Basic methods for setting up models: Difference Equations

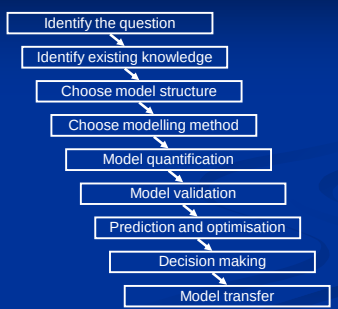
Richard White

Centre for the Mathematical Modelling of Infectious Diseases &
TB Modelling Group
LSHTM

Objectives

- By the end of this lecture you should
  - understand model development steps
  - be aware of the common model structures and types used for modelling infectious diseases
  - understand how deterministic models are set up using difference equations
  - be able to write equations for a simple model
  - be able to define key input parameters

Model development steps



(after Habbema et al, 1996)

Identify the question

- Motives for modelling infectious diseases?
  - Understand observed patterns and trends
  - Explore what happens if...?
  - Derive epidemiological parameters
- We need to set clear priorities for:
  - Determining the essential model processes
  - Making the modelling relevant
- For today, our question is:
  - *What is most likely to happen if a single measles case is introduced into a population of 100,000 susceptibles?*
    - *To keep it simple:*
      - *Over the short term (100 days)*
      - *No births or deaths*

Identify existing knowledge

- Collate existing knowledge
  - Research papers, grey literature, lab reports, existing modelling exercises...
  - Organise quantitatively by
    - Transmission
    - Epidemiology
    - Natural history
    - Control options
  - Discuss review with experts

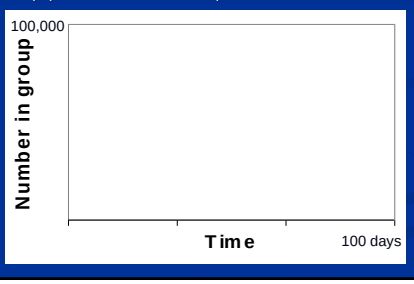
Identify existing knowledge

- We've found out
  - Measles is an infection that confers near permanent immunity
  - Each person effectively contacts 1.5 others per day
  - The pre-infectious (latent) period (infection to infectiousness) is 8 days
  - The duration of infectiousness is 7 days
  - R_0 = 14

So what happens if we introduce a single infectious case?

Identify existing knowledge

So what happens to the number susceptible, infectious and immune, if we introduce a single infectious case into a population of 100,000 susceptibles?



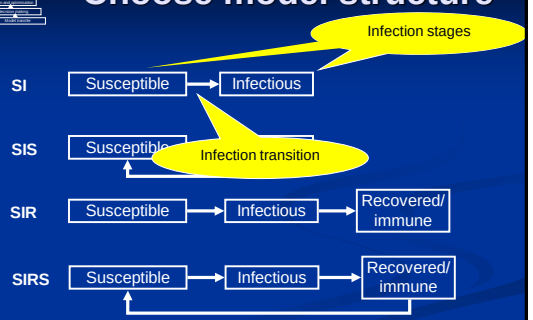
Difficult to predict purely from available data

But we can try to write down the equations for the number of susceptible, infectious and immune individuals tomorrow in terms of what we know today…

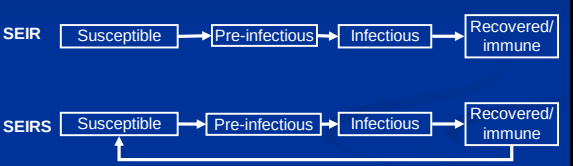
Choose model structure

- When choosing, consider
  - Infection stages
  - Infection transitions
  - Population groups
- While bearing in mind that *'models should be as simple as possible and no simpler'*

Choose model structure



Choose model structure



Which is most appropriate model structure to answer our measles question?

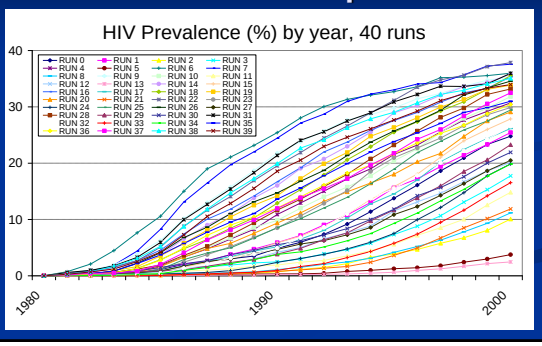
Choose model structure

- Either:
- But, the structure also depends on required accuracy of model prediction
  - Do we need a prediction to the nearest day, week or year?
    - As we are only modelling for 100 days, model to the nearest day.
    - Then the pre-infectious period(8 days) is *long* by comparison.

Choose modelling method

- *Stochastic* or *deterministic* equations
- Stochastic models
  - Use stochastic equations that incorporate chance fluctuations
  - Provide a range of outcomes per scenario
  - Many 'runs' required
  - 'Runs' averaged to give most likely outcome for scenario, along with a range of variability
  - Used where
    - chance fluctuations are important eg, small isolated populations, start or end of epidemic
    - information on the variability of the outcome is as important as average outcome
  - Stochastic models are often used to describe infection process for every individual / micro-parasite in a population

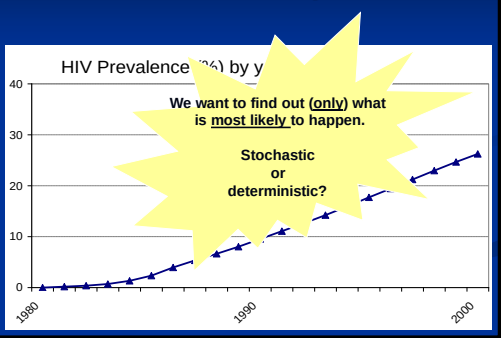
Modelling using stochastic equations



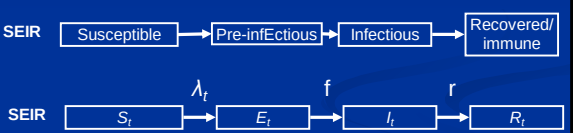
Deterministic models

- Deterministic models
  - Use deterministic equations that describe what happens on average and do not incorporate chance fluctuations
  - Provides only one outcome per scenario
  - Transitions rates are average transition rates for population
  - Most deterministic models are compartmental models in which population divided into small number of subgroups (compartments)
- Don't confuse modelling method (stochastic vs. deterministic) with 'granularity' of model structure (groups/compartmental vs. individual-level)

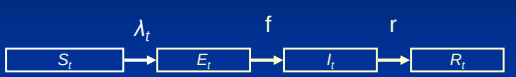
Modelling using deterministic equations



Measles model using *difference* equations



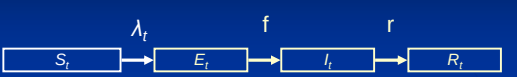
Developing our measles model



So let's write an equation to predict the number susceptible tomorrow S_{t+1} depending on the number susceptible today S_t

$$S_{t+1} = S_t - \lambda_t * S_t$$

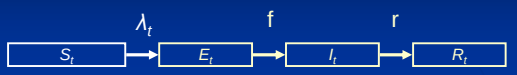
The number infected



And the number pre-infectious tomorrow E_{t+1} depending on the number pre-infectious today E_t

$$E_{t+1} = E_t + \lambda_t * S_t - f * E_t$$

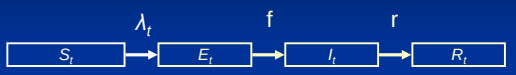
The number infectious



And the number infectious tomorrow I_{t+1} depending on the number infectious today I_t

$$I_{t+1} = I_t + f * E_t - r * I_t$$

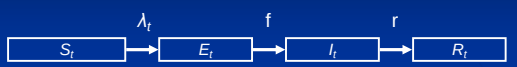
The number immune



And the number immune tomorrow R_{t+1} depending on the number immune today R_t

$$R_{t+1} = R_t + r * I_t$$

The difference equations for our measles model



$$S_{t+1} = S_t - \lambda_t * S_t$$
$$E_{t+1} = E_t + \lambda_t * S_t - f * E_t$$
$$I_{t+1} = I_t + f * E_t - r * I_t$$
$$R_{t+1} = R_t + r * I_t$$

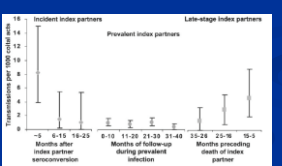
To predict what happens when we introduce measles, we simply calculate these equations for next year (tomorrow)…

This is called *solving* the equations

Can we do this now?

Model quantification

- Determining the value of input parameters and distributions
  - Transition parameters e.g. λ_t, f and r
  - Risk distributions e.g. constant, peaked or 'bathtub' risk profile?



Wawer, JID, 2005

Model quantification

- Main problem is usually lack of data, estimate using:
  - Primary data collection
  - Data analysis (statistical modelling)
  - Other modelling exercises
  - Expert opinion (?)

Model quantification

- So in our example we need to estimate
  - λ_t, the risk a susceptible becomes pre-infectious between t and t+1
  - f, the risk an pre-infectious individual becomes infectious between t and t+1
  - r, is the risk an infectious individual recovers (becomes immune) between t and t+1

The force of infection, λ_t

Assumed not to vary over time

If we assume in our population [illegible] be [illegible]es according to [illegible]

A contact that would lead to infection if it occurs between a susceptible and an infectious person

$\lambda_t = \beta * I_t$

Where:

β = the per-capita rate at which two specific individuals come into effective contact between t and t+1

I_t = the number of infectious individuals in the population at *t*

Allowed to vary over time. Why?

Model quantification - β

To estimate ß (the per-capita rate at which two specific individuals come into effective contact per unit time) we use two other relationships:

$R_0 = ecr * D$ and $ß = ecr / N$

Where:

R_0 = the number of secondary infections that would result from the introduction of an infectious case into a completely susceptible population (14)

ecr = the total number effective contacts of each person, in each day (?)

D = the duration of infectiousness, in days (7)

N = population size (12)

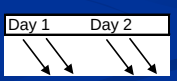
First relationship: $R_0 = ecr * D$

The definition of R_0 is:

the number of infections that would result from the introduction of one infectious case into a totally susceptible population during its entire infectious period

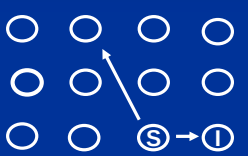
Thus, it **must** equal the number of people effectively contacted by an infectious individual per unit time *(ecr)* multiplied by the duration that they are infectiousness (*D*):

$R_0 = ecr * D$



Second relationship: ß = ecr / N

If we assume (again) the effective contact rate (total number contacts between our susceptible and any other individual per day) = 2 (per day), in a population (N) of 12 individuals:



ß (the contact rate between the specific susceptible and the specific infectious individual) is:

$ß = ecr / N$

=> ß = 2 / 12,
= 17 %

But we don't know 'ecr' so …

Model quantification - β

But if we combine the two equations, we can get rid of the unknown 'ecr':

$ß = \frac{ecr}{N}$ and, $R_0 = ecr * D$

Rearrange:

$ecr = \frac{R_0}{D}$

Substitute this into, gives $ß = \frac{R_0}{D * N}$

Model quantification - β

So if we knew R_0, *D* and *N*, we could estimate ß.

So let, $R_0 = 14$, $D = 7$ and $N = 12$

$=> ß = \frac{R_0}{D * N} = \frac{14}{7 * 12} = 17\%$ / day

Where:

R_0 = the number of secondary infections that would result from the introduction of an infectious case into a completely susceptible population

D = the duration of infectiousness

N = population size

Model quantification Risks and rates

- Transition parameters (*f* and *r*)
  - should be risks, but if <10%, are similar to rates
  - If rate is constant, then the average rate at which the event occurs is 1 over the average time to the event, ie
    =1/ (average time to event)
  - and the distribution of the times to the event is exponential

    - eg, rate of disease onset (f)
      =1/ (average latent period)
    - eg, rate of recovery/ becoming immune (r)
      =1/ (average period of infectiousness)
    - eg, mortality rate =1/ (life expectancy)

Model quantification Risks and rates

If the average pre-infectious period for measles is 8 days, the rate of disease onset

= 1 / (average pre-infectious period)

= 1 / 8 per day, or 7 / 8 per week

And, if the life expectancy of a population is 60 years, the average mortality rate is

= 1 / (average life expectancy)

= 1 / (60 years)

= 1 / 60 th **per year**, or

= 1 / (60 * 365) **per day**

Model validation

- Check outputs against independent data sets
  - Daily numbers of cases reported
  - Seroprevalence of infection
- In more complicated model
  - Demography
  - Epidemiology, age and sex patterns

Prediction and optimisation

- Should occur after model validation
- Will depends on model assumptions
- Use sensitivity analysis of input parameters

Prediction and optimisation

Days	Susceptible	Infected	Infectious	Immune
0	99,999	0	1	0
1	99,998	1	1	0
2	99,996	3	1	0
3	99,995	4	1	0
4	99,993	5	1	1
5	99,991	6	2	1
…				
95	0	73	339	99,587
96	0	64	300	99,636
97	0	56	265	99,679
98	0	49	234	99,717
99	0	43	207	99,750
100	0	38	183	99,779

Prediction and optimisation

Decision making

- Make predictions understandable
- Purely technical descriptions are not likely to be sufficient for policy makers and programme managers

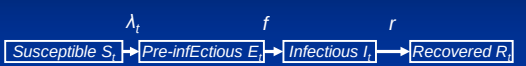
Model transfer

- Once model stable, if desired, a user friendly version can be transferred to policy makers and programme managers

Summary

- Hopefully you now
  - understand model development steps
  - are aware of the common model structures and types used for modelling infectious diseases
  - understand how deterministic models are set up using difference equations
  - could write equations for a simple model
  - could define key input parameters
- But if you only take one thing away…

If you only take one thing away...



$$S_{t+1} = S_t - \lambda_t * S_t$$
$$E_{t+1} = E_t + \lambda_t * S_t - f * E_t$$
$$I_{t+1} = I_t + f * E_t - r * I_t$$
$$R_{t+1} = R_t + r * I_t$$

Where, $\lambda_t = \beta * I_t$

Basic methods for setting up models: Difference Equations

Richard White

Centre for the Mathematical Modelling of Infectious Diseases &
TB Modelling Group
LSHTM