Imperial College London

MRC Centre for Outbreak Analysis and Modelling

Introduction to designing infectious-disease models

Dr Peter J White

(some slides from Prof Christophe Fraser & Dr Thomas Hagenaars)

Epidemiology & Control of Infectious Diseases

Models

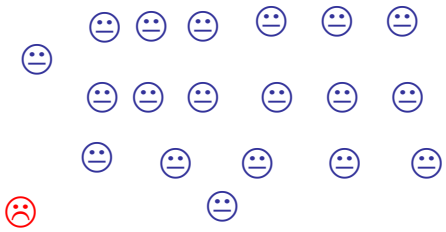
- Simplification of a system, suitable for analysis.
- They are tools for thinking.
- *"All models are wrong, but some are useful."* G.E.P. Box
  NB. applies to statistical models, too
- *"Describing complex, poorly-understood reality with a complex, poorly-understood model is not progress."* J. Maynard Smith
- Needs to capture essential behaviour of interest and incorporate essential processes.
- **Everyone uses models** in their head: making them explicit mathematically clarifies thinking & allows others to examine them.
- Mathematical models allow precise, rigorous analysis and quantitative prediction.

Models of infectious disease transmission

- We are interested in the population-level effect of processes occurring at the individual level.
- An uninfected individual's risk of becoming infected (the "force of infection") depends upon the prevalence of infectious individuals (a population-level characteristic)
  - (and rate of contact between individuals, infectiousness of infected individuals, etc).
- So transmission of infection in a population is a dynamic process and the individual risk of infection can change over time.
- Incidence depends upon prevalence…
  and of course prevalence depends upon incidence
  → non-linearity, with time-varying positive & negative feedback
- It requires dynamic models for prediction & analysis of putative programmes.

Transmission (1)

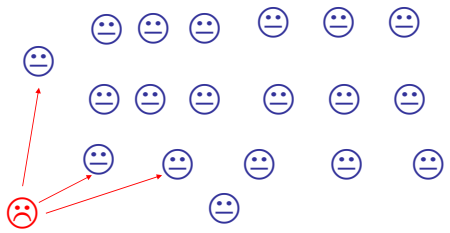
😐 Naïve → ☹ Infected → ☺ Immune



One has acquired infection

Transmission (2)

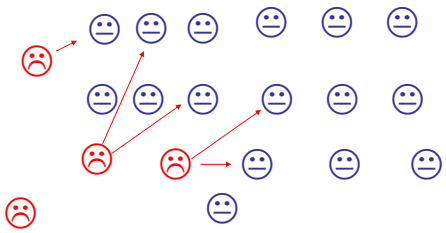
😐 Naïve → ☹ Infected → ☺ Immune



Transmission

Transmission (3)

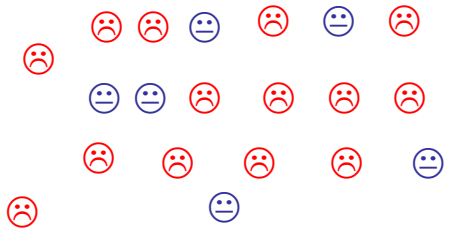
😐 Naïve → ☹ Infected → ☺ Immune



More transmission: prevalence ↑ → per-susceptible risk ↑

Transmission (4)

😐 Naïve → ☹ Infected → ☺ Immune



Transmission (5)

😐 Naïve → ☹ Infected → ☺ Immune

Recovery to immunity

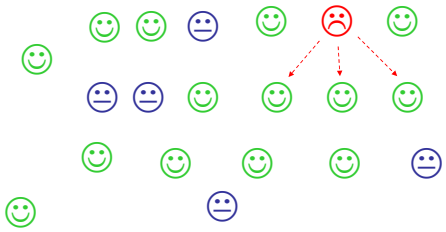
Transmission (6)

😐 Naïve → ☹ Infected → ☺ Immune

Another acquires infection from elsewhere

Transmission (7)

😐 Naïve → ☹ Infected → ☺ Immune



No transmission due to herd immunity – naives not infected

Modelling considerations

- Natural history of infection
  - e.g. latency, infectious period, immunity

  (Depends on infection)
- Transmission of infection
  - e.g. direct or indirect? What affects contact rate?

  (Depends on population; Depends on infection)
- Population structure & demography
  - e.g. stratify by age, sex, geographic location, etc?

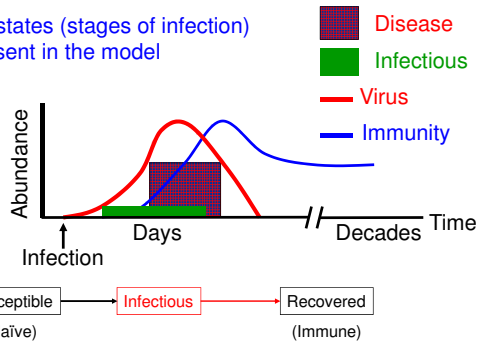
  (Depends on population)
- Interventions
  - What parts of the disease-transmission process are targeted?

  (Depends on intervention(s))

→ *there is no one 'correct' model for a particular infection*

Representing natural history of an example viral infection

Identify states (stages of infection) to represent in the model



NB. here we assume that infected = infectious

Creating a compartmental model

Modelling natural history

- Divide population into compartments (categories), containing individuals in different states (e.g. stages of infection).
  - In the model, all those in a category have the same properties
  - These are the average properties of those in the 'real world'

Compartmental modelling using differential equations (1)

Compartmental infectious disease models are constructed from two basic types of objects:

(1) **Compartments**: containing people in each infection state (stage of infection).
- These values are stored in **state variables**, which together describe the **state** of the system.

(2) **Rates of change** of numbers in compartments: e.g. birth rate, incidence of infection, recovery rate.
- These rates usually depend upon the values of one or more of the state variables, so they change as the state of the system changes – i.e. there is **feedback** – e.g. population growth rates, disease transmission rates.

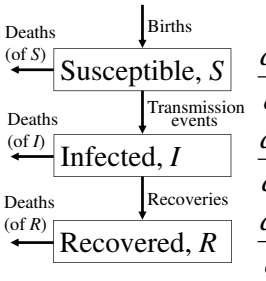
Compartmental modelling using differential equations (2)

- The relationships between the values of the **state variables** and their rates of change are expressed by **functions** (one for each rate of change).
- Each compartment has

  a **state variable**, keeping track of the number of individuals in that compartment; and

  a **differential equation**, describing the rate of change of its state variable, which is comprised of one or more of the **functions**.

Flow diagram → equations



$$\frac{dS}{dt} = \text{Births} - \underset{\text{events}}{\text{Transmission}} - \text{Deaths (of } S\text{)}$$

$$\frac{dI}{dt} = \underset{\text{events}}{\text{Transmission}} - \text{Recoveries} - \text{Deaths (of } I\text{)}$$

$$\frac{dR}{dt} = \text{Recoveries} - \text{Deaths (of } R\text{)}$$

S,*R* die at "background" (non-diseased) rate;
I die at "background" rate + disease-induced death rate (which could be 0).

Each term in these equations is a function of one or more of the state variables (*S*,*I*,*R*), meaning that the value of the term changes as the state variable(s) change(s).

(NB. *S* ≡ naïve, *R* ≡ immune)

More-complex natural histories

Examples:
- Maternal antibodies in young
- Incubation and latent periods
- Asymptomatic infection
- Infectious period with multiple stages
- Resolution of infection
  - death, immunity, return to susceptibility?
- Immunity
  - sterilising?
  - waning or permanent?
  - strain-specific or general?

But let's start simply!

- The point is to understand fundamental processes rather than to make models 'realistic' initially.
- Realistic models are complicated and hard to understand unless fundamental processes that occur are understood first.

Flows between compartments (1)

- Each flow rate is the **no. individuals entering or leaving a compartment per unit time**. It depends upon:
  - the **per-capita** rate (the hazard); and
  - the **number of individuals** subjected to that per-capita rate (i.e. exposed to the hazard).
- The **population** rate is the product of these - i.e. per-capita rate × no. individuals.

Flows between compartments (2)

- Rate of recovery from infection
  - let the **per-capita** rate of recovery be σ ;
  - since the no. Infecteds is I,
  - the **population** recovery rate is therefore σI.

State variable
• Varies intrinsically as model runs
• Not manipulated directly
• Model 'outputs'

Parameter
• Specified extrinsically
• Only changes if/when we specify
• Model 'inputs'

When we analyse the model we vary parameter values and see how this affects the state variables – i.e. how the output graphs change.

Flows between compartments (3)

- Death rate of Susceptibles, Recovereds
  - let the **per-capita** "background" death rate be b;
  - since the no. Susceptibles is S, and Recovereds, R,
  - the **population** death rate of Susceptibles is bS and of Recovereds is bR.
- Death rate of Infecteds
  - Infecteds experience the "background" death rate (b) + disease-induced death rate (α), so their **per-capita** death rate is $(b + \alpha)$, so
  - the **population** death rate of Infecteds is $(b + \alpha)I$.

Flows between compartments (4)

- Birth rate
  - Let the **per-capita** birth rate (averaged over males & females!) be a;
  - assuming individuals in all compartments give birth at the same rate, the no. giving birth is $S+I+R$, which we can simplify by defining total population size, $N = S+I+R$, so
  - the **population** birth rate is aN.
  - In practice, we often we set the per-capita birth rate equal to the per-capita background death rate to maintain a constant population size in the absence of disease.
    - *However, using different parameters means that we do not have to make them equal if we don't want to.*

Flows between compartments (5)

- Infection rate
  - per-capita rate of infection of susceptibles (the "**force of infection**") is not fixed;
  - it depends upon:
    - the number Infectious at the particular point in time, and
    - the rate of contact with Susceptible individuals, and
    - the transmission probability.
  - Therefore the **population** rate of infection depends upon the number of Infecteds as well as the number of Susceptibles.
    - Hence it is **non-linear**.

Transmission rate formula

(for direct transmission)

Consider <u>one</u> susceptible individual…

Rate of contacting other individuals: c ← Contact rate

Transmission requires contact with infectious individuals…

Rate of <u>contacting</u> infectious individuals: cI/N ← Proportion of population infectious (I = no. infectious; N = no. in total population)

Rate of <u>transmission</u> from infectious individuals: pcI/N ← Probability of transmission when an infectious individual contacts a susceptible

Force of infection, λ

Now consider <u>all</u> susceptible individuals…

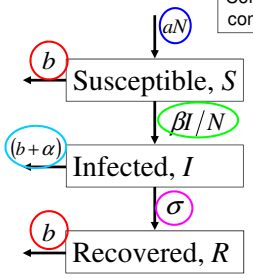
Total transmission rate in population: $pcSI/N$ ← No. susceptible individuals

(NB. *S, I, N* are state variables, which can change intrinsically, whilst *p, c* are parameters and cannot.)

Often β is written in place of pc.

Equations of an SIR model

Conventionally, the <u>per-capita rate of leaving</u> the compartment is shown on the diagram, as here.



$$\frac{dS}{dt} = aN - \beta SI/N - bS$$

$$\frac{dI}{dt} = \beta SI/N - \sigma I - (b+\alpha)I$$

$$\frac{dR}{dt} = \sigma I - bR$$

$$N = S + I + R$$

N, S, I, R are variables. a, b, α, β, σ are parameters.

Solving the equations: integration

- We are interested in plotting how the numbers in each compartment change over time.
- In our differential equations we have specified **derivatives** – i.e. defined how the <u>slopes of the lines</u> of the numbers in each compartment relate to the state of the system (i.e. the <u>numbers</u> in each compartment) at any point in time.
- To get the lines themselves, we solve the equations by **integration**. Most models cannot be solved algebraically, so use computers for **numerical integration** (fortunately for biologists!).

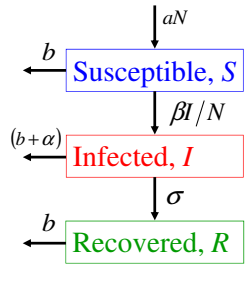
Behaviour of simple models

- The point is to understand fundamental processes rather than to make models 'realistic' initially.
- Realistic models are complicated and hard to understand unless fundamental processes that occur are understood first.
- *Analysis involves varying parameter values singly and in combination to see how model outputs change (but no time to do it here).*

Example epidemic: SIR model

- Initially in an epidemic in a naïve population,
  the rate of spread accelerates as transmission ↑ Infecteds,
  which ↑ force of infection
  which ↑ further the rate of spread.
- Then spreading slows as Susceptibles ↓ significantly,
  (even though force of infection continues to ↑).
- If Infecteds recover to become immune then the epidemic can fade out, unless
  - new naive individuals enter the population, or
  - immunity wanes, returning individuals to susceptibility.

Equations of an SIR model



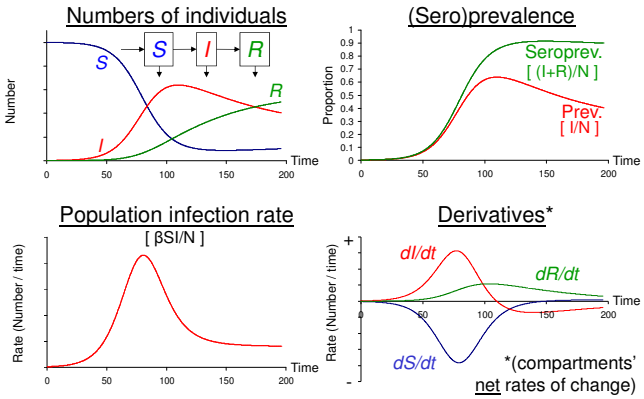
$$\frac{dS}{dt} = aN - \beta SI/N - bS$$

$$\frac{dI}{dt} = \beta SI/N - \sigma I - (b+\alpha)I$$

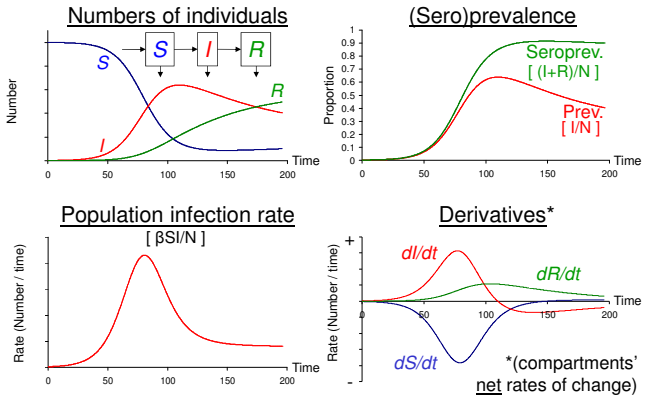
$$\frac{dR}{dt} = \sigma I - bR$$

$$N = S + I + R$$

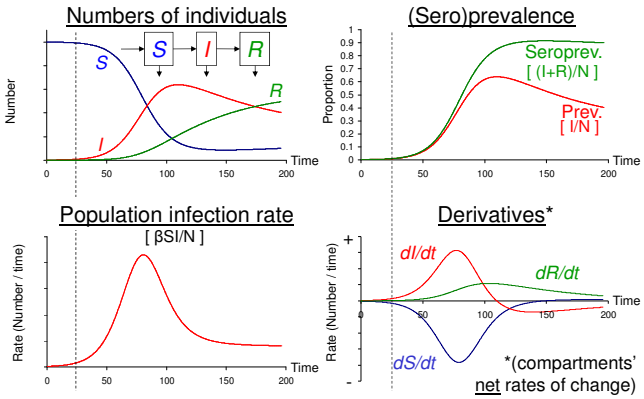
SIR model: dynamics



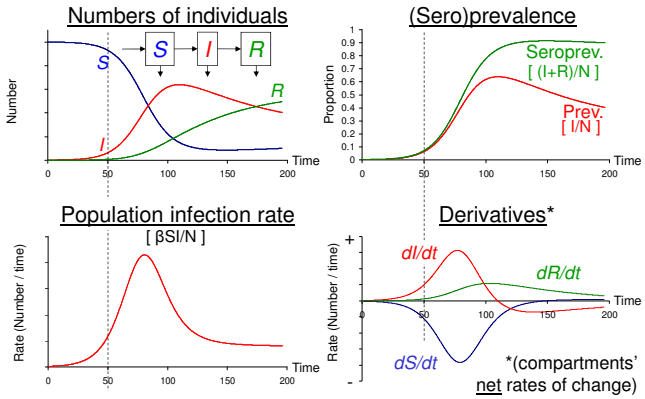
Time 0: low transmission rate, as I=1 so I/N tiny



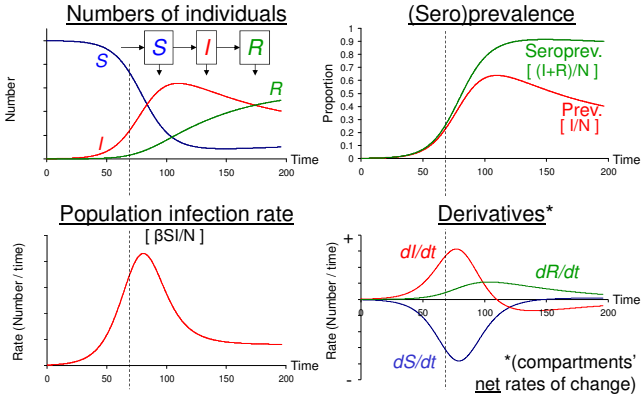
Exponential growth: transmission rate >> than at time 0 but still low, as I/N still small



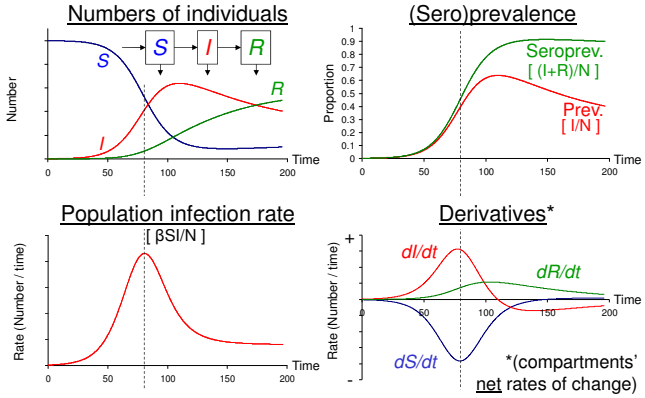
Exponential growth: transmission rate ↑, as I/N ↑



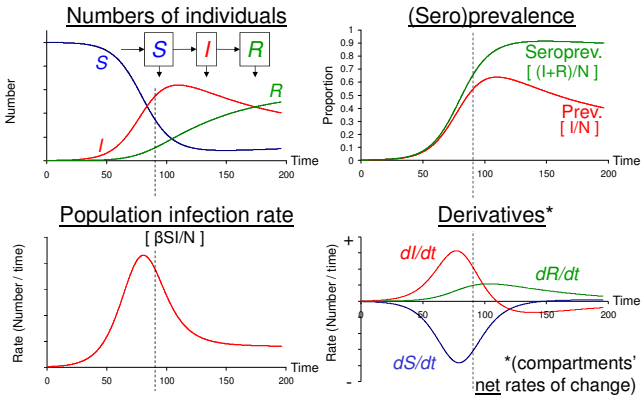
Transmission rate still growing as I/N ↑ but growth slower than exponential, as S ↓



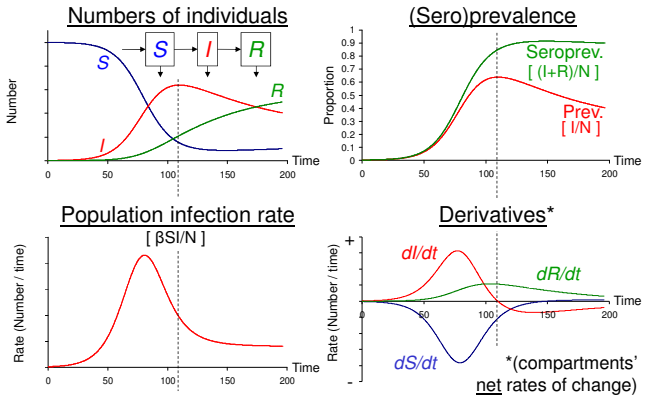
Peak transmission rate, due to S ↓: I/N still ↑



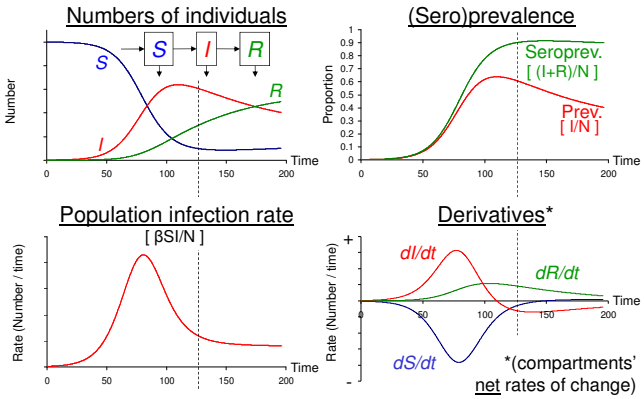
Transmission rate ↓, due to S ↓: I/N still ↑



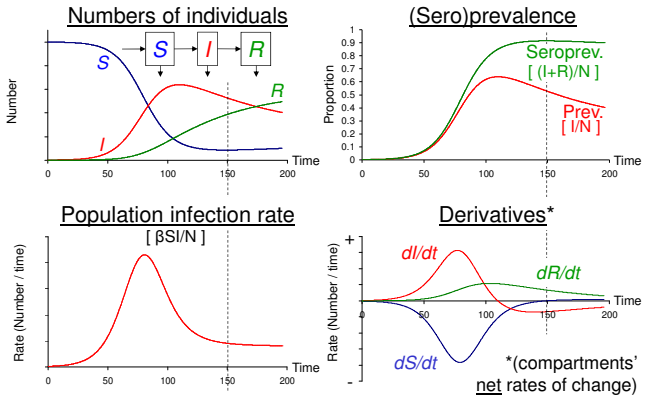
Transmission rate ↓↓, due to S ↓: I/N at its peak (i.e. force of infection is at its peak)



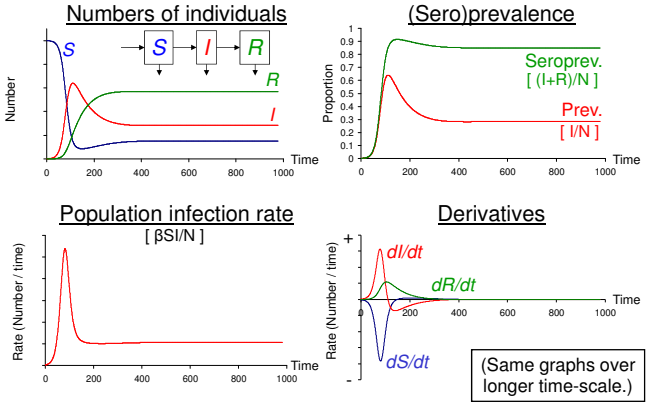
Transmission rate ↓ continues, due to S ↓ and now I/N ↓



Transmission rate ↓ continues, approaching steady state



SIR model: steady state



More-complex natural histories

- Although models should be parsimonious, they need to capture essential details of the disease, so we often have more compartments.
- Which details are included will depend on the questions addressed by the modelling, and the availability of data.

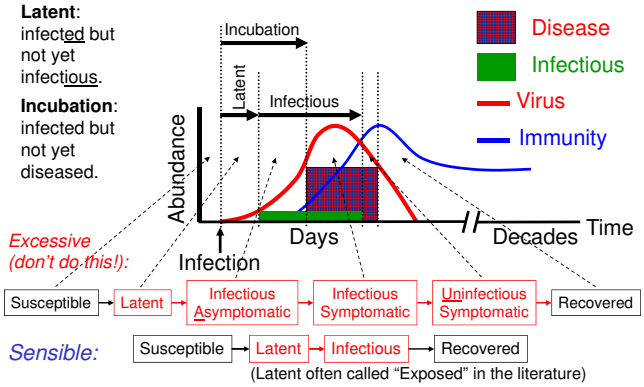
Latent period

- In simple models, we assume individuals become infectious as soon as they are infected.
- However, there may be a significant latent period between being infected and becoming infectious.
- *In modelling literature, latently-infected are often called "Exposed".*

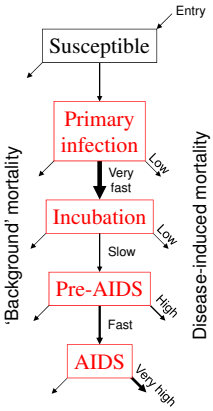
Incubation period

- This is the time between becoming infected and becoming symptomatic.
- Often infections are only treated when a person becomes symptomatic and so becomes aware that they have an infection.
- For some diseases, symptoms occur before the person is infectious (e.g. SARS) whilst for others symptoms begin after the person is infectious (e.g. HIV, possibly influenza).
- For some diseases, symptoms and infectiousness occur together (e.g. pulmonary TB).

Natural history of an example viral infection



HIV



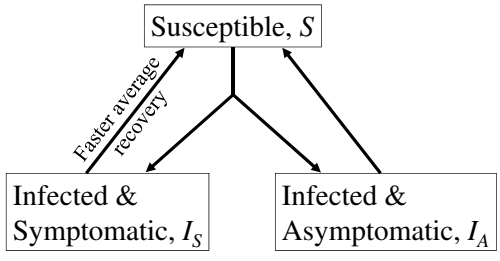
Stages differ in infectivity, duration & risk of disease-induced mortality

- Infectivity / viral load: very high
  Duration: a few weeks
  Disease-associated mortality: low
- Infectivity / viral load: low
  Duration: ~8-10 years
  Disease-associated mortality: low
- Infectivity / viral load: high
  Duration: ~1-2 years
  Disease-associated mortality: high
- Infectivity / viral load: high
  Duration: ~1-2 years
  Disease-associated mortality: very high

'Branching' natural histories

- Some infections have different natural histories in different people.
- e.g. with gonorrhoea, some people develop symptomatic infection, whilst others develop asymptomatic infection.

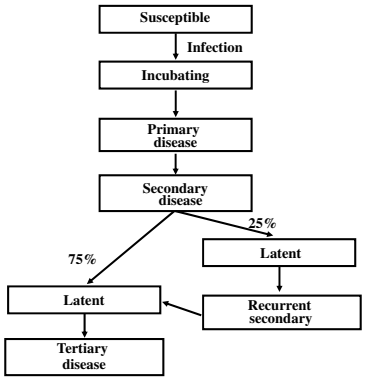
Simple model of gonorrhoea



- Note there is no immunity.
- Symptomatic infection has shorter mean duration because many/most seek care, whilst few asymptomatics do so – most of them recover through natural immune processes.

(Garnett et al. 1999)

Syphilis (without treatment)

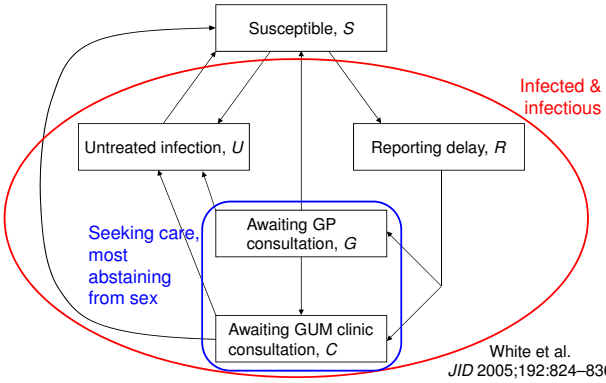


(Garnett et al. 1997)

Models incorporating behaviour

- Sometimes people are categorised according to their behavioural response to their infection status, as well as the physiological status of the infection.
- If people have the similar physiological characteristics, but different behavioural ones, then they may need separate compartments for different behavioural categories.

Gonorrhoea & treatment-seeking behaviour in the UK



White et al. *JID* 2005;192:824–836

Heterogeneity

- Compartmental models represent the population in **aggregate**.
- Those in a compartment have to have **identical** characteristics because they can't be distinguished. (Although different outcomes occur to different 'individuals', the probabilities of their occurrence are the same for each person.)
- To represent **variation**, can **stratify** the population into different groups:

- Stratification means that continuous variables have to be 'discretised' – i.e. placed into discrete categories with often arbitrary boundaries (but often the data are like this!).

Final remarks

Optimal scepticism

Some people are unduly hostile to modelling

- Riposte: "The first model builders in tuberculosis met with considerable opposition from those who maintained that many essential parameters were not established with sufficient precision, although paradoxically, those very *opponents apparently had their own intuitive models* on which to base highly assertive decisions." World Health Organization, 1973; quoted in Lietman & Blower CID 2000.
- **Everyone uses models** for decision-making – "modelling" is making them explicit, which clarifies thinking & allows others to examine them.
- Making models mathematical allow rigorous analysis and comparison with data, and *potentially* quantitative prediction.

Others are unduly enthusiastic about modelling & have unrealistic expectations

- Caution: "Mathematical model formulation in the biomedical sciences more often *acts to tell you what you do not know* rather than providing a precise predictive tool." Roy Anderson PNAS 1998.

Modelling is not alchemy!

It cannot turn base metal (poor quality* or non-existent data) into gold (an accurate, precise prediction). [*NB. quality >> quantity]

Where data are lacking or imprecise, modelling can examine scenarios based on varying a parameter within its plausible range. This can determine the importance (or not) of measuring that parameter more accurately.

It might be possible to rule-out potential interventions even without quantifying them accurately – if the best case scenario is still unimpressive.

A model is **not** a substitute for data – it is a tool to analyse data.

Parameter estimation is the hardest part

To model something mathematically its effect has to be quantified: models have parameters, which have numerical values – either:

(i) measured: usually hard – and requires a lot of high-quality data;

(ii) varied across plausible ranges – scenario analysis

Models are data-hungry because lots of factors affect transmission – this is the price of realism

Estimating parameters from data is typically the hardest part of modelling – not the designing or programming of the model

It requires expertise in modelling and understanding of the data – multidisciplinary

Most modelling analyses use models specially designed (or at least adapted) to the available data

Van Kerkhove et al. PLoS Med 2010


Policy Forum

Studies Needed to Address Public Health Challenges of the 2009 H1N1 Influenza Pandemic: Insights from Modeling