Statistical methods for epidemiological surveillance

Leo Bastos (PROCC/Fiocruz) and Luiz Max Carvalho (FGV/EMAp)

July 4, 2022

Outline

(Infectious) Disease process

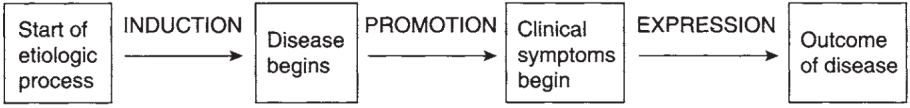


Figure 1.1 *Schematic of disease evolution.*

1. Start of etiologic process
   - Ex: Infection through a mosquito bite (dengue fever)
2. Disease begins
   - Viral replication
3. Clinical symptoms
   - Fever, headaches,...
4. Outcome of disease
   - Cure, hospitalization, death

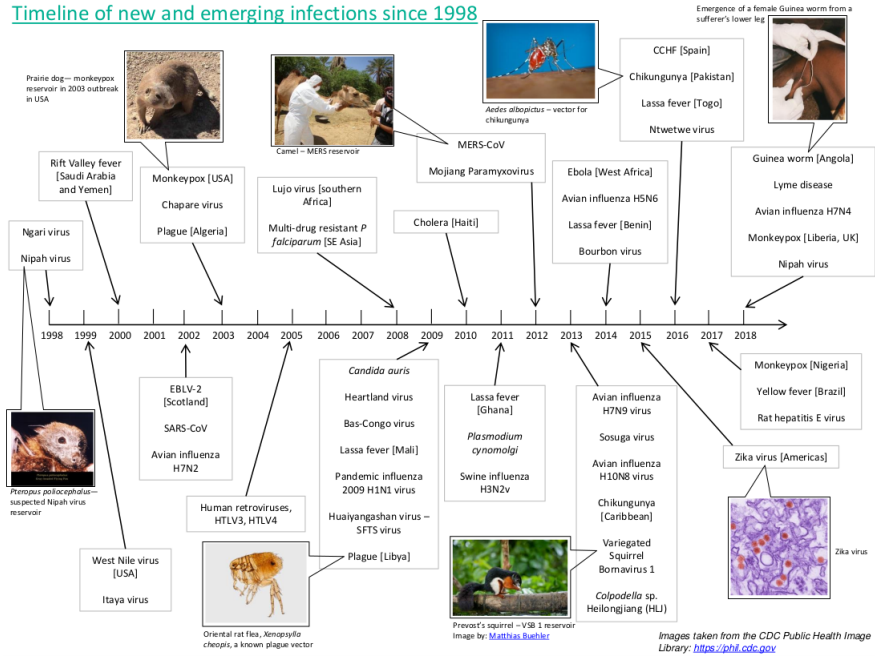
Epidemiological study of Disease process

1. Start of etiologic process
   - Can we avoid infection?
2. Disease begins
   - Is our immune system prepared?
3. Clinical symptoms
   - Can we treat or avoid evolution?
4. Outcome of disease
   - How to reduce the burden?

Infectious disease epidemiology (IDE)

- Infectious disease epidemiology (IDE) is the study of how and why infectious diseases emerge and spread among different **populations**, and what strategies can prevent or contain the spread of disease at the population level.
- Why is this important?

Timeline of new and emerging infections since 1998



Data type: In Infectious Disease Epidemiology

- Infectious disease data is mainly binary

$$Z_i = \begin{cases} 1 & \text{person } i \text{ is infected with pathogen A or has a disease D,} \\ 0 & \text{otherwise.} \end{cases}$$

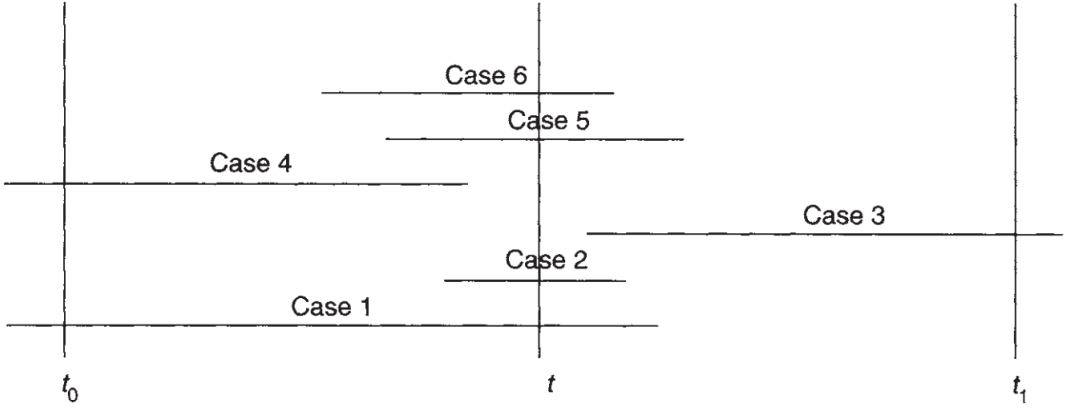
- Counts of cases through time

$$Y_t = \sum_i Z_{i,t}$$

- Induce time and spatial dependence is important, since dependence is present by definition of infectious diseases

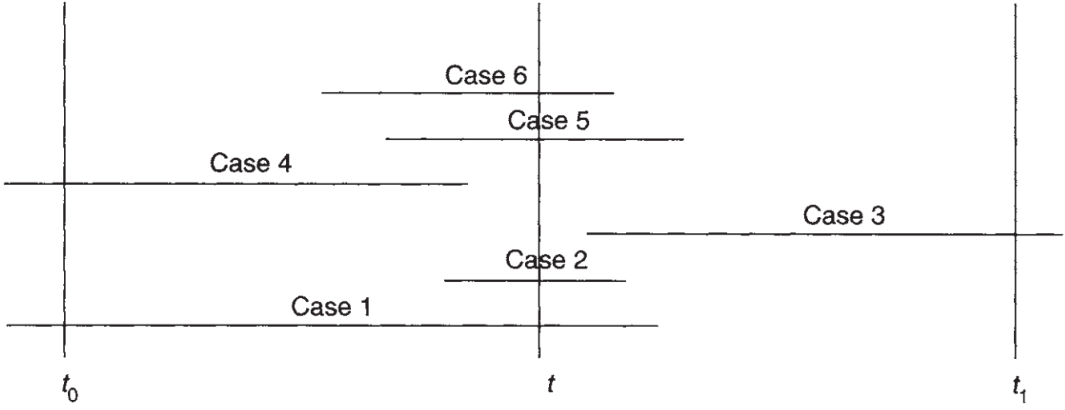
Important measures of disease occurrence

- Disease prevalence and incidence both represent proportions of a population determined to be diseased at certain times.
- Suppose there are 100 people being followed.



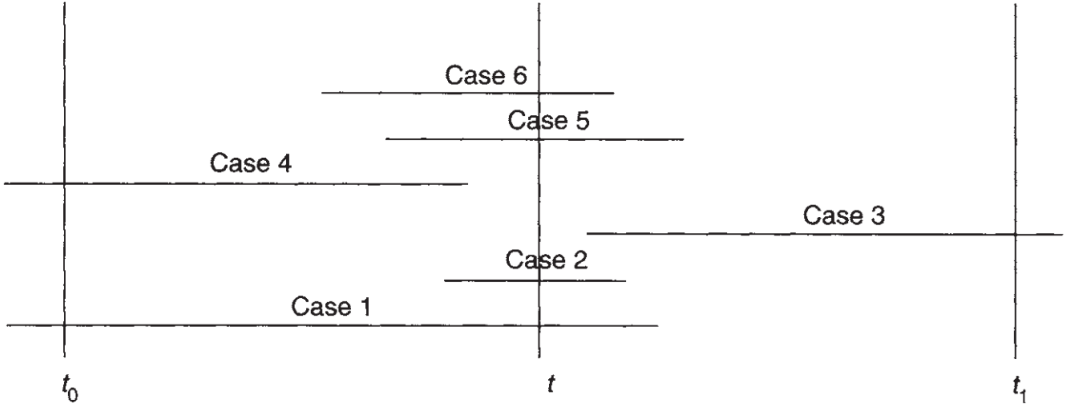
Important measures of disease occurrence

- **Disease prevalence** at time t: 4 / 100 or 4 / 99

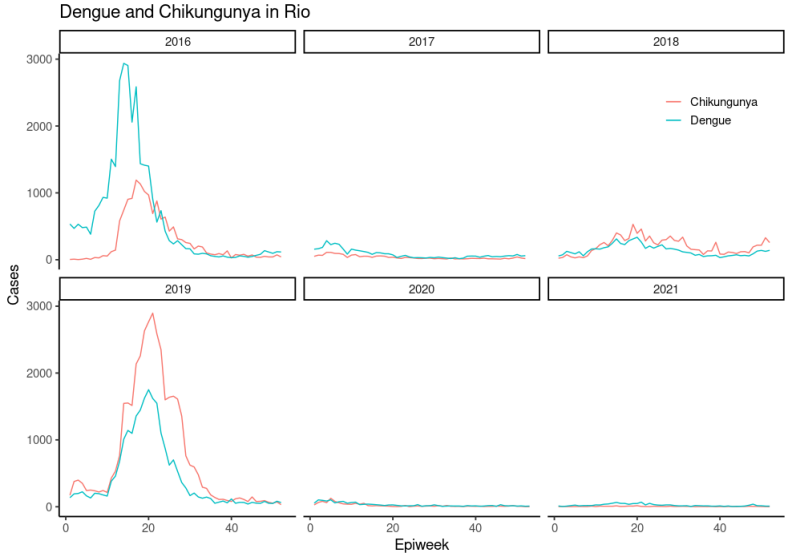


Important measures of disease occurrence

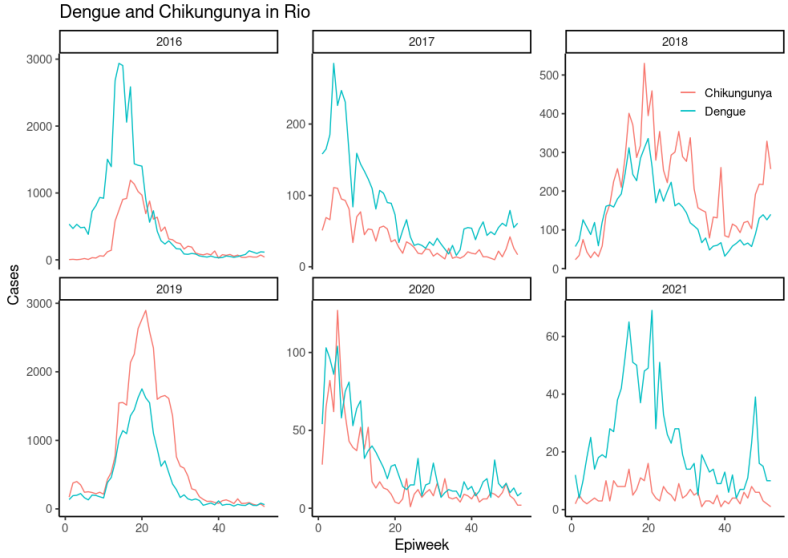
- **Disease incidence** (new cases) in time $[t_0, t_1]$: 4 / 98



Dengue and Chikungunya new cases in Rio de Janeiro



Dengue and Chikungunya new cases in Rio de Janeiro



Surtos, endemias, epidemias, pandemias

- **Surto**: Aumento repentino de casos em uma região específica.
  - Ex: Aumento de casos de sindrome gripal no Vidigal, Rio de Janeiro (Nov/21).
- **Epidemia**: Aumento repentino de casos em várias regiões distintas.
  - Ex: Aumento de casos de sindrome gripal no Rio de Janeiro (Dez/21).
- **Pandemia**: Epidemias da doença em várias regiões do planeta. (Influenza/H1N1, COVID-19)
- **Endemia**: Não tem a ver com a frequencia, dizemos que uma doença é endêmica numa região quando ela é frequente nessa região, podendo ser sazonal. (Febre amarela em algumas regiões do Brasil, malária na região norte)

Modelando surtos e epidemias

- Um modelo de séries temporais para dados de contagem pode ser usado para modelar surtos e epidemias.
- Baixo poder preditivo se o conhecimento a respeito da dinâmica da doença for pequeno e se o histórico da doença não existir (doenças emergentes)
- Doenças endêmicas com um bom histórico de dados são melhores para prever, as previsões podem ser vistas como o que é esperado para definir uma epidemia.

Definindo limiares epidêmicos

DOI:10.1111/j.1750-2659.2012.00422.x
www.influenzajournal.com
Original Article

Influenza surveillance in Europe: establishing epidemic thresholds by the Moving Epidemic Method

Tomás Vega,[a] Jose Eugenio Lozano,[a] Tamara Meerhoff,[b] René Snacken,[c] Joshua Mott,[d] Raul Ortiz de Lejarazu,[e] Baltazar Nunes[f]

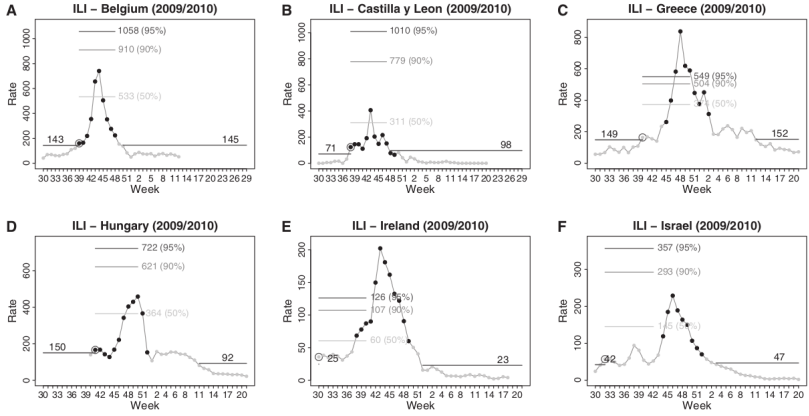
[a]Dirección General de Salud Pública, Consejería de Sanidad, Valladolid, Spain. [b]The Radboud University Nijmegen Medical Centre, Nijmegen, The Netherlands. [c]European Centre for Disease Prevention and Control, Stockholm, Sweden. [d]Division of Health Security, Infectious Diseases and the Environment, WHO Regional Office for Europe, Copenhagen, Denmark. [e]Centro Nacional de Gripe de Valladolid, Universidad de Valladolid, Spain. [f]Instituto Nacional de Saude Doctor Ricardo Jorge, Lisboa, Portugal.
Correspondence: Tomás Vega, Dirección General de Salud Pública, Consejería de Sanidad, Paseo de Zorrilla, 1. 47071 Valladolid, Spain. E-mail: vegaloto@jcyl.es

Accepted 27 June 2012. Published Online 16 August 2012.

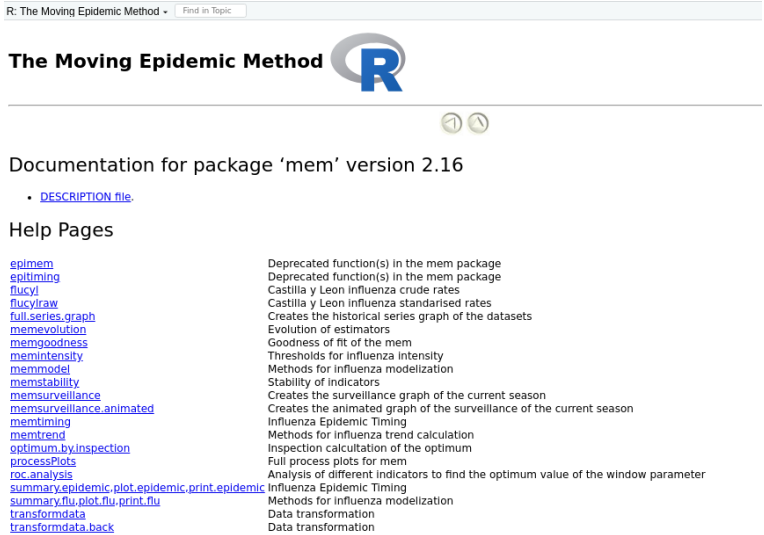
MEM: Moving Epidemic Method

- Proposto por Vega et al. (2012) para modelar Influenza
- O MEM é um algoritmo que consiste em três etapas:
  1. Determinar o início, a duração e o fim de um período epidêmico anual
  2. Calcular usando dados históricos limiares de incidência pré- e pós-epidêmicos
  3. Calcular diferentes níveis de intensidade para o período epidêmico

Definindo limiares epidêmicos



Definindo limiares epidêmicos



R: The Moving Epidemic Method — Find in Topic

The Moving Epidemic Method

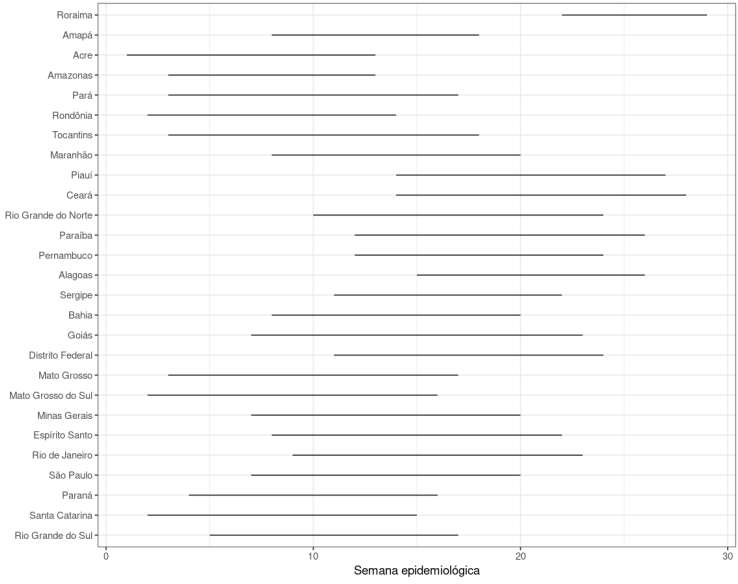
Documentation for package ‘mem’ version 2.16

- DESCRIPTION file.

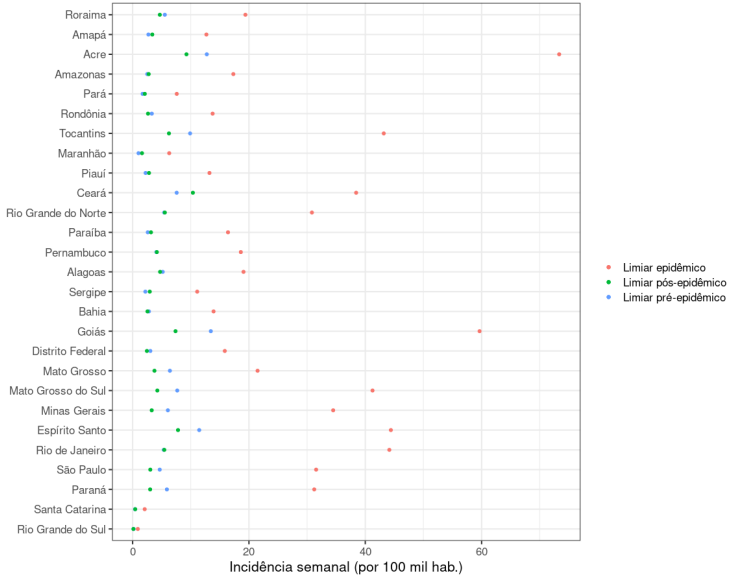
Help Pages

epimem	Deprecated function(s) in the mem package
epitiming	Deprecated function(s) in the mem package
flucyl	Castilla y Leon influenza crude rates
flucylraw	Castilla y Leon influenza standarised rates
full.series.graph	Creates the historical series graph of the datasets
memevolution	Evolution of estimators
memgoodness	Goodness of fit of the mem
memintensity	Thresholds for influenza intensity
memmodel	Methods for influenza modelization
memstability	Stability of indicators
memsurveillance	Creates the surveillance graph of the current season
memsurveillance.animated	Creates the animated graph of the surveillance of the current season
memtiming	Influenza Epidemic Timing
memtrend	Methods for influenza trend calculation
optimum.by.inspection	Inspection calcultation of the optimum
processPlots	Full process plots for mem
roc.analysis	Analysis of different indicators to find the optimum value of the window parameter
summary.epidemic,plot.epidemic,print.epidemic	Influenza Epidemic Timing
summary.flu,plot.flu,print.flu	Methods for influenza modelization
transformdata	Data transformation
transformdata.back	Data transformation
transformseries	Transformation of series of data

Início e duração do período epidêmico (dengue)

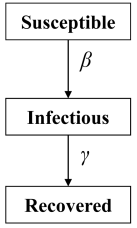


Limiares epidêmicos por UF (dengue)



A classical model

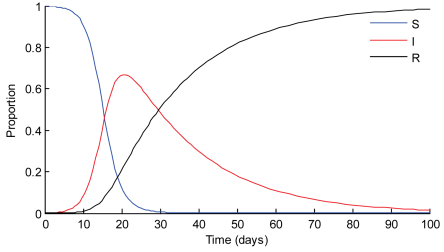
Luz, Struchiner & Galvani (2010).



$$\frac{dS}{dt} = -\beta SI$$

$$\frac{dI}{dt} = \beta SI - \gamma I$$

$$\frac{dR}{dt} = \gamma I$$



Analysing a model

We have

$$S(t) + I(t) + R(t) = N \forall t$$

The basic reproductive number is

$$\mathcal{R}_0 = \frac{\beta N}{\gamma}. \tag{1}$$

Moreover,

$$\lim_{t \to \infty} I(t) = 0 \implies (S_e, 0, R_e),$$

is the only equilibrium point. Consider the Jacobian

$$J(S, I) = \begin{bmatrix} -\beta I & -\beta S \\ \beta I & \beta S - \gamma \end{bmatrix}.$$

Analysing a model (cont.)

- **Equilibria**: The characteristic polynomial is

$$\lambda^2 - (\beta S_e - \gamma)\,\lambda = 0,$$

thus $\lambda_1 = 0$ and $\lambda_2 = \beta S_e - \gamma$. This means we have neutral stability if $S_e = \gamma/\beta$, i.e., $\lambda_2 = 0$ and instability otherwise ($\lambda_2 > 0$ or $\lambda_2 < 0$).

- **Epidemic regimes**:

$$\frac{dI}{dS} = -1 + \frac{\gamma}{\beta}\frac{1}{S},$$

gives

$$I(t) = 1 - R(0) - S(t) + \frac{\gamma}{\beta}\ln\left(\frac{S(t)}{S(0)}\right).$$

From $\lim_{t\to\infty} I(t) = 0$ we know that

$$S(\infty) = 1 - R(0) + \frac{\gamma}{\beta}\ln\left(\frac{S(\infty)}{S(0)}\right).$$

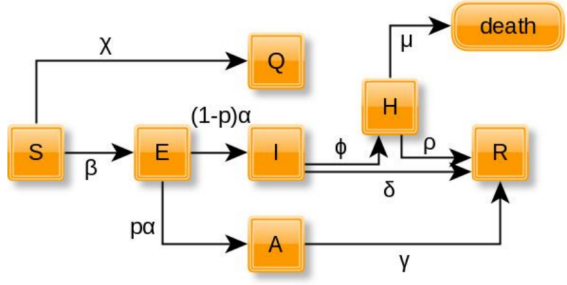
Thus **an epidemic occurs iff**

$$\frac{\beta S(0)}{\gamma} > 1.$$

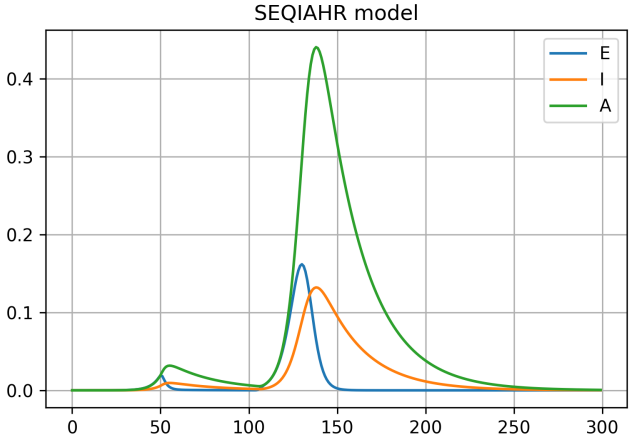
Big structured epidemic models

Coelho et al. (2020)

$$
\begin{aligned}
\frac{dS}{dt} &= -\eta[(1-\chi)S], \\
\frac{dE}{dt} &= \eta[(1-\chi)S] - \alpha E, \\
\frac{dI}{dt} &= (1-p)\alpha E - \delta I - \phi I, \\
\frac{dA}{dt} &= p\alpha E - \gamma A, \\
\frac{dH}{dt} &= \phi I - (\rho+\mu)H, \\
\frac{dR}{dt} &= \delta I + \rho H + \gamma A, \\
\eta &:= \beta(I+A).
\end{aligned}
$$

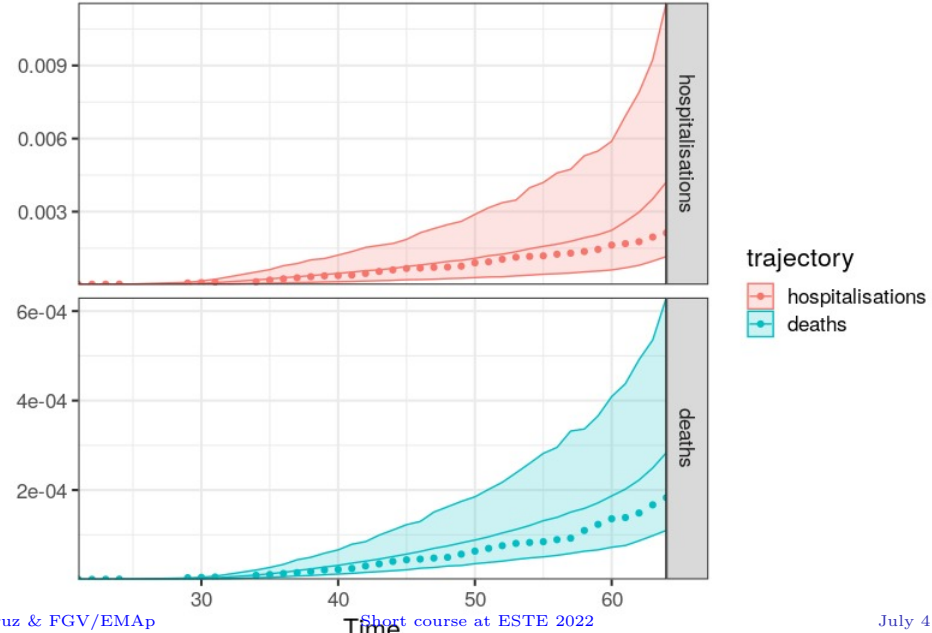


Big model (cont.)



$$\mathcal{R}_0 = \frac{\beta(1-\xi)[p(\phi+\delta)+(1-p)\gamma]}{\gamma(\delta+\phi)}$$

Why so rigid?


0.009
0.006
0.003
hospitalisations
6e-04
4e-04
2e-04
deaths
30
40
50
60
Time
trajectory
hospitalisations
deaths

Stochastic models

Description	State transition	rate
Infection	$(S, E) \to (S-1, E+1)$	$\lambda(1-\chi)S$
Exposed to I	$(E, I) \to (E-1, I+1)$	$(1-p)\alpha E$
Exposed to A	$(E, A) \to (E-1, A+1)$	$p\alpha E$
Hospitalization	$(I, H) \to (I-1, H+1)$	$\phi\ I$
Recovery of I	$(I, R) \to (I-1, R+1)$	δI
Recovery of A	$(A, R) \to (A-1, R+1)$	γA
Recovery of H	$(H, R) \to (H-1, R+1)$	ρH
Death of H	$H \to H-1$	μH

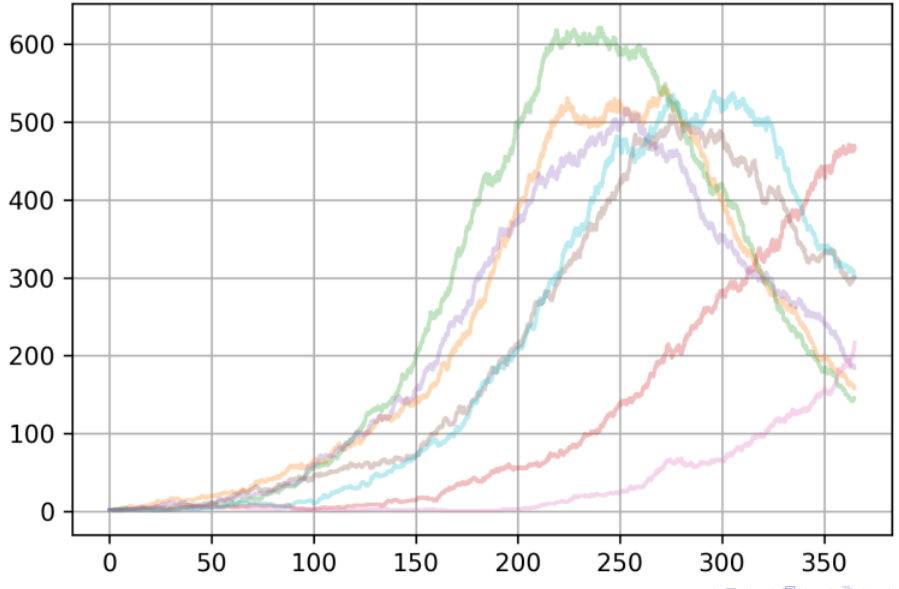
State probabilities

$$P_{s,e,i,a,h}(t) := \mathbb{P}\left(S = s, E = e, I = i, A = a, H = h\right).$$

FCK equation:

$$\begin{aligned}\frac{dP_{s,e,i,a,h}}{dt} &= P_{s+1,e-1,i,a,h}\lambda(1-\chi)(s+1) + P_{s,e+1,i-1,a,h}(1-p)\alpha(e+1) \\ &+ P_{s,e+1,i,a-1,h}p\alpha(e+1) + P_{s,e,i+1,a,h-1}\phi(i+1) + P_{s,e,i+1,a,h}\delta(i+1) \\ &+ P_{s,e,i,a+1,h}\gamma(a+1) + P_{s,e,i,a,h+1}\rho(h+1) + P_{s,e,i,a,h+1}\mu(h+1).\end{aligned}$$

Trajectories



Analysing a stochastic model

Probability-generating functions Starting with $I_i(0)$, the probability of an infected individual in state i producing offspring of type j given that $I_j(0)$ can be obtained from

$$f_i(z_1, \ldots, z_k) = \sum_{j_k=0}^{\infty} \cdots \sum_{j_1=0}^{\infty} P_i(z_1, \ldots, z_k) z_1^{j_1} \cdots z_k^{j_k}.$$

Now, define a matrix whose entries $m_{ji} = \frac{\partial f_i}{\partial u_i}|_{u=1}$ are the expected number of offspring generated in $i \to j$.

$$\mathbb{M} := \begin{bmatrix} 0 & 1-p & p \\ \frac{\beta(1-\chi)}{\beta(1-\chi)+\delta+\phi} & \frac{\beta(1-\chi)}{\beta(1-\chi)+\delta+\phi} & 0 \\ \frac{\beta(1-\chi)}{\beta(1-\chi)+\gamma} & 0 & \frac{\beta(1-\chi)}{\beta(1-\chi)+\gamma} \end{bmatrix}$$

Analysing a stochastic model (cont.)

Under some regularity conditions, we can calculate the **extinction probability**:

$$\mathbb{P}_0 = \prod_{i=1}^{3} q_i^{k_i}$$

after finding $(q_1, q_2, q_3) \in (0,1)^3$ which satisfy constraints. Here $k_1 = E(0)$, $k_2 = I(0)$ and $k_3 = A(0)$.

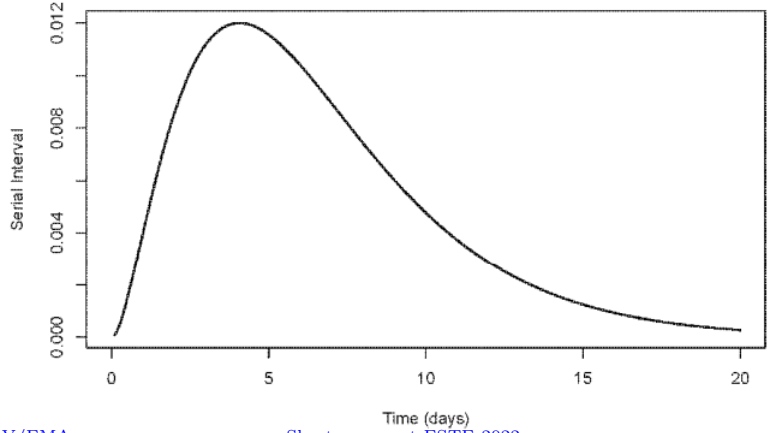
I_0	E_0	A_0	**Approx.** $\mathbb{P}_0$	**SEIAHR** $\mathbb{P}_0$	**SIR** $\mathbb{P}_0$
1	0	0	0.63	0.64	0.58
2	0	0	0.43	0.41	0.33
3	0	0	0.25	0.26	0.19
4	0	0	0.18	0.17	0.11
5	0	0	0.10	0.10	0.06

No more differential equations, please!

Start by looking at

$$R_t = \frac{I_t}{\sum_{s=1}^{t} I_{t-s} w_s}.$$

Then R_t is "the average number of secondary cases that each infected individual would infect if the conditions remained as they were at time t" (Cori et al. 2013).



Modernising the model

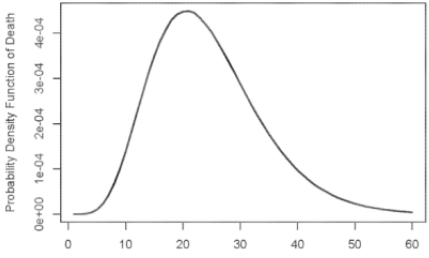
$$R_{t,m} = R_{0,m}\left(2\,\text{logit}^{-1}\left(-\sum_{k=1}^{4}(\alpha_k + \beta_{mk})X_{ktm}\right)\right)$$

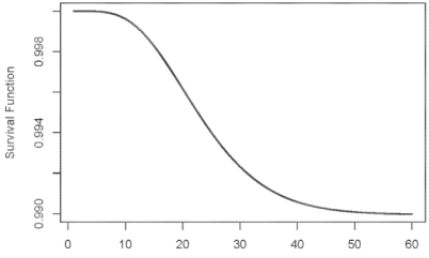
Priors:

$$\begin{aligned}
\alpha_k &\sim \text{Normal}(0, 5);\\
\beta_{m,k} &\sim \text{Normal}(0, \gamma);\\
\gamma &\sim \text{HalfNormal}(0, 5);\\
R_{0,m} &\sim \text{Normal}(3.28, \kappa);\\
\kappa &\sim \text{HalfNormal}(0, 1/2).
\end{aligned}$$

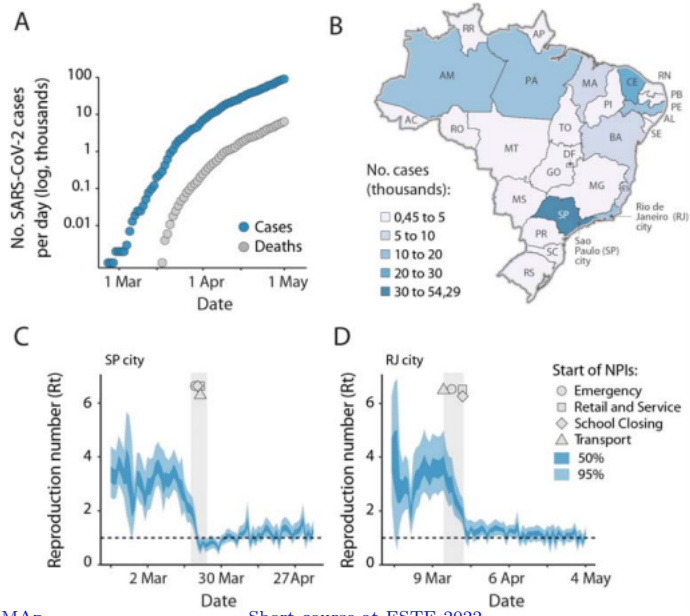
Modernising the model II: likelihood

$$
\begin{aligned}
c_{t,m} &= R_{t,m} \sum_{\tau=0}^{t-1} c_{\tau,m} g_{t-\tau}, \\
d_{t,m} &= \sum_{\tau=0}^{t-1} c_{\tau,m} \pi_{t-\tau,m}, \\
D_{t,m} &\sim \text{NegativeBinomial}\left(d_{t,m}, \phi\right), \\
\phi &\sim \text{HalfNormal}(0, 5).
\end{aligned}
$$





The end product



Frame Title

xxx