SIR and SEIR modelling for infectious disease dynamics

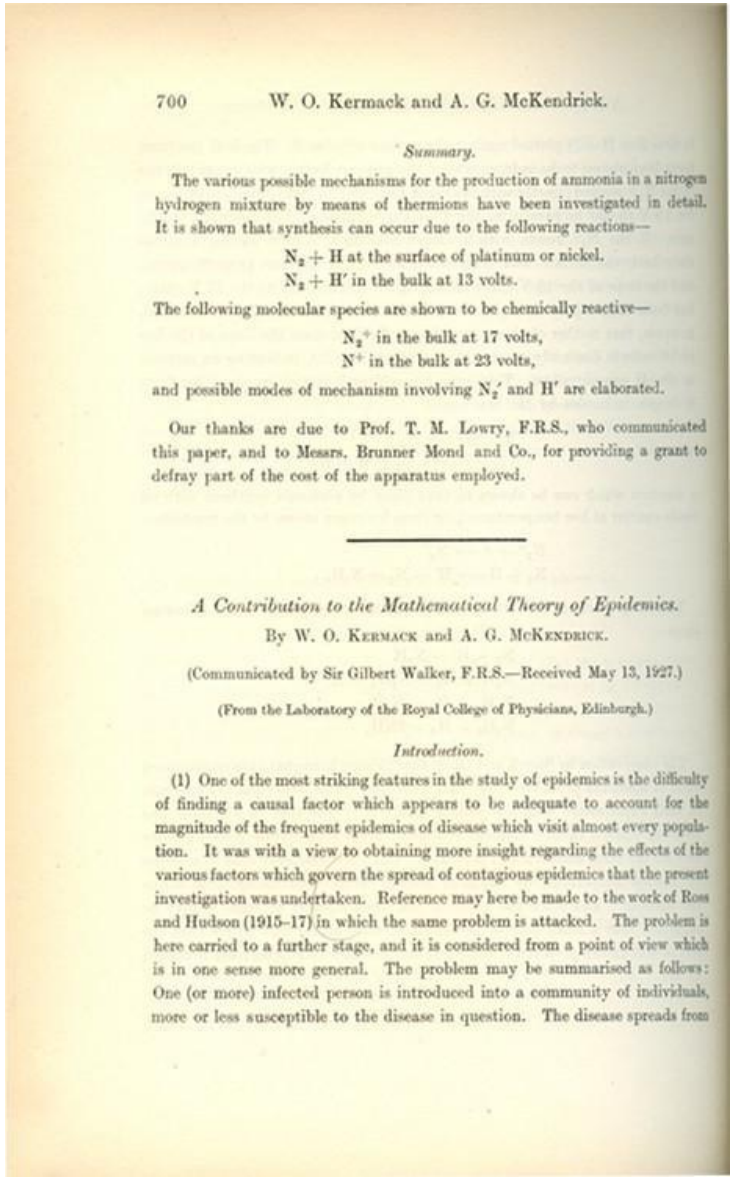
Basics, assumptions, and interpretation

10 June 2020

Julia Shen

Queries and corrections very welcome to julia.shen1@lshtm.ac.uk

Origins of Susceptible-Infected-Removed models



700 W. O. Kermack and A. G. McKendrick.

Summary.

The various possible mechanisms for the production of ammonia in a nitrogen hydrogen mixture by means of thermions have been investigated in detail. It is shown that synthesis can occur due to the following reactions—

$N_2 + H$ at the surface of platinum or nickel.
$N_2 + H'$ in the bulk at 13 volts.

The following molecular species are shown to be chemically reactive—

N_2^+ in the bulk at 17 volts,
N^+ in the bulk at 23 volts,

and possible modes of mechanism involving N_2' and H' are elaborated.

Our thanks are due to Prof. T. M. Lowry, F.R.S., who communicated this paper, and to Messrs. Brunner Mond and Co., for providing a grant to defray part of the cost of the apparatus employed.

A Contribution to the Mathematical Theory of Epidemics.

By W. O. Kermack and A. G. McKendrick.

(Communicated by Sir Gilbert Walker, F.R.S.—Received May 13, 1927.)

(From the Laboratory of the Royal College of Physicians, Edinburgh.)

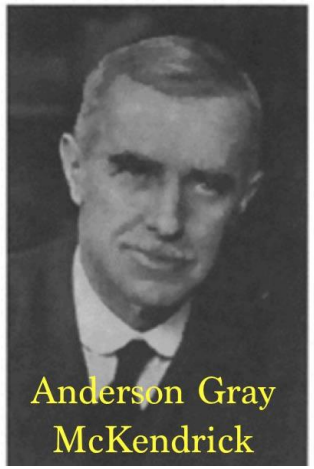
Introduction.

(1) One of the most striking features in the study of epidemics is the difficulty of finding a causal factor which appears to be adequate to account for the magnitude of the frequent epidemics of disease which visit almost every population. It was with a view to obtaining more insight regarding the effects of the various factors which govern the spread of contagious epidemics that the present investigation was undertaken. Reference may here be made to the work of Ross and Hudson (1915–17) in which the same problem is attacked. The problem is here carried to a further stage, and it is considered from a point of view which is in one sense more general. The problem may be summarised as follows: One (or more) infected person is introduced into a community of individuals, more or less susceptible to the disease in question. The disease spreads from

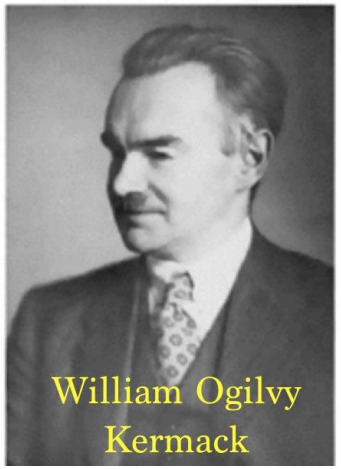
"A contribution to the mathematical theory of epidemics"

https://doi.org/10.1098/rspa.1927.0118

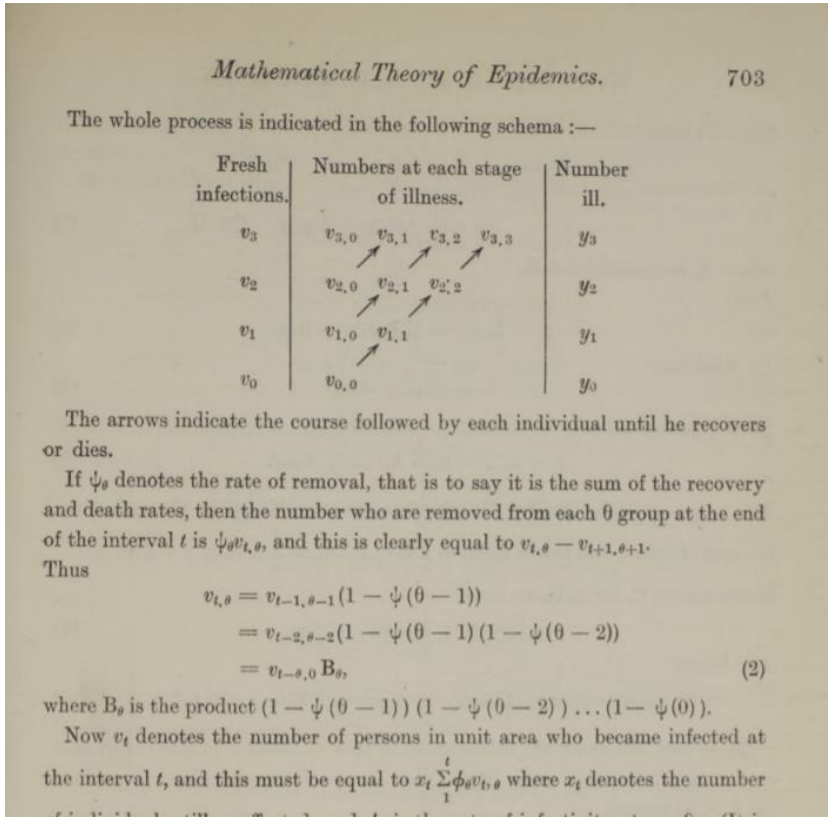
1 August 1927



Anderson Gray McKendrick



William Ogilvy Kermack



Mathematical Theory of Epidemics. 703

The whole process is indicated in the following schema :—

Fresh infections.	Numbers at each stage of illness.	Number ill.
v_3	$v_{3,0}$ $v_{3,1}$ $v_{3,2}$ $v_{3,3}$	y_3
v_2	$v_{2,0}$ $v_{2,1}$ $v_{2,2}$	y_2
v_1	$v_{1,0}$ $v_{1,1}$	y_1
v_0	$v_{0,0}$	y_0

The arrows indicate the course followed by each individual until he recovers or dies.

If ψ_θ denotes the rate of removal, that is to say it is the sum of the recovery and death rates, then the number who are removed from each θ group at the end of the interval t is $\psi_\theta v_{t,\theta}$, and this is clearly equal to $v_{t,\theta} - v_{t+1,\theta+1}$.
Thus

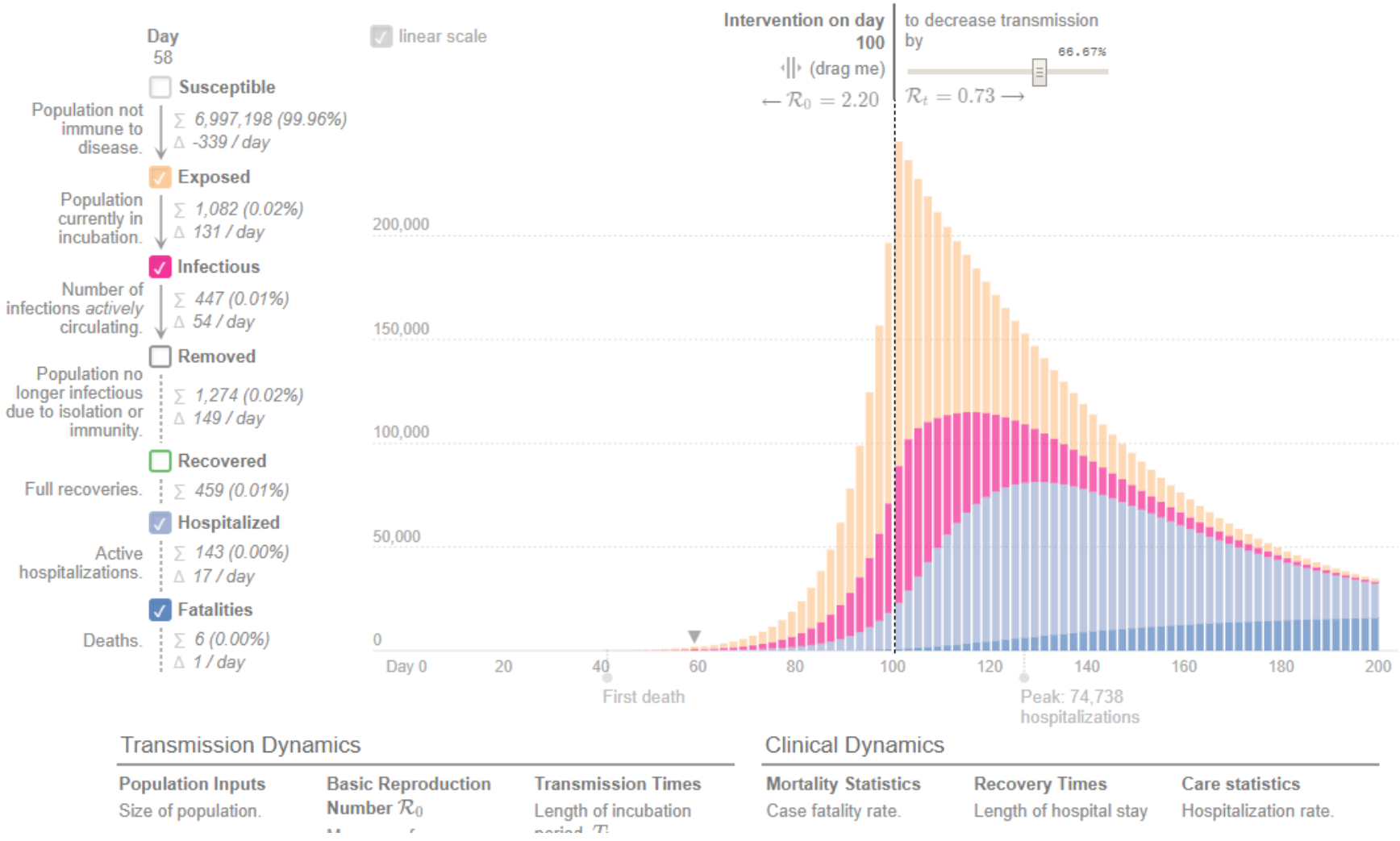
$$\begin{aligned} v_{t,\theta} &= v_{t-1,\theta-1}(1 - \psi(\theta - 1)) \\ &= v_{t-2,\theta-2}(1 - \psi(\theta - 1))(1 - \psi(\theta - 2)) \\ &= v_{t-\theta,0} B_\theta, \end{aligned} \qquad (2)$$

where B_θ is the product $(1 - \psi(\theta - 1))(1 - \psi(\theta - 2)) \ldots (1 - \psi(0))$.

Now v_t denotes the number of persons in unit area who became infected at the interval t, and this must be equal to $x_t \sum_1^t \phi_\theta v_{t,\theta}$ where x_t denotes the number

Example SIR model outputs

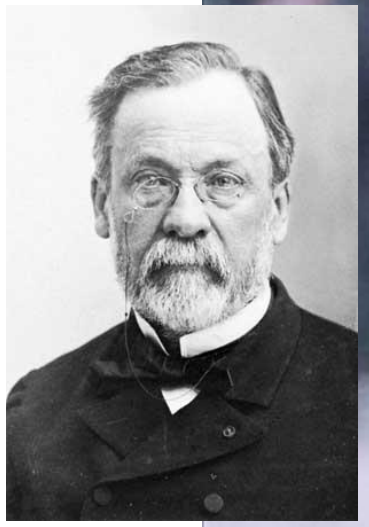
Epidemic Calculator

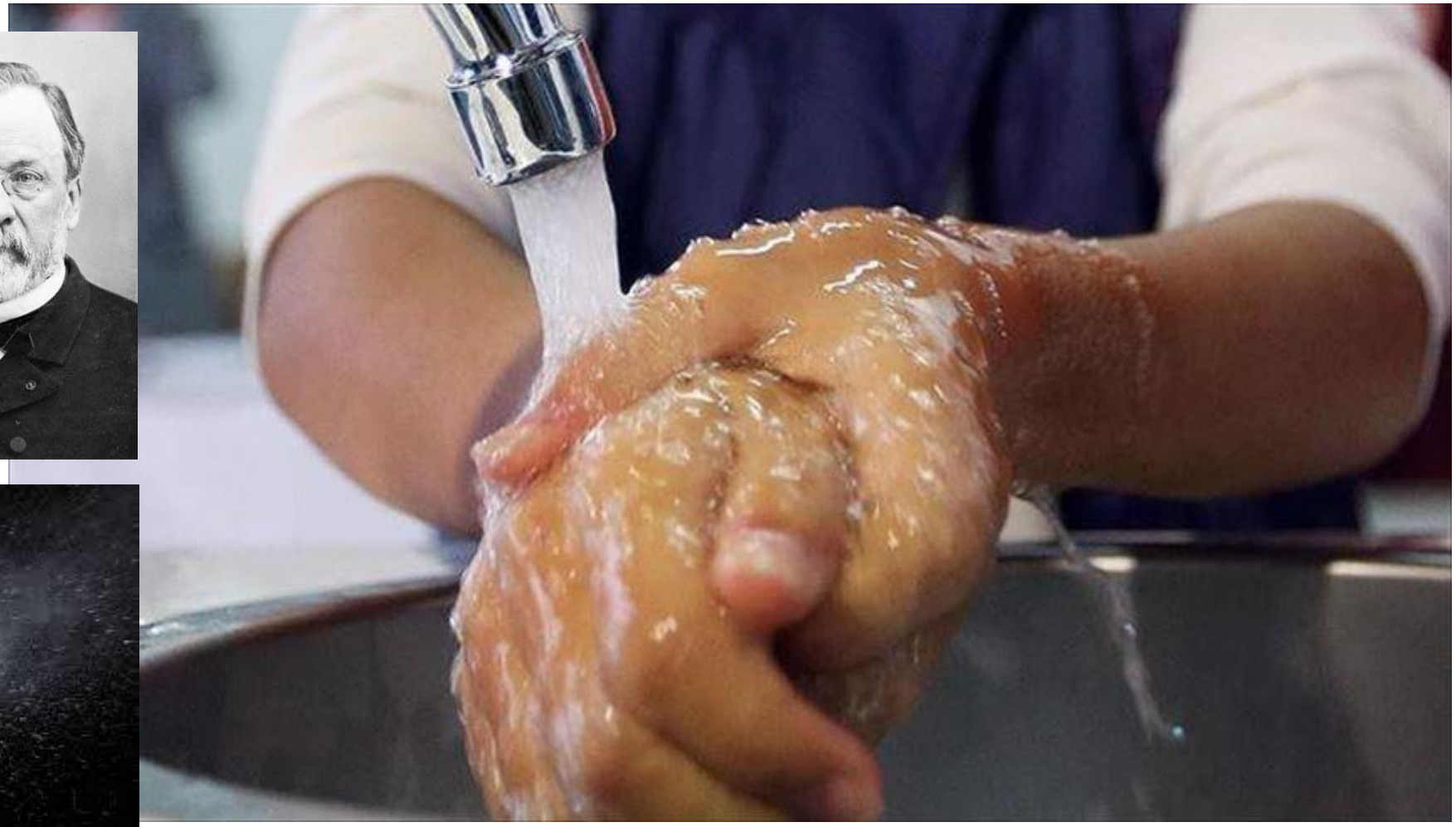


Gabriel Goh, ***Epidemic Calculator,*** https://gabgoh.github.io/COVID/

SEIR models work as long as germ theory holds

Do you trust the work of Avicenna, Louis Pasteur, Koch, Semmelweiss, and bioscience since the 1890s? If so, then you trust underlying theory that makes these models valid.





Key basic parameters in SEIR models

R_0

- **basic reproduction number**, the average number of people with secondary infections resulting from contact with an infectious individual introduced to a totally susceptible population

β

$= R_0 / ND$

- **social mixing parameter**, the rate at which two specific individuals come into effective contact per unit time
- calculated when we know R0 as above, alongside two other parameters: the population size *N* and the duration of infectiousness *D* (usually directly estimated from the 'serial interval' or observed time between linked infections)

λ(t)

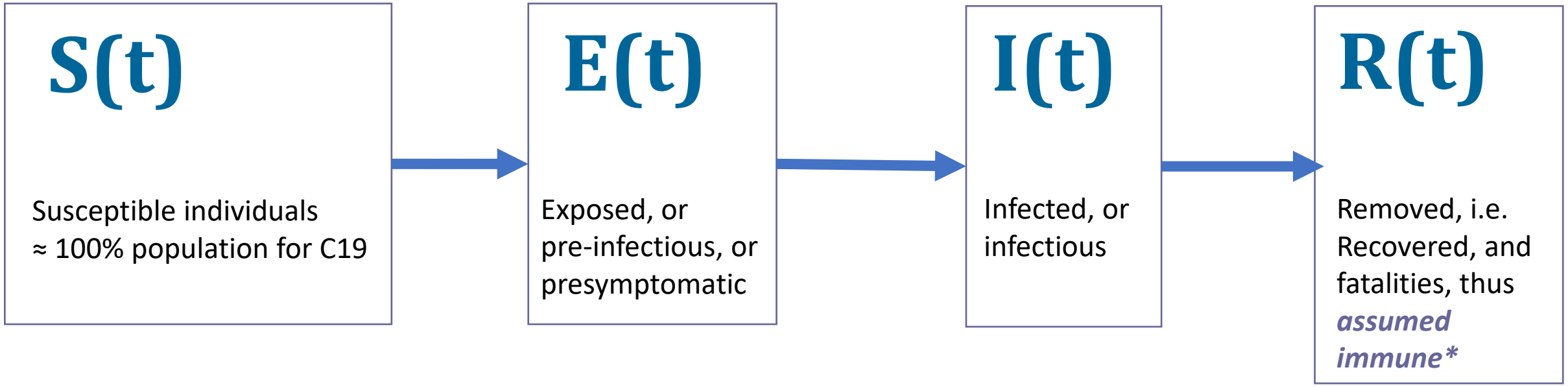
$= \beta * I(t)$

- **force of infection** at time *t*, the rate of which susceptible individuals are infected per unit time
- can be set equal to the social mixing parameter multiplied by the number of infectious individuals at time *t*, assuming homogenous random mixing in a closed population

r

- rate at which individuals recover or die after being infected

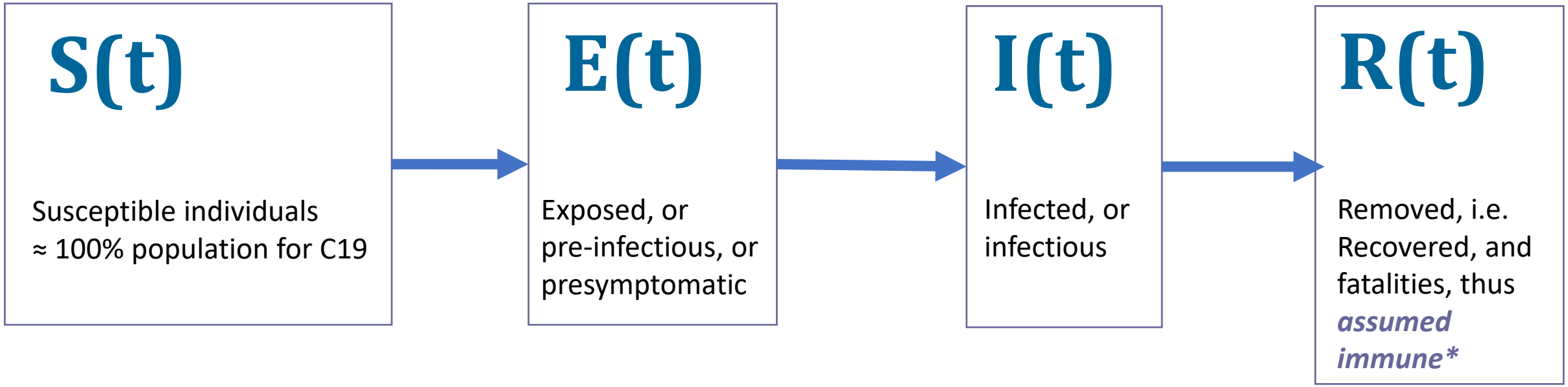
An S(E)IR model is a “deterministic compartment” model, i.e. stocks and flows



n.b. we typically **assume the time period of infectiousness corresponds to clinical symptoms**, but the true picture is more complicated than this for COVID-19!

a major assumption here *is that fatalities and recovered patients do not further infect others. This makes the model unidirectional, and thus statistically fittable to data with Monte Carlo Markov chain methods. SEIRS extension discards this assumption and requires more sophisticated mathematical tools to create reliable estimates while dealing with the ‘recursion’ of flow back to S(t).*

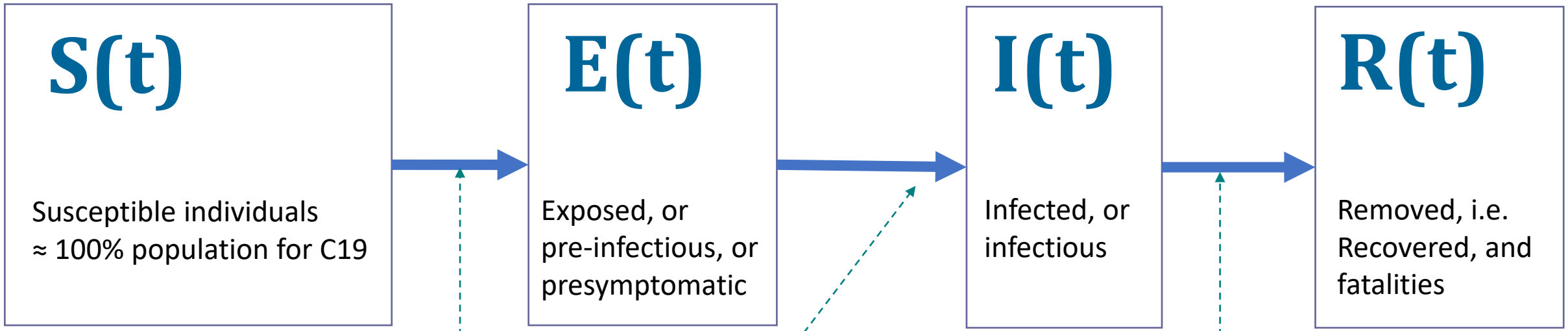
An S(E)IR model is a “deterministic compartment” model, i.e. stocks and flows



$$S + E + I + R = N$$

total population is captured in these stocks and flows and is **effectively closed** (i.e. we can ignore births, deaths, and migration)

An S(E)IR model is a “deterministic compartment” model, i.e. stocks and flows



λ(t)

β(t) * I(t)

force of infection over time from social contact

f

(as a clinical component of **λ** that can be estimated separately from social mixing, fyi only)

r

removal from
- recovery rate over time
- fatality rate over time

An S(E)IR model is a "deterministic compartment" model, and **public health interventions** modify the flows

S(t)

Susceptible individuals
≈ 100% population for C19

E(t)

Exposed, or
pre-infectious, or
presymptomatic

I(t)

Infected, or
infectious

R(t)

Removed, i.e.
Recovered, and
fatalities

Effective **vaccines** directly reduce the pool of susceptible individuals

Physical **distancing**, and **hygiene** measures such as handwashing or masks reduce exposure and infection

Antivirals and other medical treatments reduce the fatality rate and improve the recovery rate

An analogy for SIR dynamics

Susceptible

- λ(t) * S(t)
- β * I(t) * S(t)

+ β * I(t) * S(t)

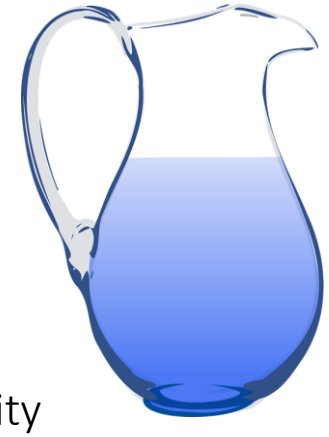
Every flow *out of* a previous compartment goes *into* the next one

Infected

- r * I(t)

The epidemic **grows** whenever flow to infections exceeds flow to removals

$$\delta I/\delta t > \delta R/\delta t$$

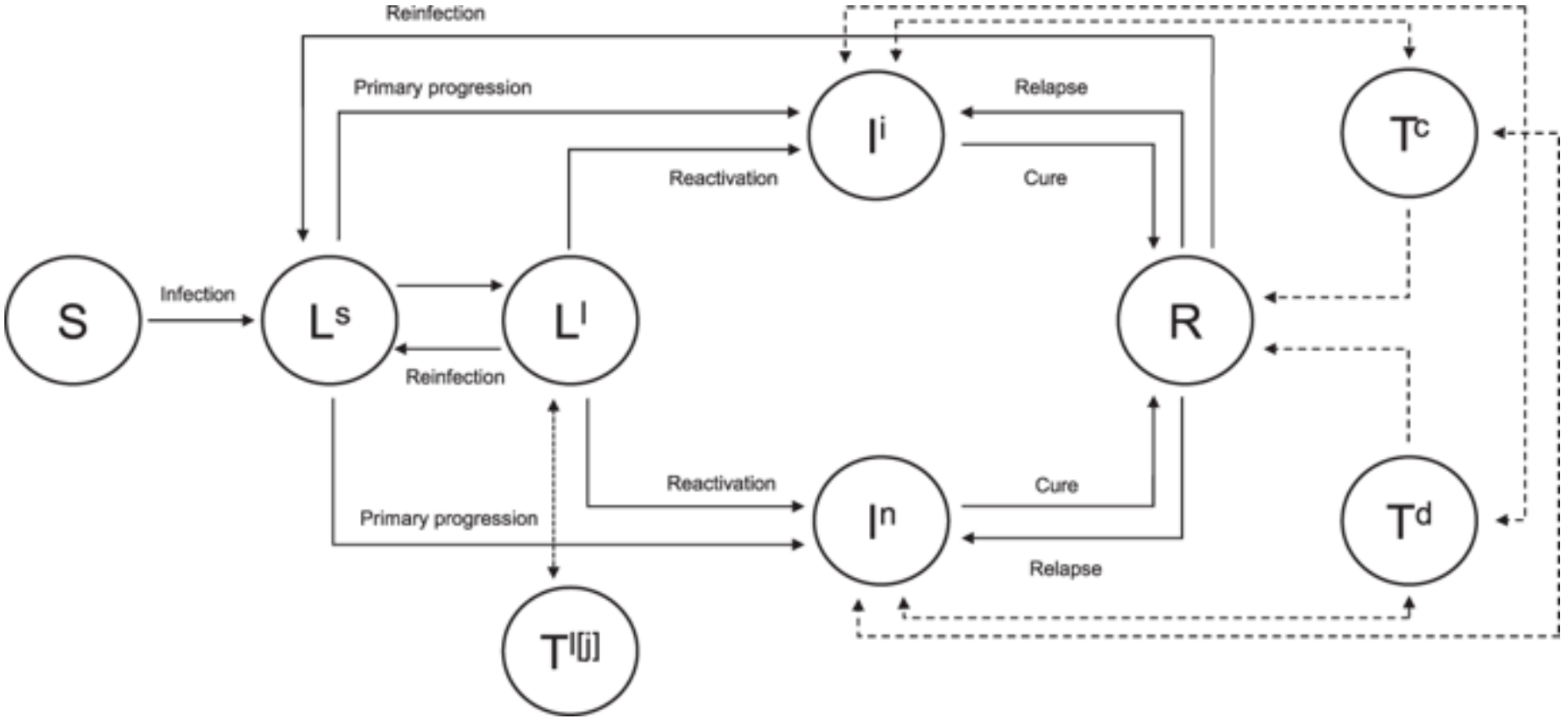


Removed

+ r * I(t)

$R_0 > 1$ whenever β * I(t) is greater than r * I(t),
i.e. β exceeds r,
i.e. rate of new infections exceeds rate of removal to immunity

SEIR models can be extended *structurally* in compartments and flows to better reflect the true epidemiology of infection



Chong KC et al, "Mathematical modelling of the impact of treating latent tuberculosis infection in the elderly in a city with intermediate disease burden", *Nature Scientific Reports*, 19 March 2019, https://www.nature.com/articles/s41598-019-41256-4

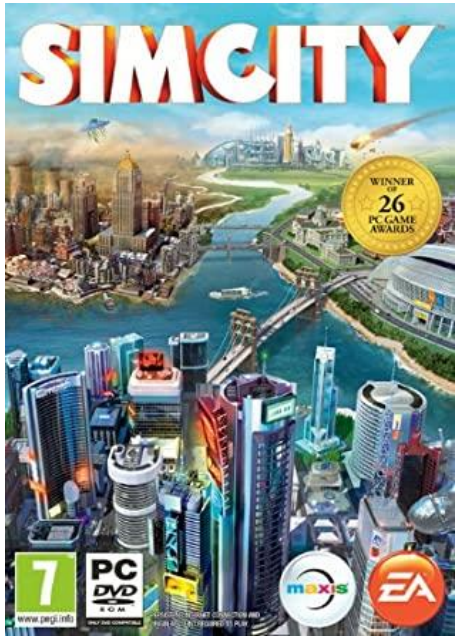
Schematic flow of the age-stratified compartmental model for TB transmissions. TB, tuberculosis; S, susceptible; L^s, latently infected (recent); L^l, latently infected (remote); I^i, infectious; I^n, non-infectious; R, recovered; T^c, treatment completion; T^d, treatment defaulted; $T^{l[j]}$, LTBI treatment for duration *j*.

Basically, epidemiology = biology + social systems (incl. medicine)

Deterministic compartment models are often critiqued…

> ***All models are wrong, some are useful.***
>
> George Box

Strategy games like Civ6 are fun models with exciting graphics…



Code Review of Ferguson's Model

6 May 2020. Updated 10 May 2020.

by Sue Denim (not the author's real name)

[Please note: a follow-up analysis is now available here.]

Imperial finally released a derivative of Ferguson's code. I figured I'd do a review of it and send you some of the things I noticed. I don't know your background so apologies if some of this is pitched at the wrong level.

613

My background. I have been writing software for 30 years. I worked at Google between 2006 and 2[illegible]re I was a senior software engineer working on Maps, Gmail and account security. I spent

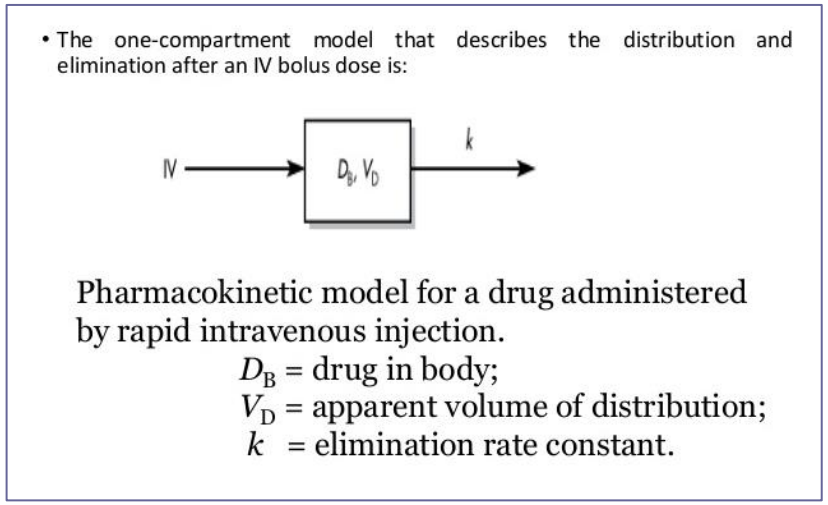
https://lockdownsceptics.org/code-review-of-fergusons-model/

SimCity without the graphics

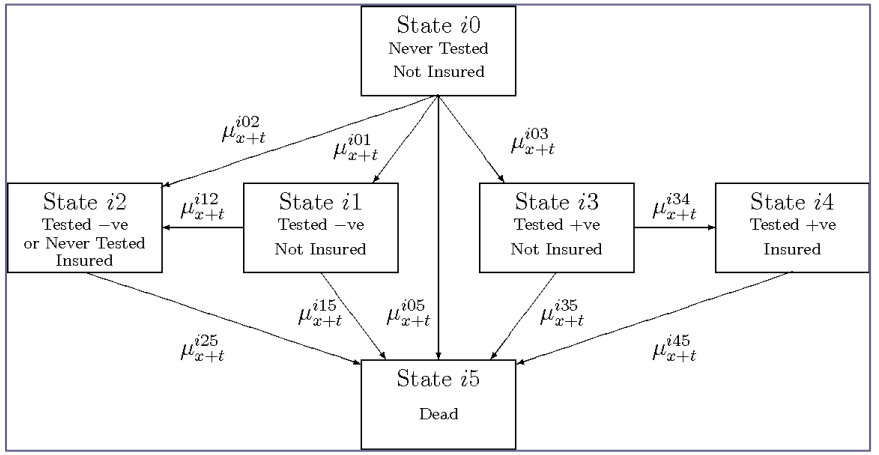
The above pseudonymously authored "analysis" critiquing Neil Ferguson et al's code from the Imperial model has gotten lots of press. In my view, **this is also unfair and misleading.**

See Phil Bull's counterargument.

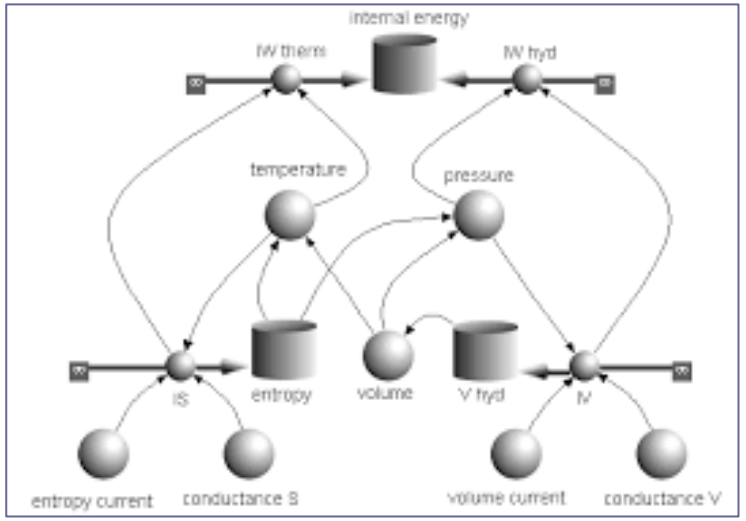
… but compartment models and ordinary differential equations have abundant scientific validity and applications



Pharmacokinetic models used to evaluate dosage safety and efficacy



Health economic & actuarial models



See also most 'business case' or financial models:

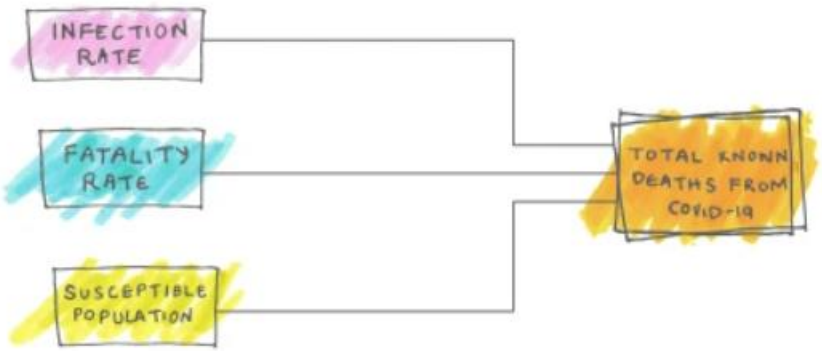
- *Pro forma income statements*
- *Investment cases*

… and other scientific examples:

- *Macroeconomic models (e.g. Computable General Equilibrium)*
- ***Physics and engineering models*** (example from Berkeley Madonna software at left for electrical system dynamics; Madonna is sometimes used similarly for SEIR model fitting)

The proof is in the pudding: good estimation of data parameters and assumptions is required for truly accurate SEIR models

So, imagine a simple mathematical model to predict coronavirus outcomes. It's relatively easy to put together — the sort of thing people on our staff do while buzzed on a socially isolated conference call after work. The number of people who will die is a function of how many people could become infected, how the virus spreads and how many people the virus is capable of killing.



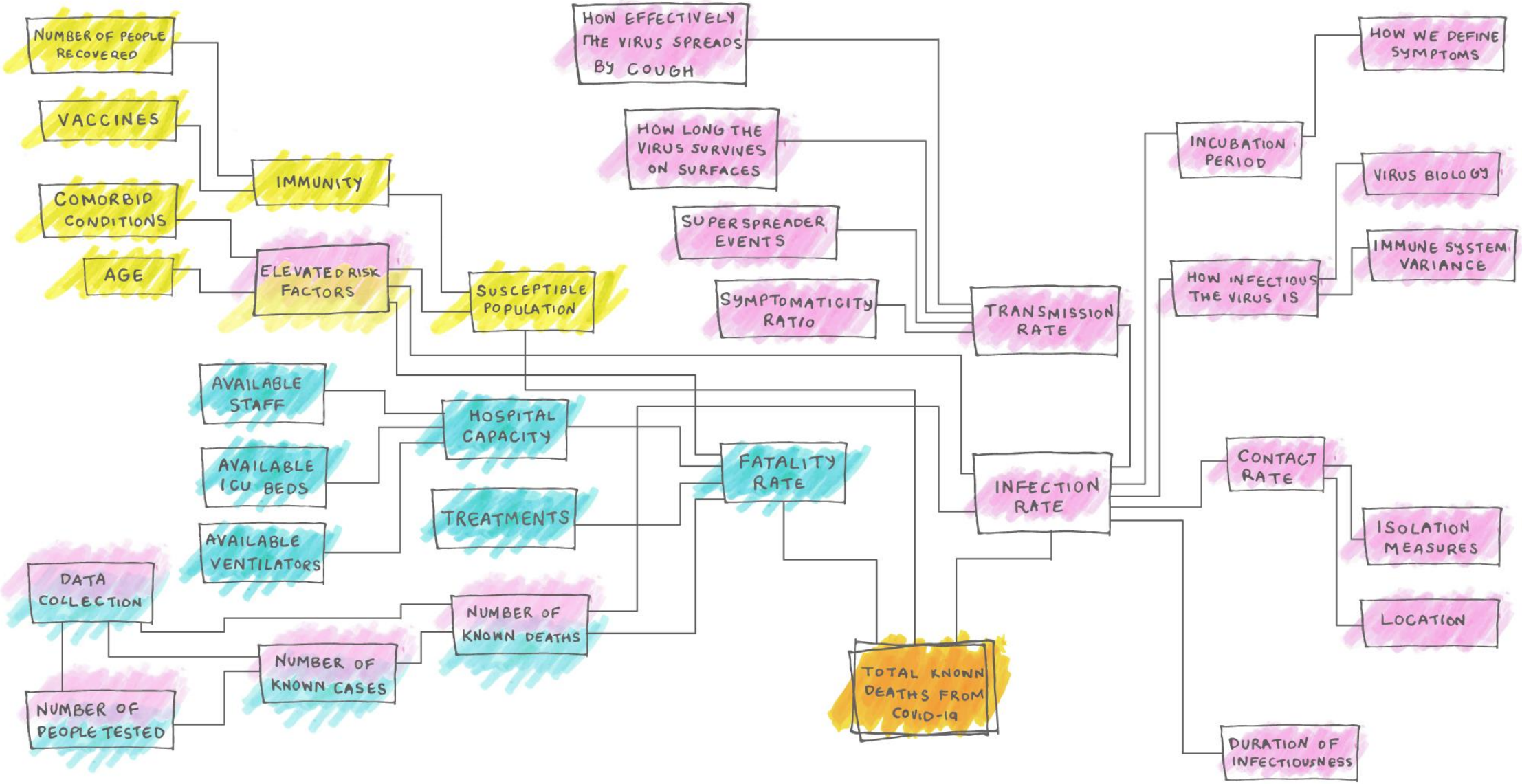
MIX IT ALL UP IN A MODEL

To make a model, then, you have to assemble all those variables (and others our editor wouldn't let us mention), account for their uncertainty, how correlated they are to each other and all sorts of other stuff. It can get messy.

And all of these factors can be affected by all the `interventions` we've tried to reduce the virus' spread — social distancing, hand-washing, school closing, reducing elective surgeries, and the like. This is the big unknown that can drastically change the shape of the outbreak — and yet it also differs by country, state and even city.

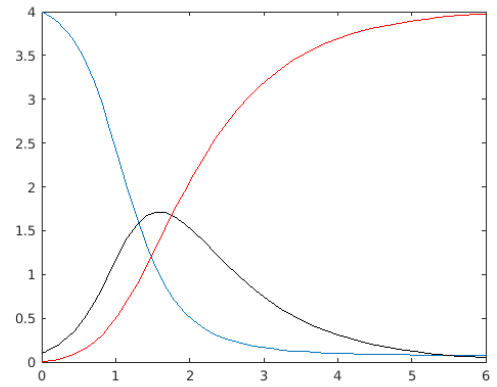
Maggie Koerth, Laura Bronner, and Jasmine Mithani. "Why it's so freaking hard to make a good COVID-19 model", *FiveThirtyEight*, https://fivethirtyeight.com/features/why-its-so-freaking-hard-to-make-a-good-covid-19-model/

Parameters become complicated to measure very quickly



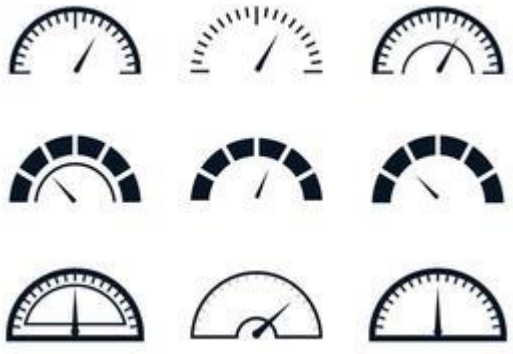
Maggie Koerth, Laura Bronner, and Jasmine Mithani. "Why it's so freaking hard to make a good COVID-19 model", *FiveThirtyEight*, https://fivethirtyeight.com/features/why-its-so-freaking-hard-to-make-a-good-covid-19-model/

The importance of *fitting the data* based on real information



Every SEIR model basically gives the same output.

What we care about is the *numerical estimates.*



Does the model show predictive power?

(i.e. *parameterisation* and *structure*)

Does random homogeneous mixing over time actually hold? (i.e. *ergodicity*, averages over population and time are equivalent)

Critical appraisal of SEIR-type models

- Check the **structural** assumptions.

  e.g. *If immunity is lost over time (endemic disease), then the appropriate form is SEIR**S**.*

- Check the **parameters** including data sources and validation, and how the modelers depict **uncertainty** around the estimates and predictions.

  e.g. *Social mixing patterns – and physical distancing responses – differ a lot by setting!*

- Consider implications of **heterogeneity** in the real world.

  Key differences in susceptibility by age, occupation, deprivation may need depiction.

- Consider **generalisability** to your decision problem in light of the above.

The principle of SEIR models is usually sound, but the devil is in the details of implementation and *which decisions* you are trying to inform with *which data.*

SEIR modelling strengths and weaknesses

Strengths

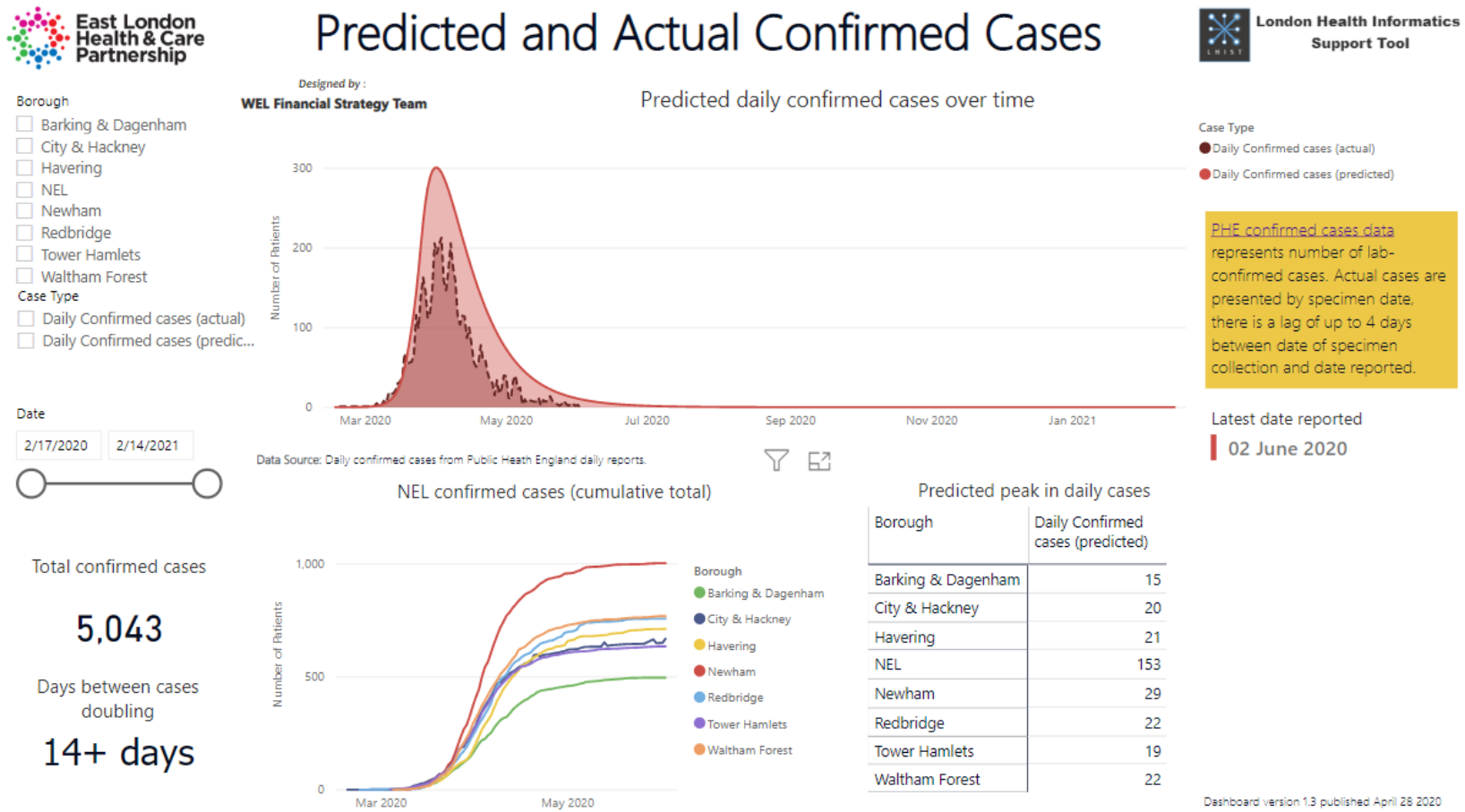
- **Parsimony** (simplicity or 'elegance'): fewer parameters = less data needed for useful estimates
- **Flexibility**: relevance to predicting nearly all infectious diseases, for population-level risk snapshots and scenarios
- **Sufficient transparency**: tractable computation, troubleshooting, and interpretation in most settings

Weaknesses

- **Strong assumptions**: structural choices make big differences and more accuracy requires more modelling complications (e.g. to estimate hospitalisation flow over time)
- Greater **reliance on a few key parameters:** can lead to large uncertainty of final estimates where these are not yet well-established (as is true in COVID-19)
- **Limited utility for operational or micro-scale questions:** prefer ***agent-based microsimulation***, *geospatial modelling, network modelling, queuing theory*)

Example of SEIR modelling in action: NEL dashboard

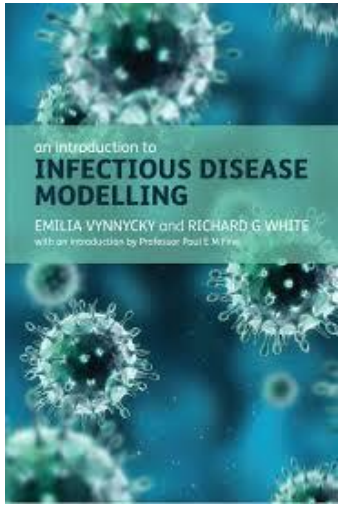
Main test: do the model predictions fit the real-world data?
If not, need to tune and fix the model. **NEL fit below is great!**



East London Health & Care Partnership, **NEL COVID-19 Demand Modelling Dashboard,** *available at* https://app.powerbi.com/view?r=eyJrIjoiNmI2MjM4OTAtYTBmYS00MGNhLTgzOGEtYjJhNTg0NGY0ZWU4IiwidCI6ImQyMjc2ODJmLWFiNWEtNDlmNi04NzNhLThlZmQ1MDQ1ZjBmNCJ9

Further resources: SEIR theory and applications

Emilia Vynnycky and Richard G White, ***An Introduction to Infectious Disease Modelling***. London: Oxford U P, 2011. Available online at http://anintroductiontoinfectiousdiseasemodelling.com/



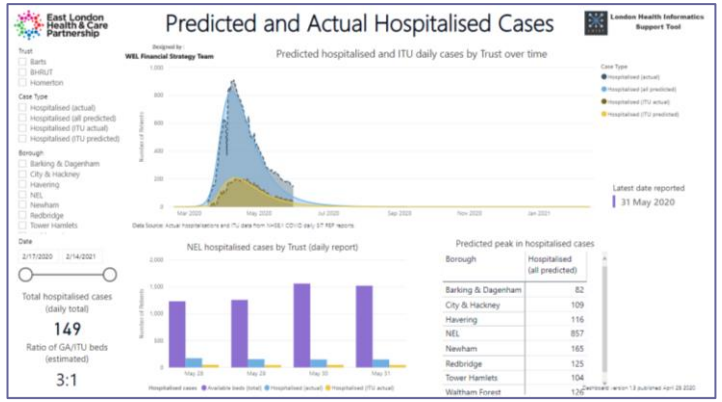
Davies N et al, LSHTM CMMID nCoV working group COVID-19 Transmission App, available at https://cmmid.github.io/visualisations/covid-transmission-model

More from the LSHTM COVID-19 repository (Apps and research notes) at https://cmmid.github.io/topics/covid19/

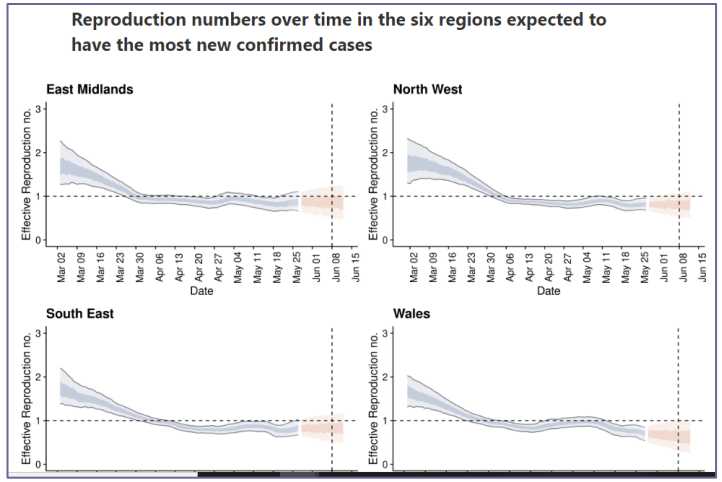
These slides + simple Excel template: https://github.com/7j7j/SEIR-model-basics

Nice examples of SEIR and R_0 subnational modelling in England

1. East London Health & Care Partnership – WEL Financial Strategy Team, "NEL COVID-19 Demand Modelling Dashboard", available at https://app.powerbi.com/view?r=eyJrIjoiNmI2MjM4OTAtYTBmYS00MGNhLTgzOGEtYjJhNTg0NGY0ZWU4IiwidCI6ImQyMjc2ODJmLWFiNWEtNDlmNi04NzNhLThlZmQ1MDQ1ZjBmNCJ9.



2. Funk S et al, LSHTM CMMID nCoV Working Group, "Temporal variation in transmission during the COVID-19 outbreak – National and subnational estimates for the United Kingdom", https://epiforecasts.io/covid/posts/national/united-kingdom/



References and further reading

[25min video, nb notation of parameters is slightly different: *r* replaces *β* in these slides and *a* replaces *r*]
Crawford T, "Oxford mathematician explains SIR disease model for COVID-19 (coronavirus)", https://www.youtube.com/watch?v=NKMHhm2Zbkw

Adam D. "The simulations driving the world's response to COVID-19", *Nature* magazine, 2 April 2020, https://www.nature.com/articles/d41586-020-01003-6.

Bull P, "Why you can ignore reviews of scientific code by commercial software developers",10 May 2020, https://philbull.wordpress.com/2020/05/10/why-you-can-ignore-reviews-of-scientific-code-by-commercial-software-developers/

Goh G, "Epidemic Calculator", https://gabgoh.github.io/COVID/

Kelly J, "Is the 'science' behind the lockdown any good?" *FT Alphaville*, 28 May 2020, https://ftalphaville.ft.com/2020/05/21/1590091709000/It-s-all-very-well--following-the-science---but-is-the-science-any-good--/

Kermack WO and AG McKendrick, "A contribution to the mathematical theory of epidemics", Aug 1927, https://royalsocietypublishing.org/doi/10.1098/rspa.1927.0118

Koerth M et al, "Why it's so freaking hard to make a good COVID-19 model", *FiveThirtyEight*, https://fivethirtyeight.com/features/why-its-so-freaking-hard-to-make-a-good-covid-19-model/.

Kissler S et al, "Projecting the transmission dynamics of SARS-CoV-2 through the postpandemic period," *Science*, 9 Apr 2020, https://doi.org/10.1126/science.abb5793 and more accessible companion piece Kupferschmidt K and Cohen J, "The beast is moving very fast'. Will the new coronavirus be contained – or go pandemic?", *Science* magazine, 5 Feb 2020, https://www.sciencemag.org/news/2020/02/beast-moving-very-fast-will-new-coronavirus-be-contained-or-go-pandemic.