Block 1:
Basic methods and dynamics of infectious diseases

Lecture notes

Introduction to Infectious Disease Modelling and its Applications – 2018

Session 1: Introduction to the epidemiology of infections
Lecture

This session is devoted to a broad introduction to the epidemiology of infections and infectious diseases. The handout summarises terms and concepts which arise in considering the subject. Many will appear repeatedly in subsequent sessions.

I WHAT ?

1. General characteristics of infections

a) The necessary (+/- sufficient) aetiological factors of infectious diseases are infectious agents (eg prions, viruses, mycoplasma, rickettsia, bacteria, fungi, protozoa, helminths, arthropods) which "live" on or in a larger organism, a "host".

b) Symbiosis is a general term covering all intimate associations between different types of organisms, including the association between infectious agent and host. Symbiotic associations are sometimes classified as being either mutualistic, or commensal or parasitic - parasitic associations being those in which one organism (the parasite) "benefits" at the expense of the other organism (the "host").

c) All infectious diseases may be considered as host-parasite associations. Unfortunately, just to make things confusing, many authors and institutions use the terms "parasite" and "parasitology" with reference to just 3 groups of infectious agents: protozoa, helminths, and arthropods.

2. Origins and evolution of infectious agents

a) A few infections of man may have evolved along with the human species from lower primates. Others have probably been "picked up" by human populations from non-human animal species, during historical time (eg influenza, measles, smallpox etc...) (Burnet, 1962; McNeill, 1976).

b) Infectious agents are "living" organisms, and as such undergo evolution along with their host species. There are many examples of variation and natural selection among infectious agents:

eg

1) Antigentic drift and shift of influenza viruses, responsible for periodic influenza epidemics.

2) Drug resistant mutants, such as penicillin-resistant *Gonococcus*, chloramphenicol resistant *S typhi* and chloroquine resistant *Plasmodium* spp.

3) Selection of low-virulence strains of myxomatosis virus.

c) Host species may also evolve, in response to selection pressure of infections - eg HbS in man, and perhaps the HLA system as well.

d) It is widely held that host-parasite associations tend to evolve (genetically) towards increased resistance in the host and decreased pathogenicity of the agent - towards less and less severe clinical disease (Smith, 1934). A favourite example is viral myxomatosis in rabbits in Australia (Fenner and Ratcliffe, 1965). There are many exceptions - eg infections which depend upon causing symptoms (lesions, cough) for transmission.

3. Terminological paradoxes

a) Disease due to microbial toxins produced outside the host, but ingested by the host, are generally included among the "infectious diseases" - though actual infection of the host does not occur (eg botulism, staphylococcal exotoxin food poisoning).

b) Though deep in our language, the phrase "communicable disease" is itself a misnomer, in that it is infections, not diseases, which are transmitted. The same criticism may be levelled at such phrases as "vector-borne - " or "sexually transmitted diseases". Careful epidemiologists try to avoid such terms. (thus the increasing use of “sexually transmitted infection” (STI) today

c) There has been an unfortunate tendency in recent years to speak of "chronic" as distinct from "infectious" disease epidemiology. Given the chronicity of many infectious diseases (eg leprosy, filariasis ...) and the increasing recognition of infections in the aetiology of cancers and degenerative neurological diseases, this separation is somewhat absurd and probably counter-productive.

II WHO ?

1. Characteristics of host individuals

a) Naturally resistant - not able to become infected, because of genetic constitution. (Man is naturally resistant to tobacco mosaic virus.)

b) Susceptible - uninfected, but able to become infected, if exposed.

c) Infected - implies presence of the infectious agent, in some form, within the host. Does not necessarily mean "diseased".

d) Colonised - implies presence of a microbial infectious agent on an epithelial surface of the host, but without tissue invasion.

e) Infested - often used in preference to "infected", or "colonised", for a host carrying arthropod ectoparasites - eg lice.

f) Incubating - infected, but not yet manifesting clinical signs.

g) Latent infection - (silent infection, subclinical infection) - implies presence of infectious agent but absence of clinical disease.

h) Carrier - implies a protracted infected state with shedding of the infectious agent. Carriers may be diseased, convalescent or healthy.

i) Patent infection - implies the presence of sufficient infectious agents (in the blood, tissues, stools, etc of the host) for them to be demonstrable microscopically or by culture.

j) Infectious (infective) - implies the infected individual is in a state to transmit the agent, either directly (by contagion), or indirectly (eg via an insect vector).

k) Contagious - implies that an infected individual can transmit the infectious agent, by actual contact or by aerosol, directly to a susceptible individual. All contagious individuals are infectious; but not all infectious individuals are contagious.

l) Diseased - implies the presence of clinical signs of pathology. Not synonymous with "infected".

m) Immune - possessing cell-mediated or humoral antibody protection against an infection.

n) Passive immunity - due to transfer of maternal IgG across the placenta, or to serum prophylaxis/therapy.

o) Active immunity - due to a prior sensitisation of the individual to the antigens (or similar antigens) of the infectious agent.

p) Contact - an individual who has been exposed to a source of infection and who thus may be infected.

q) Infection intensity - a term used to describe helminthic or arthropod infections, indicating the number (or load, or burden) of parasites in or on the infected individual.

r) Case - an infected individual. In some contexts this term is restricted only to those manifesting clinical signs of the infection, ie diseased individuals.

2. Characteristics of host populations

a) Host range (species) - those species or organisms which are naturally susceptible to a given infectious agent.

b) Reservoir host (species) - a species, or population, or complex of populations, which is capable of maintaining an infectious agent indefinitely.

c) Definitive host (species) - term applied to infections with agents which have an obligatory sexual phase in their life cycle (many helminths and protozoa) - the definitive host being that in which the agent undergoes sexual reproduction. (eg man is a definitive host of *Schistosoma haematobium*).

d) Intermediate host (species) - term applied to infections whose life cycle involves two different host species - the intermediate host being (formally) that in which asexual reproduction of agent occurs. (eg snails are intermediate hosts of schistosomes).

e) Vector host (species) - an invertebrate species which serves to carry an infectious agent from one to another individual of a vertebrate host species. This may be a purely mechanical ("flying pin") transport function (mechanical transmission by vectors) or the infectious agent may actually multiply and develop within the vector (biological transmission). Houseflies may be mechanical vectors of *Entamoeba histolytica*. Mosquitoes are biological vectors of malaria *Plasmodium* spp.

f) Amplifying host (species) a host species which undergoes periodic epidemics of an infection, thereby providing a large population of infectious agents which may "spill over" into other species not usually exposed. (Swine are amplifying hosts for Japanese B encephalitis virus).

g) Dead end host (species) - a species whose individuals may become infected, but which do not become functionally infectious, and thus do not transmit on the infection. (Man is, in effect, a dead end host for *Brucella* spp.).

3. Host population size requirements

Like all living species, infectious agents depend on the availability of certain environments, and certain nutrients, for their continued existence. This means a dependence on a supply of appropriate susceptible hosts. This essential "supply" can be quantified.

a) Critical population size - a theoretical minimum host population size required for the continued survival of a population of infectious agents. The critical population size varies markedly according to the infectious agent involved, the demographic structure and conditions (hygiene, etc) of the host population. The critical human population size for measles virus is approximately 500,000; but for cytomegalovirus the number may be less than 1,000.

b) Threshold population size (epidemic threshold) - a certain ("threshold") number of susceptibles must be available in order for an epidemic to occur. There has been much theoretical discussion of this point, leading to the formulation of "threshold theorems" for certain infections (Kermack and McKendrick, 1927 - 33).

c) Herd immunity - refers to that proportion of a host population which is immune to an infection. If herd immunity is high, perhaps as a consequence of an effective immunisation campaign, the population of susceptibles may be insufficient to maintain

the infectious agent (and lead to eradication, as in smallpox) or at least insufficient to support large scale epidemics (as in control of diphtheria or rubella). Implies concept that the presence of immune individuals protects those whose are not themselves immune. The herd immunity threshold to reduce incidence is $1–1/R_0$, where R_0 is the average number of transmissions per case in a susceptible population - see below (Fine, 1993).

III WHERE ?

a) Like all living organisms, infectious agents have strict environmental requirements (eg malaria transmission depends on *Anopheles* mosquitoes which depend on climatic conditions).

b) Study of the geographic distribution of infections is a branch of ecology (the study of the relationship between organisms and their environment). The niche concept is especially appropriate (eg the association of the cholera *Vibrio* with the Ganges delta). This field is sometimes called "geographic pathology" or "landscape epidemiology" (the latter term being associated with older Russian workers, Pavlovsky in particular).

c) Spatial and temporal clustering is a prominent feature of many infections, especially of common source outbreaks. Actual spatial patterns will depend on the distribution and mobility of the host population, and on the mechanism of transmission (Snow, 1854; Budd, 1874).

IV WHEN ?

1. Time course of a single infection

a) Incubation period - the time from initial infection to the onset of clinical signs. Incubation periods are variable for any infection but generally have a positively skewed frequency distribution (almost log normal). This means that a range and a median or geometric mean are the best statistics to describe incubation periods. Knowledge of the frequency distribution of incubation periods is essential for contact tracing and for quarantining of individuals exposed to infection (Sartwell, 1966).

b) Latent period - there is some disagreement over this term, but some authors have used it to describe the time from initial infection to onset of infectiousness (Hope-Simpson, 1948). For many infections, *this pre-infectious period* is slightly shorter than the incubation period. Important for contact tracing.

c) Duration of infectiousness, patency or communicability Important for isolation of cases, and assessing contact risk period. It varies greatly between infections, from a few days (eg many acute respiratory viruses) to years (eg leprosy, filariasis, many STIs...).

d) Relapse, recrudescence or reactivation of disease = recurrence of clinical signs after a period of silent (latent, subclinical) infection (as with *Plasmodium vivax* malaria, *Rickettsia prowazeki*, *Herpes simplex*, *Herpes varicella-zoster* or *Mycobacterium tuberculoisis*.

2. Time relationship between cases

a) Index case. The first *recognised case* in a defined outbreak or epidemic.

b) Primary case. The first case in a defined outbreak or epidemic (may only be recognized in retrospect.

c) Secondary cases (also tertiary, quarternary cases...etc). Secondary cases owe their infections to the primary cases...etc.

d) Serial interval The time between homologous stages (*eg onset of clinical symptoms*) of two successive cases in the chain of transmission. This provides a measure of the duration of what is sometimes called the "life cycle" of the infection. The serial interval may be greater or lesser than the incubation period, depending on whether a case becomes infectious before or after the onset of clinical signs. This depends as much on the "latent" period and duration of infectiousness as on the incubation period.

e) Generation interval. The time between *infection* of successive cases ina chain of transmission. This is rarely observed, but an important concept in modelling.

f) Time clustering - indicates that the clustered cases were exposed to a source of infection during a restricted period of time.

3. Middle-term time trends

a) Endemic - the maintenance of a relatively "constant" incidence and prevalence rate over a "long" period of time (years...). The term is also used to describe the long-term persistence of an agent within a single population (even with dramatic changes in incidence and prevalence - ie "measles is still endemic in Great Britain"). NB: "endemic" is an adjective in English.

b) Epidemic - a sharp and significant rise in incidence (and prevalence?) over the expected level for a given time period. (Both noun and adjective)

c) Outbreak - a sudden epidemic, usually of short duration. Used especially for point-source episodes for food-borne or zoonotic infections. The term is less emotive than "epidemic".

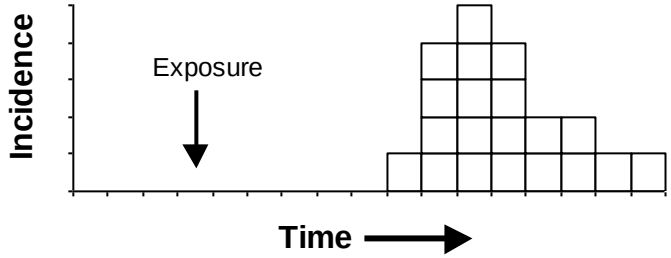
d) Pandemic - a world-wide epidemic. Classical pandemic infections are plague, cholera and influenza because of their periodic appearance and disappearance. Now add AIDS (?).

4. Epidemic curves are graphical representations of incidence (eg of clinical disease) against time.

Three of the more common shapes are as follows:

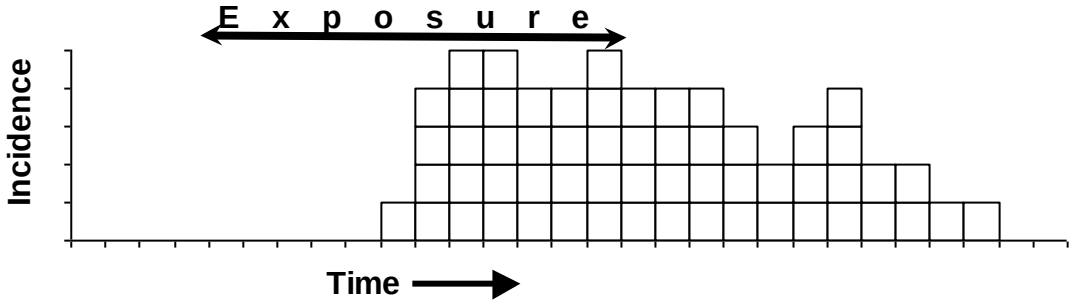
a) Common source

I Point Source



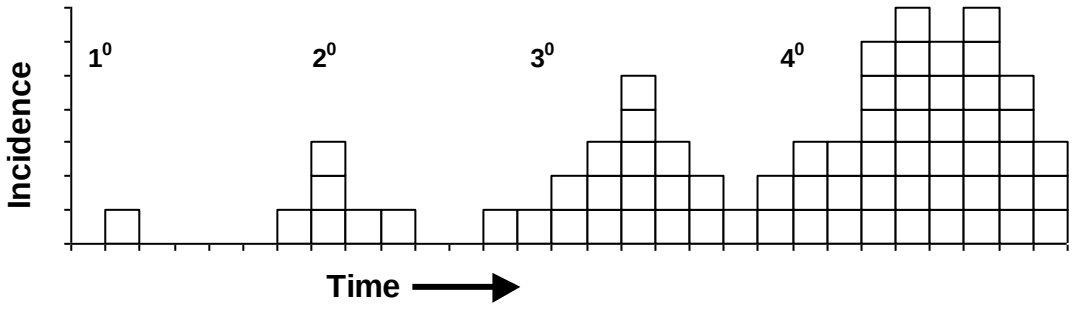
Such a "curve" is expected if all cases were exposed at a "point" in time. The shape of the curve then reflects the frequency distribution of incubation periods. (Typical of food-borne outbreaks).

II Extended source



Such a "curve" indicates the onset and continuation of exposure to an environmental source of infection.

b) Propagated source



Such a "curve" indicates successive "generations" of cases after introduction of an index case into a susceptible population. The time between successive peaks is a measure of the average serial interval. (Typical of measles or chickenpox in institutions.)

5. Cyclical phenomena

a) Seasonal cycles - eg many respiratory infections in winter or dry season, arthropod-borne and gastrointestinal infections in summer or wet season.

b) Biennial, triennial etc cycles - especially prominent with the acute contagious infections of childhood, such as measles (typically 2-year cycle) or pertussis (3-4 year cycle). Cycle period is a function of the duration of the latent and infectious periods, and the time required for sufficient susceptibles (eg children) to accumulate up to the epidemic threshold.

c) Irregular cycles - as with influenza, in which case successive pandemics are due mainly to the appearance of "new" antigenic types in the virus population.

6. Long-term secular trends - Gradual changes in incidence and prevalence over many years - for example the disappearance of leprosy from Europe, the decline of tuberculosis and the changing pattern of poliomyelitis. Such trends may be due to demographic, social, behavioural or nutritional changes in the host population, to climatic or environmental changes, or to public health intervention.

V HOW ?

1. Transmission "cycles" are qualitative descriptions of the mechanisms of transmission of infection.

TYPE	EXAMPLES
Soil reservoir ----> human	Histoplasmosis, tetanus
Animal reservoir --> human	Zoonoses, eg brucellosis, rabies
Non-human animals <----------> human	"Cyclozoonoses" - eg beef and pork tapeworms
Human <----------> human (vector mediated)	Malaria, filaria
Human <----------> human (fomites)	Cholera, typhoid
Human <----------> human (contact, airborne, venereal)	Influenza, leprosy, gonorrhea,

2. **Vertical and horizontal transmission**. These concepts have entered the epidemiology literature from the field of tumor virology. Vertical transmission implies the direct transfer of an infectious agent from a parent to his or her progeny. This includes the transmission of genetically integrated (eg viral) material. Horizontal transmission of an infection implies the direct transfer of infection from someone other than a parent (Gross, 1951).

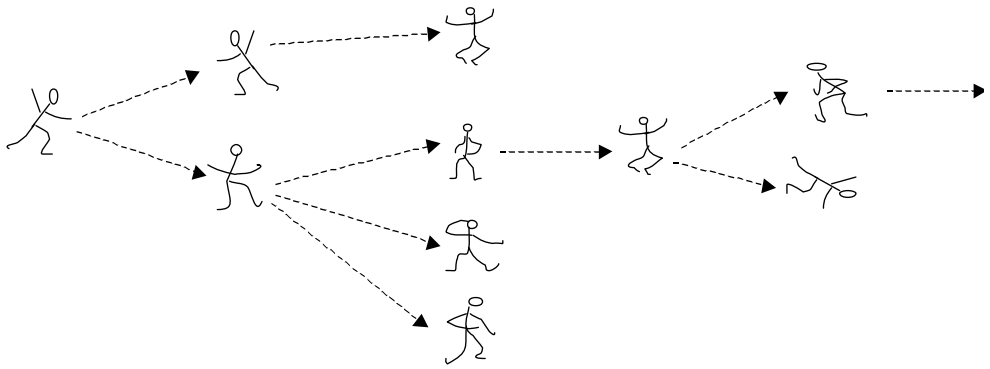
3. **"Direct" versus "indirect" contact** - There is some confusion in the literature over the definitions of these terms, and whether direct contact implies actual physical touching (as with venereal infections), or includes short-distance airborne transfer (eg by "droplets") or more distant airborne transfer (eg by "droplet-nuclei" aerosol). This has been further confused by the use of such phrases as "prolonged and intimate contact" (eg in the TB and leprosy literatures).

4. **Probability of effective contact** - This term (due originally to W H Frost) describes the probability that two members of a host population have "the sort of contact required for transmission of the infection", during a specified period of time. Such contact probabilities vary with contagiousness of the infection and with the density, mixing and hygienic conditions of the population.

5. **Attack rates** - Attack rates are actually *risks* during specified epidemic periods. A "secondary attack rate" is the incidence risk of secondary cases among all those exposed to a primary or index case (precisely defined as all cases with onset between 1 minimum and 1 maximum incubation period after exposure to a primary case). Such "secondary attack rates" are often calculated for household exposure, and provide a standard measure of the transmissibility, or contagiousness, of an infection (Hope-Simpson, 1952; Kemper, 1980).

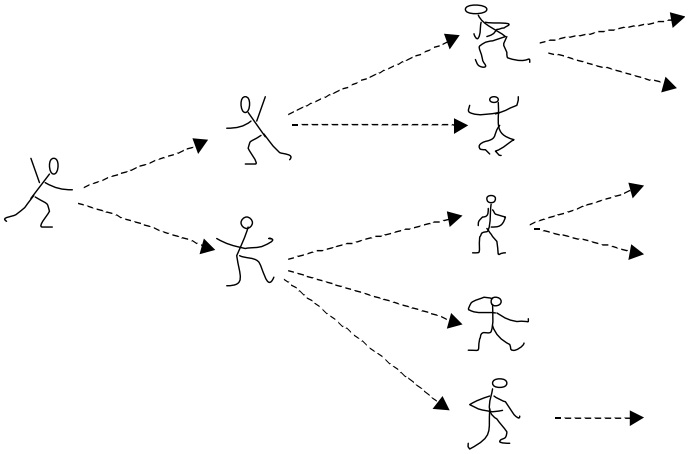
6. **Case reproduction number (R)**- this is the average number of secondary cases which are attributable to a single case of infection. Note that in an endemic situation, the "net case reproduction number" should be, on average, approximately R = 1. (In the older literature these were called reproduction “rates”)

In an endemic situation:



R = (number of transmissions)/(number of sources) = 10/10 = 1

But in the growth phase of an epidemic, the net case reproduction number must exceed 1:



R = 12/8 = 1.5

The case reproduction number should be maximum for an index case introduced into a totally susceptible population. This is sometimes called the "basic reproduction number R_0" (Macdonald, 1957). R_0 must exceed unity for an infection to persist. For a randomly mixing population, it is related to the average age at infection (A) and the (average) life expectancy (L) through the formulae: $R_0=L/A$ ("developed country" or "type I" populations) and $R_0=1+L/A$ ("developing country" or "type II" populations).

From this perspective, the object of an infectious disease control programme is to reduce the net case reproduction rate below unity, thereby causing a progressive decrease in incidence.

7. Dynamics It is intuitively reasonable that the risk of infection in a community is a function of the prevalence of infectious cases and the amount and pattern of mixing within the population. There is a considerable body of theory attempting to quantify these relationships, the most important concept being the so-called "mass action" principle as applied to epidemiology: ie that incidence is a function of the product of the number of infectious cases times the number of susceptibles (Hamer, 1906; Ross, 1911, Fine, 1979).

8. Microparasites versus **macroparasites** These terms call attention to important distinctions between infections by monerans (viruses, bacteria and protozoa, the "microparasites") and those by metazoans (helminths and arthropods, the "macroparasites"). Not only are the microparasites invisible to the naked eye, but they are typically present in such vast numbers that enumeration focuses upon numbers of infected hosts, rather than on numbers of parasites themselves. The opposites hold for the macroparasites, which are typically large enough to be seen with the naked eye, and which can themselves be counted. The distributions of macroparasites within hosts are typically highly skewed, with many hosts harbouring few and a few hosts harbouring large numbers of parasites.

9. Stability and instability - These concepts were originally developed with reference to malaria, by Macdonald (1957). They are especially applicable to chronic infections, but may be generalised to all infectious disease systems.

Stable epidemiological systems have transmission mechanisms which are far more efficient than what is required for mere maintenance of the infectious agent. These

are typically manifested by constant incidence patterns (the patterns themselves may be cyclical) and are relatively resistant to control and eradication efforts. The classical example of epidemiological stability is malaria in the holoendemic regions of equatorial Africa.

Unstable epidemiological systems have transmission mechanisms which are just sufficient to maintain the infectious agent. These systems typically manifest irregular incidence patterns and are relatively easy to control or eradicate (eg hypoendemic malaria in the Mediterranean region).

10. Molecular epidemiology - This fashionable term describes the incorporation of molecular biological techniques within epidemiological studies - for example, we are now able to recognize particular genotypic strains of infectious agents, eg by restriction fragment length polymorphism (RFLP) "fingerprinting", or whole genome sequencing and thereby to trace chains of transmission more precisely than ever before.

VI ABOUT VACCINES

Definitions

a) Vaccination - the administration (by mouth or injection) of antigenic material with intent to bring about an active immunological response in the recipient.

b) Immunisation (active) a vaccination which succeeds in bringing about a (hopefully protective) immunological response.

c) Vaccine efficacy or effectiveness (= protective efficacy, or effectiveness, of a vaccine) - defined as the percent reduction in risk, attributable to vaccination, among vaccinated individuals compared to equally exposed non-vaccinated controls. Thus

$$VE = (R_{nv} - R_v)/R_{nv}$$

where R_{nv}, R_v = risks in non-vaccinated and vaccinated groups. By current preferred usage, “vaccine efficacy” refers to the ideal as measured in a trial, whereas the term “vaccine effectiveness” is used for the analogous measure in routine programmes.

VII SO WHAT ?

What is special about the epidemiology of infections?

1. **Importance:** “Communicable diseases” are responsible for much morbidity and mortality - eg 46 % of the disability-adjusted life years lost globally in 1990 (71 % in Africa, 10 % in established market economies). This calculation (World Bank, 1993) considers cancers as "noncommunicable diseases", despite fact that almost a quarter of the cancers in the world are now known to be attributable to infections.

2. **Novelty:** Recent years have seen the appearance or recognition of new infectious agents and diseases, such as *Helicobacter*, hantavirus, HIV, a new virus for Kaposi's sarcoma,... We have not yet seen the end of the list.

3. **Understanding:** The mechanisms involved in many infectious diseases are very well understood, all the way from the genetic and molecular aspects of the agents to the genetic, cellular, ... behavioral and demographic levels of the host populations. This understanding allows a level of prediction, prevention and treatment beyond what is possible for most non-infectious diseases.

4. **Population dynamics:** Infectious agents are themselves living organisms, with their own population dynamics which depend in turn on the population dynamics of their hosts. This combined dynamics results in complex but predictable temporal patterns.

5. **Dependency:** For transmissible infections, future incidence is a function of current prevalence. Thus a case may also be a risk factor (ie a source of infection for another individual). This may have implications in statistical modelling - ie, cases may not be independent events.

6. **Immunity:** Immunological recognition of infection experience may provide a long-lasting historical record of past exposure (let alone the implications for prevention via vaccination).

7. **Elimination** and **eradication:** Our ability to reduce transmission risks and to destroy infectious agents provides opportunities for elimination (reduction to near zero) or eradication (reduction to zero) of individual infectious agents and infectious diseases. This has occurred for smallpox, is programmed for guineaworm and polio, and ....

References

Publications cited in this handout emphasise classical and historical studies. For more recent and detailed references, see texts cited in introductory pages and the reference lists for specific lectures.

1. Budd W (1874). Typhoid Fever: Its nature, mode of spread and prevention. Republished by Arno Press 1977.

2. Burnett FM. Natural History of Infectious Disease. Cambridge University Press, 1966.

3. Fenner F and Radcliffe FN. Myxomatosis. Cambridge University Press, 1965.

4. Fine PEM. John Brownlee and the measurement of infectiousness: an historical study in epidemic theory. J Roy Stat Soc A 1979; **142:** 347-362.

5. Fine PEM. Herd immunity: history, theory, practice. Epidemiologic Rev 1993; **15:** 265 - 302.

6. Frost WH. Some conceptions of epidemics in general. Am J Epid 1976; **103**: 141-151.

7. Gross L. The "vertical" transmission of mouse mammary carcinoma and chicken leukaemia. Its possible implications for human pathology. Cancer 1951; **4**: 626-633.

8. Hamer WH. Epidemic disease in England - the evidence of variability and of resistance of type. Lancet 1906; **2**: 733-739.

9. Hope-Simpson RE. The period of transmission in epidemic diseases. Lancet 1948; **2**: 755-760.

10. Hope-Simpson RE. Infectiousness of communicable diseases in the household. Lancet 1952; **2**: 549-554.

11. Kemper JT. Error sources in the evaluation of secondary attack rates. Am J Epid 1980; **112**: 457-464.

12. Kermack WO and McKendrick AG. Contributions to the mathematical theory of epidemics. Proc Roy Soc A 1927 - 33; **115**: 700-721 (part I, 1927), **138**: 55-83 (Part II, 1932), **141**: 94-122 (part III, 1933).

13. Macdonald G. The Epidemiology and control of Malaria. Oxford University Press, 1957.

14. McNeill WH. Plagues and Peoples. Anchor Press/Doubleday, 1974.

15. Nash A, Dalzeil R, Fitzgerald J. The Pathogenesis of Infectious Disease. Academic Press, 6th Edition, 1915.

16. Nelson KE, Williams CM. Infectious Disease Epidemiology. Jones and Bartlett Learning 2014

17. Panum PL. Observations Made During the Epidemic of Measles on the Faroe Islands in the Year 1846. Republished by the American Public Health Association in 1940.

18. Pavlovsky EN. Natural Nidality of Transmissible Diseases. (Translated by Plous and Levine) University of Illinois Press, 1966.

19. Ross R. The Prevention of Malaria, 2nd Ed., with addendum on the theory of happenings, London: John Murray, 1911.

20. Sartwell P. The incubation period and the dynamics of infectious disease. Amer J Epid 1966; **84**: 204-216.

21. Smith T (1934). Parasitism and Disease. Republished by Hafner in 1963.

22. Snow J (1855). On the Mode of Communication of Cholera. Reprinted as "Snow on Cholera" by Hafner in 1965

Introduction to Infectious Disease Modelling and its Applications – 2018

Session 2: Why bother with modelling?

Lecture

This introductory session will consist of a lecture delivered by Eduardo Massad (see accompanying slides) providing a general overview of why mathematical modelling is important.

Introduction to Infectious Disease Modelling and its Applications – 2018

Session 3: Basic methods for setting up models I: Difference equations

Lecture

Objectives

By the end of this lecture, you should:

- Understand the steps involved in developing models;
- Be aware of the common model structures and methods used for modelling the transmission of infections;
- Understand how deterministic models are set up using difference equations;
- Be able to write down simple models of the transmission dynamics of an infection using such equations;
- Be able to define key input parameters.

Steps in the development and use of models

Before we set up a model describing the transmission of an infection it is worth reflecting on why we are modelling it and the wider considerations that may be relevant in creating and using the model.

Figure 1 shows a comprehensive list of the steps that could be involved in developing infectious disease models (after Habbema, et al. 1996). The list is exhaustive because it was written after experience of the development of large microsimulation models for the control of infectious diseases, with the eventual aim of providing a tool for use by policy makers and programme mangers. Therefore, it may be pertinent for work you may be doing in the future. The whole process is highly iterative and the development of a large microsimulation model may take many years.

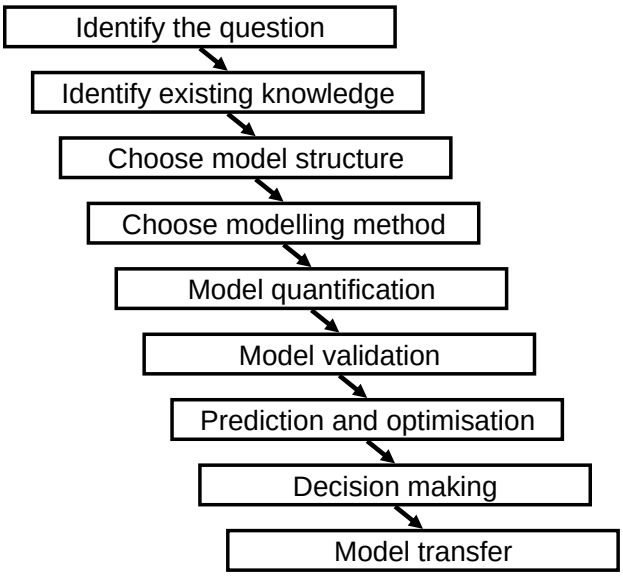
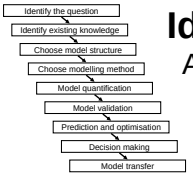


Figure 1 Steps in the development and use of a model (after Habbema et al, 1996)

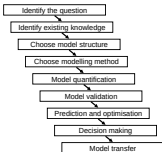
In this lecture, we will now work through the steps outlined above using a simple model of measles to illustrate the development of a '*deterministic'* model using '*difference'* equations.



Identify the Question

As for any research project, the first step is *identifying the research question*. As you saw in the previous lecture, modelling has frequently been used to get a better understanding of the natural history and epidemiology of infections. Later, when these are reasonably well understood modelling is commonly used to explore whether a particular control strategy will be effective, and if so how effective, especially in comparison with other control strategies.

Measles example: *Suppose we would like to know the following: If one measles case is introduced into a population of 100,000 susceptible individuals, how will the average number of people who are susceptible, infectious and immune to measles change over time? To keep it simple for now, we will ignore births and deaths.*



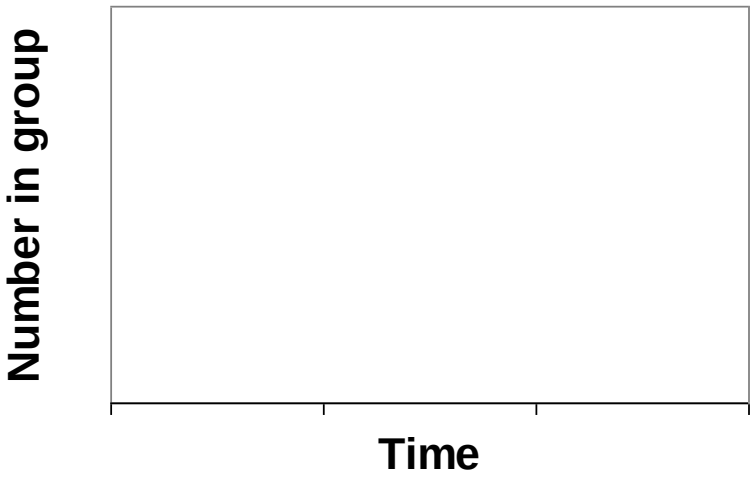
Identify existing knowledge

A *review* should follow, that is quantitatively orientated to collect data on the demographic, epidemiological and biological characteristics of the infection and the population. If control options are to be considered then data should also be collected.

Measles example: *We will need some information about measles to start thinking about what model design would be appropriate. However very few data are available:*

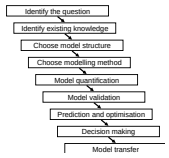
- Measles is an infection that confers near permanent immunity
- Each person effectively contacts 1.5 others per day
- The pre-infectious ("latent") period is 8 days
- The duration of infectiousness is 7 days

Draw in the following graph how you might expect the number of people who are susceptible, "pre-infectious" (infected, but not yet infectious), infectious and immune to measles to change over time after the introduction of a single infectious individual into a population of susceptibles:



It is difficult to draw an accurate graph of the way we expect the numbers in each group to vary using these few data.

A better approach is to write down equations for the number of susceptibles, infectious and immune individuals at some later time (tomorrow for example) in terms of the number now (today for example) and then to *solve* those equations for each time (today, tomorrow, the next day…).



Choose Model Structure

In choosing the *structure* of your model, the infection categories and transitions need to be described as well as important categories in the population itself. With Einstein's quote '*models should be as simple as possible and no simpler*' in mind, consider what infection categories and population categories are really necessary to address the research question.

Figure 2 shows some of the more common categories of models used for modelling infections. At this stage, it also useful to consider the format of model output to ensure that it can be compared to available data to test and validate the model later on.

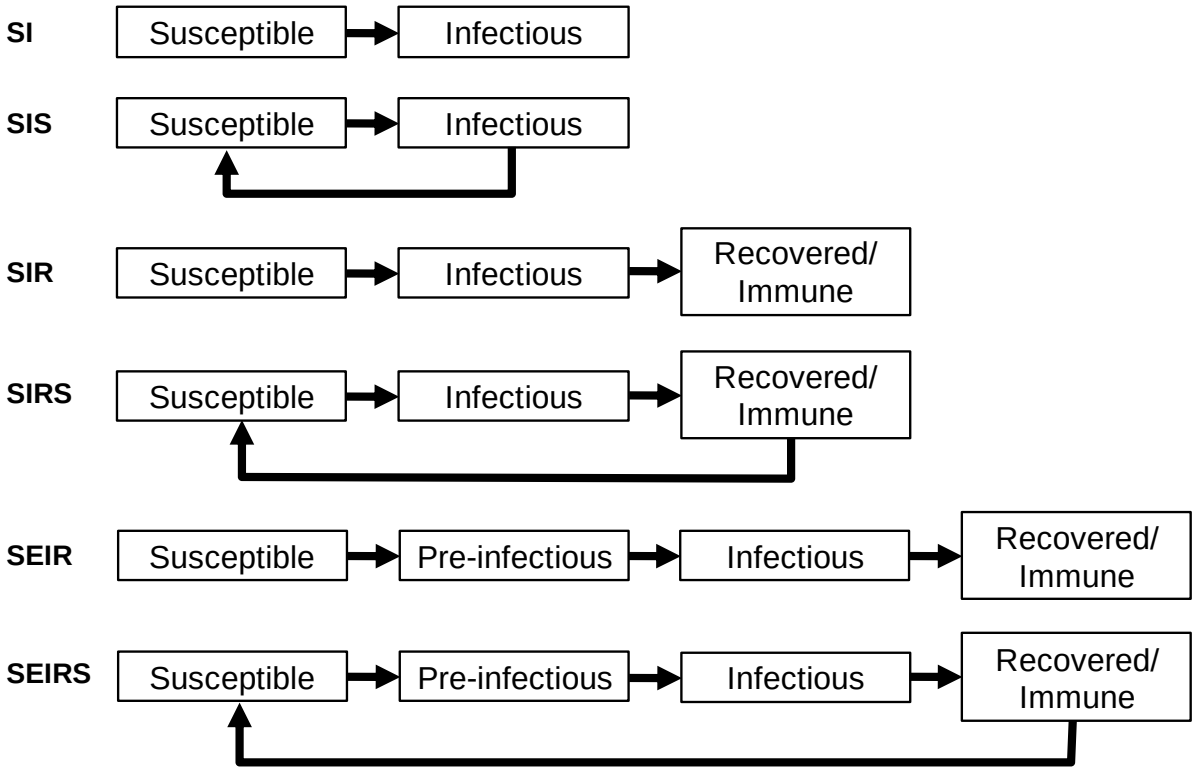


Figure 2 Common categories of infectious disease models. Note that in the modelling literature, the "Pre-infectious" individuals in SEIR and SEIRS models are frequently referred to as "Infected" or "Exposed" individuals. However, this terminology can be misleading, since all individuals can be considered to be exposed; likewise, infectious individuals can be considered to be infected. To reduce confusion, we shall refer to these individuals as being "pre-infectious" during the course.

Measles example: *Which model structure(s) shown in Figure 2 would be the most appropriate for modelling infections resulting in near permanent immunity such as measles?*

_____

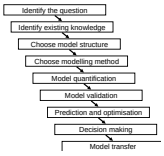
The structure you choose also depends on how accurate your estimates need to be e.g. to the nearest day, the nearest week, the nearest month etc. when compared to the time period that individuals spend in each of the categories.

For example, if the estimates need to be accurate to the nearest day, then as the pre-infectious period for measles is *long* on this time scale (8 days) the model would need to account for the fact that some individuals may be infected, but not yet infectious. In this case the *SEIR* design may be the most appropriate. Alternatively, if they need to be accurate to the nearest month, the *SIR* design may be adequate, given that newly infected individuals can be infectious within 8 days, which is much less than a month.

The model population may also need to be split into various subgroups. For example, if you needed to be able to predict the number of cases by age, then the model population would need to be split by age. In the extreme, models can be set up to model each person individually (so called 'individual-based' models).

Further, if you need to describe the long-term (rather than the short-term) dynamics then the model may need to incorporate births, deaths, possibly seasonal mixing patterns (e.g. resulting from school attendance) and in- or out-migration.

Measles example: *As the objective of our measles exercise is very simple in this case - to forecast only the short-term dynamics in a population without age-structure, the model can be kept simple. We can assume random mixing, no seasonal variation and no age-dependencies.*



Choose the type of modelling method

Models can be implemented using at least two modelling methods: *deterministic* or *stochastic* equations. These are based on difference/ differential or stochastic equations respectively. In reality models can be set up using both equation types - some largely deterministic models also incorporate stochastic elements and nearly all largely stochastic models incorporate some deterministic elements (e.g. *STDSIM*, Korenromp et al. 2000).

Stochastic modelling

Stochastic models incorporate chance variation into the transmission process and will provide a *range* of possible outcomes (Bailey 1975). They are used when chance fluctuations are important, such as for small or isolated populations, or for modelling the beginning or end of epidemics. They are used in situations in which the variance of the outcome is as important as the average behaviour. Stochastic models are often used to describe the infection process for every individual or even every (macro-) parasite in a population.

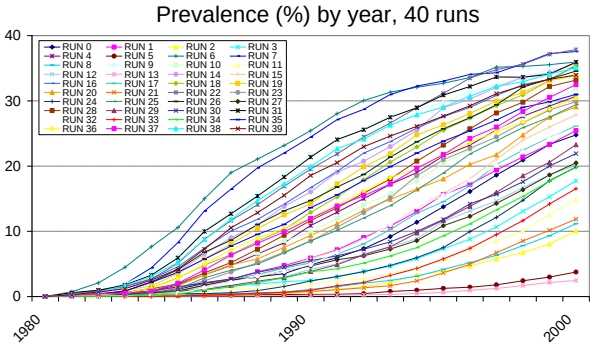


Figure 3 Example of the output from a model using stochastic equations

Deterministic modelling

Most deterministic models are so-called "compartmental" models. The term 'compartmental' comes from the fact that the model population is stratified into broad subgroups (compartments), such as those who are susceptible, pre-infectious, infectious or immune Deterministic models aim to describe what happens "on average" in a population. These describe the transitions between categories by applying average transition rates.

Most of the models in the literature have been deterministic, largely because they are relatively easy to set up and for many purposes knowing what happens on average is sufficient.

Note, don't confuse, model equation method (eg stochastic or deterministic) with the level / granularity at which we model structure in the model (eg compartmental or individual-level).

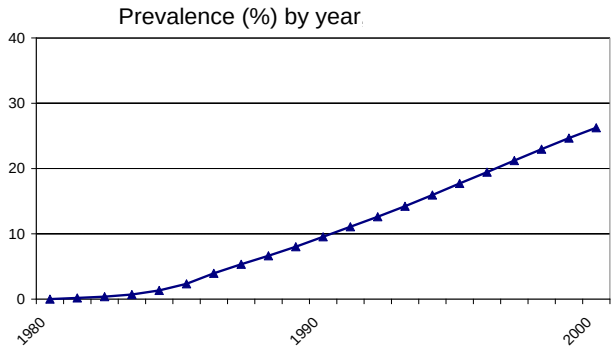


Figure 4 Example of the output from a model using deterministic equations

Measles example: *Which model type, stochastic or deterministic would be the most appropriate for our research question?*

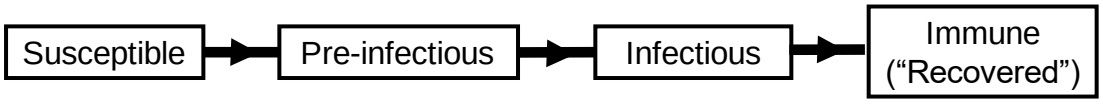
Can you think of any reasons why the other type should be used?

Deterministic models can be set up using *difference* or *differential* equations. The distinction between the two will become clear later in this and next lecture. For now, the main difference between the two lies in the time step used in the modelling. Difference equations use a time step that is an appreciable period of time, such as a day or a year, whereas differential equations use vanishingly small time steps (continuous time). In this lecture and the following practical, we will use difference equations. You will cover the differential equations approach in the next lecture and practical.

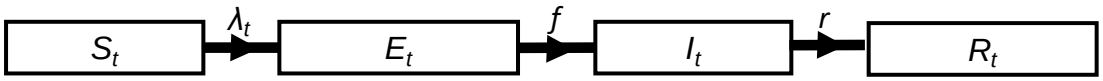
Deterministic modelling using difference equations

Difference equations aim to describe the transitions between different disease categories using discrete time steps (daily, for example) by expressing the number of cases at a given time t+1 (tomorrow, for example), in terms of that at an earlier time t, (e.g. today).

For our measles example, we have decided to use the SEIR model structure:



Adding symbols so that the boxes and arrows are easier to refer to, this becomes:



Using the following notation, we will now write down equations describing how the number of susceptible, pre-infectious, infectious and recovered/ immune individuals change over time.

***S*$_t$** for the number susceptible at time *t;*
***E*$_t$** for the number pre-infectious (i.e. infected but not yet infectious) at time *t;*
***I*$_t$** for the number infectious at time *t;*
***R*$_t$** for the number who are recovered (i.e. are immune) at time *t*;
λ_t is the *incidence risk* or the risk of a susceptible individual becoming infected between *t* and *t*+1. For historical reasons it is also called the *force of infection* by modellers (this is discussed in later sessions);
f is the risk of a pre-infectious individual becoming infectious between time *t* and *t*+1;
r is the risk of an infectious individual of recovering (become immune) between time *t* and *t*+1.

Susceptible

The number who are susceptible at time *t*+1 (S_{t+1}) is given by:

S_{t+1} = the number who were susceptible at time *t* (S_t)
- the number newly infected between time *t* and *t*+1

The number newly infected between time *t* and *t*+1 is the product of the risk of infection per unit time (the force of infection) λ_t and the number at risk of becoming infected (S_t)

the number newly infected between time *t* and *t*+1 = $\lambda_t * S_t$

So, substituting this equation into the expression for S_{t+1} above, gives

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

Pre-infectious

Similarly, the number in the pre-infectious category at time *t*+1 (E_{t+1}) is given by:

E_{t+1} = the number who were in the pre-infectious category at time *t* (E_t)
+ the number newly infected between time *t* and *t*+1 ($\lambda_t * S_t$)
- the number who became infectious between time *t* and *t*+1

The second term must equal the number leaving the susceptible category ($\lambda_t * S_t$), but as they are entering the pre-infectious category, the sign in front of this term is '+', rather than '-'.

The last term in this expression is just given by the product of the risk of becoming infectious between time *t* and *t*+1 and the number who are at risk of becoming infectious at time *t* (E_t) = $f * E_t$,

The number in the pre-infectious category at time t+1 is given by:

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

Infectious

Similarly, the number who are infectious at time t+1 (I_{t+1}) is given by:

I_{t+1} = the number who were infectious at time t (I_t)

+ the number who became infectious between time t and t+1 ($f * E_t$)
- the number who recovered (became immune) between time t and t+1 ($r * I_t$)

So,

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

Recovered/ immune

Finally, the total number of individuals who are immune at the time t+1 (R_{t+1}) is given by:

R_{t+1} = the number who were immune at time t (R_t)
+ the number who became immune between time t and t+1 ($r * I_t$)

So,

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

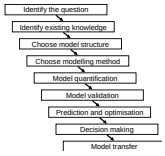
In summary:

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

These equations provide a good method for predicting the numbers of individuals who are susceptible, pre-infectious, infectious and immune over time. For example, if we know how many individuals are susceptible, pre-infectious, infectious and immune at the start (e.g. on day 0) and if we know the values for λ_t, f and r, then we can calculate the numbers of individuals in these categories on day 1. We can then substitute the values that we calculated for day 1 into the above equations to obtain the numbers of individuals who are susceptible, pre-infectious, infectious and immune on day 2 and so on. This is can be done easily using a spreadsheet as we shall see in the following practical.

The reliability of the difference equation method depends on the time step used. For example, the number of susceptibles newly infected between time t and t+1 (given by $\lambda_t S_t$) is expressed in terms of the number susceptible or infectious at time t, so if the time step used is large (3 days for example), this expression may under- or overestimate the true number newly infected. More reliable estimates may then be obtained using smaller time steps, so we allow values to change more often, for example twice a day t+½ , t+1, t+1½… etc.

Taken to the limit, the differential equation approach describes the transmission dynamics in "continuous time" (i.e. vanishingly small time steps δt). This is covered in the next lecture.



Model Quantification

In parallel with the model development, the rates and distributions of the flows between the model states (susceptible, infectious etc.) need to be *quantified*. The main problem at this stage is often the lack of reliable data. Sometimes primary data collection, data analysis (statistical models) or other modelling exercises will be required to quantify the value of specific parameters. Parameter estimation is also the subject of forthcoming lectures.

Specifying the input parameters

To solve these difference equations, we need to estimate the transition parameters. In the simple example described above, we need to estimate:

- λ_t – the risk that a susceptible becomes infected between time t and t+1 (the incidence risk or force of infection)
- $\boldsymbol{f}$ - the proportion of pre-infectious individuals who become infectious between time t and t+1, and
- $\boldsymbol{r}$ - the proportion of infectious individuals who recover (i.e. become immune) between time t and t+1

Calculating the risk (or force) of infection λ_t

The risk that a susceptible individual becomes infected during a given time step depends on two factors:

1. The number of infectious individuals present in the population at that time, and
2. The rate at which the susceptible individual comes into *effective contact* with an infectious individual. An *effective contact* is defined, as by Abbey (1952), as one sufficient to lead to infection if it occurs between a susceptible and an infectious individual.

The simplest assumption that we could make about contact is that individuals mix randomly i.e. the probability that two individuals meet or have contact is the same, regardless of their age or other characteristics. A simple method for writing down this assumption is to use the "mass-action" equation, i.e. to say that the risk of infection is proportional to the number of infectious individuals as follows:

$$\lambda_t = \beta^* I_t$$

The precise definition of β is "the *per capita* rate at which two specific individuals come into effective contact per unit time"; however, for simplicity, we shall refer to it as the "rate at which two specific individuals come into effective contact per unit time".

We therefore need to estimate the rate at which a specific infectious and susceptible individual come into effective contact between time *t* and *t*+1 (*ß)*. Of all the above parameters, *ß* is the most difficult to estimate (and to conceptualize). Methods for estimating *ß* are shown in later sessions. For now, we will use an approximation for *ß* that relates it to the following two quantities:

a) The **effective contact rate (*ecr),*** the number of individuals that each person 'effectively contacts' per unit time, and the population size, ***N***
b) The ***basic reproduction number (R_0),*** the number of secondary infectious persons that would result from the introduction of an infectious person into a completely susceptible population, the duration of infectiousness (***D****)* and the effective contact rate (***ecr***)*.*

Considering
(a)
If in a hypothetical population, of size ***N,*** each person 'effectively contacts' ***ecr*** other people per unit time, then the *per capita* rate at which the two specific people will come into effective contact per time step (i.e. *ß)*, must be

$$ß = ecr / N$$

For example, in a population with 12 individuals in which each individual effectively contacts two others per day, the ***per capita*** rate at which two specific individuals come into effective contact per unit time is the number of effective contacts per day (2) divided by the population size, i.e. 12, so:

$$ß = ecr / N$$
$$= 2 / 12 \text{ per day}$$
$$\sim 17\% \text{ per person per day}$$

(b)
R_0 is the number of secondary infectious persons that would result from an infectious person, during the entire infectious period, when introduced into a totally susceptible population. Therefore for infections for which all of those infected become infectious, it is related to the average effective contact rate, the number of individuals effectively contacted by each person per unit time (***ecr***) and the duration of the infectious period (***D***), as follows:

$$R_0 = ecr * D$$

After rearranging this expression, we obtain the following:

$$ecr = R_0 / D$$

For example, if the number of secondary infectious persons that would result from the introduction of an infectious person during the entire infectious period when they are introduced into a totally susceptible population (R_0) is 14, and the duration of the infectious period *D* is 7 days then the average number of individuals effectively contacted by each person during each day (*ecr*) is:

$$ecr = R_0 / D$$
$$= 14 \text{ infectious persons/ } 7 \text{ days}$$

= 2 effective contacts during each day

Substituting the expression ***ecr = R_0 / D*** into the equation ***ß = ecr / N***, gives the expression we need for *ß*, using parameters that we can estimate (R_0, *N* and *D*):

ß = R_0 / (N * D)

Consequently, if we did not know *ecr,* but we did know R_0 = 14 and *D* = 7 days, then for the population of 12 individuals considered above, we can calculate the rate at which an infectious and a susceptible individual come into effective contact each day (*ß*) to be

ß = R_0 / (N * D)
= 14 / (12 * 7)
~ 17 % per person per day

<u>Estimating transition parameters (rates vs. risks)</u>

Technically, the transition parameters which are used in difference equations should be transition *risks*, i.e. the number of pre-infectious individuals at time t who become infectious by time *t*+1 is given by the number in the pre-infectious category at time *t* multiplied by the *proportion* who become infectious between time *t* and *t*+1. However, if the rate is small (<10%), then risks and rates are approximately the same (see the appendix for the derivation).

As a result, difference equations tend to use transition *rates* rather than *risks*, since they are relatively easy to calculate from data on duration of an event (or average time to an event), as shown below. In general, the rate at which an event occurs can be derived from the average time to an event using the expression (see the appendix for derivation):

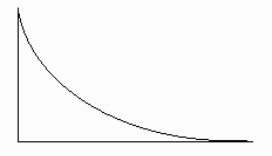
Rate at which the event occurs = 1 / (average time to event)

The rate at which individuals become infectious can thus be derived from the average pre-infectious period, the recovery rate can be obtained from the average infectious period, and the mortality rates can be derived from the average '*period of life*' (life expectancy), as follows:

Rate of onset of infectiousness = 1 / (average pre-infectious period)

Recovery rate = 1 / (average duration of infectiousness)

Mortality rate = 1 / (average life expectancy)



The above equations assume that the distributions of the time until each event occurs follow the exponential distribution, as shown on the right. Though this assumption may not be entirely realistic, it is generally adequate for most modelling studies, especially those looking at the long-term dynamics of acute immunizing infections. Illustrations of where this assumption may be inappropriate are provided later in the course.

When defining the model in practice, you should ensure that the units are consistent with the size of the time step. Therefore, if the equations use *1-day* time steps, then the rate of onset of infectiousness should be a *daily* rate.

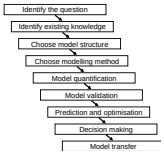
For example, the average pre-infectious period for measles is 8 days, the rate of onset of infectiousness is:

Rate of onset of infectiousness = 1 / (pre-infectious period)
= 1 / 8 per day, or 7 / 8 per week

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

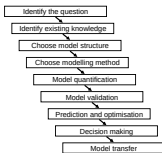
Mortality rate = 1 / (average life expectancy)
= 1 / (60 years)
= 1 / 60th per year
= 1 / (60 * 365) per day

In the practical after this lecture, you will be consolidating these ideas by calculating the input parameters and setting up difference equations in Excel.



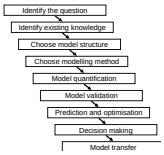
Model Validation

A fully quantified model is not likely to be a valid model. Thus *model validation* involves checking the model outputs against independent data sets, for example in and model of heterosexual HIV transmission you could check the sexual behaviour you are modelling is consistent with the data you have. At this stage, the model outputs should also be exposed to criticism from experts in the field.



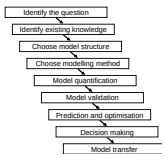
Prediction and optimisation

Assessment of the effectiveness of alternate control strategies should only occur after model validation. *Prediction* is likely to depend critically on the model assumptions, some of which may not be accurately known. These assumptions can be tested in a *sensitivity analysis* during which the model is run with different plausible combinations of parameters. For example, if you are comparing the effectiveness of various control strategies, it is likely that a sensitivity analysis will provide a range in the predicted effectiveness of the various strategies. However if one strategy is always the most effective in all sensitivity analysis scenarios, you could report this with more confidence, even thought the absolute level of effectiveness would be uncertain.



Decision Making

Predictions that are actually useful in *decision-making* should be understandable to field workers and managers. Thus, a purely technical-mathematical description of the model will not be sufficient.



Model Transfer

If the model design and quantification are not likely to change in the near future the model can be transferred to managers of control programmes. For this to be successful, a user-friendly interface will be required along with user manuals.

Further reading

Vynnycky E and White RG (2010) An introduction to infectious disease modelling. Oxford University Press. Chapter 2.

Bibliography

Abbey H (1952 An examination of the Reed-Frost theory of epidemics *Human Biology*, 1952;24:201-2339

Anderson, R. and R. May (1991). Infectious Diseases of Humans: Dynamics and Control. Oxford, Oxford University Press. Chapters 1 and 2)

Bailey, N. T. J. (1975). "The mathematical theory of infectious diseases and its applications. 2nd ed.".

Boily, M.-C. and B. Masse (1997). "Mathematical models of disease transmission: A precious tool of the study of sexually transmitted diseases." Canadian Journal of Public Health **88**: 255-265.

Habbema, J. D. F., et al. (1996). "The microsimulation approach to epidemiologic modeling of helminthic infections, with special reference to schisosomiasis." Am. J. Trop. Med. Hyg. **55**(5): 165-169.

Korenromp, E. L., C. Van Vliet, et al. (2000). "HIV spread and partnership reduction for different patterns of sexual behaviour - a study with the microsimulation model STDSIM." Mathematical Population Studies **8(2)**: 135-173.

Appendix

The following provides the methods for deriving some of the results mentioned in the lecture.

Derivation of the relationship between risks and rates

Suppose that uninfected individuals are infected at a rate λ per unit time and that there are S(0) susceptible individuals at the start. As you will see in the differential equations lecture, the number of susceptible individuals who are infected by time t is given by:

$$S(0)(1-e^{-\lambda t})$$

Thus the proportion of individuals who were susceptible at the start who are infected by time t (i.e. the risk of infection by time t) is given by

$$S(0)(1 - e^{-\lambda t})/S(0)$$
$$= 1 - e^{-\lambda t}$$

To obtain the risk of infection per unit time, substitute t=1 into this expression. This leads to the expression:

$$risk\ per\ unit\ time = 1 - e^{-\lambda} \qquad (D)$$

As discussed in the maths refresher, $e^{-\lambda} = (1 - \lambda + \lambda^2/2! - \lambda^3/3! + \lambda^4/4! - + (-\lambda)^n/n! + ...)$

If the rate λ is sufficiently small (e.g. <0.1), then the terms $\lambda^2/2!$, $\lambda^3/3!$, $\lambda^4/4!$ etc contribute a negligible amount to the expression $e^{-\lambda}$ and hence

$$e^{-\lambda} \approx 1 - \lambda$$

Substituting this approximation into expression (D) for the risk of infection/per unit time, we see that:

$$\text{risk} = 1- e^{-\lambda} \approx 1- (1 - \lambda) = \lambda$$

Derivation of relationship between (for example) the rate of recovery from being infectious and the average duration of infectiousness

Suppose r is the rate at which individuals recover from being infectious and there are *I(0)* infectious individuals at the start. We provide 2 illustrations of how you derive this relationship.

The most rigourous proof uses the logic described in the lecture, namely that the number of infectious individuals at time t, I(t), is given by $I(0)e^{-rt}$. The average duration of infectiousness is given by the expression:

$$\int I(0)e^{-rt}rtdt \;/\; \int I(0)e^{-rt}rdt$$

which can be shown to equal 1/r.

An alternative "proof" uses the fact that the proportion of individuals who recover per unit time is approximately given by the recovery rate. The average duration of infectiousness is then given by:

Suppose that r is the rate at which infectious individuals recover, and I(0) is the number who are infectious at the start. The average duration of infectiousness (or the average time until individuals recover) is given by the expression:

$$= \frac{\sum_t t \times \text{Number who recover on day t}}{\sum_t \text{Number who recover on day t}} \qquad \text{(equation A1)}$$

Thus, the number who recover on day t

= (No infectious until t-1)*(risk of recovering on day t)

= (Number who escaped recovery for t-1 days)* risk of recovery on day t

This can be expressed as: $I(0)\ (1-r)^{t-1}\ r$ (here, we're using the terms risks and rates interchangeably, since we're assuming that the rate is small and therefore rate ≈risk).

After substituting the expression for the number who recover on day t into equation A1, we see that the expression for the average duration of infectiousness becomes:

$$= \frac{\sum_t t\ r(1-r)^{t-1}}{\sum_t r(1-r)^{t-1}} = \frac{1+2(1-r)+3(1-r)^2+4(1-r)^3+\ldots}{1+(1-r)+(1-r)^2+(1-r)^3+(1-r)^4+\ldots} \qquad \text{(equation A2)}$$

Note (see derivation below) that the denominator of equation A2 is given by the expression 1/r and the numerator of equation A2 is given by $1/r^2$. After substituting in these expressions for the denominator and the numerator back into equation A2, we see that the average time until individuals recover is given by $(1/r^2)\div(1/r) = 1/r$

Proof that the denominator of equation A2 is given by 1/r

Note that the denominator of equation A2 is

$$1+(1-r)+(1-r)^2+(1-r)^3+(1-r)^4$$

We know that in general, the sum to the sum to infinity ($S_\infty(x)$) of an expression

$$S_\infty(x) = 1 + x + x^2 + x^3 + x^4 \ldots\ldots.$$

is given by $1/(1-x)$ (equation A3),

In this expression, x can be anything (even another expression, such as 1-r). So if we substitute for x=1-r into equation A3, we see that

$$S_\infty(1-r) = 1 + (1-r) + (1-r)^2 + (1-r)^3 + (1-r)^4 \ldots\ldots = (1/(1-(1-r)) = 1/r$$

Note that the highlighted part of the equation above is identical to the denominator of equation A2, so this is saying that the denominator of equation A2 equals 1/r.

Proof that the numerator of equation A2 is given by $1/r^2$

Note that the numerator of equation A2 is

$$1+2(1-r)+3(1-r)^2+4(1-r)^4$$

We know that in general (see below), the sum to infinity of an expression

$$1 + 2x + 3x^2 + 4x^3 + 5x^4 \ldots\ldots = 1/(1-x)^2. \quad \text{(equation A4)}$$

Again, x can be anything here, even another expression (e.g. 1-r).

If we substitute for x=1-r into equation A4, we see that:

$$1 + 2(1-r) + 3(1-r)^2 + 4(1-r)^3 + 5(1-r)^4 + \ldots = 1/(1-(1-r))^2 \ldots = 1/r^2 \ldots . \text{(equation A5)}$$

But the highlighted part of equation A5 is the same as the numerator of equation A2, so this is saying that the numerator of equation A2 equals $1/r^2$.

Proof that the infinite sum $1 + 2x + 3x^2 + 4x^3 + 5x^4 \ldots\ldots = 1/(1-x)^2$

Note that the expression $1 + 2x + 3x^2 + 4x^3 + 5x^4 + \ldots\ldots$

is just the derivative of the expression $1 + x + x^2 + x^3 + x^4 \ldots\ldots$ (result X)

However, we know that $1 + x + x^2 + x^3 + x^4 \ldots\ldots = 1/(1-x)$.

Given that $1 + 2x + 3x^2 + 4x^3 + 5x^4 + \ldots\ldots$ is the derivative of the expression $1 + x + x^2 + x^3 + x^4 \ldots\ldots$, and that $1 + x + x^2 + x^3 + x^4 \ldots\ldots =$

$1/(1-x)$, it follows that $1 + 2x + 3x^2 + 4x^3 + 5x^4 +$ …must equal the derivative of $1/(1-x)$.

It can be shown that the derivative of $1/(1-x)$ is $1/(1-x)^2$. Hence $1 + 2x + 3x^2 + 4x^3 + 5x^4 + \ldots$ must also equal $1/(1-x)^2$.

Introduction to Infectious Disease Modelling and its Applications – 2018

Session 4: Basic methods for setting up models II: differential equations

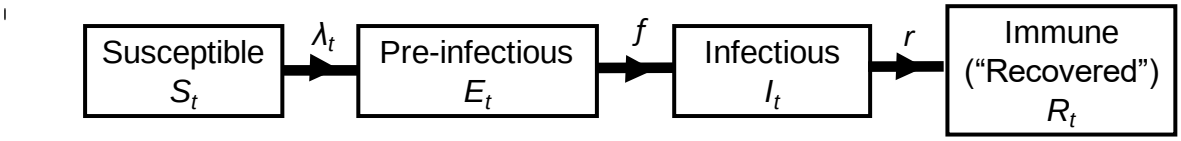
Lecture

Objectives

By the end of this lecture you should:

- Understand how models of the transmission dynamics of an infection are set up using differential equations;
- Be able to write down simple models of the transmission dynamics of an infection using differential equations;
- Understand the relationship between difference and differential equations;
- Understand the key input parameters which go into differential equations

Summary of the difference equation model of measles



In the last lecture, we saw how we could set up the above model of the transmission dynamics of a simple immunizing infection, such as measles, in a closed population using the following *difference* equations:

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

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

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

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

By assuming that individuals mix randomly, ie that $\lambda_t = \beta * I_t$, the first two equations can be rewritten as:

$$S_{t+1} = S_t - ß I_t S_t$$

$$E_{t+1} = E_t + ß I_t S_t - fE_t$$

where

- S_t , E_t , I_t and R_t are the number of susceptible, pre-infectious (infected, but not infectious), infectious and immune (recovered) individuals respectively at time *t*,
- λ_t is the risk that a susceptible becomes infected between time t and *t*+1
- *ß* is the rate at which a specific infectious and susceptible individual come into effective contact per unit time,
- *f* is the proportion of pre-infectious individuals who become infectious between time *t* and *t*+1, and
- *r* is the proportion of infectious individuals who recover (become immune) between time *t* and *t*+1.

Model predictions for the evolution of the measles epidemic

These difference equations express the number of individuals in a given category at a time t+1 in terms of the number present in that category at a time t and the number who leave that category between time t and t+1.

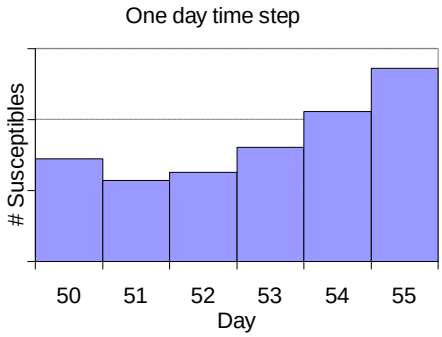
We successfully used these equations to predict the evolution of the measles epidemic. However, as you saw in the practical, the *reliability* of models based on difference equations depends on the *size of the time step used*. For example, in the measles model used in the practical relatively small (daily) time steps were needed for the model to predict the familiar cycles in infection incidence observed in many populations, as shown in the figure below.

Using a one-day time step

The following shows the model predictions for the 50th–55th days after the introduction of an infectious case into the population, if a 1 day time step were used.

Time step = 1 day; $\beta = 1.86\times10^{-5}$ per day.

Day	Number of individuals who are:		
	Susceptible	Infectious	Newly infected by the end of the current time step ($\beta\times S_t\times I_t$)
50	9.45	24,044	4.22
51	9.14	22,381	3.80
52	9.26	20,735	3.56
53	9.61	19,131	3.41
54	10.11	17,586	3.30
55	10.72	16,114	3.21



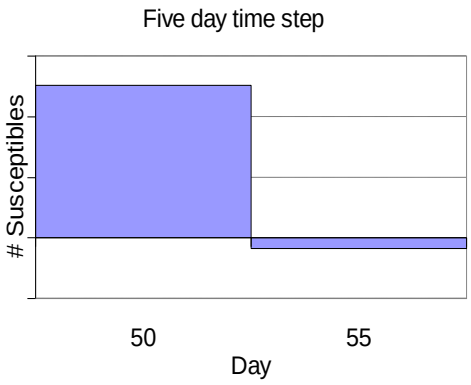
In this situation, the total number of individuals who are newly infected between 50th and 55th day was 18.29 (i.e. 4.22+3.80+3.56+3.46+3.30).

Using a five-day time step

When the time step was larger, 5 days for example, the model predicted "nonsense" results for the 50th and 55th day, as shown in table and the figure below.

Time step = 5 days; $\beta = 9.29\times10^{-5}$ per 5 days.

Day	Number of individuals who are:		
	Susceptible	Infectious	Newly infected by the end of the current time step ($\beta\times S_t\times I_t$)
50	50387	11,530	53,948
55	-3,551	24,265	-8,002



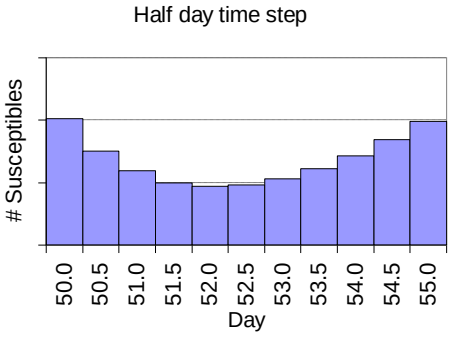
In this instance, the model predicted that 53,948 individuals are newly infected between 50th and 55th day. This number is greater than the number of susceptibles present in the population on the 50th day (50,387), and it predicted a negative number of susceptibles on day 55 which is clearly impossible.

General rule

As a general rule, the larger the time step used in the difference equations, the greater the inaccuracy in the model predictions. In the above example, the accuracy improves even more if the size of the time step is reduced further e.g. to 0.5 days:

Time step = ½ day;

Day	Number of individuals who are:		
	Susceptible	Infectious	Newly infected by the end of the current time step ($\beta \times S_t \times I_t$)
50	13.02	20,402	2.47
50.5	12.51	19,644	2.28
51	12.19	18,896	2.14
51.5	12.00	18,161	2.02
52	11.94	17,440	1.93
52.5	11.96	16,735	1.86
53	12.06	16,046	1.80
53.5	12.22	15,375	1.74
54	12.43	14,723	1.70
54.5	12.69	14,089	1.66
55	12.98	13,474	1.62



In this instance, 21.22 (=2.47+2.28+2.14+ .. +1.66) individuals are predicted to be newly infected between 50th and 55th day.

As illustrated in the above example, the most precise description of the transmission dynamics is obtained using time steps which are as small as possible. If we extend this idea to its logical limit, the most accurate description of the transmission dynamics is obtained using time steps that become vanishing small, or put another way, modelling time flowing *continuously,* rather than in distinct time step.

Simplifying the equations: Using λ_t rather than βI_t

From now on in this lecture, to keep the equations as simple as possible we will be writing the equations in terms of the *force of infection* at time *t,* λ_t, rather than β and I_t. Remember if our population mixes randomly, $\lambda_t = \beta * I_t$.

To describe the transmission dynamics of an infection in which individuals move between categories *continuous*ly, we need to use *differential* rather than *difference* equations. The equations for the model of the transmission dynamics of measles described above are very similar to the difference equations used earlier.

Importantly, *differential* equations describe the *rate of change of the number* of individuals in each category at time t, while *difference* equations describe the total *number* of individuals in each category at time t.

The following summarizes the two sets of equations for comparison:

Difference equations	**Differential equations**
Describing the *number* of susceptibles, pre-infectious… individuals at time *t*	Describing the *rate of change in the number* of susceptibles, pre-infectious… individuals at time *t*
$S_{t+1} = S_t - \lambda_t S_t$	$\frac{dS(t)}{dt} = -\lambda(t)S(t)$
$E_{t+1} = E_t + \lambda_t S_t - fE_t$	$\frac{dE(t)}{dt} = \lambda(t)S(t) - fE(t)$
$I_{t+1} = I_t + f E_t - r I_t$	$\frac{dI(t)}{dt} = fE(t) - rI(t)$
$R_{t+1} = R_t + r I_t$	$\frac{dR(t)}{dt} = rI(t)$

Where:

Difference		**Differential**	
S_t	the *number* of susceptible individuals at time *t*	$\frac{dS(t)}{dt}$	The *rate of change in the number* of susceptible individuals at time *t*
E_t	the *number* of pre-infectious individuals at time *t*	$\frac{dE(t)}{dt}$	The *rate of change in the number* of pre-infectious individuals at time *t*
I_t	the *number* of infectious individuals at time *t*	$\frac{dI(t)}{dt}$	The *rate of change in the number* of infectious individuals at time *t*
R_t	the *number* of recovered (immune) individuals at time *t*	$\frac{dR(t)}{dt}$	The *rate of change in the number* of recovered (immune) individuals at time *t*
λ_t	the *risk* of a susceptible individual becoming infected between *t* and *t*+1	$\lambda(t)$	the *rate* at which susceptible individuals becoming infected per unit time, at time t
f	the *risk* of an individual in the pre-infectious category becoming infectious between *t* and *t*+1	*f*	the *rate* at which individuals in the pre-infectious category become infectious per unit time
r	the *risk* of an infectious individual recovering between *t* and *t*+1	*r*	the *rate* at which infectious individuals recover (become immune) per unit time

When we use *differential* equations, the symbol for time '*t'* is enclosed in brackets *(t).* This is to highlight that we are describing time flowing continuously. For example the number of susceptible, pre-infectious, infectious and immune individuals at exact time *t* are written as

S(t), E(t), I(t) and *R(t)* when writing *differential* equations, but as S_t, E_t, I_t and R_t when writing *difference* equations.

Note also that the definitions of λ, *f* and *r* used in the *differential* equations differ slightly from their definitions in the *difference* equations. In difference equations we should use risks, and in differential we should use rates. However, as was shown in the appendix of the last lecture, risks and rates are numerically similar if they are small (typically less than about 10%).

What are differential equations?

Differential equations describe the rate of change of a given quantity relative to something else, for example the rate of change of the number of susceptibles over time. To describe the transmission dynamics of measles over time, we would need to write down the expressions for the rate of change in the number of susceptible, pre-infectious, infectious and immune individuals over time.

The notation used for the rate of change in the number of susceptibles over time is $\frac{dS(t)}{dt}$; the notation for the rate of change in the number of pre-infectious, infectious and immune individuals is $\frac{dE(t)}{dt}$, $\frac{dI(t)}{dt}$ and $\frac{dR(t)}{dt}$ respectively.

The expression for the rate of change e.g. in the number of susceptible, over time can be written down either using an *intuitive* argument or using the more formal *mathematical* argument.

We first use the intuitive argument.

1. Intuitive derivation of the differential equation expressions describing the transmission dynamics of measles

We begin by noting that the *rate of change* in the number of individuals in a given category over time is given by:

= the number who *enter* the category per unit time
- the number who *exit* the category per unit time

As no-one enters the susceptible group, the **rate of change in the number of *susceptible* individuals over time** is given by:

$\frac{dS(t)}{dt}$ = - the number of susceptible individuals who become infected per unit time

The *rate of change* in the number of susceptible individuals over time can be written as:

$$\frac{dS(t)}{dt} = -\lambda(t)S(t)$$

The **rate of change in the number of *pre-infectious*** individuals over time is given by:

$\frac{dE(t)}{dt}$ = the number of susceptible individuals who are infected per unit time

- the number of pre-infectious individuals who become infectious per unit time

The latter at a given time t is given by *fE(t)*, and thus the *rate of change* in the number of pre-infectious individuals over time is then written as:

$$\frac{dE(t)}{dt} = \lambda(t)S(t) - fE(t)$$

The **rate of change in the number of *Infectious* individuals** over time is given by:

$\frac{dI(t)}{dt}$ = the number of pre-infectious individuals who become infectious per unit time

- the number of infectious individuals who recover (become immune) per unit time

The latter at a given time t is given by *rI(t)*, and thus the *rate of change* in the number of infectious individuals over time is then written as:

$$\frac{dI(t)}{dt} = fE(t) - rI(t)$$

The **rate of change in the number of *Immune* individuals** over time is given by:

$\frac{\mathrm{dR(t)}}{\mathrm{dt}}$ = the number of infectious individuals who recover (become immune) per unit time

and thus is given by

$$\frac{dR(t)}{dt} = rI(t)$$

Exercise:
Write down the *differential* equations describing the transmission dynamics of the infections on the next page. The letters above the arrows represent the rate at which individuals move from one category to the next. The first infection is a simplified representation of *hookworm* and the second is a simplified representation *Haemophilus Influenzae* type B.

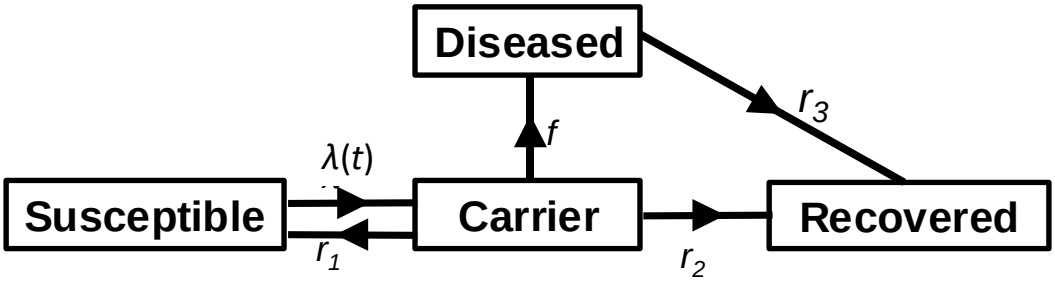
Hookworm

$$\frac{dS(t)}{dt} =$$

$$\frac{dE(t)}{dt} =$$

$$\frac{dR(t)}{dt} =$$

Haemophilus Influenzae type B



$$\frac{dS(t)}{dt} =$$

$$\frac{dC(t)}{dt} =$$

$$\frac{dD(t)}{dt} =$$

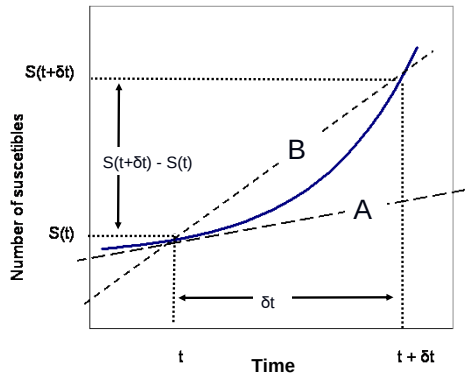
$$\frac{dR(t)}{dt} =$$

2. "Mathematical" derivation of the differential equation expressions describing the transmission dynamics of measles

The rate of change in the number of susceptibles over time

The mathematical definition of the rate of change in the number of susceptibles at time *t* is "the value of the expression $\frac{S(t+\delta t)-S(t)}{\delta t}$ as the size of the time step δt approaches zero (i.e. becomes very small)". This is illustrated in the figure.

The number of susceptibles at any time *t* is shown by the solid line. The *rate of change in the number* of susceptibles at time t, is the *gradient* of the line A.



This can be calculated by noting that the rate of change in the number of susceptibles at time t is the value of

$$\frac{dS(t)}{dt} = \frac{S(t+\delta t) - S(t)}{\delta t} \qquad \text{as } \delta t \to 0 \qquad (1)$$

ie, *the rate of change in the number* of susceptibles at time *t* is the gradient of line B as $\delta t \to 0$, or as the gradient of line B tends to the gradient of line A.

Looking at it another way:

The rate of change in the number of susceptible individuals over time can be derived using a similar logic to that used to derive the difference equations expressions you used in the last lecture, namely using the fact that the number of susceptible individuals at time $t+\delta t$ is given by

$S(t+\delta t) =$ the number of susceptible individuals at time *t, S(t)*
- the number who were newly infected between time *t* and $t+\delta t$ (2)

When the size of the time interval δt is sufficiently small, the number of susceptible individuals who are infected between time *t* and $t+\delta t$ is given by the number of individuals who are infected per unit time, $\lambda(t)S(t)$, multiplied by the size of the time interval δt. This is = $\lambda(t)S(t)\delta t$. A proof of this is provided in the Appendix.

Using this fact in expression (2), we see that the number of susceptible individuals at time $t+\delta t$ is given by:

$$S(t+\delta t) = S(t) - \lambda(t)S(t)\delta t$$

Subtracting the number of susceptibles at time t, S(t), from both sides, we see that

$$S(t+\delta t) - S(t) = -\lambda(t)S(t)\,\delta t$$

Dividing both sides of the equation by δt, we see that

$$\frac{S(t+\delta t) - S(t)}{\delta t} = -\lambda(t)S(t)$$

If δt is taken to be "infinitesimally" small, then, using equation (1) above, the left hand side of this equation can be replaced by $\frac{dS(t)}{dt}$, and we obtain the expression given on page 4 for the *rate of change in the number* of susceptibles over time.

$$\frac{dS(t)}{dt} = \frac{S(t+\delta t) - S(t)}{\delta t} = -\lambda(t)S(t) \qquad \text{as } \delta t \to 0$$

i.e. the *rate of change in the number* of susceptibles at time t is equal to (minus) the product of the force of infection at time t, and the number of susceptibles at time t:

$$\frac{dS(t)}{dt} = -\lambda(t)S(t)$$

Similarly, the expression for the rate of change in the number of pre-infectious individuals ($\frac{dE(t)}{dt}$) can be derived in a similar way to that for the rate of change in the number of susceptibles. For example, the *number* of pre-infectious individuals at time $t+\delta t$ is given by:

E(t+δt) = the number of pre-infectious individuals at time *t*
- the number of pre-infectious individuals who become infectious between time t and t+δt
+ the number of susceptible individuals who are newly infected between time t and t+δt (3)

Following the logic described above, when the time interval δt is sufficiently small, the number of pre-infectious individuals who become infectious between time *t* and t+ δt is given by $fE(t)\delta t$. Substituting this expression into expression 3, we see that

$$E(t+\delta t) = E(t) + \lambda(t)S(t)\delta t - fE(t)\delta t$$

Subtracting *E(t)* from both sides, we see that

$$E(t+\delta t) - E(t) = \lambda(t)S(t)\delta t - fE(t)\delta t$$

Dividing both sides of the equation by δt, we see that

$$\frac{E(t+\delta t) - E(t)}{\delta t} = \lambda(t)S(t) - fE(t)$$

And again, if δt is taken to be "infinitesimally" small, the left hand side of this equation can be replaced by $\frac{dE(t)}{dt}$, and we obtain the expression given on page 4 for the *rate of change in the number* of pre-infectious individuals over time.

$$\frac{dE(t)}{dt} = \lambda(t)S(t) - fE(t)$$

Using this approach we can also derive expressions for *the rate of change in the number* of infectious $\frac{dI(t)}{dt}$ and immune $\frac{dR(t)}{dt}$ individuals over time, shown on page 4.

Solving differential equations

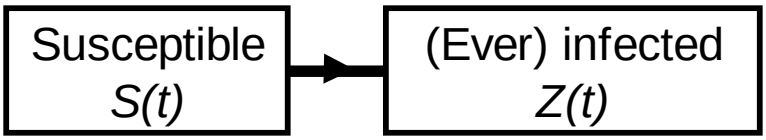
Specialist packages are generally required to manipulate differential equations to obtain eg the number of infectious cases at a given time. These packages vary from the fairly user-

friendly such as Berkeley Madonna or Stella that allow you to input the model by drawing it using the 'flowchart', to the less user-friendly such as Matlab, Mathematica, Maple that require you to type in the equations yourself. Berkeley Madonna (which is used on this course) is unique, in that it allows users to set up models using models using either using flowcharts, or by using equations. The output generated is identical irrespective of the approach used.

Each package converts the differential equations to difference equations using different techniques (e.g. Euler, Runge-Kutta, Bulirsch-Stoer etc), which make special adjustments for possible errors during each time step. For this reason, differential equations are often considered to be more rigourous in their approach than difference equations. However, for many practical purposes (especially if the time step used in the difference equations is sufficiently small) difference and differential equations give similar results to the same problem.

For some simple models it's possible to get explicit expressions e.g. for the number of susceptible individuals in terms of the input parameters, without using specialist packages to manipulate the differential equations.

Example:
The following is the possible structure of a model which we might use to track how the number of susceptible individuals in a cohort changes over time, as they become infected at a constant rate λ.



The differential equations for this model (assuming that no individuals are born or die in the population) are:

$$\frac{dS(t)}{dt} = -\lambda S(t)$$

$$\frac{dZ(t)}{dt} = \lambda S(t)$$

This model belongs to a special category of models for which the number of individuals in a given compartment can be expressed in terms of the mathematical constant, *e*, which equals 2.71828... The definition of *e* is provided in the maths refresher. In this instance, $S(t)$ is given by the equation:

$$S(t) = S(0)e^{-\lambda t} \qquad \text{Equation 1}$$

where $S(0)$ is the number of susceptible individuals at the start. Since no individuals are assumed to die from this cohort, the number of individuals at time *t* who have ever been infected is given by the difference between the number present at the start $S(0)$ and the number susceptible at time *t,* i.e.

$$Z(t) = S(0) - S(0)e^{-\lambda t} \qquad \text{Equation 2}$$

$= S(0)(1- e^{-\lambda t})$

For example if there were 100 susceptible individuals individuals at the start and 90 are still susceptible after 10 days, then the number who would have been infected by the 10th day is 100-90 = 10.

To verify that the expression for *S(t)* satisfies the differential equations given above, try differentiating it (see Appendix - from your background knowledge of calculus, you may remember that the derivative of $e^{-\lambda t}$ is $-\lambda e^{-\lambda t}$).

Note that if we divide Equations 1 and 2 by the size of the cohort, *S*(0) we obtain the following expressions for the proportion of individuals in the cohort that are susceptible or who have ever been infected, denoted by *s(t*) and *z(t*), respectively*:*

$$s(t) = e^{-\lambda t}$$
$$z(t) = 1 - e^{-\lambda t}$$

We will be using these equations in Block 2 of the course.

Partial differential equations (optional)

The equations described in this lecture do not account for changes in the age of individuals and thus do not describe the transmission dynamics of a population stratified by the age. When susceptible, pre-infectious, infectious and immune individuals are stratified by age, we also need to account for the fact that the age of individuals changes over time and we would then use "partial differential equations" to describe the transmission dynamics. The equations for the simple measles model would then be written as follows:

$$\frac{\partial S(t)}{\partial a} + \frac{\partial S(t)}{\partial t} = -\lambda(t)S(a,t)$$
$$\frac{\partial E(t)}{\partial a} + \frac{\partial E(t)}{\partial t} = \lambda(t)S(a,t) - fE(a,t)$$
$$\frac{\partial I(t)}{\partial a} + \frac{\partial I(t)}{\partial t} = fE(a,t) - rI(a,t)$$
$$\frac{\partial R(t)}{\partial a} + \frac{\partial R(t)}{\partial t} = rI(a,t)$$

The expression on the left-hand side of these equations e.g. $\frac{\partial S(t)}{\partial a} + \frac{\partial S(t)}{\partial t}$ denotes the fact that the number of individuals in a given category changes according to age and over time. Unfortunately, these equations are more difficult to solve than are ordinary differential equations and generally require use of a programming language. (e.g. Berkeley Madonna is not set up to deal with partial differential equations). One method of overcoming this problem is to change the model so that individuals are stratified e.g. by the *year of birth*

rather than age; the transmission dynamics within annual birth cohorts can then be described using ordinary differential equations.

Summary

In this lecture, you were introduced to the limitations of the difference equations and learnt about how you would describe the transmission dynamics of infections using differential equations. These assume that individuals move between different categories (e.g. from the susceptible to the pre-infectious category) continuously, rather than at discrete time intervals. The expressions for the differential equations can be derived using either an intuitive argument or from the mathematical definition of what a differential equation represents. The accompanying practical provides with practice in setting up models based on differential equations in a specialist package, Berkeley Madonna.

Further reading (optional)

Vynnycky E and White RG (2010) An introduction to infectious disease modelling. Oxford University Press. Chapter 3 and Basic maths, section B.5.

Appendix

The following provides the methods for deriving some of the results mentioned in the lecture.

Derivation of the proportion of individuals who are e.g. infected in a short time interval δt

Suppose that a proportion λ of uninfected individuals are infected per unit time. The most straightforward method of deriving the proportion who are infected during a time interval δt is to calculate it as:

1 - proportion who escape infection during this time interval,

Note that the proportion of individuals who escape infection per unit time is given by

$$1-\lambda$$

The proportion of susceptible individuals who escape infection during a time interval δt is given by

$$(1-\lambda)^{\delta t}$$

(This expression follows from the fact that individuals have to escape infection for δt lots of time units.) Thus the proportion of susceptible individuals who are newly infected per unit time is given by

$$1- (1-\lambda)^{\delta t} \qquad \text{(C)}$$

The expression for $(1-i)^{\delta t}$ can be expanded to give

$$(1-\lambda)^{\delta t} = 1- \lambda\delta t + (\lambda\delta t)^2/2! - (\lambda\delta t)^3/3! + ...$$

If the size of the time step δt is sufficiently small, then all terms involving δt^2, δt^3 etc contribute only a small amount to the expression for $(1-\lambda)^{\delta t}$. Hence the proportion who escape infection during a small time interval δt is given by:

$$(1-\lambda)^{\delta t} \approx 1- \lambda\delta t$$

Substituting this approximation into expression C, we obtain the result that the proportion of individuals who are newly infected during a small time interval δt is approximated by:

$$1- 1- \lambda\delta t = \lambda\delta t$$

Revision of differentiation

Considering the example in the notes, suppose that the number of susceptibles S(t) in your population decreases exponentially over time at a rate of λ/day and that you have S(0) susceptible to begin with. The expression for S(t) is given by the expression $S(t) = S(0)e^{-\lambda t}$. We will differentiate this expression to derive the rate of change in the number of susceptibles over time.

Note that the derivative of a function, eg $y=x^n$ is just nx^{n-1}. For example, note that the derivative of a function is the same as its gradient ("rate of change"). The gradient of any straight line

$$y=mx + c$$

is just m. What will you get if you differentiate the expression for the straight line with respect to x?

You may remember that the expression for e^t is as follows:

$$e^t = 1+ t + t^2/2! + t^3/3! + t^4/4! + ... t^n/n! + ...$$

$e^{-\lambda t}$ is given by the expression:

$$e^{-\lambda t} = 1+ (-\lambda t) + (-\lambda t)^2/2! +(-\lambda t)^3/3! + (-\lambda t)^4/4! + ... (-\lambda t)^n/n! + ...$$
$$= 1-\lambda t + \lambda^2t^2/2! – \lambda^3t^3/3! + \lambda^4t^4/4! + .. + (-\lambda)^nt^n/n! + ...$$

Applying the rules for differentiating a function, you should see that the derivative of $e^{-\lambda t}$ is given by:

$$= 0 -\lambda + 2\lambda^2t/2! – 3\lambda^3t^2/3! + 4\lambda^4t^3/4! + .. + n(-\lambda)^nt^{n-1}/n! + ...$$

Cancelling out common terms in the denominator and numerator of each term, this expression simplifies to:

$$-\lambda + \lambda^2t – \lambda^3t^2/2! + \lambda^4t^3/3! + .. + (-\lambda)^nt^{n-1}/(n-1)! +...$$

Taking out a common factor of $–\lambda$ from each expression, this expression becomes:

$$- \lambda(1-\lambda t + \lambda^2t^2/2! – \lambda^3t^3/3! + .. - (-\lambda)^{n-1}t^{n-1}/(n-1)! +...)$$

which, using the expression for $e^{-\lambda t}$ above, gives:

$$-\lambda e^{-\lambda t}$$

Hence, if you differentiate the expression $S(t) = S(0)e^{-\lambda t}$, you should get

$$dS/dt = -\lambda S(0)e^{-\lambda t}$$
$$= -\lambda S(t)$$

Introduction to Infectious Disease Modelling and its Applications – 2018

Session 5: The natural dynamics of infectious diseases

Lecture

Overview and Objectives

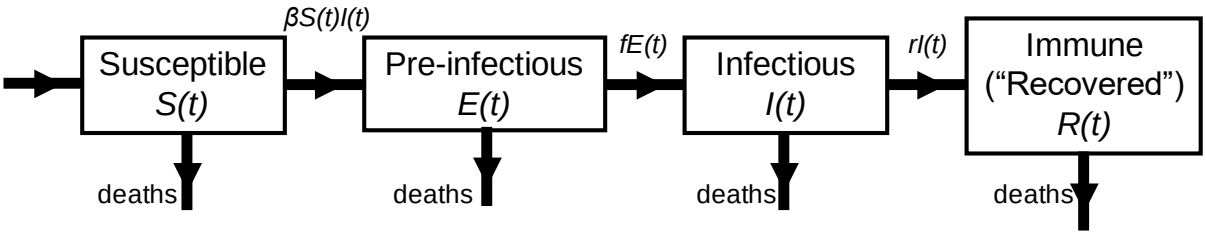
In the last two sessions, you learned about the methods for setting up simple models of the transmission dynamics of immunizing infections such as measles. In this session, we will discuss the insights into the dynamics of infections that these models provide, and relate these model predictions to data.

By the end of this lecture you should:

- Understand what determines whether the numbers of infectious individuals increases or decreases;
- Be aware of methods for calculating R_0 for an infection from the growth rate of an outbreak or epidemic;
- Understand the theory and factors underlying the cycles in the occurrence of immunizing infections;
- Be able to calculate the inter-epidemic period for an immunizing infection
- Know some of the insights into the epidemiology of immunizing infections which are provided by these models.

Introduction

The following diagram summarizes the structure of the models that you have used during the last two sessions.



These models are amongst the simplest models which can be used to describe the long-term transmission dynamics of an infection. For example, they assume that individuals mix randomly, they did not stratify individuals according to their age and sex, and the population size was assumed to remain unchanged over time.

Some of these assumptions are obviously unrealistic. Despite these simplifications, the models were able to reproduce (at least, for a while!) the familiar epidemic cycles in measles that are observed in many populations and are illustrated in Figure 1. The fact that the model's general predictions are reasonably consistent with observed data suggests that the models, whilst approximations, may be used to help us understand the occurrence of epidemics.

However, the fact that the peaks in the epidemic became progressively weaker ("damped") over time (see Figure 2) but that we do not see this in the actual data, suggests that other factors that are not in the model are needed to sustain the epidemic cycles.

Before discussing the cycles and the damping in further detail, we discuss some of the other insights into the dynamics of infections provided by this model, specifically:

1. *What determines whether or not the number of infectious individuals increases following the introduction of an infectious person into a totally susceptible population?*
2. *How fast might we expect the number of infectious individuals to increase following the introduction of an infectious person into a totally susceptible population and what can we infer from it?*
3. *Why does the incidence of an immunizing infection cycle over time?*
4. *What other factors lead to the epidemic cycles?*
5. *What inter-epidemic period might we expect to see for an immunizing infection?*

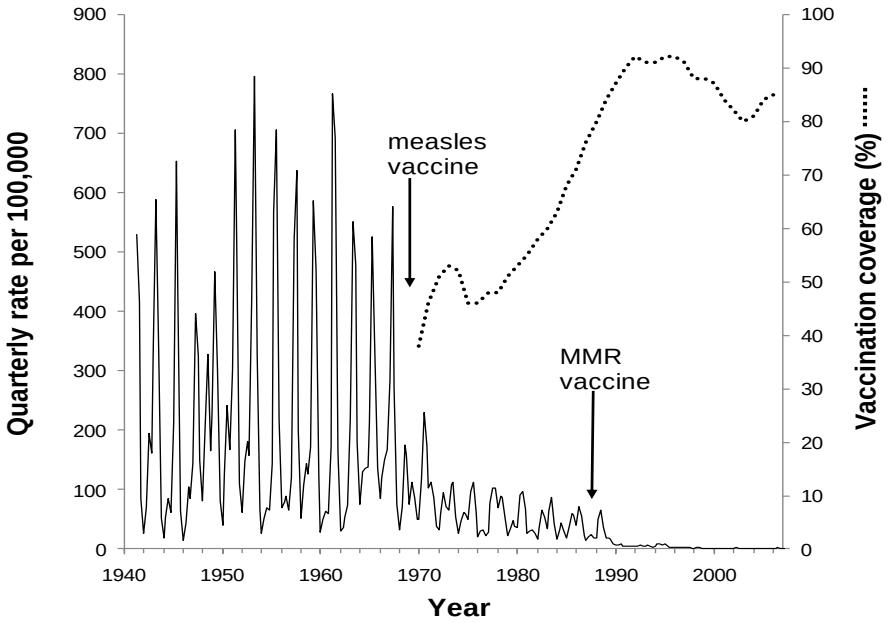


Figure 1: Quarterly measles notifications in England and Wales, 1941-2010 (Data source: The Office for National Statistics and Public Health England (formerly, Health Protection Agency)

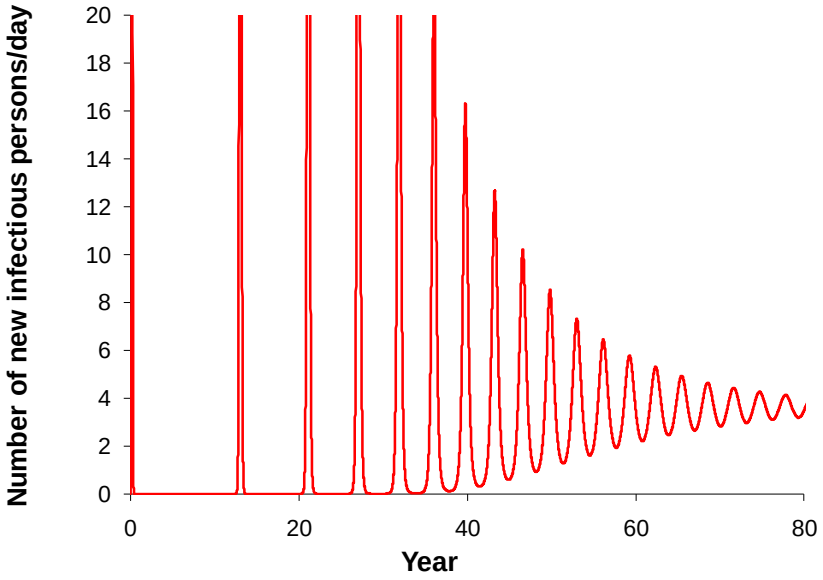


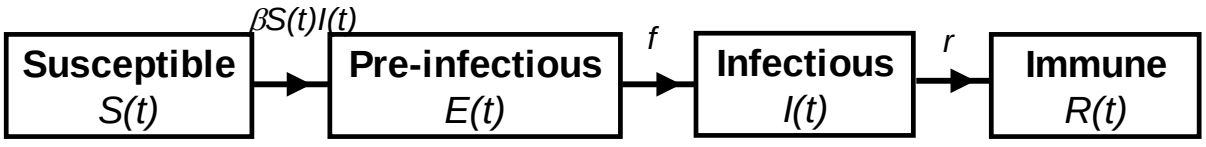
Figure 2: Predictions of the numbers of infectious individuals with measles per day following the introduction of an infectious person with measles into a totally susceptible population, assuming that R_0=13, pre-infectious period=8 days, infectious period=7 days, total population size=100,000, average life expectancy=70 years, birth rate=death rate.

Properties of the basic model

1. What determines whether or not the number of infectious individuals increases following the introduction of an infectious person into a totally susceptible population?

You may remember from your previous epidemiological training that if the basic reproduction number (R_0) of a pathogen is greater than one, then if one infectious person is introduced into a population, the number of new infectious individuals would be expected to increase and the infection would be expected to persist in the population.

In fact, this criterion can be derived relatively easily from the differential equations describing the transmission dynamics of immunizing infections which you have used hitherto. The logic is outlined below. For simplicity, we illustrate the proof considering a "closed" population, i.e. in which there are no births or deaths. The model therefore has the following structure:



Reminder:
From the difference equations lecture, you may recall that

$$\beta = R_0 / \{\text{population size } (N) \times \text{duration of infectiousness } (D)\}$$

This expression can be rearranged to give the following expression for the basic reproduction number:

$$R_0 = \beta ND$$

Using the result that the average rate at which infectious individuals recover (r) and the duration of infectiousness are related through the equation: $r=1/D$ or, equivalently, $D=1/r$, we obtain the following equation for R_0:

$$R_0 = \frac{\beta N}{r}$$

We will show that for the number of infectious individuals to increase following the introduction of an infectious person into a totally susceptible population $\frac{\beta N}{r}$, or equivalently, βND, must be bigger than 1.

We begin by noting that if the number of infectious individuals increases following the introduction of an infectious person into a totally susceptible population, then the rate of change in the number of pre-infectious and infectious individuals must be "positive" i.e. $\frac{dE}{dt} > 0$ and $\frac{dI}{dt} > 0$.

In the last lecture, we saw that the expressions for the rate of change in the number of pre-infectious and infectious individuals are given by:

$$\frac{dE}{dt} = \beta S(t)I(t) - fE(t) \qquad \text{Eq 1}$$

$$\frac{dI}{dt} = fE(t) - rI(t) \qquad \text{Eq 2}$$

where f is the rate at which pre-infectious individuals become infectious and r is the rate at which infectious individuals recover to become immune.

For the rate of change in the number of pre-infectious individuals (described in equation 1) to be positive, the number of new infections which occur in the population per unit time ($\beta S(t)I(t)$) has to exceed the number of pre-infectious individuals who become infectious per unit time ($fE(t)$). i.e.

$$\beta S(t)I(t) > fE(t)$$

Similarly, for the rate of change in the number of infectious individuals to be positive, the number of pre-infectious individuals who become infectious per unit time ($fE(t)$) has to exceed the number of infectious individuals who recover per unit time ($rI(t)$). i.e.

$$fE(t) > rI(t)$$

By this logic, the number of individuals who are newly infected per unit time must also be larger than the number of individuals who recover per unit time. i.e.

$$\beta S(t)I(t) > rI(t)$$

Dividing both sides of this expression by the number of infectious individuals in the populations (*I(t)*) we see that for the rate of change in number of infectious individuals to be positive, and therefore for the number of infectious individuals to increase, the following condition has to hold:

$$\beta S(t) > r \qquad \text{Eq 3}$$

Notice that when the infection is introduced into the population, the number of individuals who are susceptible is the same as the total population size (given the symbol *N*). So for the number of infectious individuals to increase following the introduction of an infectious person into a totally susceptible population, the following has to hold:

$$\beta N > r$$

If we divide this expression by *r*, we see that the condition:

$$\frac{\beta N}{r} > 1 \qquad \text{Eq 4}$$

also has to hold for the number of infectious individuals to increase following the introduction of infectious into a totally susceptible population. As discussed on page 4, the left-hand side of this expression is just the basic reproduction number. We can also use an intuitive argument to see that $\frac{\beta N}{r}$ is the basic reproduction number. e.g. we can say that each infectious person contacts βN other individuals per unit time; multiplying βN by the infectious period, which is just 1/r, we obtain the total number of individuals contacted during the infectious period.

Using this verbal definition for R_0 in expression 4, we can conclude that for the introduction of an infectious person into a totally susceptible population to lead to an increase in the number of infectious individuals (i.e. for an epidemic to occur) the basic reproduction number has to be greater than 1.

Note that if we assume that individuals mix randomly, the rate at which susceptible individuals are newly infected is directly proportional to the number of infectious individuals (i.e. $\lambda(t) = \beta I(t)$). Therefore if the number of infectious individuals increases (or decreases) the infection incidence would also be expected to increase (or decrease).

The above logic can also be used to obtain the result that the proportion of individuals who need to be susceptible in the population for the number of infectious individuals (and therefore for the infection incidence) to increase at a given time has to be greater than $1/R_0$.

Exercise

Derive the condition that proportion of individuals who need to be susceptible in the population for the number of infectious individuals to increase at a given time has to be

greater than $1/R_0$. Hint: divide both sides of equation (3) by the total population size (N) and β and use the result that $R_0 = \frac{\beta N}{r}$

These threshold theorems were discussed by Kermack and McKendrick in 1927, although they weren't expressed in terms of the basic reproduction number (which was first defined by Macdonald in 1955).

Note

Notice that these equations do not account for the fact that pre-infectious or infectious individuals may die during the pre-infectious or infectious periods. See Anderson and May (1992) (ch 1-4) for details of expressions which take this into account. For most common immunizing infections, the pre-infectious and infectious periods are generally a few days, whereas the life expectancy (at least in industrialized populations) is about 70 years. As a result, since the mortality rate is much smaller than the rate at which individuals become infectious and recover, the effects of these adjustments to the above expression for the R_0 are relatively small. In any case, the above logic can be extended to obtain the equations for R_0 to account for mortality or for other infections, e.g. those which do not confer immunity.

2. How fast might we expect the number of infectious individuals increase following the introduction of an infectious person into a totally susceptible population and what can we infer from it?

It can be shown that, following the introduction of an infectious person into a totally susceptible population, the number of infectious individuals would be expected to increase at a rate (usually denoted by the symbol Λ) given approximately by:

$$\Lambda \approx \frac{R_0 - 1}{D}$$

where D is the average duration of infectiousness. The mathematical derivation of this expression is described in Anderson and May (1992) and in Lipsitch et a (2003) and has been extended recently by Wearing et al (2005).

The above expression can be rearranged to give the following expression for R_0 in terms of the rate of increase:

$$R_0 \approx \Lambda D + 1$$

Therefore, given empirical estimates of the rate of increase in the infection prevalence (Λ), it should be possible to infer the R_0 of the pathogen. The above expression (and its variants) have been used to derive estimates of the basic reproduction number for HIV during the early stages of the HIV epidemic in Kenya and Uganda (Anderson et al (1988)), which ranged between 4 and 11. Such estimates can then be used to infer future trends in the infection incidence (see sessions in block 4).

A variant of this expression was used in studies estimating the basic reproduction number of SARS in 2003, using the average serial interval (calculated as 8.4 days using data from Singapore) and the growth rate in the cumulative numbers of cases (Lipsitch et al (2003)). Those analyses implied that the R_0 for SARS was in the range 2.0-3.6.

The drawback of this formula is that it is only reasonably reliable if the rate of increase in the prevalence of infection with the pathogen (Λ) is calculated during the early stages of an epidemic, and if the pathogen has only recently been (re)introduced into the population.

3. Why does the incidence of an immunizing infection cycle over time?

Explanations for the cycles

During the first part of the 20th century, there were 2 main explanations for the cycles in measles notifications, namely that i) they reflected cycles in the infectivity of the measles virus or ii) they resulted from changes in the prevalence of susceptible individuals, as a result of a constant influx of susceptibles born into the population, and susceptibles becoming immune after becoming infected (Hamer, 1906). Experimental studies in mice carried out during the 1930s found no evidence for changes in the infectivity, (see e.g. Fine (1979)) and thus the second argument has become accepted.

Before discussing the mechanism by which epidemic cycles for immunizing infections occur, we first revise the relationship between the net reproduction number, the trend in the number of new infectious individuals and the proportion of individuals in the population who are susceptible.

From your previous training, you may recall that the net reproduction number (R_n) is defined as the average number of secondary infectious individuals resulting from each infectious person in a given population, i.e. in which some individuals may be immune. It is related to the basic reproduction number through the following equation:

$$R_n = R_0 \, s$$

where s is the proportion of the population that is susceptible. When the number of new infectious individuals is increasing, $R_n>1$; when the number of new infectious individuals is decreasing, $R_n<1$ and when the number of new infectious individuals is stable, $R_n=1$.

Exercise: *The following shows model predictions of an epidemic curve following the introduction of an infectious person into a totally susceptible population. Given that the net reproduction number is related to the proportion of the population which is susceptible through the equation:*

$$R_n = R_0 \times \{\text{proportion susceptible}\}$$

what can we say about the proportion of the population which is susceptible when the daily number of new infectious individuals is: i) increasing ii) decreasing iii) at a peak?

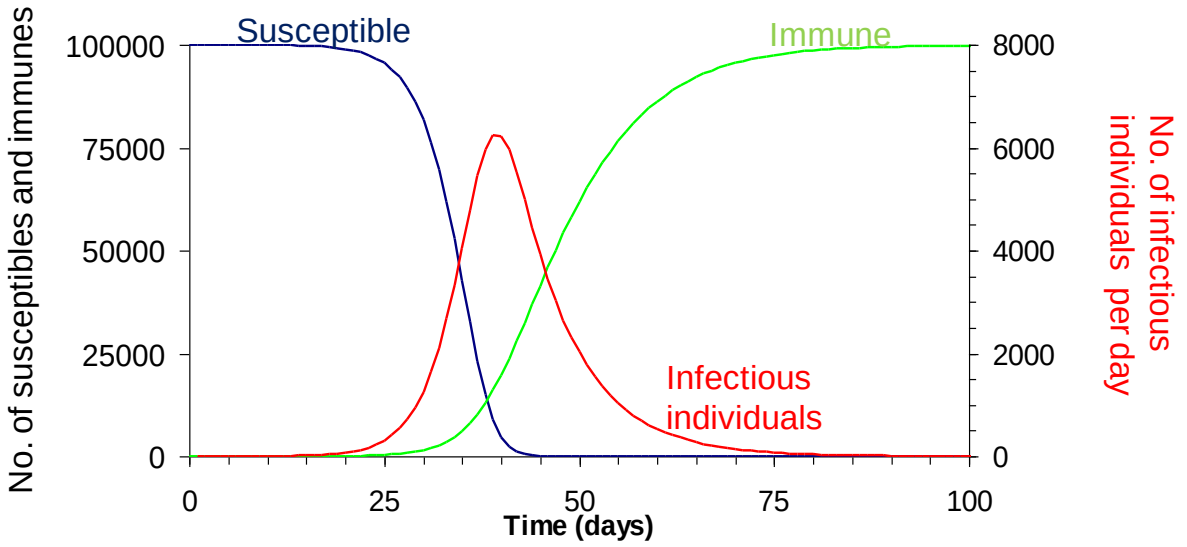


Figure 3: Predictions of the daily number of new infectious individuals and the number of susceptible and immune individuals following the introduction of an infectious person with measles into a totally susceptible population, assuming that R_0=13, pre-infectious period=8 days, infectious period=7 days, total population size=100,000.

The above exercise highlights the fact that the proportion of the population which is susceptible has to be above a certain threshold value ($1/R_0$) for the number of infectiious individuals to increase, and it has to be below the same threshold value for this to decrease. When the proportion of individuals who are susceptible equals this threshold value, the number of new infectious individuals per day is stable.

Without the entry of new susceptibles into the population, e.g. as a result of births, the proportion of susceptible individuals would remain consistently below the threshold level (i.e. $<1/R_0$) and transmission would eventually cease.

The entry of new susceptibles as a result of new births into the population means that the following occurs (see also figure 4):

1. The proportion of susceptible individuals will eventually start to increase once a sufficient number of births have been added (point A in figure 4). At some point, the proportion of susceptible individuals is sufficiently large ($>1/R_0$) for each infectious person to start to lead to >1 infectious person (point B in figure 4). Once this occurs, the number of new infectious individuals/unit time starts to increase.
2. At some point, the number of susceptibles who are removed by infectious individuals exceeds the number added to the population through new births. Once this occurs, the increase in the proportion of susceptible individuals slows and then reverses (point C in figure 4)**.**
3. Once the proportion of susceptible individuals in the population has declined to reach $1/R_0$, the number of new infectious individuals/unit time peaks (point D in figure 4)
4. Eventually, the proportion of susceptible individuals becomes sufficiently low ($<1/R_0$) and each infectious person leads to <1 secondary infectious person, and the number of new infectious individuals starts to decrease (point E in figure 4).

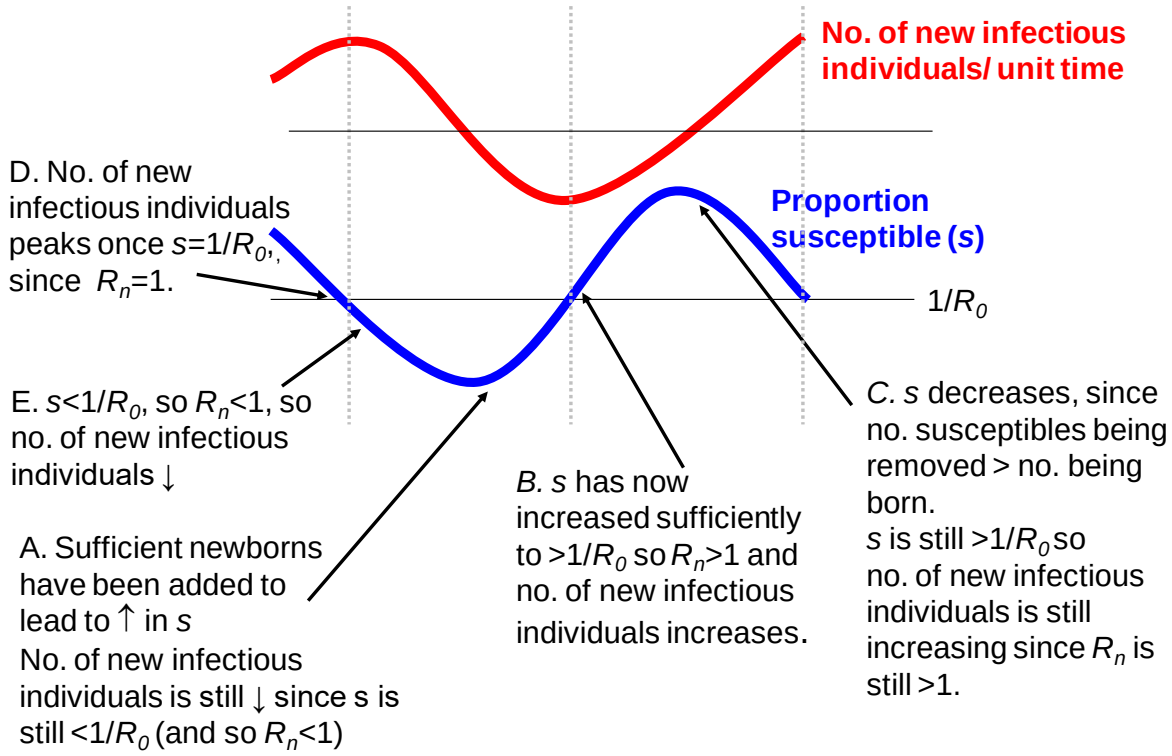


Figure 4: Summary of changes in the number of new infectious individuals/unit time and the proportion of individuals who are susceptible in the population during the epidemic cycles for an immunizing infection

Exercise:

a) What is the proportion of individuals who are susceptible in the population at the points when the number of new infectious individuals/unit time is either at a maximum or a minimum (B or D)?

b) On the above figure, plot the proportion of individuals who are immune in the population over time.

An argument similar to that above was first used by Hamer to explain the cycles in measles cases. His argument used the "mass action" principle, and was based, implicitly, on the following difference equation model:

$$C_{t+1} = kC_tS_t$$
$$S_{t+1} = B + S_t - C_{t+1}$$

where

- C_t and S_t are the number of cases and susceptibles respectively at time t,
- k is the proportion of the total possible contacts that actually lead to new cases (Fine and Clarkson (1982))
- B is the number of births in each time step.
- The time step used in this original model was just the serial interval for the infection.

The difference between this model and the models which you have been using so far, is that the generations of cases in this model do not overlap, i.e. all cases have onset and infect each other at times steps of 1 serial interval.

Hamer argued that when the number of susceptibles reached a maximum or a minimum, the number of new measles cases occurring per serial interval equalled the number of births into the population.

Exercise:
How might you use the above equations to show this?

Figure 5 shows the number of susceptibles and cases predicted by this model over time

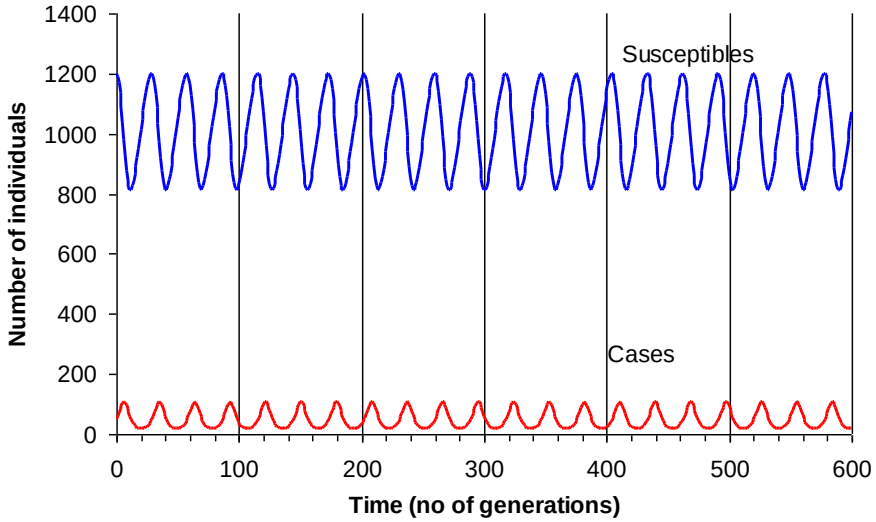


Figure 5 Number of cases and susceptibles over time predicted using the mass-action model, assuming that C_0=50, S_0=1200, B=50 and k=0.001 per serial interval. See massact.xls for details of this model.

In contrast with the model used in the practicals, this model predicts that the cycles in the numbers of infectious individuals should never damp out. Subsequent theoretical elaborations by Soper (1929) found that when the simple mass-action model was changed to assume that the infectious period was not fixed and followed a variable distribution, the cycles damped out. Thus the simple mass action model was too simplistic to fully explain the epidemic cycles.

4. What other factors lead to the epidemic cycles?

The damped cycles in the numbers of infectious individuals predicted by the model are inconsistent with the observed data and has led to suggestions that other factors must be important in sustaining these cycles. These factors are listed below:

a) Seasonality in transmission
Seasonal transmission is the most obvious factor likely to lead to the epidemic cycles in immunizing infections and there have been many modelling and observational studies of its

effect. Analyses by Fine and Clarkson (Fine and Clarkson (1982)) have found evidence for seasonal transmission in England and Wales, occuring as a result of intense mixing between children during the school terms, and less intense mixing during the school holidays (see Figure 6) These analyses used a model-based approach, calculating the transmission parameter in the simple mass action model (i.e. the proportion of the total possible contacts that actually lead to new cases) using the ratio between the number of cases observed in the UK in each week over the time period 1950-1977, and the number of susceptible individuals.

In contrast, seasonal transmission appears to be less important in sustaining the epidemic cycles for pertussis than for measles, as the epidemics occur at different times each year (Figure 7). This has led to suggestions that other e.g. climatic factors may be important in sustaining the cycles.

b) Age-dependent mixing

Work by Schenzle (1984) has illustrated that if the population in models are set up so that transmission is confined to individuals within specific annual birth cohorts (corresponding to classes within school years), then the models predict regular cycles in the numbers of infectious individuals; other assumptions about age-dependent mixing (such as those discussed in the heterogeneous mixing session) did not lead to predictions of regular epidemic cycles.

c) Stochastic effects

The models described so far have been deterministic and do not take account of the fact that individuals come in integer, and not fractional, quantities. Work by Bartlett (1956) and Anderson and May (1986) have illustrated that models which are amended to deal with discrete (integer) numbers of individuals and which allow for stochastic variation in transmission predict regular epidemic cycles (as long as the population size was sufficiently large).

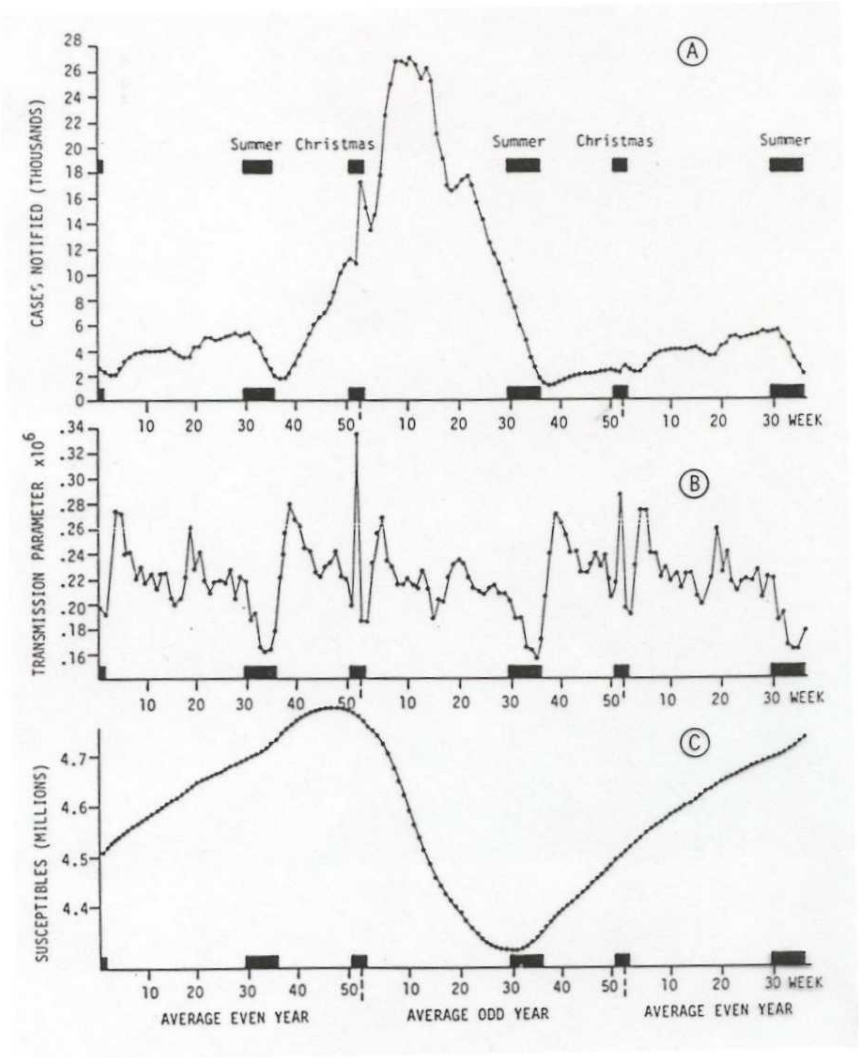


Figure 6 Analysis of average biennial measles pattern, based on data from 1950 to 1965 in the UK a) Average number of cases notified per week; b) Calculated weekly transmission parameters; c) Estimated numbers of susceptibles. Shaded blocks indicate school summer and Christmas holiday periods. (Fine and Clarkson (1982))

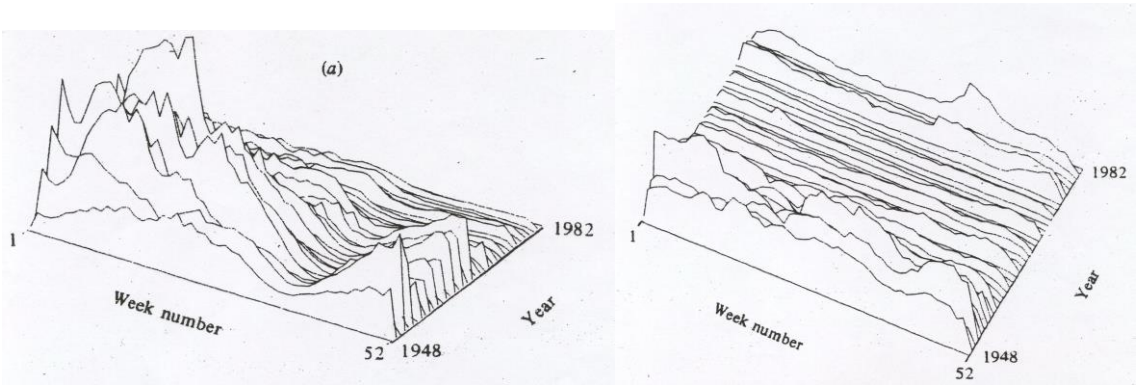


Figure 7: Seasonal patterns in measles and pertussis notifications (left and right-hand figures respectively), based on weekly case reports in England and Wales, 1948-82. Source: Anderson and May (1992).

5. What inter-epidemic period might we expect to see for an immunizing infection?

It can be shown that the inter-epidemic period (*T*) for immunizing infections predicted by the simple models used in the last two sessions is given by the expression:

$$T \approx 2\pi\sqrt{A(D+D')} \qquad \text{Eq 5}$$

where *A* is the average age at infection, *D*' and *D* are the average pre-infectious and infectious periods respectively.

In the absence of vaccination or control, the average age at infection *A*, the average life expectancy and the basic reproduction number are related through the expression $R_0 = 1+L/A$, which can be rearranged to give the expression $A=L/(R_0-1)$. Substituting this expression for *A* into the above expression gives the following expression for the inter-epidemic period in the absence of control:

$$T \approx 2\pi\sqrt{\frac{L(D+D')}{R_0-1}} \qquad \text{Eq 6}$$

As shown in Table 1, predictions of the inter-epidemic period are generally consistent with those observed.

Table 1: Estimates of the observed and predicted inter-epidemic periods for different infections in various locations (extracted from Anderson and May (1992))

		Inter-epidemic period	
Infection	**Location**	**Calculated**	**Observed**
Measles	England and Wales 1948-68	2	2
	Aberdeen, Scotland 1883-1902	2	2
	Baltimore, USA 1900-27	2	2
	Paris, France 1880-1910	2	2
	Yaounde Cameroon, 1968-75	1-2	1
	Ilesha, Nigeria, 1958-61	1-2	1
Rubella	Manchester, UK 1916-83	4-5	3.5
	Glasgow, Scotland, 1929-64	4-5	3.5
Mumps	England and Wales 1948-82	3	3
	Baltimore, USA 1928-73	3-4	2-4
Polio	England and Wales, 1948-65	4-5	3-5
Smallpox	India, 1868-1948	4-5	5
Chickenpox	New York City, USA, 1928-72	3-4	2-4
	Glasgow, Scotland, 1929-64	3-4	2-4
Scarlet fever	England and Wales, 1897-1978	4-5	3-6
Diphtheria	England and Wales, 1897-1979	4-5	4-6
Pertussis	England and Wales, 1970-82	3-4	3-4

Exercise:
Why might the inter-epidemic period associated with a pathogen which has a high basic reproduction number be short?

The first expression for the inter-epidemic period above highlights the fact that the introduction of vaccination into a population may lead to increases in the inter-epidemic period. For example, vaccination reduces the prevalence of infectious individuals in the population and therefore postpones infection for some individuals until later in life. This increases the average age at infection and according to the above expression, should lead to an increase in the inter-epidemic period.

The above expressions have been useful for highlighting unexpected features of the transmission of pathogens. For measles, for example, the inter-epidemic period after the introduction of vaccination has been shorter than that predicted (see Anderson and May (1992)) and this has led to discussions that mixing between individuals may be age-dependent. After the introduction of pertussis vaccination in the UK, the inter-epidemic period for pertussis, was also shorter than might have been expected, leading to suggestions that vaccination does not protect against infection, but does protect against disease following infection (Fine and Clarkson (1982)).

Extending the logic to other pathogens...

The discussion above has related to simple immunizing infections, such as measles, mumps and rubella, for which the pre-infectious and infectious periods are short relative to the lifetime of individuals. The logic is not easily extendible to infections which do not confer immunity against reinfection. Similarly the logic is not easily extendible to diseases such as tuberculosis for which the interval between infection and disease may be long (e.g. decades) and thus conditions (such as the number of individuals contacted by each infectious person) has probably declined over time, at least in many Western populations.

In the accompanying practical, you will be exploring some of the concepts covered here using the models you have developed hitherto.

Further reading

Vynnycky E and White RG (2010) An introduction to infectious disease modelling. Oxford University Press. Chapter 4.

Some of the issues raised in this lecture are discussed in the paper Anderson and May (1982) Science; 215:1053-1060, which is provided at the end of these notes. The papers by Fine and Clarkson (1982) and Hamer (1906) are also provided as supplementary reading material.

Bibliography

Anderson RM (1982) Directly transmitted viral and bacterial infections in man. in *The population dynamics of infectious diseases: theory and applications* (RM Anderson, ed) Chapman and Hall, London

Anderson RM and May RM (1986) The invasion, persistence and spread of infectious diseases within animal and plant communities *Phil Trans R Soc* B314, 533-70

Anderson, RM, Grenfell BT and May RM (1984) Oscillatory fluctuations in the incidence of infectious diseases and the impact of vaccination: time series analysis *J Hyg (Camb)*, 93: 587-608

Anderson RM, May RM, McLean AR (1988) Possible demographic consequences of AIDS in developing countries *Nature,* 332,228-234

Black FL (1966) Measles endemicity in insular populations: critical community size and its evolutionary implication. *J Theor Biol*, 11, 207-11

Fine PEM (1979) John Brownlee and the measurement of infectiousness: an historical study in epidemic theory. *J Roy Statist Soc*, 143(2): 347-62

Fine PEM and Clarkson JA (1982) Measles in England and Wales I – an analysis of factors underlying seasonal patterns. *Int J Epidemiol* 11(1):5-14

Fine PE, Clarkson JA (1982) The recurrence of whooping cough: possible implications for assessment of vaccine efficacy. *Lancet.* 1982 Mar 20;1(8273):666-9

Kermack WO and McKendrick AG (1927) A contribution to the mathematical theory of epidemics Proc Roy Soc A 115:700-21

Lipsitch M, Cohen T, Cooper B, Robins JM, Ma S, et al. Transmission dynamics and control of severe acute respiratory syndrome. Science. 2003 Jun 20;300(5627):1966-70.

Panum PL (1940) Observations Made During the Epidemic of Measles on the Faroe Islands in the Year 1846. Republished by the American Public Health Association in 1940.

Schenzle (1984) An age-structured model of pre- and post-vaccination measles transmission. *IMA J Math Appl Med Biol* 1:169-91

Wearing HJ, Rohani P & Keeling MJ. Appropriate models for the management of infectious diseases. *PLoS Med* 2, e174 (2005).

Introduction to Infectious Disease Modelling and its Applications – 2018

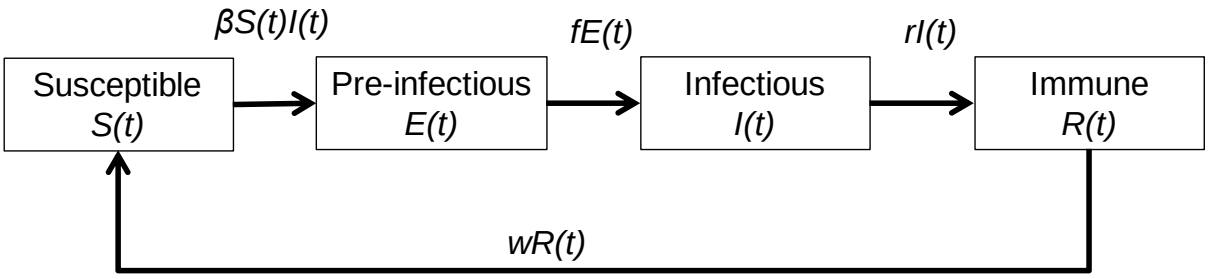
Session 6: Review

Lecture

This is an optional session going over the key points arising in the sessions in block 1. Please bring any questions that you may have about the material in this block to this session.

The next few pages include some multiple choice questions that you may like to try to test your understanding. We will go over the questions during the review session on Wednesday.

Q1. The following is a diagram of the transmission dynamics of *Mycoplasma pneumoniae*. The expressions above the arrows reflect the number of people moving from one category to the next per unit time. Which of the following statements is correct? Note that there may be more than one correct answer.



a) The model will predict that the total population size will decrease over time.

b) The model will predict that the total population size will increase over time.

c) The model will predict that the total population size will remain unchanged over time.

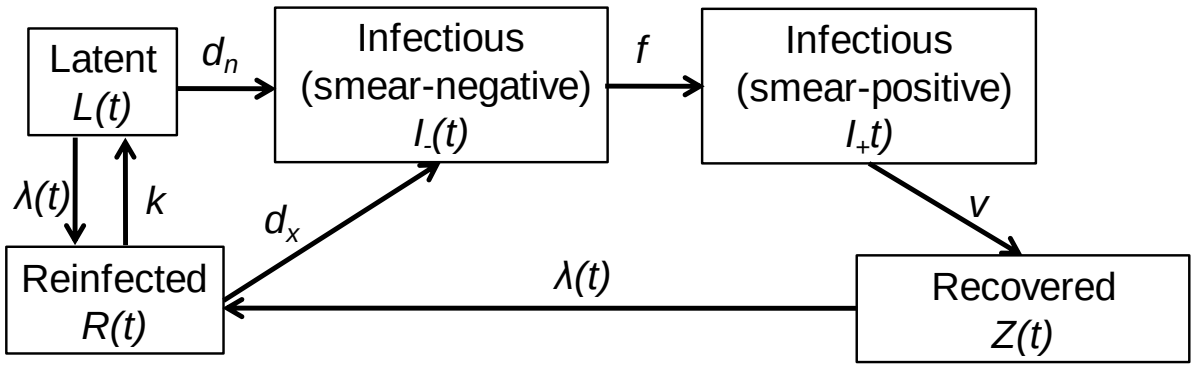
d) The rate of change in the number of Immune individuals is given by the following equation:

$$\frac{dR(t)}{dt} = r - w$$

e) The rate of change in the number of Immune individuals is given by the following equation:

$$\frac{dR(t)}{dt} = rI(t) - wR(t)$$

Q2. The following is a simplified diagram of the transmission dynamics of *M tuberculosis* in a closed high transmission setting. Smear-negative and smear-positive individuals are both infectious; however, smear-negative individuals are 25% less infectious than are smear-positive individuals. The *per capita* rate at which smear-positive individuals come into effective contact with others is denoted by the symbol *ß*.



Which one of the following statements is correct?

a) The rate of change in the total number of infectious individuals ($I(t) = I_+(t) + I_-(t)$) is given by the following equation:

$$\frac{dI(t)}{dt} = d_n L(t) + d_x R(t) - vI_+(t)$$

b) The rate of change in the total number of infectious individuals ($I(t) = I_+(t) + I_-(t)$) is given by the following equation:

$$\frac{dI(t)}{dt} = d_n L(t) + d_x R(t) - fI_-(t) - vI_+(t)$$

c) There are no errors in the following equation:

$$\frac{dL(t)}{dt} = -\lambda(t)R(t) + kL(t) - d_n L(t)$$

d) The force of infection is given by the following equation:

$$\lambda(t) = \beta I(t)$$

e) The force of infection is given by the following equation:

$$\lambda(t) = \beta I_+(t) + 0.25\beta I_-(t)$$

Q3. The inter-epidemic period for rubella in some populations was roughly 4 years before the introduction of vaccination. The value for R_0 for rubella was about 7. Which one of the following statements is correct?

a) The introduction of rubella vaccination among newborns was likely to lead to a reduction in the inter-epidemic period,

b) The introduction of rubella vaccination among newborns was likely to lead to an increase in the inter-epidemic period.

c) When the incidence of rubella was at a peak, on average, 25% of the population was likely to be susceptible.

d) When the incidence of rubella was increasing, the proportion of the population that was susceptible was less than 14%

e) The inter-epidemic period of rubella was likely to be less than that for measles, for which the value for R_0 was about 13.