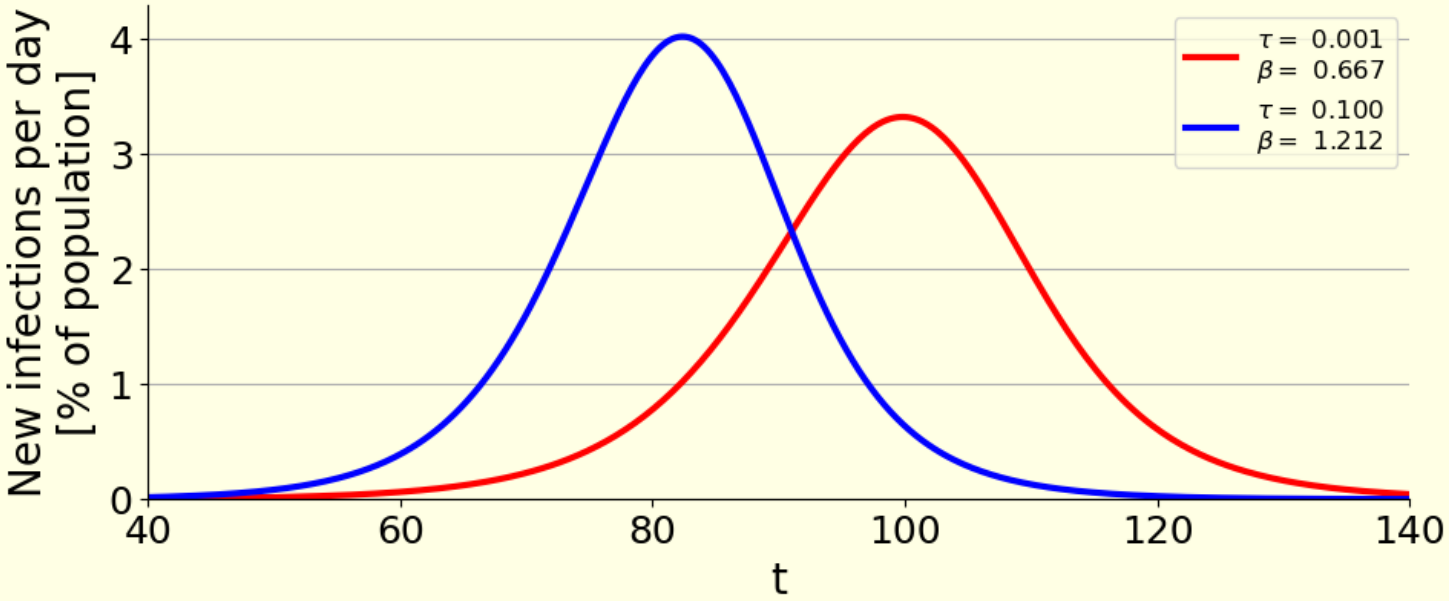
The model

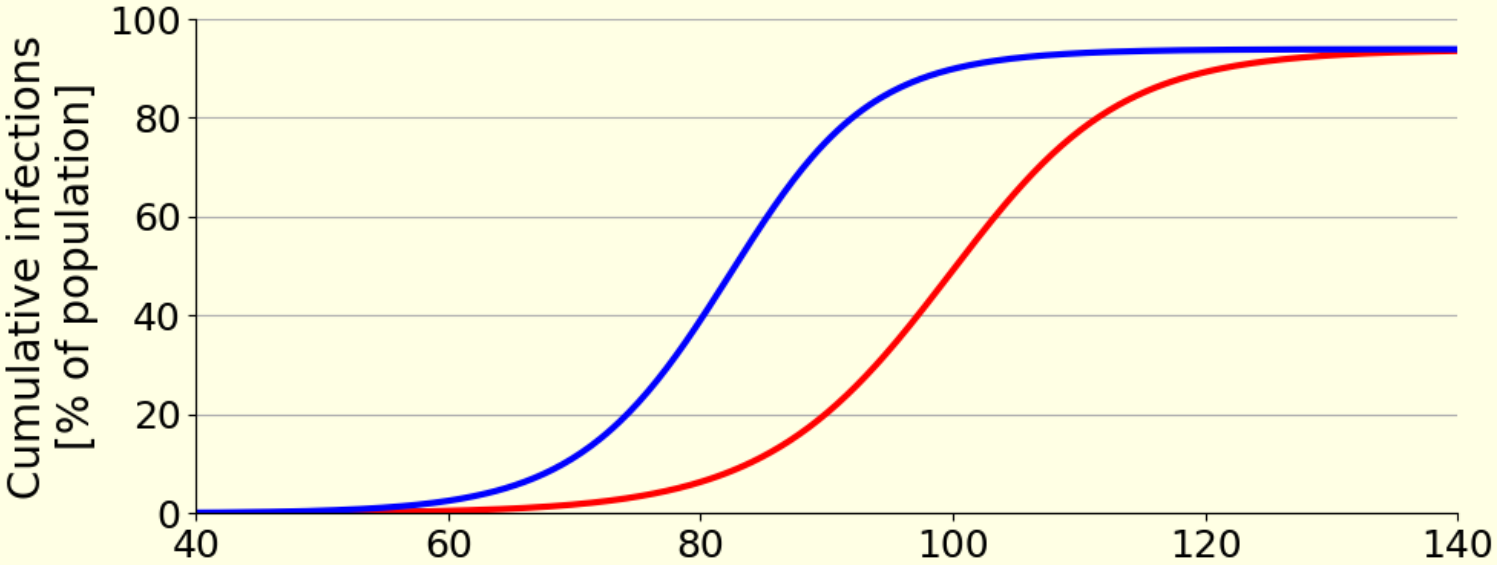


$$\dot{S} = -\beta S(P+A)$$
$$\dot{E}_1 = \beta S(P+A) - \gamma E_1$$
$$\dot{E}_2 = \gamma E_1 - (\gamma+\tau)E_2$$
$$\dot{P} = \gamma E_2 - (\gamma+\tau)P$$
$$\dot{I} = \gamma\rho P - \nu I$$
$$\dot{A} = \gamma(1-\rho)P - (\nu+\tau)A$$
$$\dot{Q} = \tau(E_2+P+A) - \nu Q$$
$$\dot{R}_p = \nu Q + \nu I$$
$$\dot{R}_n = \nu A$$

τ: Testing-rate. β: Infectivity.

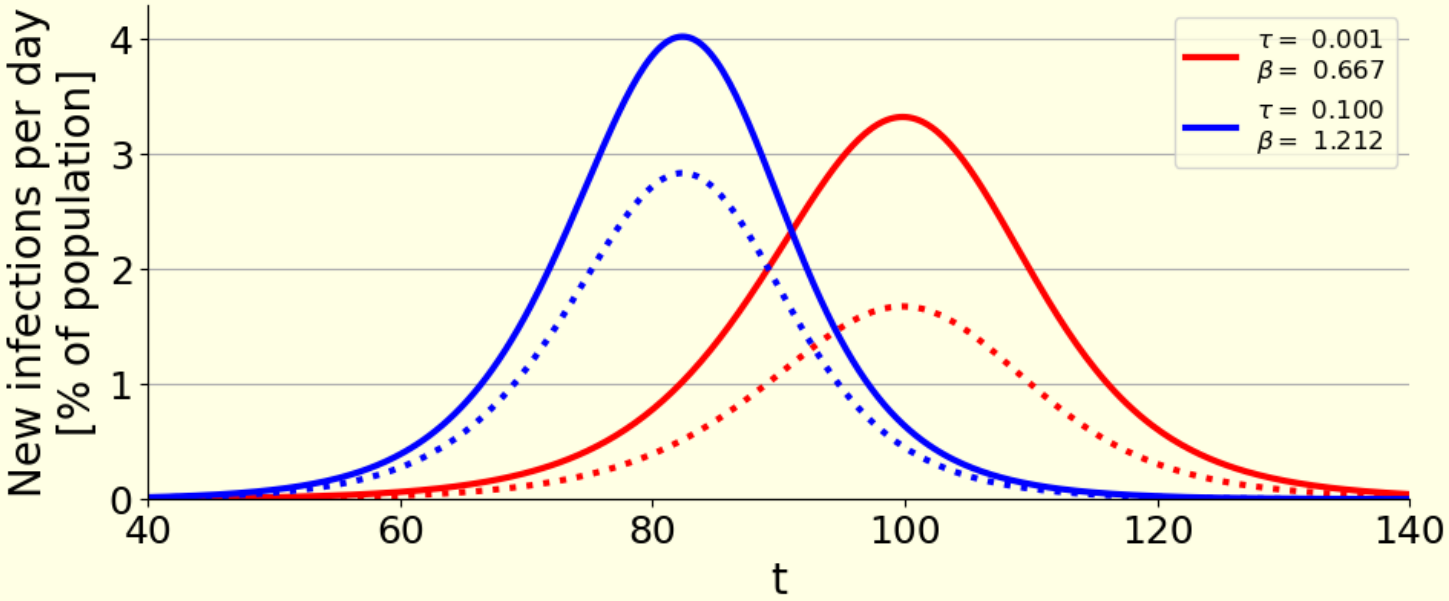
General model dynamics

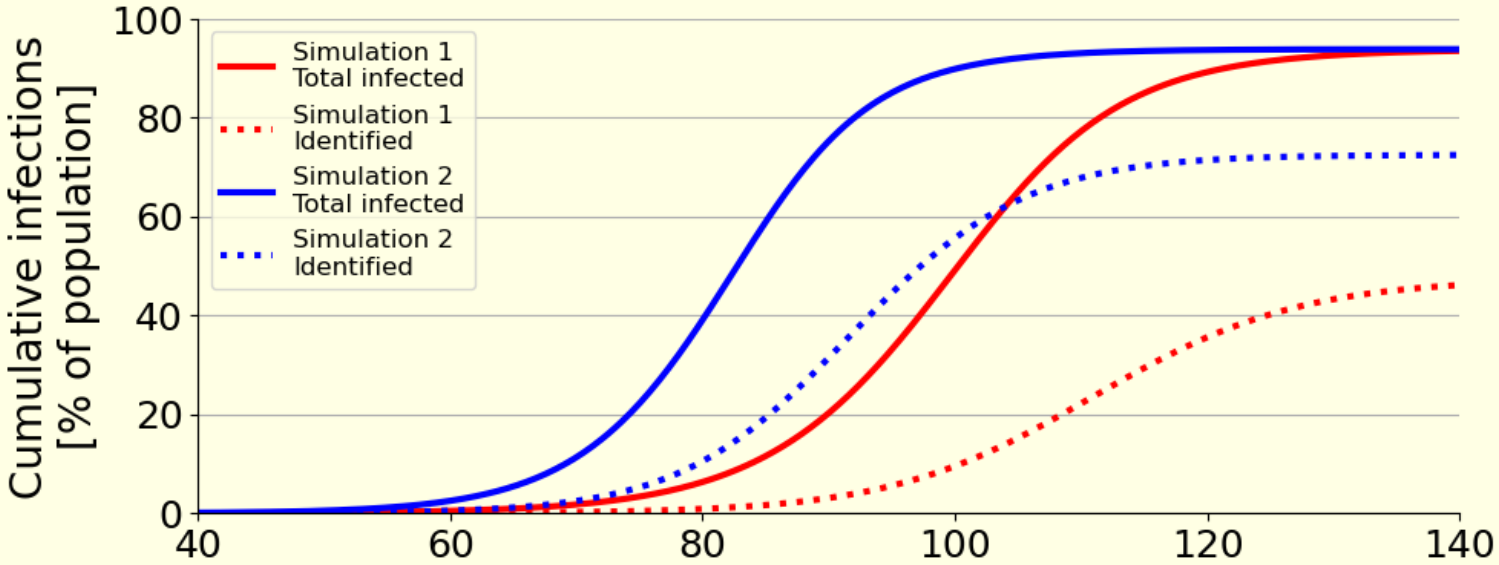




Other parameters: $\gamma = \nu = \frac{1}{3}$ and $\rho = \frac{1}{2}$.

General model dynamics





Other parameters: $\gamma = \nu = \frac{1}{3}$ and $\rho = \frac{1}{2}$.

Model analysis and correction ratio

We describe the correction factor as ratio between all cases and cases identified:

$$\frac{R_n(t) + R_p(t)}{R_p(t)} \tag{1}$$

Model analysis and correction ratio

We describe the correction factor as ratio between all cases and cases identified:

$$\frac{R_n(t) + R_p(t)}{R_p(t)} \tag{1}$$

Inspired by previous work on epidemic final size[1] (and after a lot of analysis and calculation), we find that as $t \to \infty$

$$\frac{R_p}{R_n + R_p} = 1 - \left(1 - \frac{\tau}{\gamma + \tau}\right)\left(1 - \frac{\gamma\rho + \tau}{\gamma + \tau}\right)\left(\frac{\nu}{\nu + \tau}\right) \tag{2}$$

[1] (Andreasen, 2018, *Bull. Math. Biol.*)

Model analysis and correction ratio

We describe the correction factor as ratio between all cases and cases identified:

$$\frac{R_n(t) + R_p(t)}{R_p(t)} \tag{1}$$

Inspired by previous work on epidemic final size[1] (and after a lot of analysis and calculation), we find that as $t \to \infty$

$$\frac{R_p}{R_n + R_p} = 1 - \left(1 - \frac{\tau}{\gamma + \tau}\right)\left(1 - \frac{\gamma\rho + \tau}{\gamma + \tau}\right)\left(\frac{\nu}{\nu + \tau}\right) \tag{2}$$

Note that this is independent of β.

[1] (Andreasen, 2018, *Bull. Math. Biol.*)

Calculating the correction-factor

$$\frac{R_p}{R_n + R_p} = 1 - \left(1 - \frac{\tau}{\gamma + \tau}\right)\left(1 - \frac{\gamma\rho + \tau}{\gamma + \tau}\right)\left(\frac{\nu}{\nu + \tau}\right)$$

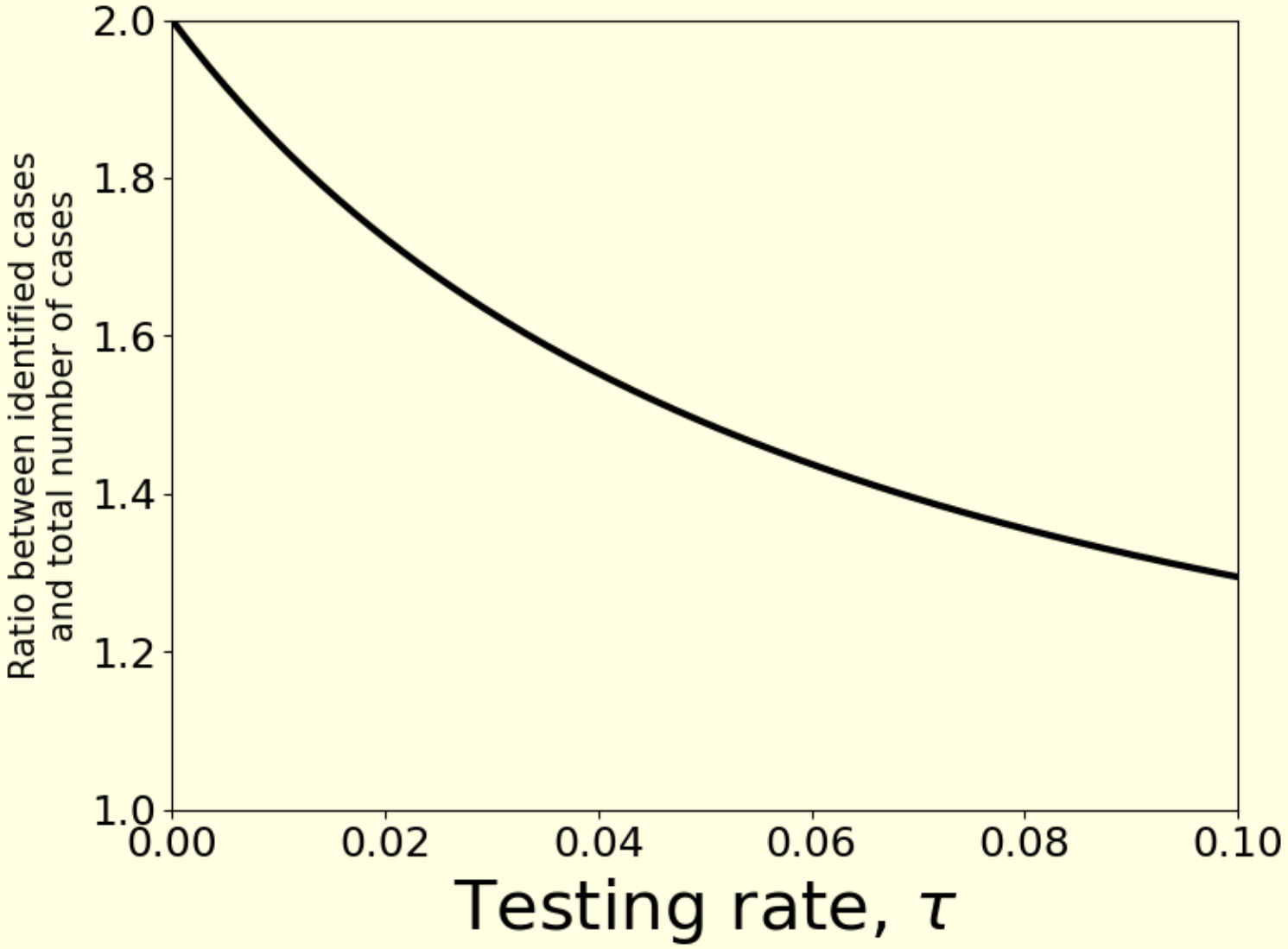
With parameters: $\gamma = \nu = \frac{1}{3}$ and $\rho = \frac{1}{2}$.

Calculating the correction-factor

$$\frac{R_p}{R_n + R_p} = 1 - \left(1 - \frac{\tau}{\frac{1}{3} + \tau}\right)\left(1 - \frac{\frac{1}{3}\cdot\frac{1}{2} + \tau}{\frac{1}{3} + \tau}\right)\left(\frac{\frac{1}{3}}{\frac{1}{3} + \tau}\right)$$

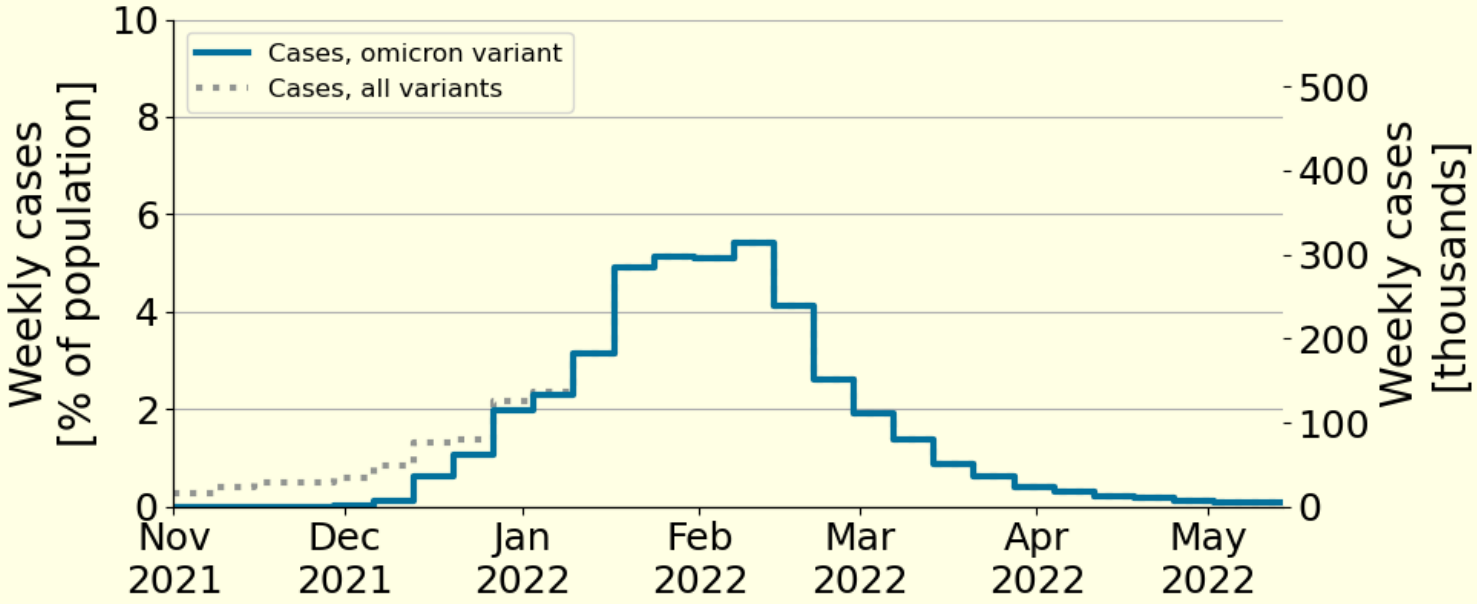
With parameters: $\gamma = \nu = \frac{1}{3}$ and $\rho = \frac{1}{2}$.

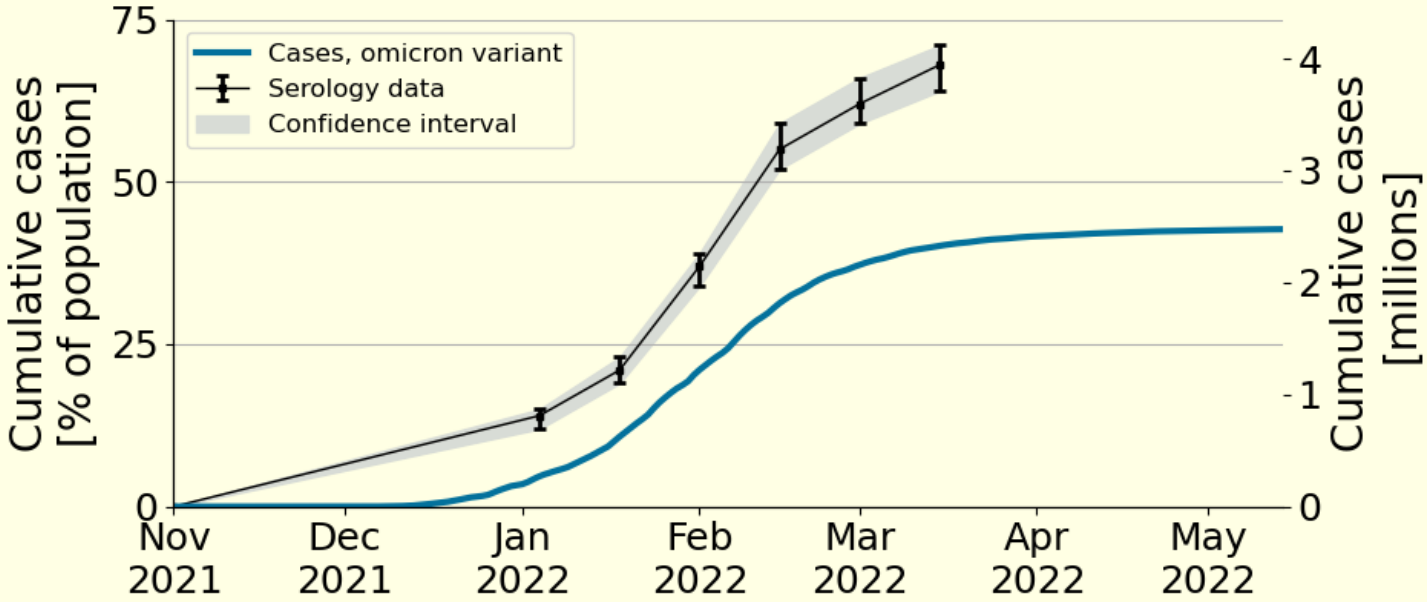
Calculating the correction-factor



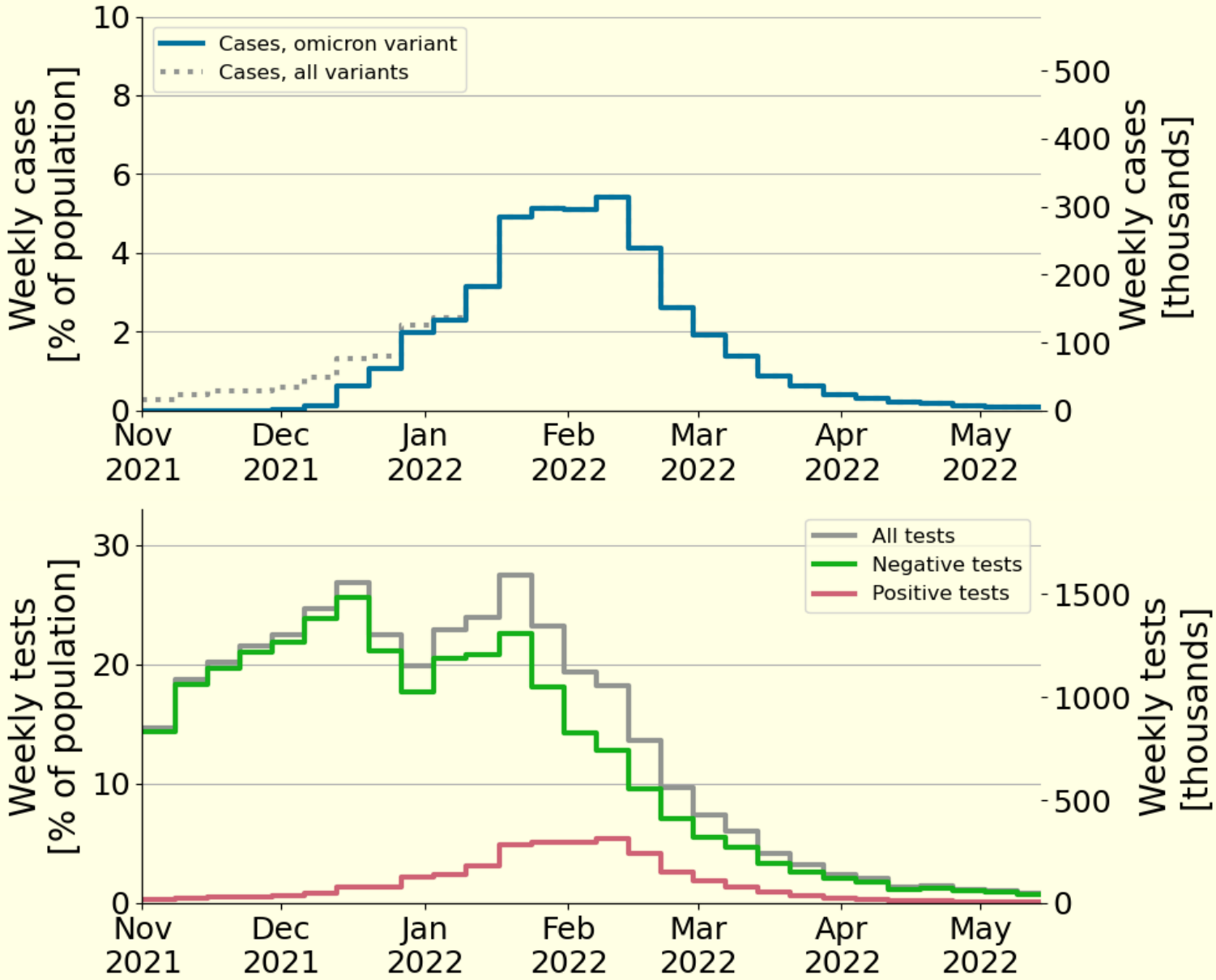
With parameters: $\gamma = \nu = \frac{1}{3}$ and $\rho = \frac{1}{2}$.

The Danish data

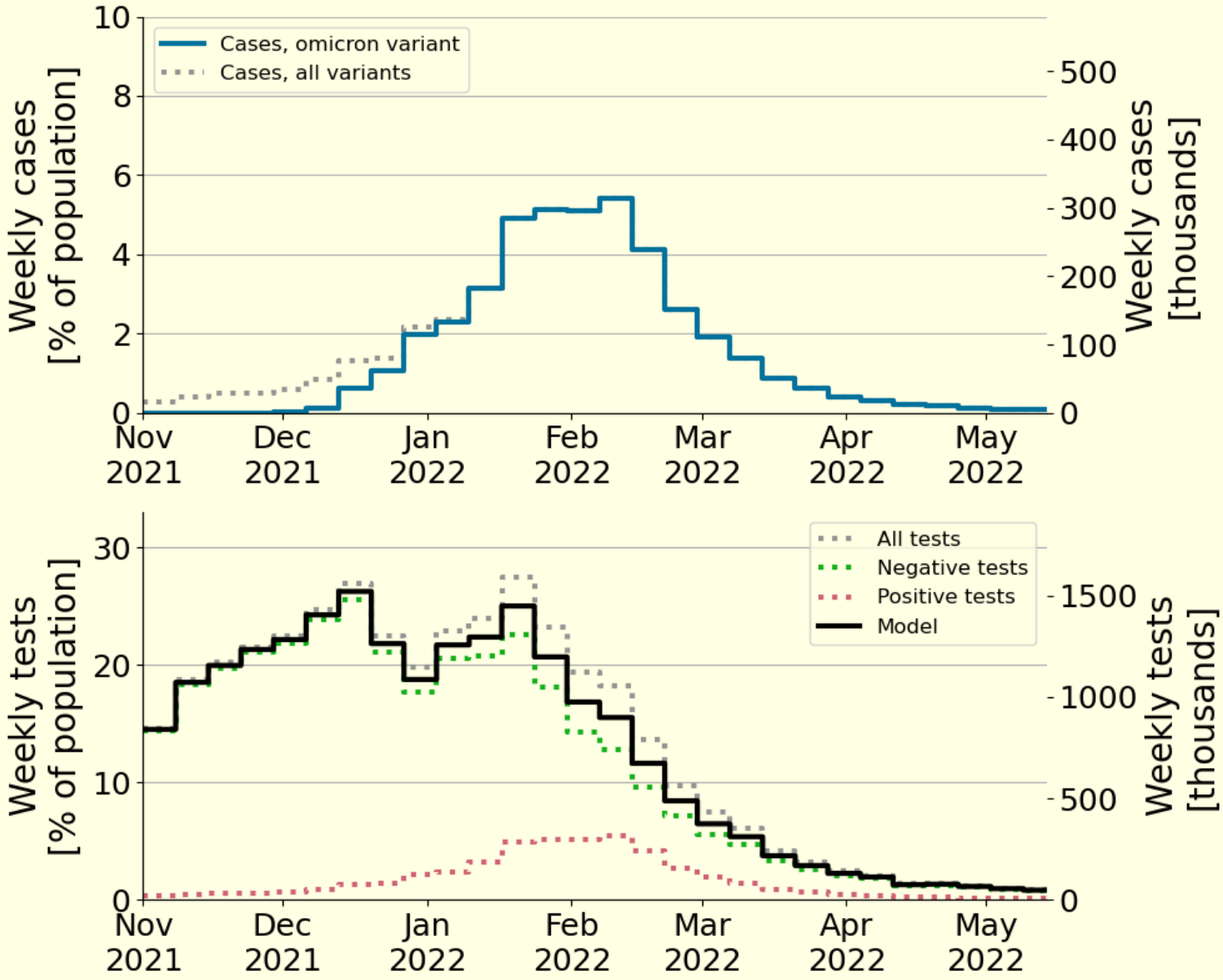




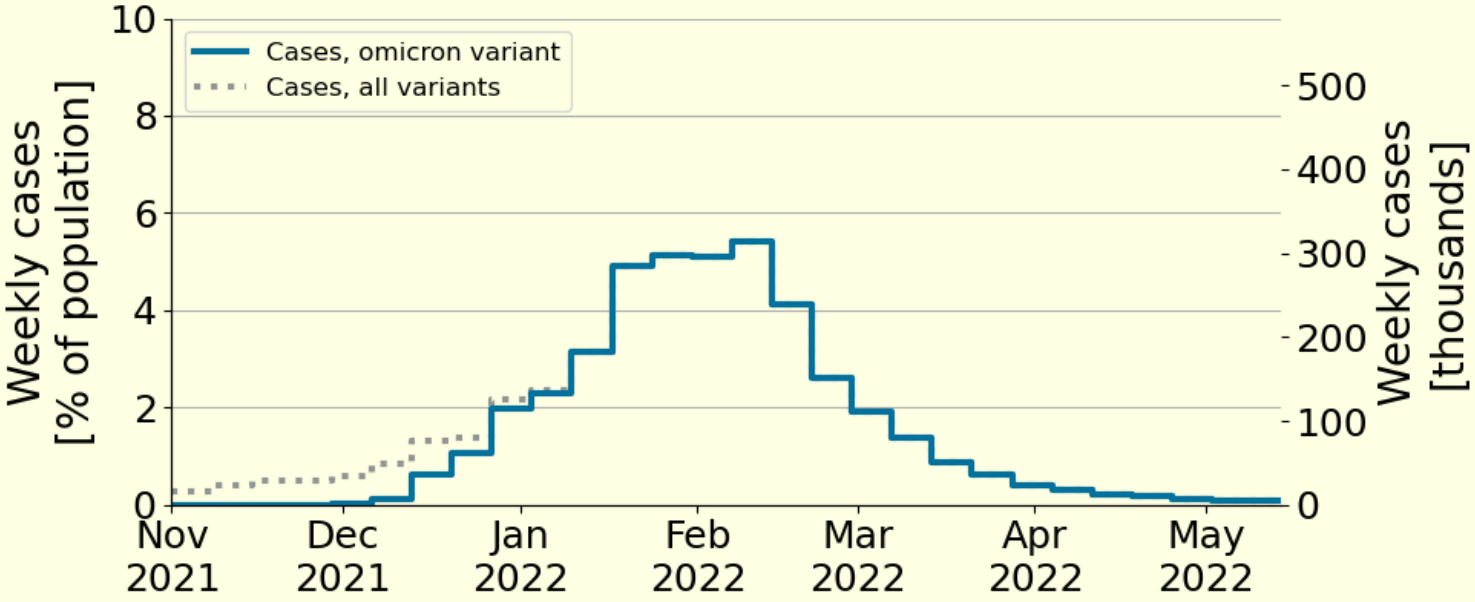
The Danish data

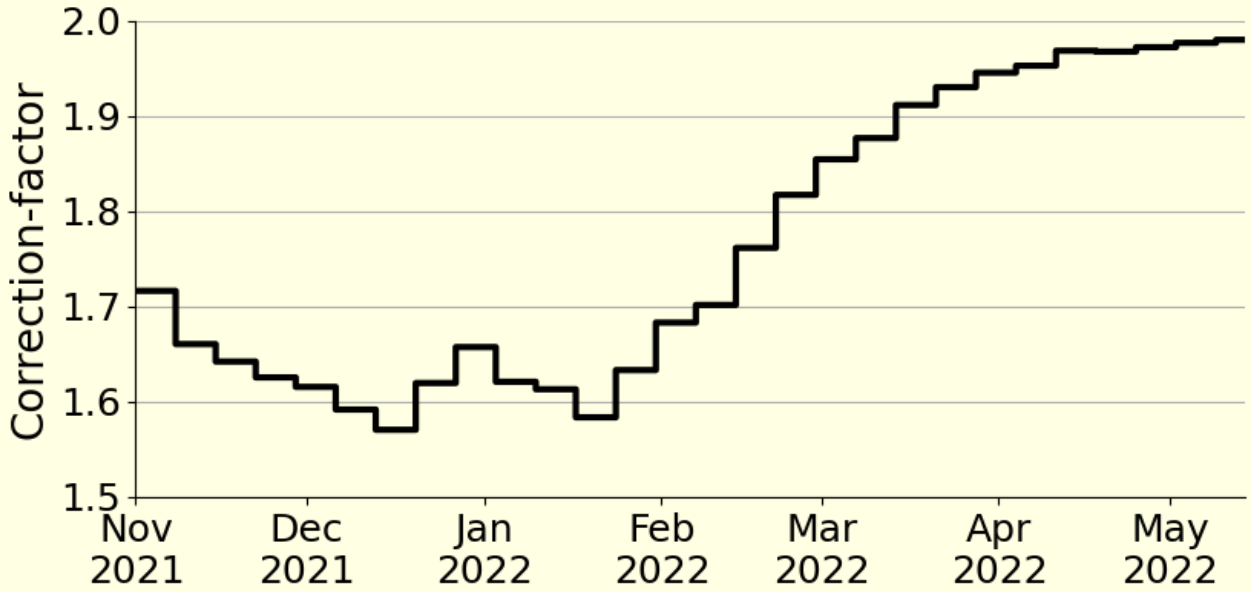


Relating our results to data

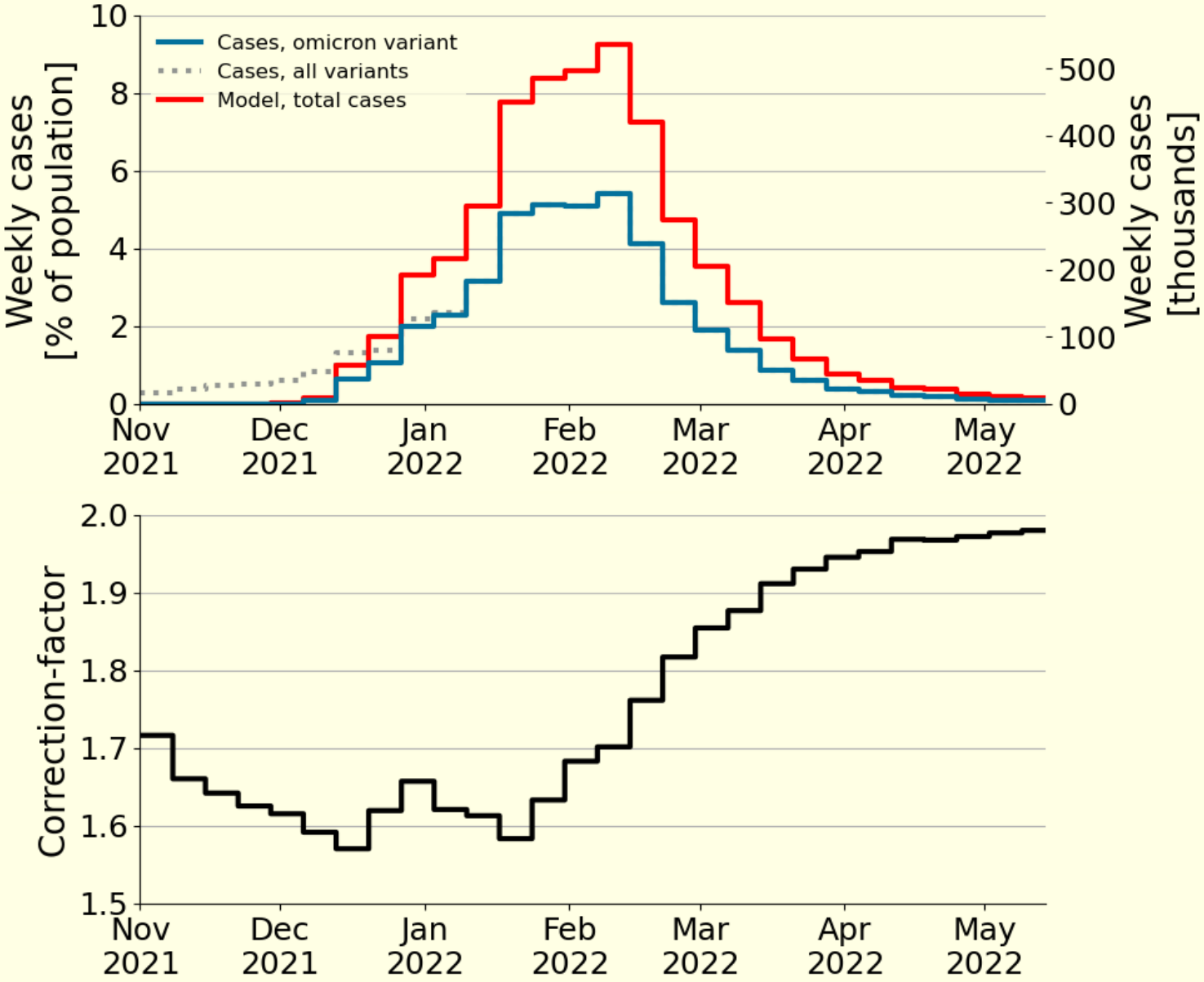


Relating our results to data

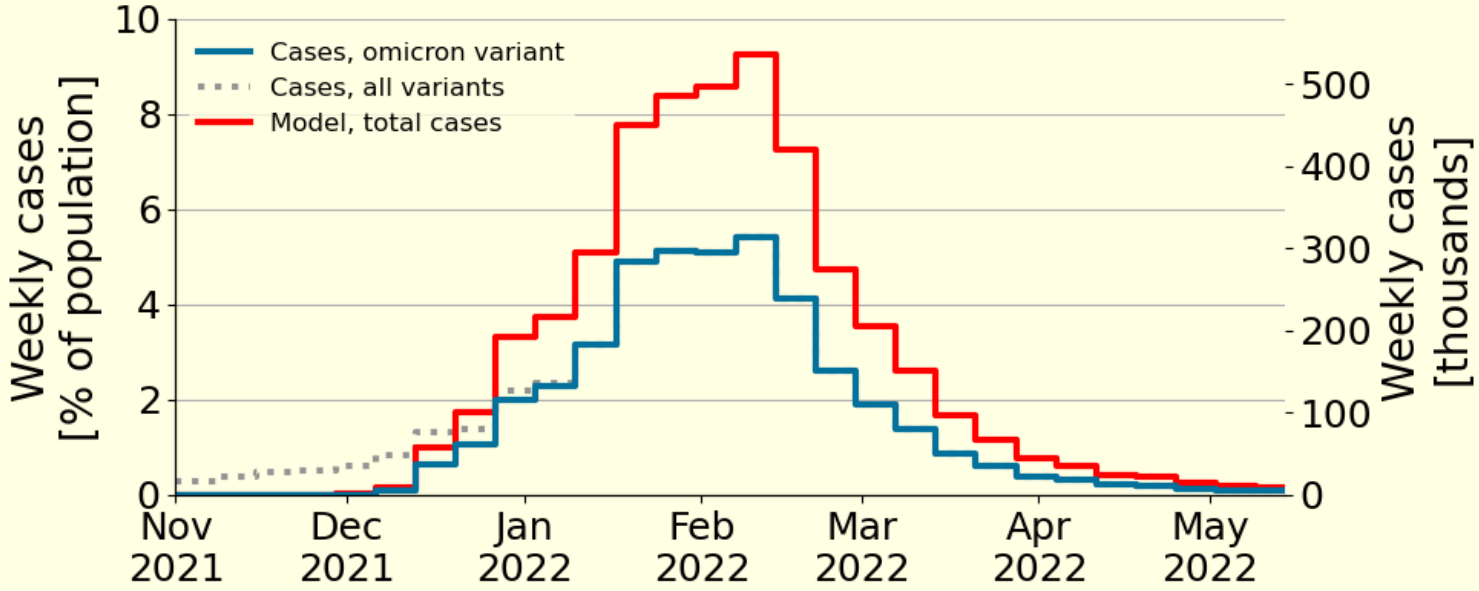


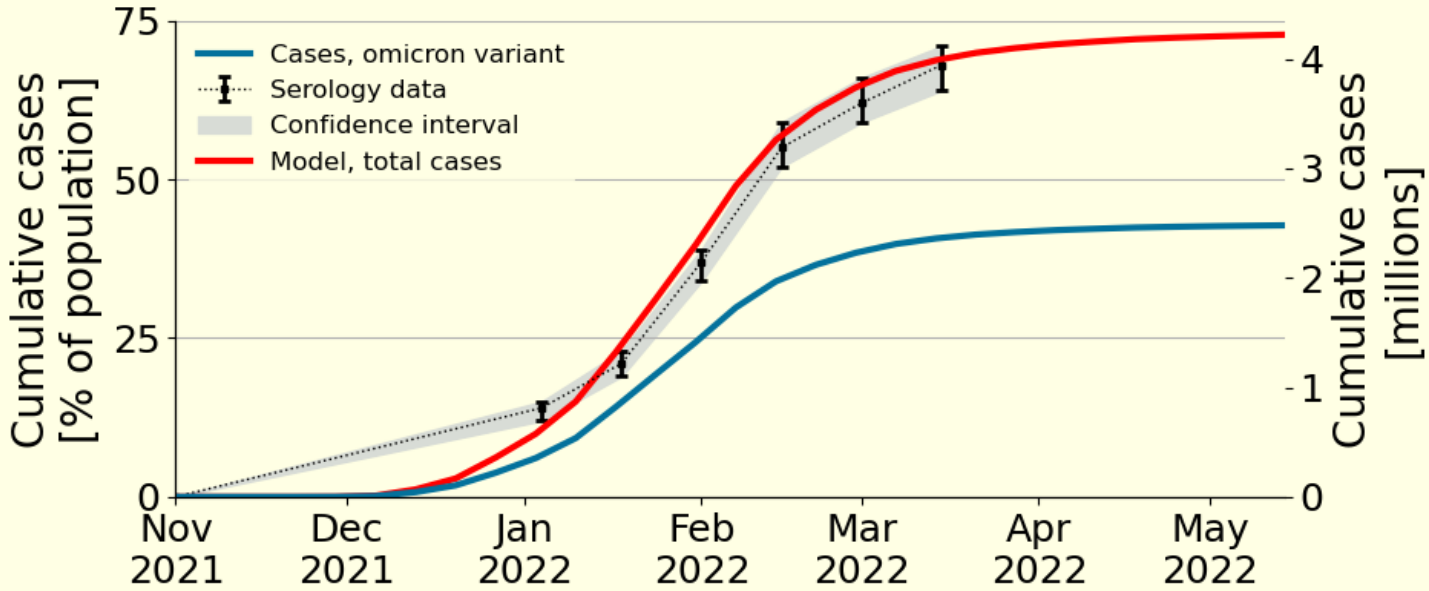


Relating our results to data



Relating our results to data





Calculation of assertion ratio

A.3 Final Size Calculations

As $t \to \infty$, the model system approaches a steady state without any active cases. In this section, we derive an analytic expression for the value that the fraction of cases identified, $K(t)$, approaches as $t \to \infty$.

To obtain an expression for K_F, we follow the methodology previously considered by **?**.

For notational purposes, we define for each variable x, the integral over the full epidemic as $T_x = \int_0^\infty x dt$.

From the system of differential equations given in equations (1), we write up the following quantities:

$$\dot{S}/S = -\beta(P + A) \tag{3a}$$
$$\dot{S} + \dot{E}_1 + \dot{E}_2 = -(\gamma + \tau)E_2 \tag{3b}$$
$$\dot{S} + \dot{E}_1 + \dot{E}_2 + \dot{P} = -(\gamma + \tau)P - \tau E_2 \tag{3c}$$
$$\dot{S} + \dot{E}_1 + \dot{E}_2 + \dot{P} + \dot{A} = -(\nu + \tau)A - (\gamma\rho + \tau)P - \tau E_2 \tag{3d}$$

As t approaches infinity, the stability of the systems implies that all variables apart from S, R_p and R_n are zero. We denote that final size of these variables as $S(t) \underset{t\to\infty}{\longrightarrow} \sigma$, $R_p(t) \underset{t\to\infty}{\longrightarrow} r_p$ and $R_n(t) \underset{t\to\infty}{\longrightarrow} r_n$

Integrating equations (3) from $t = 0$ to $t = \infty$ yields:

$$\log\sigma = -\beta(T_P - T_A) \tag{4a}$$
$$\sigma - S_0 - E_{1,0} - E_{2,0} = -(\gamma + \tau)T_{E_2} \tag{4b}$$
$$\sigma - S_0 - E_{1,0} - E_{2,0} - P_0 = -(\gamma + \tau)T_P - \tau T_{E_2} \tag{4c}$$
$$\sigma - S_0 - E_{1,0} - E_{2,0} - P_0 - A_0 = -(\nu + \tau)T_A - (\gamma\rho + \tau)T_P - \tau T_{E_2} \tag{4d}$$

Where X_0 denote the initial condition for variable X.

Furthermore, observe that the equations for $\dot{R}_n$ and $\dot{R}_p$, equations (1i) and (1h) respectively, when integrated from $t = 0$ to $t = \infty$ yields:

$$r_p - R_{p,0} = \nu T_Q + \nu T_I \tag{5}$$
$$r_n - R_{n,0} = \nu T_A \tag{6}$$

In general, we consider initial conditions such that the vast majority of the population is initially susceptible, $S_0 \approx 1$, and the initial number of cases is low, $0 < E_{1,0} \ll 1$. In the limit where $S_0 \to 1$, with $E_{1,0} \to 0$, $E_{2,0} \to 0$, $P_0 \to 0$ and $A_0 \to 0$, equations (4) become:

$$\log\sigma = -\beta(T_P - T_A) \tag{7a}$$
$$\sigma = 1 - (\gamma + \tau)T_{E_2} \tag{7b}$$
$$\sigma = 1 - (\gamma + \tau)T_P - \tau T_{E_2} \tag{7c}$$
$$\sigma = 1 - (\nu + \tau)T_A - (\gamma\rho + \tau)T_P - \tau T_{E_2} \tag{7d}$$

Assuming $T_P + T_A \neq 0$, this can be written as:

$$\beta = \frac{-\log\sigma}{T_P + T_A} \tag{8a}$$
$$T_{E_2} = \frac{1}{\gamma + \tau}(1 - \sigma) \tag{8b}$$
$$T_P = \frac{1}{\gamma + \tau}(1 - \sigma - \tau T_{E_2}) \tag{8c}$$
$$T_A = \frac{1}{\nu + \tau}(1 - \sigma - (\gamma\rho + \tau)T_P - \tau T_{E_2}) \tag{8d}$$

We define $K_F = \frac{r_p}{r_p + r_n}$ and note that at steady state $\sigma = 1 - r_p - r_n$ must hold. This implies that $K_F = \frac{r_p}{1-\sigma}$. Combining equations (8) with equations (5) and (6) under the assumptions $R_{p,0} = 0$ and $R_{n,0} = 0$ and simplifying yields:

$$K_F = \frac{r_p}{r_p + r_n} = \frac{r_p + r_n - r_n}{r_p + r_n} = 1 - \frac{r_n}{r_p + r_n} = 1 - \frac{r_n}{1-\sigma} = 1 - \frac{\nu}{1-\sigma}T_A \tag{9}$$
$$K_F = 1 - \left(\frac{\nu}{\nu + \tau}\right)\left(1 - \frac{\tau}{\gamma + \tau}\right)\left(1 - \frac{\gamma\rho + \tau}{\gamma + \tau}\right) \tag{10}$$

For initial conditions sufficiently close to the case where $S_0 = 1$ and all other variables are zero, K_F is an approximation of the final size of $K(t)$ as $t \to \infty$.

Note that the expression for K_F, equation (10) is independent of σ and β.

Furthermore, in the absence of tests, i.e. for $\tau = 0$, we have $K_F = 1 - 1(1-0)(1-\rho) = \rho$. This is expected, as only the symptomatic cases, I, are found in the situation where $\tau = 0$, and the symptomatic cases make up exactly ρ of all cases.

In the situation where all cases are symptomatic, $\rho = 1$, we obtain $K_F = 1$, that is, all cases are identified.

We note that equation (8a) describes a relation between β and σ. Since T_P and T_A are described in terms of γ, τ, ν and σ, it is possible to use equation (8a) to determine a value of β that yields a particular σ.

$$\frac{R_p}{R_n + R_p} = 1 - \left(1 - \frac{\tau}{\gamma + \tau}\right)\left(1 - \frac{\gamma\rho + \tau}{\gamma + \tau}\right)\left(\frac{\nu}{\nu + \tau}\right)$$

Simplified method, Flow-considerations

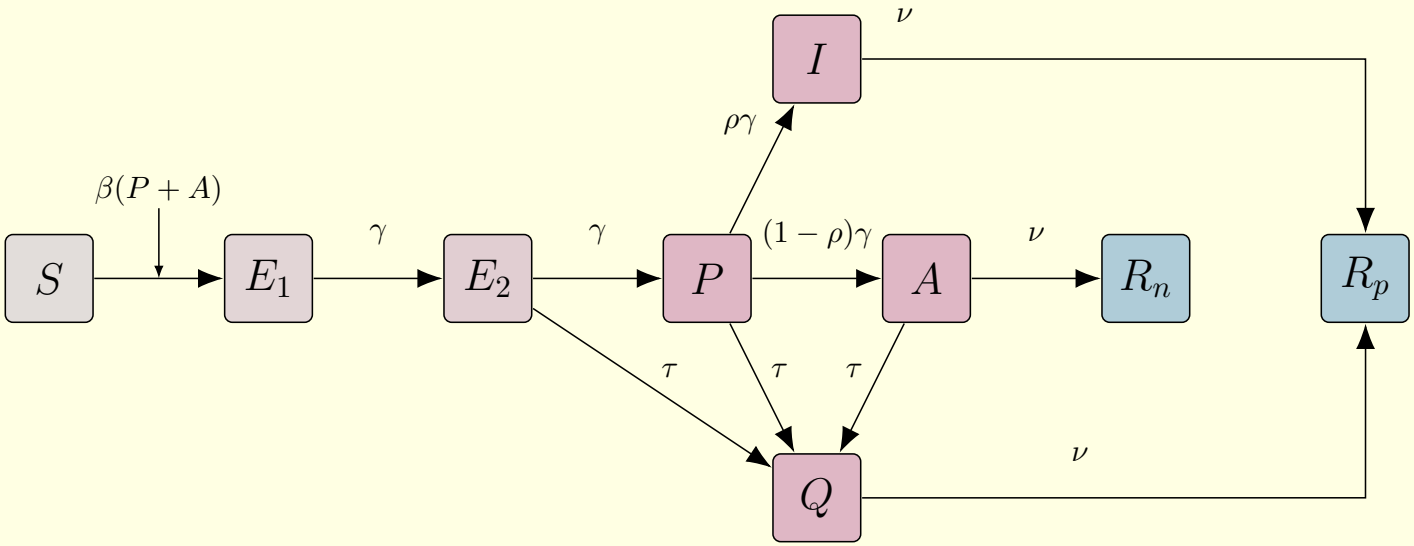
$$\frac{R_p}{R_n + R_p} = 1 - \left(1 - \frac{\tau}{\gamma + \tau}\right)\left(1 - \frac{\gamma\rho + \tau}{\gamma + \tau}\right)\left(\frac{\nu}{\nu + \tau}\right)$$

can be rewritten as:

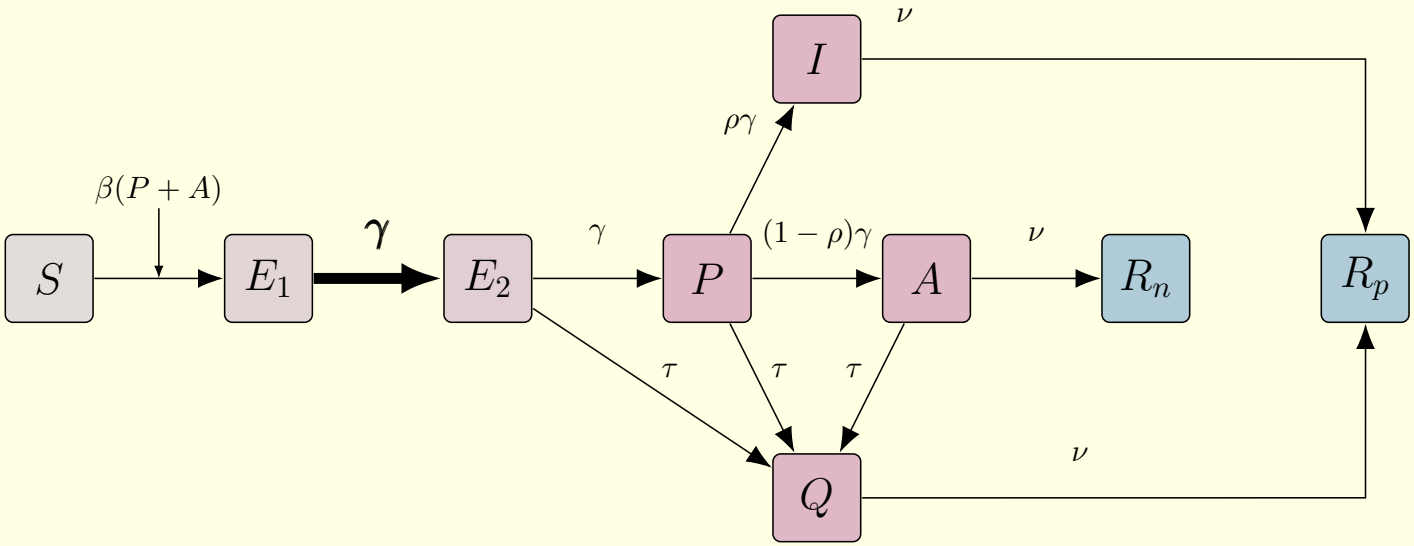
$$1 - \frac{R_p}{R_n + R_p} =$$

$$\frac{R_n}{R_n + R_p} = \left(\frac{\gamma}{\gamma + \tau}\right)\left(\frac{(1-\rho)\gamma}{\gamma + \tau}\right)\left(\frac{\nu}{\nu + \tau}\right)$$

Simplified method, Flow-considerations

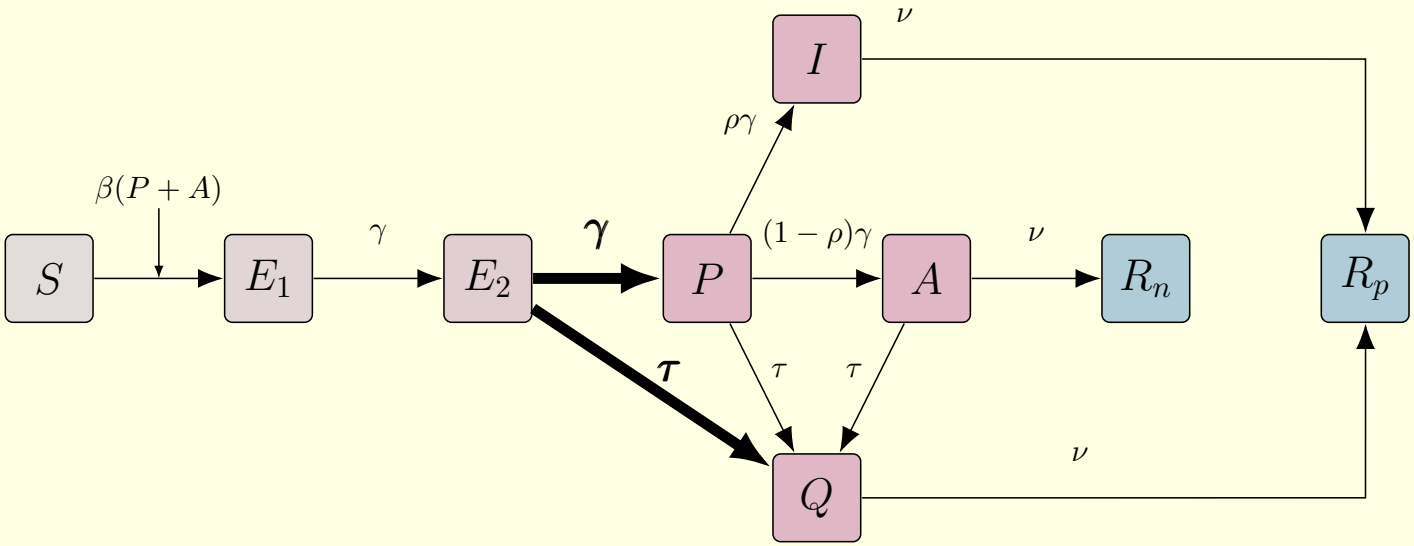


Simplified method, Flow-considerations



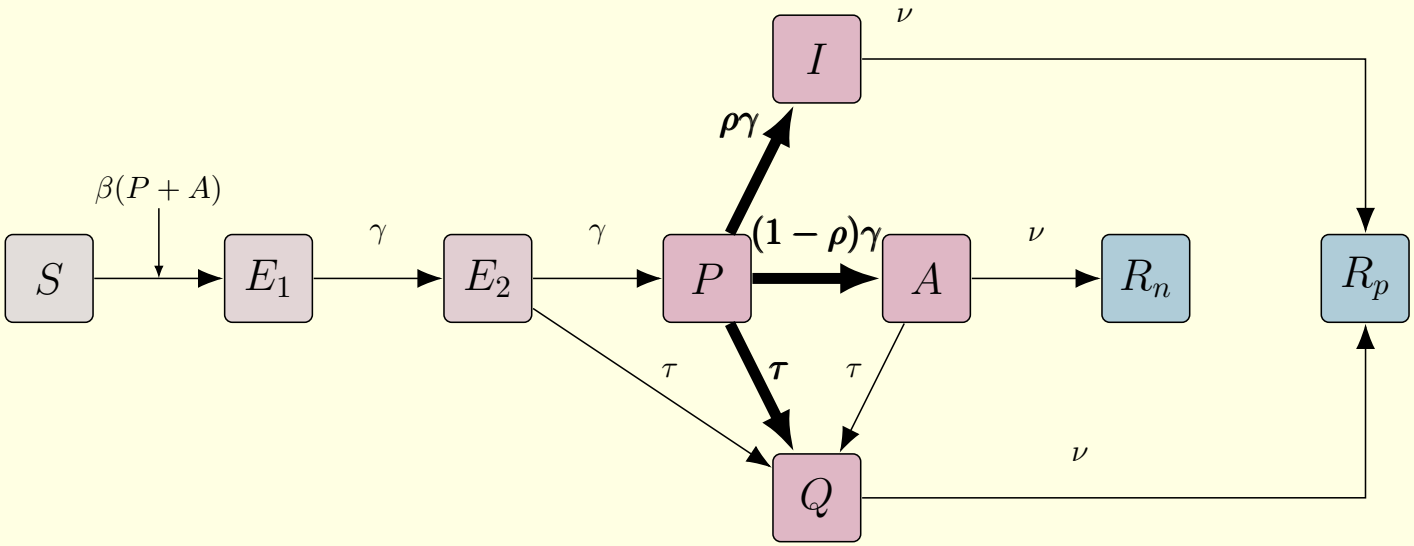
$$\begin{pmatrix} \gamma \\ \hline \gamma \end{pmatrix}$$

Simplified method, Flow-considerations



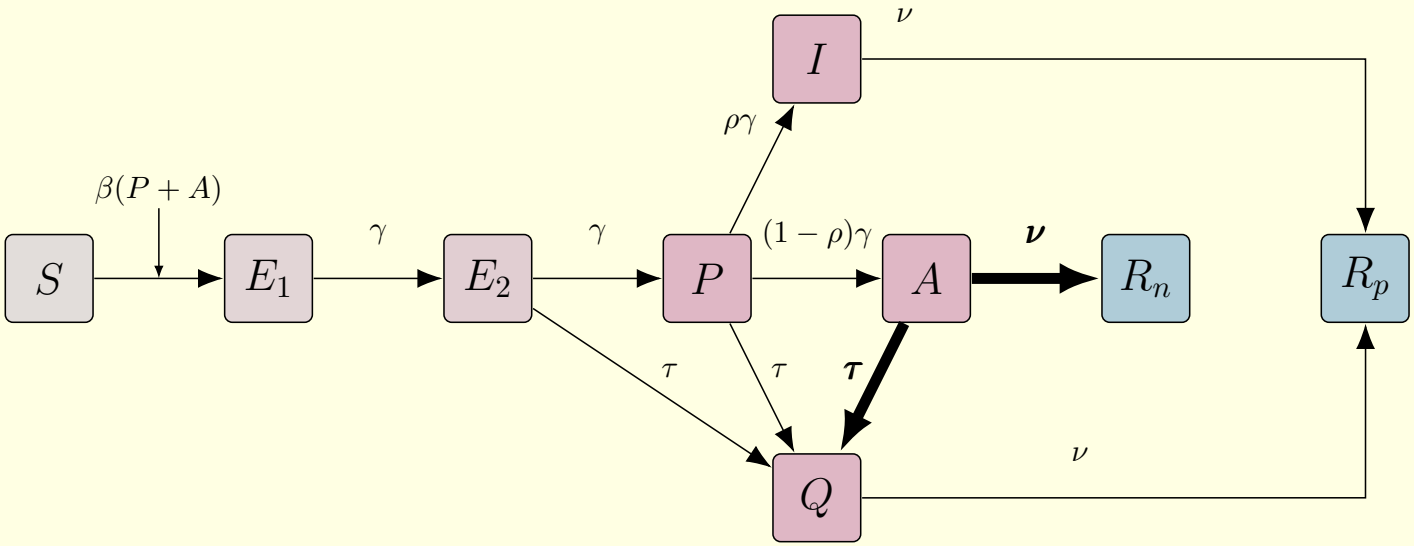
$$\left(\frac{\gamma}{\gamma}\right) \left(\frac{\gamma}{\gamma+\tau}\right)$$

Simplified method, Flow-considerations



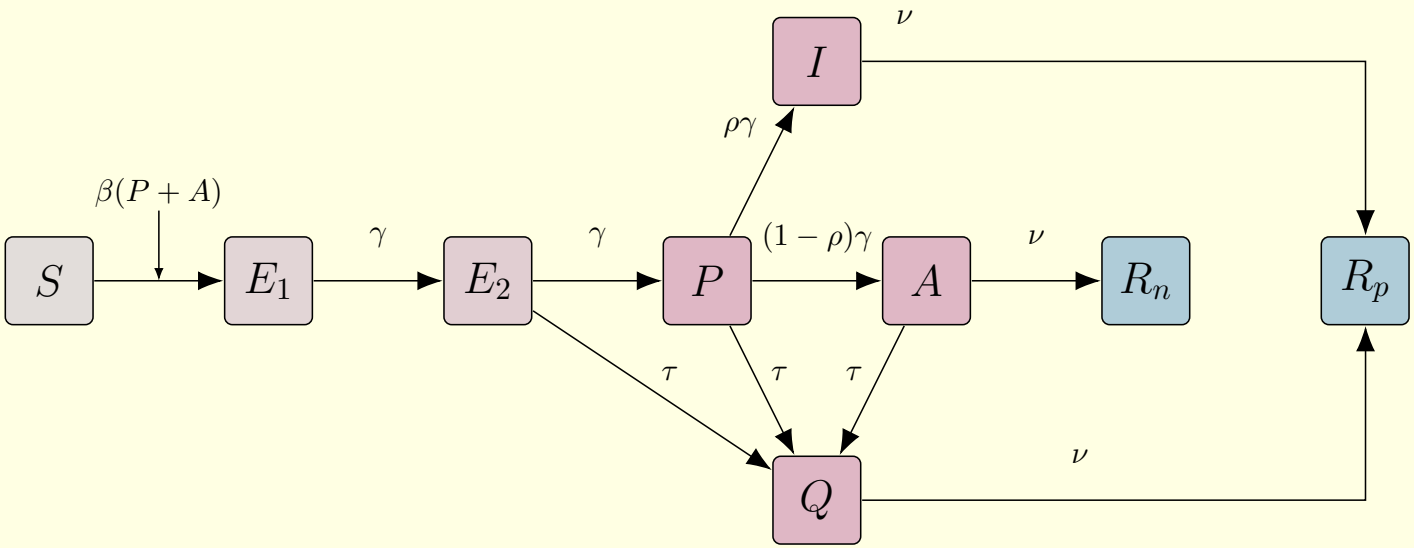
$$\left(\frac{\gamma}{\gamma}\right)\ \left(\frac{\gamma}{\gamma+\tau}\right)\ \left(\frac{(1-\rho)\gamma}{\gamma+\tau}\right)$$

Simplified method, Flow-considerations



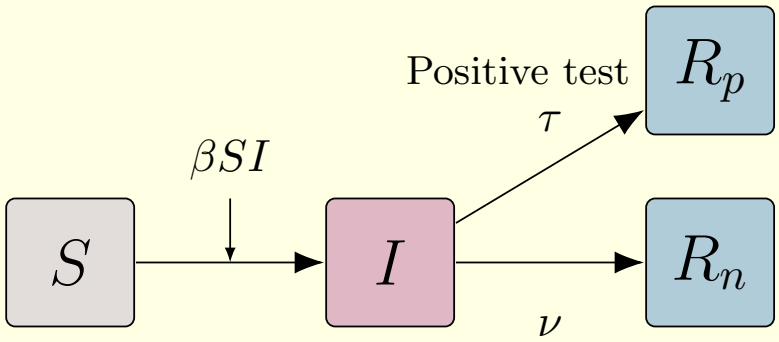
$$\left(\frac{\gamma}{\gamma}\right)\ \left(\frac{\gamma}{\gamma+\tau}\right)\ \left(\frac{(1-\rho)\gamma}{\gamma+\tau}\right)\ \left(\frac{\nu}{\nu+\tau}\right)$$

Simplified method, Flow-considerations



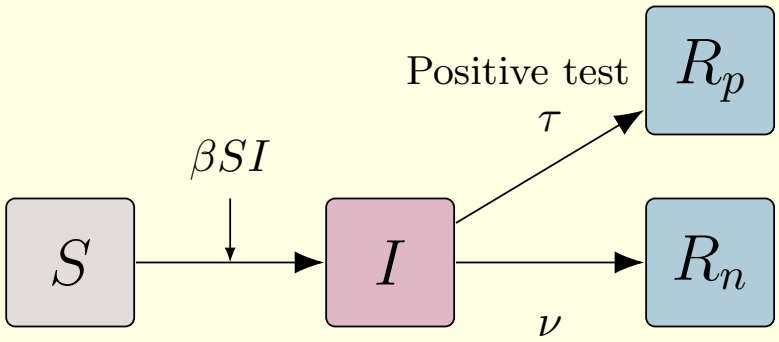
$$\frac{R_n}{R_n + R_p} = \left(\frac{\gamma}{\gamma+\tau}\right)\left(\frac{\gamma(1-\rho)}{\gamma+\tau}\right)\left(\frac{\nu}{\nu+\tau}\right)$$

Extension to other models, example 1



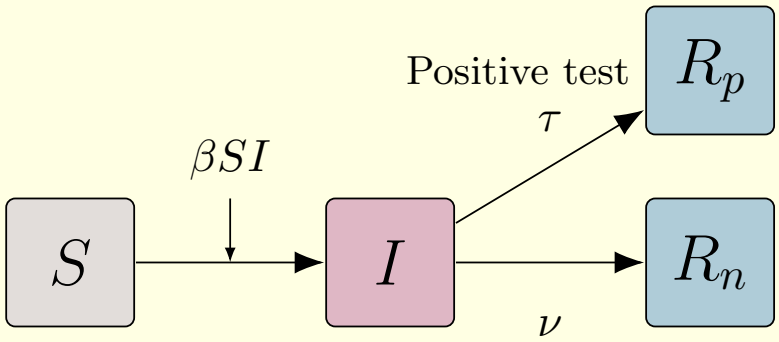
Flow-considerations:

Extension to other models, example 1



Flow-considerations: $\frac{R_n}{R_n + R_p} = \frac{\nu}{\nu + \tau}$

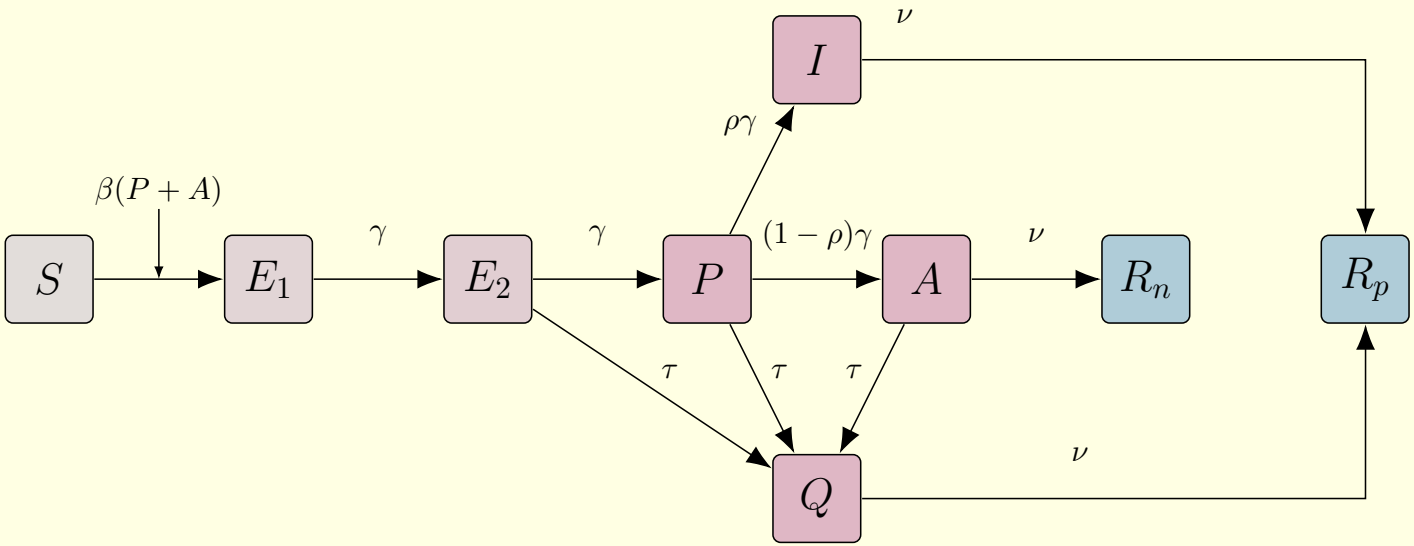
Extension to other models, example 1



Flow-considerations: $\frac{R_n}{R_n + R_p} = \frac{\nu}{\nu + \tau}$

Correction factor: $\frac{\nu + \tau}{\tau}$

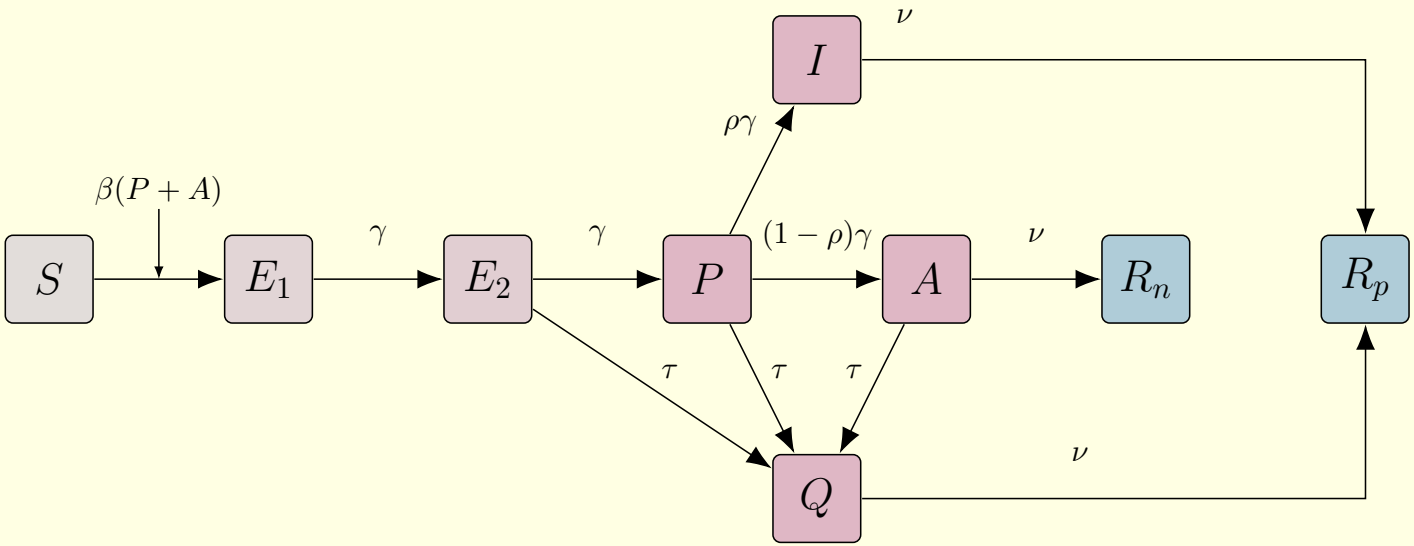
Simplified method, Matrix-form



For SIR-type models[2], the inverse of a matrix V describing flows in the "infected sub-system" is typically computed to determine the reproduction number $\mathcal{R}_0$.

[1] See (van den Drische and Watmough, 2002) for definition and derivation.

Simplified method, Matrix-form

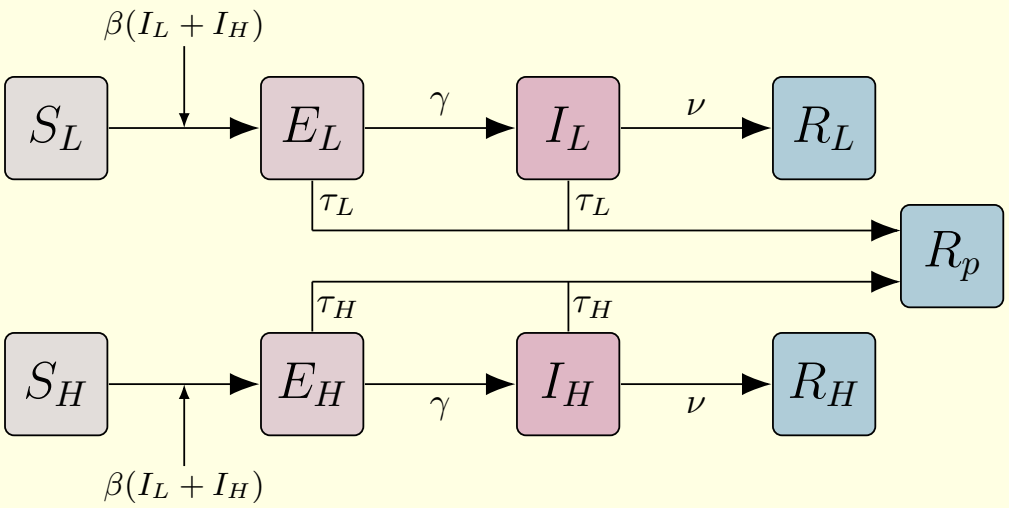


For SIR-type models[2], the inverse of a matrix V describing flows in the "infected sub-system" is typically computed to determine the reproduction number $\mathcal{R}_0$.
With sub-system $x = (E_1, E_2, P, I, A)$ and matrix V, we consider "inputs" $\alpha = (1, 0, 0, 0, 0)$ and "outputs" $\omega = (0, 0, 0, 0, \nu)$, and find that:

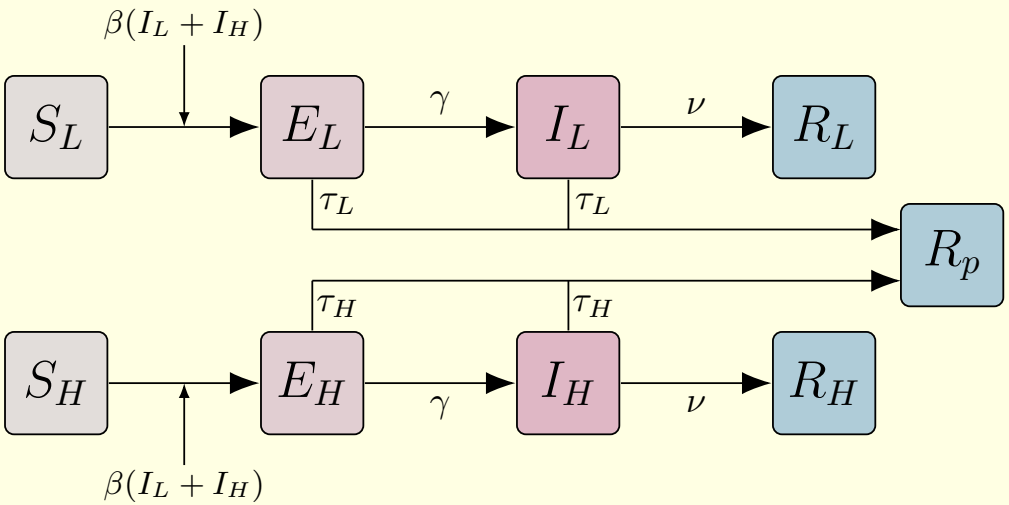
$$\frac{R_n}{R_n + R_p} = \omega V^{-1} \alpha^T$$

[1] See (van den Drische and Watmough, 2002) for definition and derivation.

Extension to other models, example 2

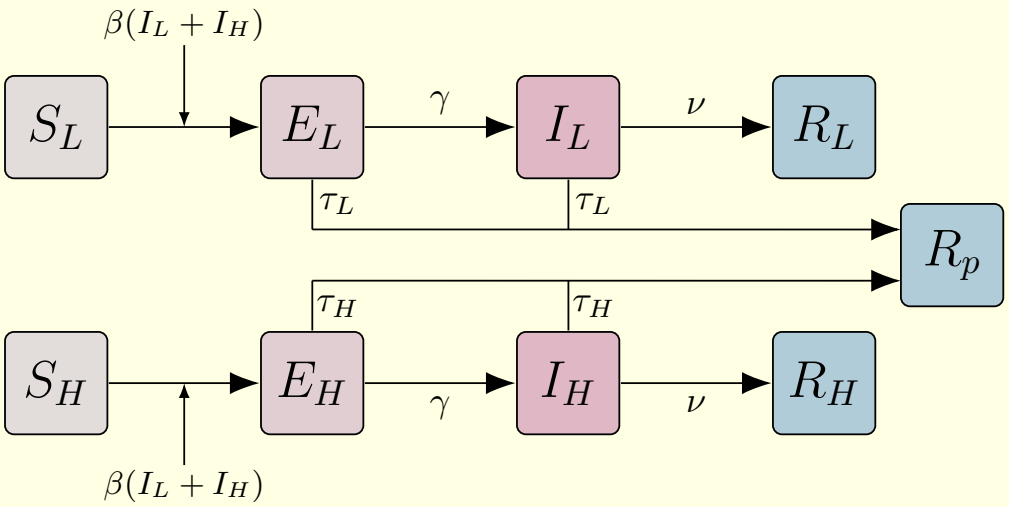


Extension to other models, example 2



$$x = (E_H, E_L, I_H, I_L), \quad V = \begin{pmatrix} \gamma + \tau_H & 0 & 0 & 0 \\ 0 & \gamma + \tau_L & 0 & 0 \\ \gamma & 0 & \nu + \tau_H & 0 \\ 0 & \gamma & 0 & \nu + \tau_L \end{pmatrix}$$

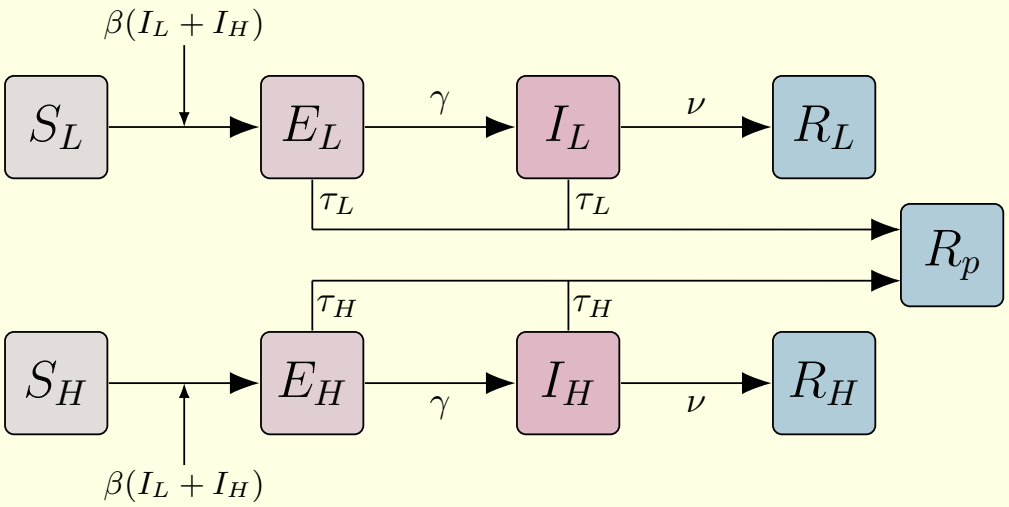
Extension to other models, example 2



$$x = (E_H, E_L, I_H, I_L), \quad V = \begin{pmatrix} \gamma + \tau_H & 0 & 0 & 0 \\ 0 & \gamma + \tau_L & 0 & 0 \\ \gamma & 0 & \nu + \tau_H & 0 \\ 0 & \gamma & 0 & \nu + \tau_L \end{pmatrix}$$

$$\alpha = \left(\frac{S_H}{S_H + S_L}, \frac{S_L}{S_H + S_L}, 0, 0 \right) \text{ and } \omega = (0, 0, \nu, \nu)$$

Extension to other models, example 2



$$x = (E_H, E_L, I_H, I_L), \quad V = \begin{pmatrix} \gamma + \tau_H & 0 & 0 & 0 \\ 0 & \gamma + \tau_L & 0 & 0 \\ \gamma & 0 & \nu + \tau_H & 0 \\ 0 & \gamma & 0 & \nu + \tau_L \end{pmatrix}$$

$$\alpha = \left(\frac{S_H}{S_H + S_L}, \frac{S_L}{S_H + S_L}, 0, 0 \right) \text{ and } \omega = (0, 0, \nu, \nu)$$

Hence: $\omega V^{-1} \alpha^T = \frac{\nu\gamma}{S_H + S_L} \left(\frac{S_H}{(\nu + \tau_H)(\gamma + \tau_H)} + \frac{S_L}{(\nu + \tau_L)(\gamma + \tau_L)} \right)$

Discussion

- We determine a relation between observed COVID-19 cases and total new cases, as a function of testing-rate.

Discussion

- We determine a relation between observed COVID-19 cases and total new cases, as a function of testing-rate.
- This relation may help us compare incidence between countries.

Discussion

- We determine a relation between observed COVID-19 cases and total new cases, as a function of testing-rate.
- This relation may help us compare incidence between countries.
- Our initial analysis was model-specific and based on calculations of final-size of variables.

Discussion

- We determine a relation between observed COVID-19 cases and total new cases, as a function of testing-rate.
- This relation may help us compare incidence between countries.
- Our initial analysis was model-specific and based on calculations of final-size of variables.
- Our new method follows from well-known results from the litterature, and requires only observing the model diagram or inverting a matrix...

Discussion

- We determine a relation between observed COVID-19 cases and total new cases, as a function of testing-rate.
- This relation may help us compare incidence between countries.
- Our initial analysis was model-specific and based on calculations of final-size of variables.
- Our new method follows from well-known results from the litterature, and requires only observing the model diagram or inverting a matrix, but also extends to a wider family of SIR-type models.

Discussion

- We determine a relation between observed COVID-19 cases and total new cases, as a function of testing-rate.
- This relation may help us compare incidence between countries.
- Our initial analysis was model-specific and based on calculations of final-size of variables.
- Our new method follows from well-known results from the litterature, and requires only observing the model diagram or inverting a matrix, but also extends to a wider family of SIR-type models.
- *My lesson from this:* When working on modelling problems, look for simpler answers to the problems...

Discussion

- We determine a relation between observed COVID-19 cases and total new cases, as a function of testing-rate.
- This relation may help us compare incidence between countries.
- Our initial analysis was model-specific and based on calculations of final-size of variables.
- Our new method follows from well-known results from the litterature, and requires only observing the model diagram or inverting a matrix, but also extends to a wider family of SIR-type models.
- *My lesson from this:* When working on modelling problems, look for simpler answers to the problems, before throwing yourself at the analysis and simulation!

Thank you for your attention.

Feel free to email me
with questions or comments

Website:
rasmuspedersen.com
Email: rakrpe@ruc.dk