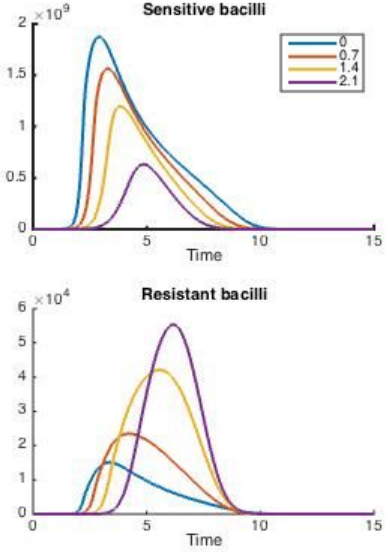
Modelling – A theory, not just a theory, not even a theory

Caroline Colijn
Simon Fraser University

Modelling has its enthusiasts… and sometimes detractors



A scientist using models to explore the consequences of assumptions

An expert whose main area of interest is not emphasized in the model

Meanings of the word “model” vary

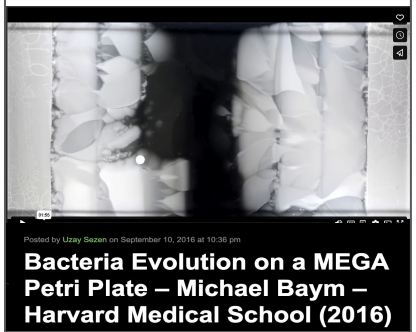
- A *physical* object -- a material model (e.g. a model of a skeleton, a model organism)
- A *fictional* object -- like Sherlock Holmes (e.g. the Bohr model of the atom, a frictionless pendulum, a Wright-Fisher population)
- (philosophy of science: set-theoretic structures)

- A “**stylized description of a target system**” (Achinstein 1968; Black 1962); an **analogy** (usually to a simpler, formalized system).

  This is probably the concept we all share as “a model”.

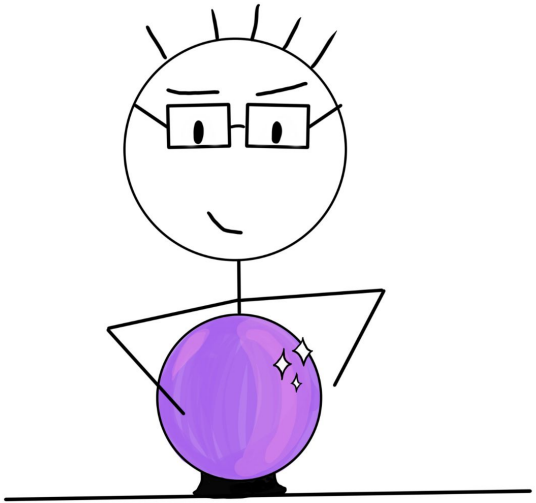
- Here, I focus on mechanistic, mathematical models

Some “models”



What do models offer?

Two extreme (and surprisingly frequently encountered) views:



Look, it has time on the x axis.
Models can predict the future!

Models are nothing but a fun little game you play in your computer. They aren't useful for anything.

Is a model “a theory”?

A model is not exactly the same thing as a theory.

- The same *theory* could admit many different models.
- Often one model cannot embody *all* of “a theory”.

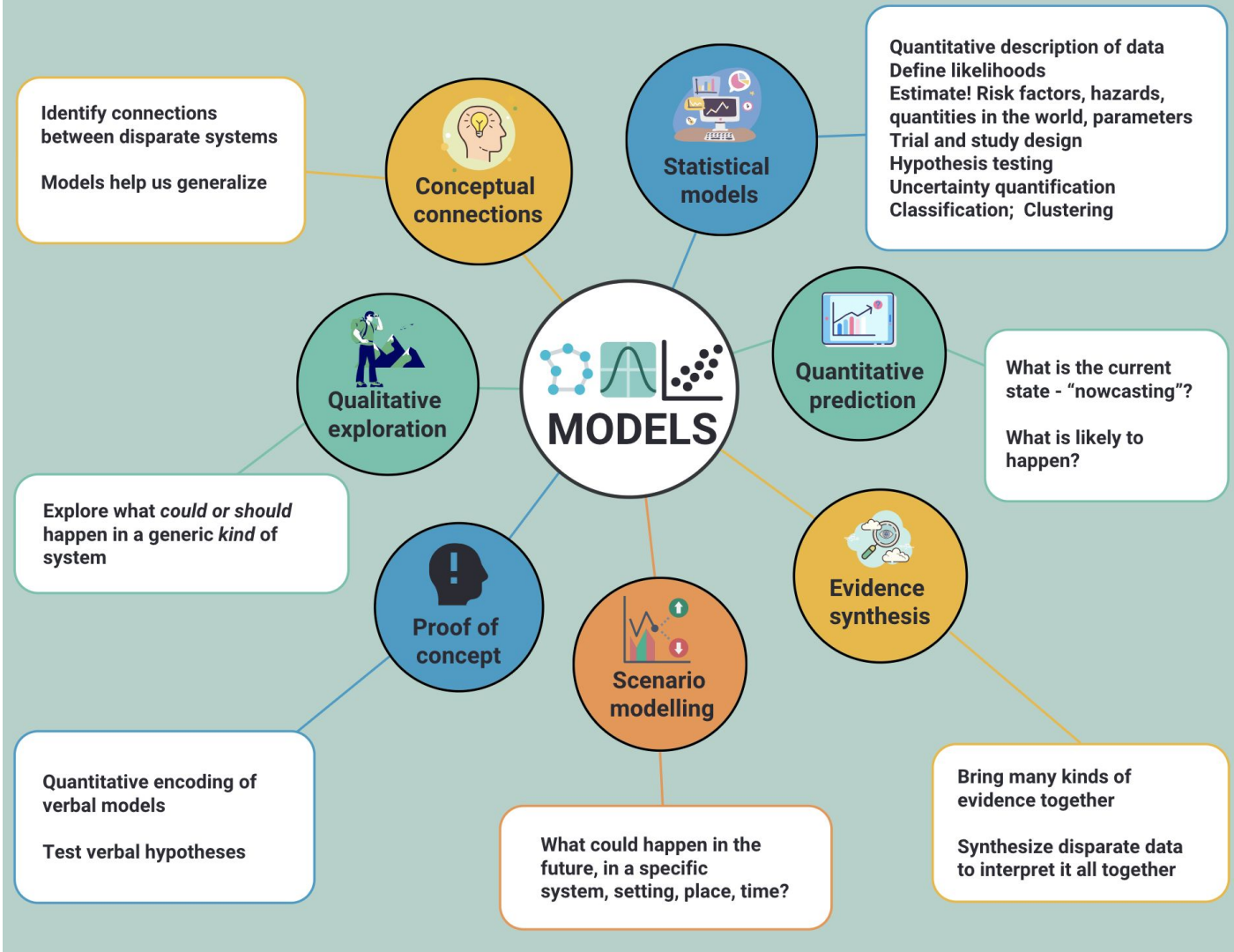
But models can be essential elements of a theory, and a suite of models may comprise a theory.

Why do we want a theory? (very brief!)

- A theory helps us understand a system, explain phenomena, reason about a system, extrapolate to new circumstances, make predictions for the future, categorize phenomena, generalize to new systems, identify systems that share an essential feature, design new systems… and many more.

The many uses of modelling

Build new theory



Questions for the audience

1. Could there be a (good, useful) model that cannot make a prediction for *any* specific system?

2. Could there be a (good, useful) model that does not need *any* data? (not necessarily in a formal model fitting sense)

3. Does a useful model have to have *relevance* for *some* specific system? (and if so, what the notion of relevance?)

“A theory”: some well-known uses of models

Statistical modelling; Quantitative prediction; Scenario modelling

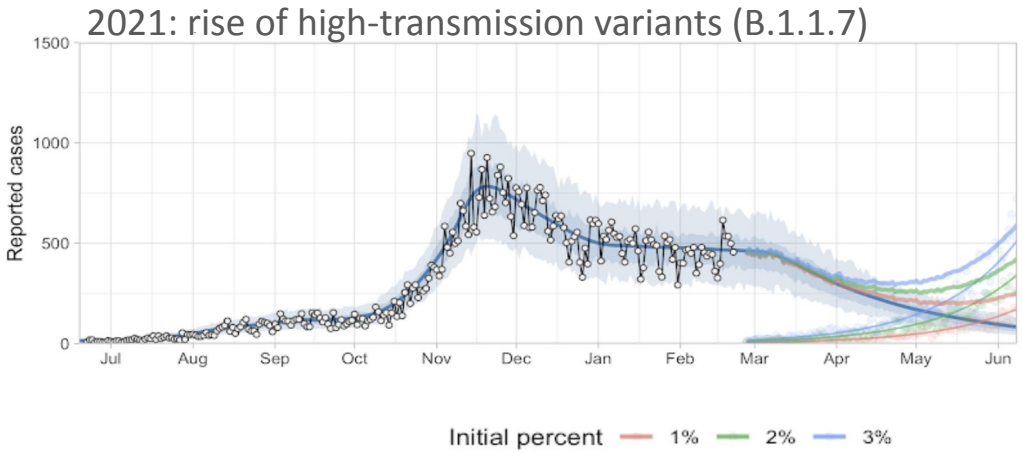
Classic view of (mechanistic) modelling:

- Develop a model that captures the key aspects of a system
- Parameterize the model with data
- Test and validate the model with more data
  - does it capture what we know already?
  - does it predict the present given only data about the past up to some time ago?
- Use the model
  - forecast given the current state
  - explore scenarios about the impact of policies or circumstances
  - quantify uncertainty

Not *just* a theory: lesser-known uses of models

- Evidence synthesis
  - Models bring disparate data together, producing a unified system that can be interrogated in a way that is consistent with many different pieces of information
- Proof of Concept:
  - models test verbal hypotheses. Models can help to fully formulate and specify a theory, so that the theory can be interrogated. (Servedio et al, PLOS Biol, below)
- Qualitative Exploration
  - Models identify consequences of a (generic) phenomenon: what might we expect as a “baseline”.
- Conceptual Connection
  - Formal similarities reveal connections between very disparate systems

Modelling can be a form of evidence synthesis



Population data: testing; contact patterns. This includes international data

Virology: viral load, immunity; duration of infectiousness, vaccine effectiveness

Epidemiological data: who is getting infected, how quickly? Vaccination data.

Contact tracing data: contact pairs and transmission links

Sequencing surveillance data: rise of VOC strains with high transmission. Includes international data.

FAMOUS COVID-19 OUTBREAKS

Morbidity and Mortality Weekly Report (*MMWR*)

CDC

High SARS-CoV-2 Attack Rate Following Exposure at a Choir Practice — Skagit County, Washington, March 2020

Following a 2.5-hour choir practice attended by 61 persons, including a symptomatic index patient, 32 confirmed and 20 probable secondary COVID-19 cases occurred (attack rate = 53.3% to 86.7%); three patients were hospitalized, and two died. Transmission was likely facilitated by close proximity (within 6 feet) during practice and augmented by the act of singing.

Days After a Wedding, a Dead Groom and Dozens of Coronavirus Cases

Officials in India have opened an investigation into the gathering after more than 100 wedding guests tested positive for the coronavirus.

News

Almost an entire class of students caught coronavirus at a Trois-Rivières school

Outbreak amid preventive measures illustrates the difficulty of controlling the spread of COVID-19 in a classroom of young children.

Matthew Lapierre • Local Journalism Initiative Reporter

News | Coronavirus pandemic

After one infected 16 at Berlin nightclub, coronavirus fears grow

Panic spreads across Germany, with officials calling on people to avoid concerts, nightclubs and football games.

A SIMPLE MODEL FOR EVENT TRANSMISSION

The following very simple model *synthesizes* data about event durations, contact patterns during events and transmission probability (if we had it).
Consider an event that lasts a total time T.

- If an infectious individual attends and is in contact with a single susceptible individual for a time τ with a constant per unit time probability of transmission β, then the probability that the susceptible individual becomes infected is $(1-e^{-\beta\tau})$.
- If they contact *k* others, the expected number of new infections is $k\,(1-e^{-\beta\tau})$.
- If they mix around and contact T/τ groups of *k* people, the expected number is

$$R_{\text{event}} = \frac{kT}{\tau}\left(1 - e^{-\beta\tau}\right)$$

Reducing transmission

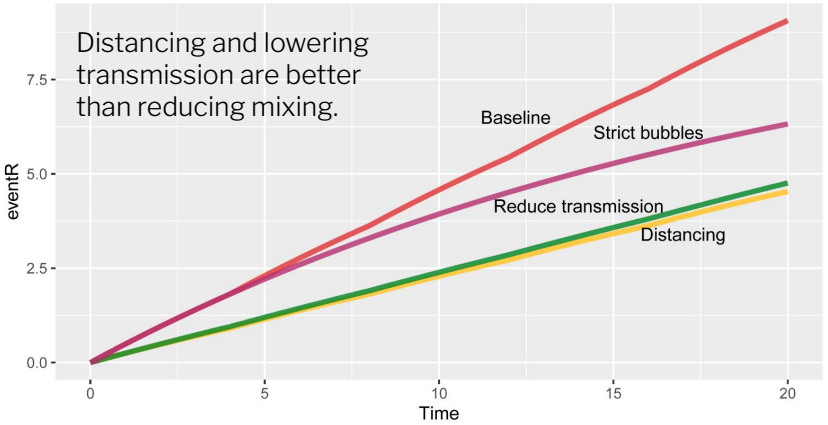
We can reduce R_{event} by changing:

- the crowding (k): reduce capacity or density
- the level of "mixing" T/τ : use "social bubbles"; people circulate less. In the model this means increasing τ . τ occurs twice in the equation.
- the duration of the event
- the per-person, per-unit-time, transmission rate: e.g. with masking

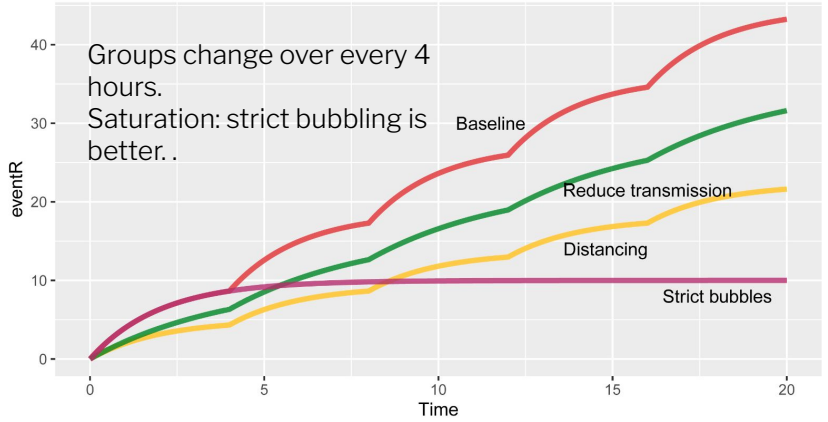
$$R_{\text{event}} = \frac{kT}{\tau}\left(1 - e^{-\beta\tau}\right)$$

If we know beta we can decide what's more effective

Low β: linear regime.



High β: saturating regime



Living evidence synthesis

(not model-based in general)

People do systematic literature searches and summaries.

Evidence is assessed for quality.

Results are tabulated and summarized.

But what does it mean in context?

Vaccine* effectiveness** against Omicron based on number of doses, time since last dose and age

Outcome (and vaccine)	Number of doses	Age	Time since last dose (days)	Vaccine Effectiveness
Any infection				
Pfizer	1	5 to 11	60	4%
		12 to 17	21 to 48	16 to 34%
			28 to 56	58%
			49 to 76	-1 to 17%
			77	-13 to -5%
			56 to 84	64%
	2	0 to 4	56 to 83	63%
			84 to 111	64%
			112 to 139	64%
			140	64%
		5 to 11	14 to 82	31%
			29 to 84	21 to 29%
			60	26%
			70	23%
		12 to 15	14 to 149	59%
		12 to 17	28 to 69	35 to 63%
			56 to 83	48 to 58%
			84 to 111	41 to 51%
			112 to 139	38 to 46%
			70	8%
		16 to 17	63	23%
	3	5 to 11	14	70%
		12 to 17	14	56 to 72%
			7 to 13	80%
			35 to 69	30%
Moderna	2	0 to 4	56 to 83	64%
			84 to 111	60%
			112 to 139	54%
			140	48%
		12 to 17	35 to 69	29%
			70	20%
Symptomatic Infection				
Pfizer	1	12 to 17	28 to 69	23 to 49%
			70 to 83	16 to 27%
			84	17 to 26%
			14 to 98	19%
		16 to 17	105	13%

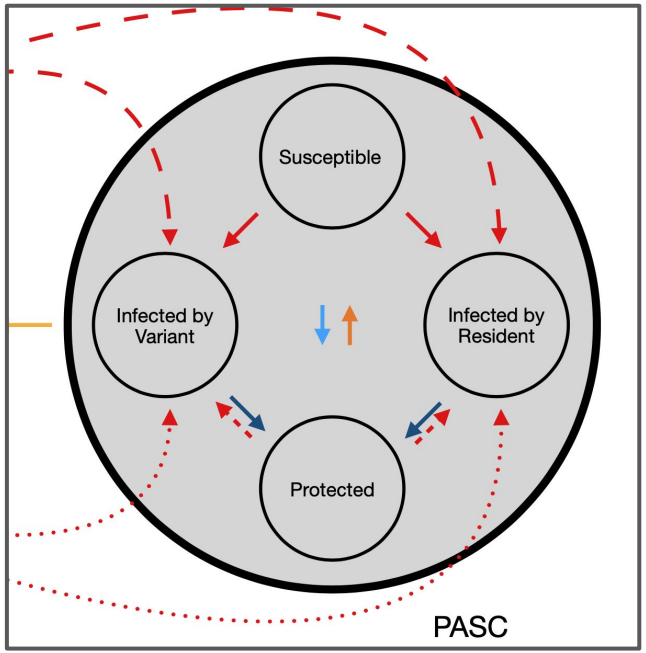
Outcome (and vaccine)	Number of doses	Age	Time since last dose (days)	Vaccine Effectiveness
Symptomatic Infection (continued)				
Pfizer	2	5 to 11	30 to 90	29%
			30 to 59	60%
			60	43%
			90	35%
		12 to 15	30 to 90	17%
			60 to 120	10%
		12 to 17	7 to 69	32 to 77%
			14 to 149	34 to 45%
			56 to 120	10 to 38%
			14 to 98	65%
		16 to 17	70	23%
	3	12 to 17	7	62 to 87%
			0 to 60	56%
	2 doses + mRNA vaccine	12 to 17	14 to 98	63%
Transmission				
No evidence available				
Admission to the intensive care unit				
No evidence available				
Multisystem inflammatory syndrome in children (MIS-C)				
Pfizer	2	12 to 18	28	92%
Severe Disease (may include death for some studies)				
Pfizer	2	5 to 11	90	100%
		12 to 17	7 to 60	76 to 84%
			60 to 120	82 to 86%
			60	74%
			98	83%
Death				
No evidence available				

* This infographic includes evidence about vaccines available in Canada.
** The values represent "range of means" and single values mean the result is derived from a single study.

Modelling and living evidence synthesis

Evidence synthesis: vaccine efficacy against infection, long COVID (PASC), severe disease.

Other data informs model: partial immunity, duration, transmissibility, VOC emergence



Model can:
Help **interpret** living evidence synthesis: what does this collection of evidence *mean* for a question?

Help **direct** living evidence synthesis: what do we most need to know?

Proof of concept models

Not Just a Theory—The Utility of Mathematical Models in Evolutionary Biology

Maria R. Servedio, Yaniv Brandvain, Sumit Dhole, Courtney L. Fitzpatrick, Emma E. Goldberg, Caitlin A. Stern, Jeremy Van Cleve, D. Justin Yeh

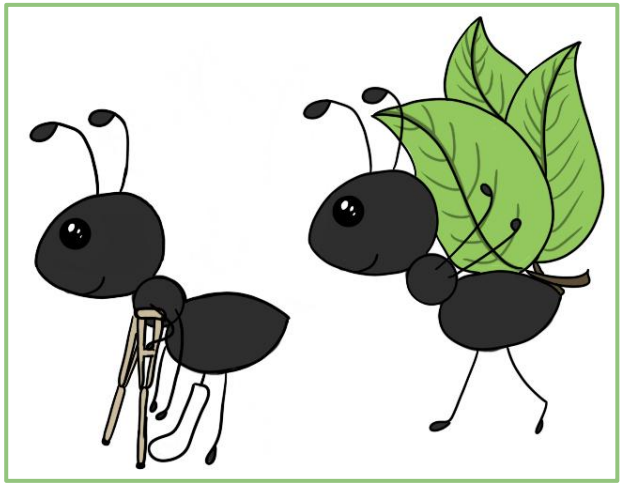
Published: December 9, 2014 • https://doi.org/10.1371/journal.pbio.1002017

https://journals.plos.org/plosbiology/article?id=10.1371/journal.pbio.1002017

- Proof of concept models test verbal hypotheses, or verbal theories.
- Parallel to “model systems” in a laboratory: modelling is a form of experimenting
- Do these models need “validation” or to be “tested against data”? They *are* the test.
- But they do need “validation”: are they well set up to test the verbal hypotheses?

A proof (actually a *test*) of concept modelling story

Hamilton (*JTB* 1964) argued that interactions with kin could favour altruistic behaviour. Limited spatial dispersal could increase the probability of interacting (mostly, more) with kin. Therefore population "viscosity" could favour the evolution of altruism.



Taylor (*Evol. Ecol.* 1992) used a simple mathematical model to try this out. But the costs of competition with kin counteract the benefit of interacting with them (altruistically).

Kin competition, life history, time of dispersal are all relevant.

Hamilton's original verbal model is not sufficient.

Summarized from Servedio et al

Models as qualitative exploration

https://journals.aps.org/prl/abstract/10.1103/PhysRevLett.90.028701

- Infectious disease on a network with a scale-free (power law) degree distribution: no matter how low the transmission rate, the infection is not eliminated.
- Contrast to the SIR++ suite of epidemic models: they have a basic reproduction number R0; usually if R0 < 1 the infection cannot spread.
- Great paper as a **qualitative exploration**. It is not a prediction for any specific system. It is not a proof of concept (it does not test a verbal hypothesis).

“Null” or “neutral” models

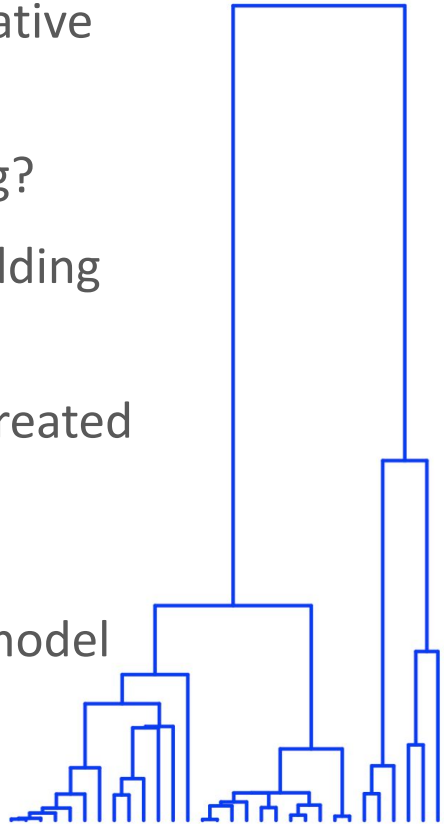
Null models are often both proof-of-concept models and qualitative exploration models.

They ask: what would we see if nothing interesting is happening?

They provide a scaffold for interpreting observations, and for adding complexity -- a bland starting point, to which we can add spice.

Wright-Fisher model: Constant population. Each generation is created by sampling the previous one uniformly with replacement. No selection, no mutation, no migration, random mating -- simple!

Coalescent theory: what do genealogical trees look like in this model (and many that reduce to it)?



Models can find and formalize conceptual connections

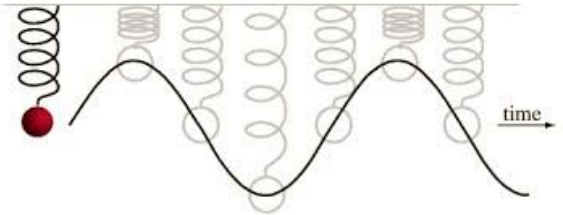
A physics example: **the harmonic oscillator**.

Describes electricity, pendulum, springs, motion in space, oscillatory dynamics in cells.

Knowledge of the mechanisms that can cause oscillations helps to theorize about oscillating population dynamics, evolution and more.

Contrasts with “test-of-concept” modelling where a model can’t, or typically doesn’t, produce a phenomenon.

Here, if a model *does* produce a phenomenon, generically, that might help us understand conceptual connections.



The power law $f(x) \sim ax^{-k}$

Same quantitative pattern, *many* diverse phenomena: contact patterns, city sizes, forest fire sizes, etc.

Can be obtained mechanistically e.g. by *preferential attachment:* more popular cities attract more people; more abundant family names have more descendants

Barabasi & Albert's growth model of a random graph: probability to attach to a node is proportional to the node's degree.

A useful null model: does not require any special property of the high-degree nodes, large cities.

Human sexual contact network: Swedish survey 1996

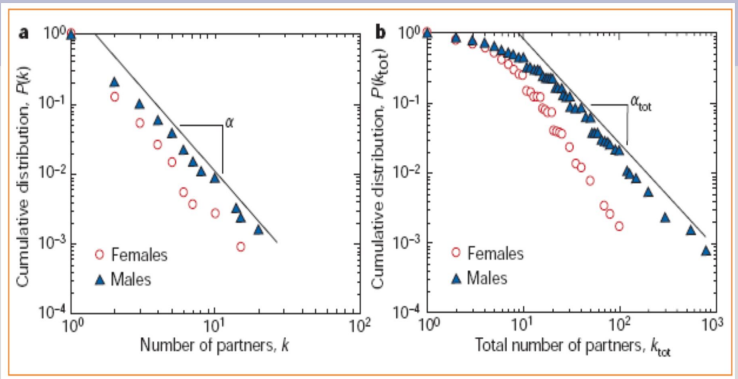
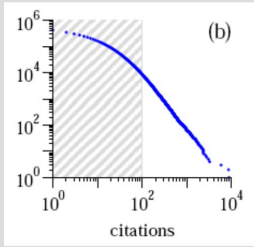
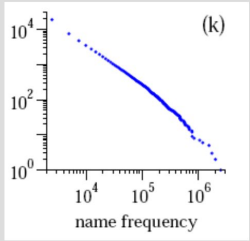


Figure 2 Scale-free distribution of the number of sexual partners for females and males. **a**, Distribution of number of partners, k, in the previous 12 months. Note the larger average number of partners for male respondents: this difference may be due to 'measurement bias' — social expectations may lead males to inflate their reported number of sexual partners. Note that the distributions are both linear,

Lilieros, Nature 2001.

Numbers of citations on scientific papers





Frequency of names

Newman 2005

Not *even* a theory (the controversial part)

There is a large volume of modelling work that does not do any of the functions of modelling.

It lands in a “un-useful valley” -- not new mathematics (ie not of utility to mathematicians; not new mathematical ideas) and not a useful contribution to science

Why?

- It’s fun to play with models. We like models.
- Publication incentives
- Poor understanding of how modelling can be good science (and how to articulate why it’s good science)
- Poor understanding of how modelling can fail to be good science


NEW SCIENCE 500M
NICE NEW MATH 1000M
Unuseful Valley

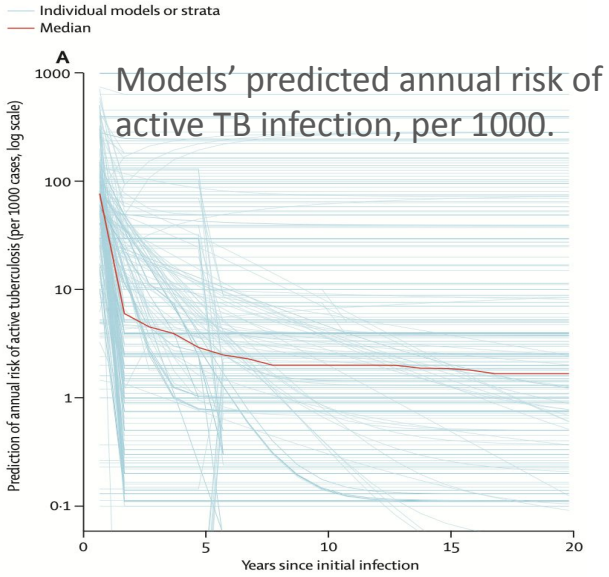
Progression from latent infection to active disease in dynamic tuberculosis transmission models: a systematic review of the validity of modelling assumptions

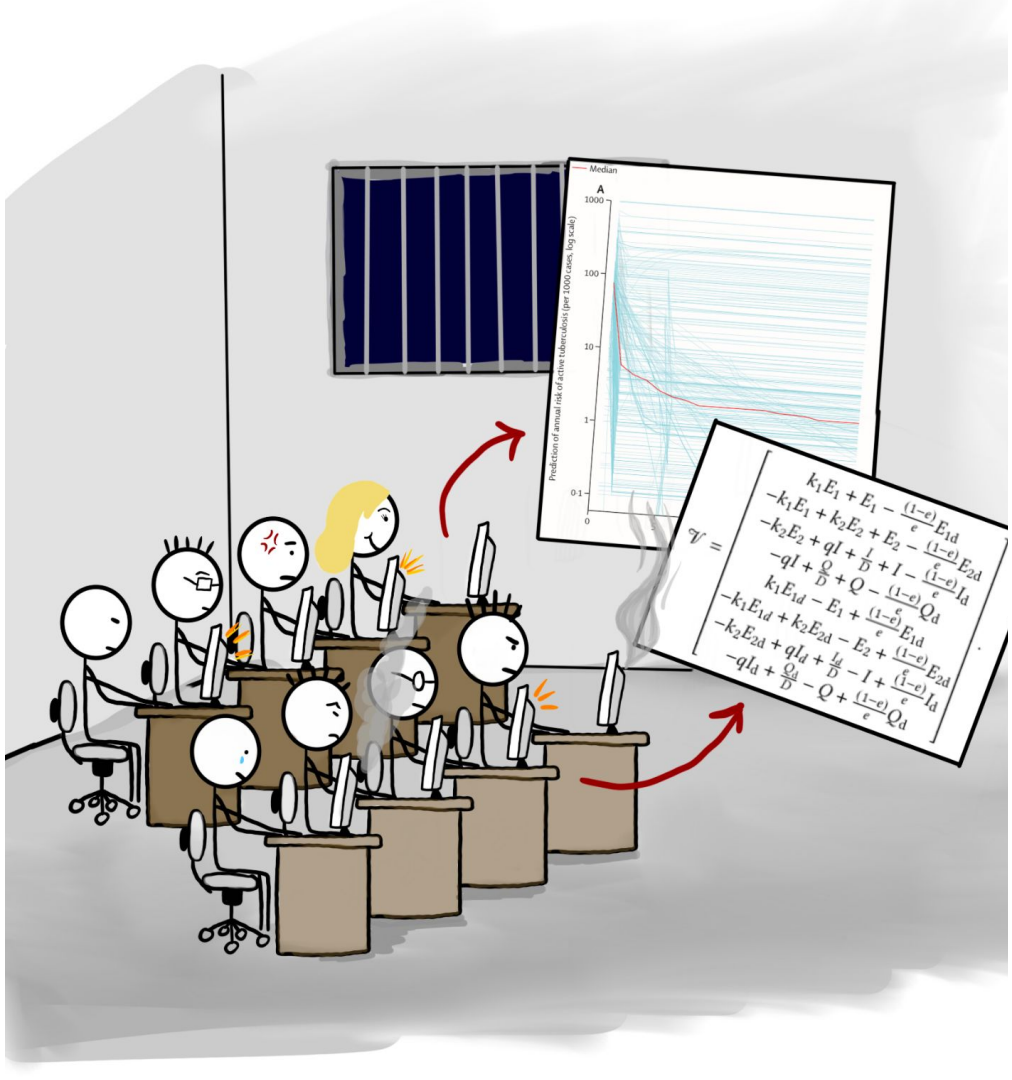
Nicolas A Menzies, Emory Wolf, David Connors, Meghan Bellerose, Alyssa N Sbarra, Ted Cohen, Andrew N Hill, Reza Yaesoubi, Kara Galer, Peter J White, Ibrahim Abubakar, Joshua A Salomon

produced because of the non-linear relationship between parameters and modelled outcomes. Some studies reported adjusting parameter values as part of model calibration, but did not report these adjusted values, and in these cases we used the original (unadjusted) values reported in the paper. In some models, individuals

is 50–100% greater than both empirical estimates. For 10-year cumulative incidence, only 60% of modelling results were within a factor of two of either empirical point estimate, and only 77% were within a factor of five. 10-year cumulative incidence was greater than 50% for 15% of all modelling results, and less than 1% for 4·6% of results.

- The authors reviewed and implemented 312 published TB transmission models.
- Key parameters are not known, so models make assumptions. Results are all over the map.
- Model results are strongly shaped by these assumptions (ie, putting *in* the answers, not getting them *out*!)
- Menzies et al said it's likely that "a substantial proportion of models adopted assumptions that were incorrect"
- It matters for policy questions: downstream impact on estimates of the likely benefits of different policies, or the costs.

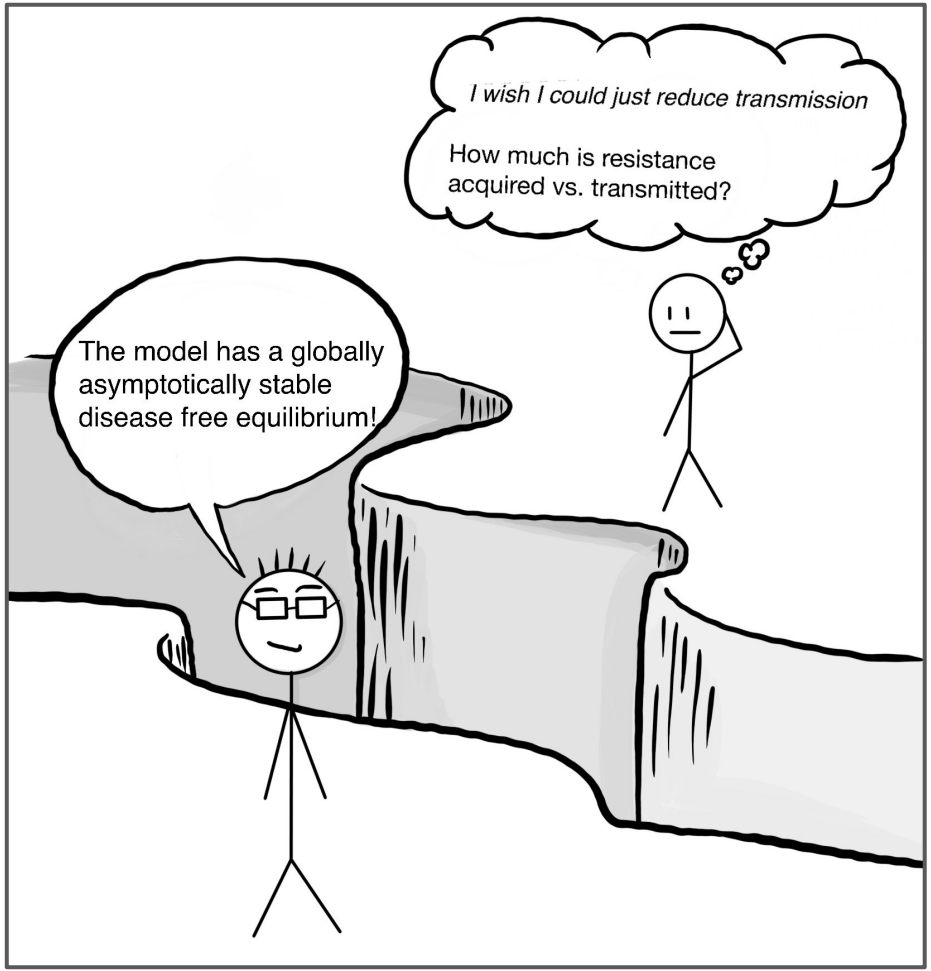



Median
A
1000
100
10
1
0·1
0
Prediction of annual risk of active tuberculosis (per 1000 cases, log scale)
$\mathscr{V} = \begin{bmatrix} k_1E_1 + E_1 - \frac{(1-e)}{e}E_{1d} \\ -k_1E_1 + k_2E_2 + E_2 - \frac{(1-e)}{e}E_{2d} \\ -k_2E_2 + qI + \frac{I}{D} + I - \frac{(1-e)}{e}I_d \\ -qI + \frac{Q}{D} + Q - \frac{(1-e)}{e}Q_d \\ k_1E_{1d} - E_1 + \frac{(1-e)}{e}E_{1d} \\ -k_1E_{1d} + k_2E_{2d} - E_2 + \frac{(1-e)}{e}E_{2d} \\ -k_2E_{2d} + qI_d + \frac{I_d}{D} - I + \frac{(1-e)}{e}I_d \\ -qI_d + \frac{Q_d}{D} - Q + \frac{(1-e)}{e}Q_d \end{bmatrix}$

Not even a theory: Problem 1

Models may answer irrelevant questions

- focus on properties of models (great, if those are of interest to mathematicians, but this is rare)
- poor analogy to the world → low generalizability to the world and poor utility for further theory
- risk erroneous conclusions
- lack of humility: over-interpretation of the scope/relevance of the model: this happens in my model, so it must happen in the world!

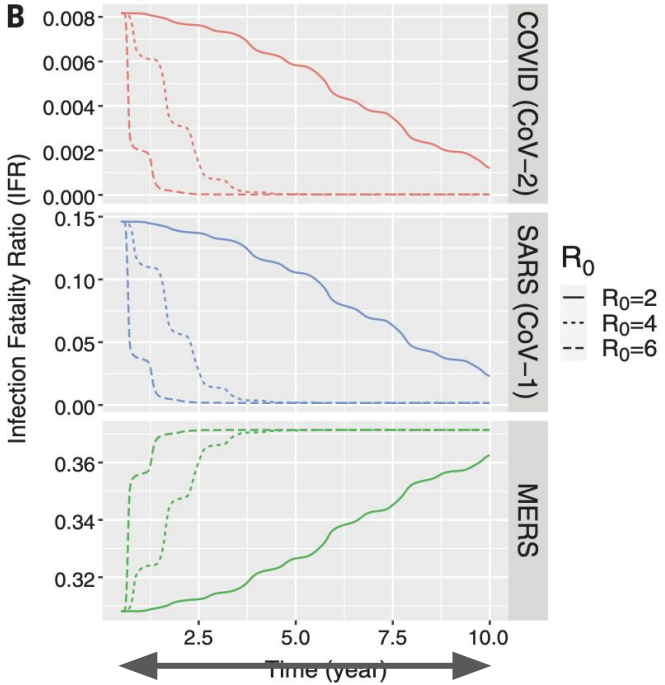


Too much belief in models: COVID-19 virulence

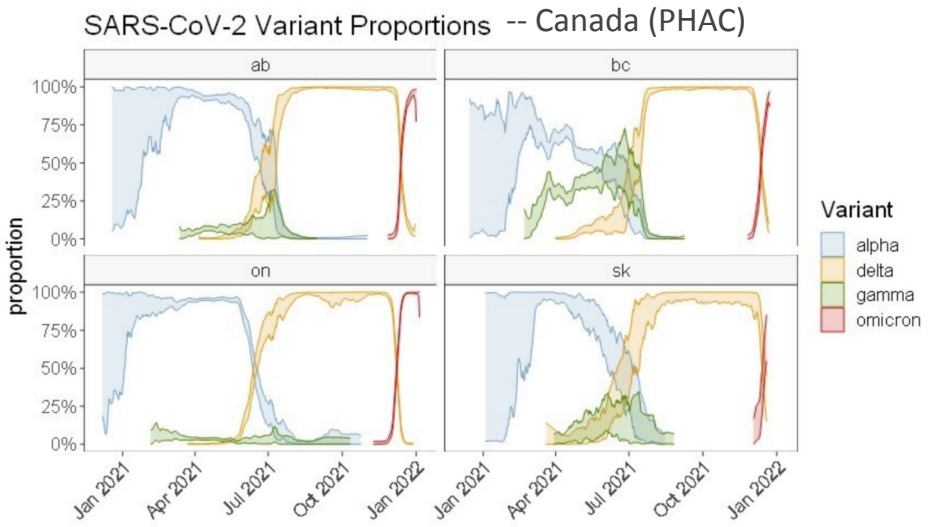
- There is a large literature on the evolution of virulence (infection-induced mortality). Many models assume a trade-off between transmission and virulence (e.g. Anderson and May, 1982; Frank, 1996; Alizon *et al.* 2009).
- People assumed that as COVID-19 became more transmissible it would also become more mild.
- But severity **increased**: first with alpha, then with delta (among others).
- There is little selection on severity for SARS-CoV-2. Severity can go either way. We were lucky with omicron.

Too much confidence - COVID without evolution?!

Modelling without evolution predicted milder and milder disease following re-exposures



10 years



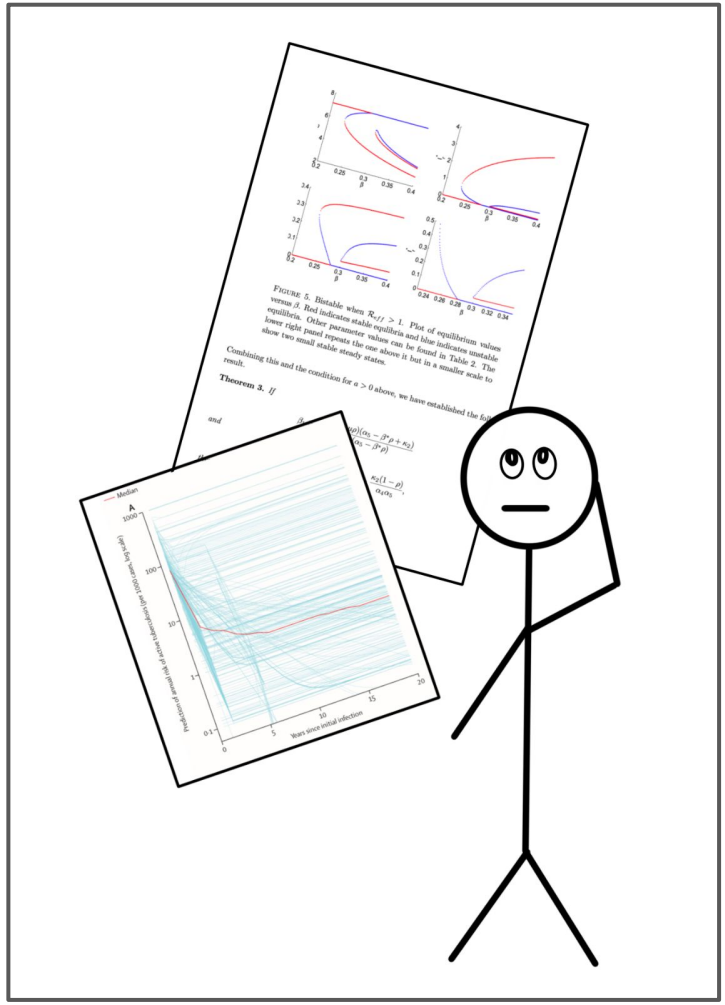
1 year

[Lavine et al, *Science*, 2021]

Not even a theory: Problem 2

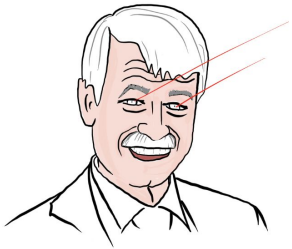
The model is not correctly set up to ask a relevant question or test a hypothesis

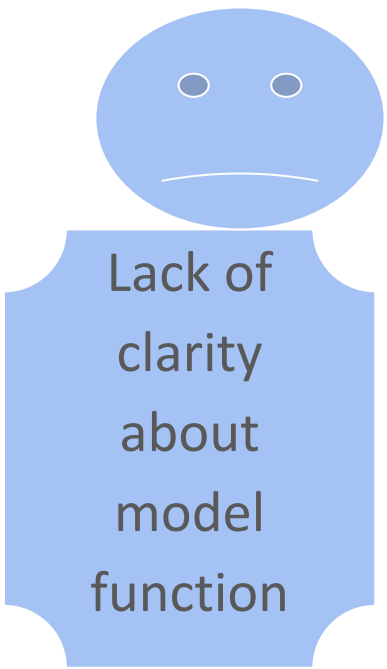
- underlying data insufficient to constrain models
- models need strong assumptions; these determine answers to questions
- assumptions may be subtle; not readily examined - particularly a problem in large, complex models
- lack of humility: erroneous results, or results that depend almost entirely on ad hoc assumptions



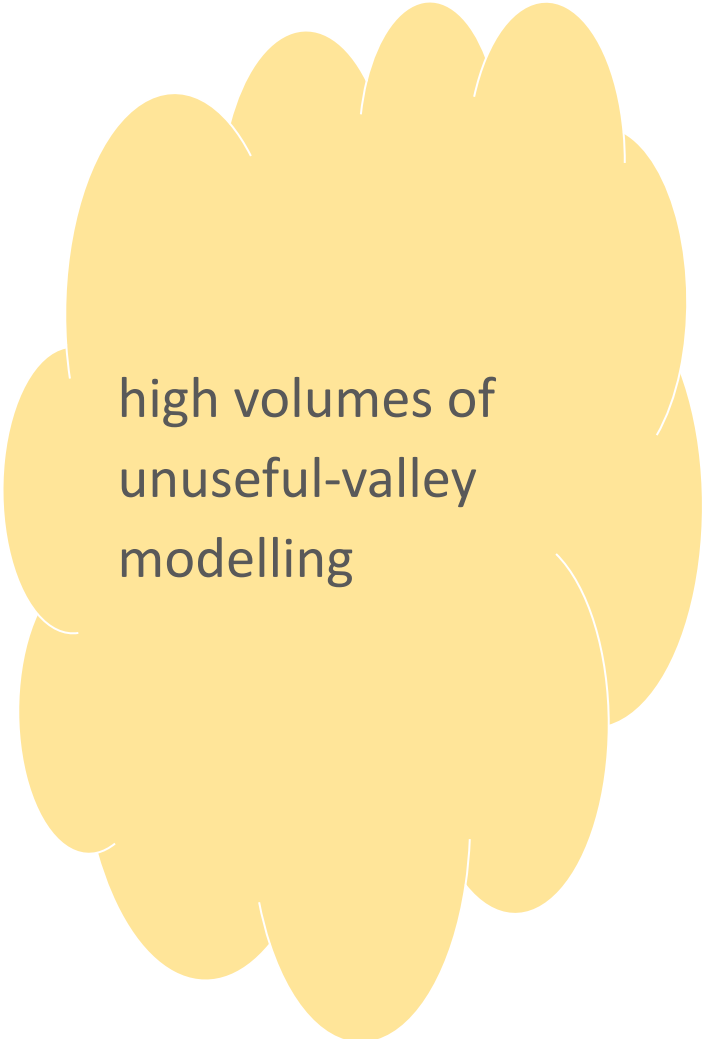
Not-even-a-theory models can do harm

- People see (high volumes of) weak or irrelevant models and disengage, or don’t engage in the first place
- This limits models’ ability to do what models are good at!
- It reduces motivation to:
  - gather data to inform models
  - share data with modellers
  - collaborate with modellers
- It gives credibility and talking points to people who are skeptical of models, or anti-modelling, or sometimes even anti-science!




Lack of clarity about model function

+


high volumes of unuseful-valley modelling

→


tenure committees, funders, collaborators and end users reject really good science

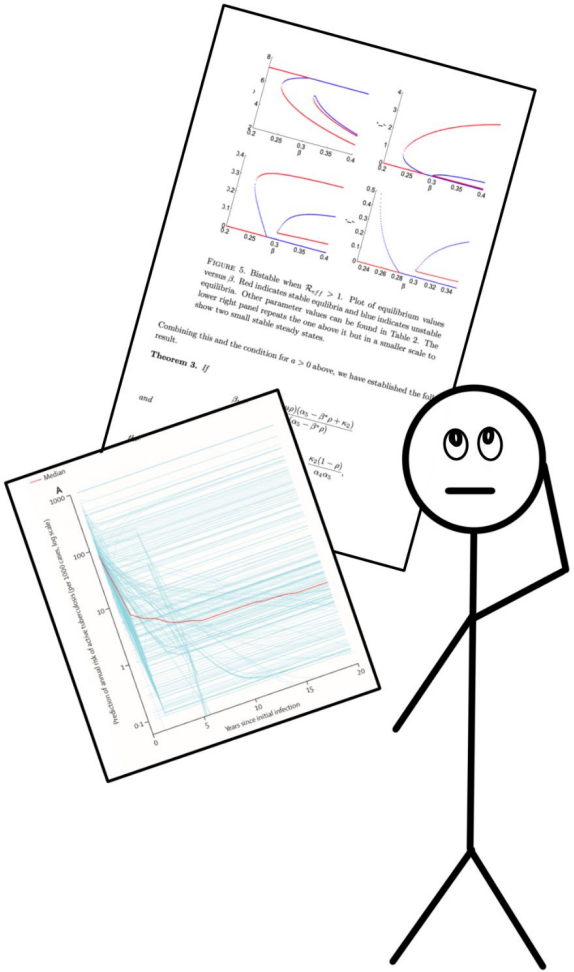


FIGURE 5. Bistable when $\mathcal{R}_{eff} > 1$. Plot of equilibrium values versus β. Red indicates stable equilibria and blue indicates unstable equilibria. Other parameter values can be found in Table 2. The lower right panel repeats the one above it but in a smaller scale to show two small stable steady states.
Combining this and the condition for $a > 0$ above, we have established the fol
result.
Theorem 3. *If*
and
Median
A
1000
100
10
1
0.1
0
5
10
15
20
Years since initial infection


Models are irrelevant.
They aren't crystal balls.
We don't need them.
We can do what we've always done.
I already know what to do.

Questions for discussion

1. Could there be a (good, useful) model that cannot make a prediction for *any* specific system?
2. Could there be a (good, useful) model that does not need *any* data? (not necessarily in a formal model fitting sense)
3. Does a useful model have to have *relevance* for *some* specific system? (and if so, what the notion of relevance?)

4. As modellers, how can we avoid pitfalls of modelling and the resulting harms?

Suggestions

“Good modelling” guidance?

- “Models must be created in consultation with stakeholders” -- no
- “Models must be fit to data” -- no
- “Model predictions must be tested against data” -- no
- “Models must include all relevant groups and processes” -- no

The problem: often guidance is rooted in the “scenario projection” function only.

Alternatives ideas:

- Think about what information informs the model (even if not “fit to data”!).
- Articulate why the model is a good analogy, and what its limitations are.
- Is the work proof-of-concept modelling? A null model? A qualitative exploration? An evidence synthesis?
- If so, what phenomenon or exploration? Why is it important? How can a model help? How does the model avoid assuming the answers to the questions being asked? What else needs to be done before it’s reliable, beyond modelling?

Concluding remarks

- Modelling's many purposes are often poorly understood by those who need to know (tenure committees, policy-makers, funders, data controllers, sometimes even researchers ourselves)
- Modelling isn't always good science. There is an "un-useful valley" out there.
- Modelling played a huge role in the pandemic. "Rt" was a household phrase…
- We can articulate the use for a particular model, with reference to models' distinct functions.
- We can articulate how models are informed by data, why a model is a good analogy for the world, and how (strongly) it is relevant.
- We should think clearly about how and why our model is good science, and communicate that to our audiences.

Thank you

Ted Cohen, Yale University, TB epidemiologist

Jennifer McNichol, SFU, PhD student, cartoonist

Maud Menten Institute -- thank you for having me!

The new **Society for Modelling in Theoretical Population Biology**

Photo from a BIRS meeting where I gave an shorter version of this talk