The ABC of terms used in mathematical models of infectious diseases

Sharmistha Mishra,[1,2] David N Fisman,[3] Marie-Claude Boily[2,4]

[1]Division of Infectious Diseases, St. Michael's Hospital, University of Toronto, Toronto, Canada
[2]School of Public Health, Imperial College, London, UK
[3]Dalla Lana School of Public Health, University of Toronto, Toronto, Canada
[4]Unité de recherche en santé des populations de l'Université Laval Centre hospitalier *affilié* universitaire de Québec, Laval, Canada

Correspondence to
Dr Marie-Claude Boily, Division of Epidemiology, Public Health and Primary Care, Imperial College, G30 16 South Wharf Road, St Mary's Campus, London W2 1PG, UK; mc.boily@imperial.ac.uk

ABSTRACT
Mathematical models that incorporate a dynamic risk of infection figure prominently in the study of infectious diseases epidemiology as a tool to inform public health policy. In recent years, their use has expanded to address methodological questions, inform and validate study design and evaluate interventions. This glossary briefly highlights the applications of transmission dynamics modelling, explains different modelling methodologies and defines commonly encountered terms to provide an introductory and conceptual understanding of the vocabulary and frameworks used in the literature.

Over the past century, the science of mathematical epidemiology has grown rapidly, providing theoretical advances and important insights into population-level characteristics of infection due to individual-level behaviour and biology.[1] Unlike chronic disease epidemiology, the study of infectious diseases require specific tools—namely transmission dynamics models—that capture the dynamic nature and spread of disease from infectious to susceptible individuals (a case is also a risk factor) and incorporate positive and negative feedback characteristics of infectious processes.[2][3] Although the term 'mathematical model' describes any 'model' based on a system of equations that summarise observed data with a goal to predict an outcome of interest, here we use it to refer to transmission dynamics models that capture the communicability of infectious diseases. In contrast to mathematical models of tumour growth or disease progression, or statistical models used for estimation and prediction, the models discussed in this review assume that the incidence of infection depends on the prevalence of infection over time—a feedback system that represents the non-linear dynamics of infectious diseases at the population level.

APPLICATION OF TRANSMISSION DYNAMICS MODELS
Mathematical models provide an insight into the spread of infections, patterns of disease and their spatiotemporal or host/pathogen-related (eg, age) determinants. They enable us to estimate important parameters, highlight key gaps in data, formulate new hypotheses and predict population-level effectiveness (or potential adverse consequences) of interventions.[1][4][5]

Early work based on simple mathematical models succinctly summarised the spread of infections in populations, providing theoretical frameworks and key epidemiological concepts such as the basic reproductive rate and eradication fraction or herd immunity threshold.[1][6–16] More recently, increasingly complex models have captured the influence of heterogeneity (in contact patterns, infectivity, or susceptibility to infection) on the spread and persistence of infection, with important implications for control and outbreak preparedness.[5][17–19]

Models help evaluate interventions, as disease trends following an intervention may be erroneously attributed to the intervention if the natural dynamics of an epidemic are not accounted for (figure 1).[4] They also assess potential short and long-term risks of the negative effects of an intervention, such as increasing age at first infection following vaccination[20] and the phenomenon of 'risk compensation' behaviour.[21–23] In conjunction with economic analyses, models help optimise control strategies in the context of limited healthcare budgets.[24]

Models contribute valuable insights at different stages of new preventive product development. They can be used to compare and determine the most beneficial vaccine candidates in the development pipeline,[25–27] guide licensing,[28] and validate and improve study design and analysis for complex research questions.[29][30]

Herein, we introduce commonly used terms the reader may encounter in infectious diseases epidemiology papers that utilise mathematical models, focusing primarily on compartmental models of microparasitic diseases.[1][2][31–33]

BASIC DEFINITIONS
Host
The living entity of interest in the study (eg, human, animal, or even a vector) so long as the pathogen can infect this host and be transmitted onwards. Any type of host (definitive host (in which pathogen multiplication occurs) or intermediate host (pathogen progresses from one life stage to another)) can serve as the 'host' of interest. For example, we may be interested in the proportion of mosquitoes that are infected at any given time when dealing with a question of malaria in a human population—in this case, both mosquito and human serve as hosts in the model (although the former will usually be called a vector). At any given time, the host resides in one of several stages of infection/disease specific to the pathogen.

Pathogen
An organism (eg, prion, bacteria, virus, helminth) capable of infecting (or colonising) the host such that the organism (or its progeny) may be transmitted between infected and susceptible hosts. The mode of transmission may be direct (eg, from human to human or animal to animal) or indirect (with an intermediate host such as a vector, or
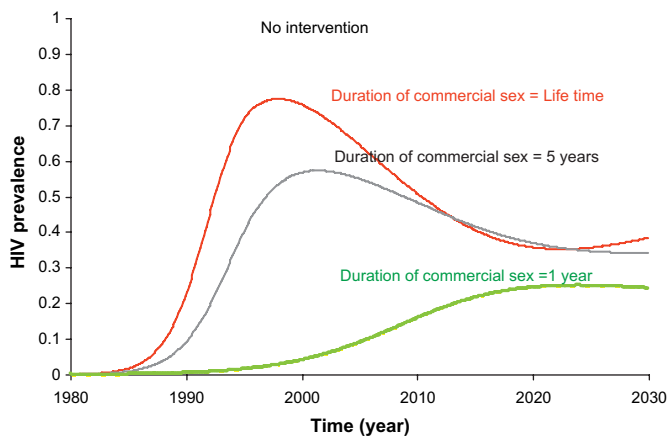
Figure 1 HIV prevalence among female sex workers (FSW) over time in the absence of intervention. In these model simulations the decline in HIV prevalence after 2000 is due to AIDS differential mortality (removal of high-risk individuals from the susceptible population). FSW with many commercial sex partners become rapidly infected with HIV, have a higher mortality rate than non-FSW, but mortality occurs faster than the entry rate of new FSW into the population. Therefore, over time the overall level of risk (ie, mean number of unprotected sex) in the population is reduced, which reduces the potential for HIV spread as well as other co-factor sexually transmitted infections. The magnitude of the HIV epidemic and the speed of the decline is larger if FSW remain in sex work for life compared with the cessation of sex work after an average of 1 or 5 years (all with population-level replacement with new entrants into sex work). Therefore, interpretation of prevalence trends following an intervention can be difficult, and mathematical models offer an opportunity to help evaluate the impact of intervention in the face of the infection's natural dynamics.[4]

inanimate object in which the agent may survive (even replicate) for sufficient periods of time, such as hospital surfaces.

Microparasitic diseases
Classically, infection caused by bacteria, viruses and occasionally parasites (commonly malaria) in which, for the purposes of the model, transmission is assumed to manifest as an all or nothing process and the outcome of interest is the presence of absence of infection in the host.[1]

Macroparasitic diseases
Classically, infection caused by helminths in which transmission (and progression to other biological stages and/or morbidity) is a process dependent on intensity (parasite level) of infection in the infected host and the outcome of interest is the distribution of worm burden in the host.[1]

Model
A simplified representation of a complex real-world event or structure. The goal is to explain or predict the observed phenomena, and in doing so glean the fundamental processes (the essential features) that underpin the system. The use of mathematical theory enables clarity in understanding and communicating these generalisable processes. Different types of models are used to capture the essential behavioural and demographic characteristics of a population and the biological features (natural history) of the infection.

Transmission dynamic model
A model that describes the force of infection as a function of the prevalence of infection and therefore time. The term 'dynamic' describes models in which populations fluctuate as a result of birth, death and migration,[1] but also the manner in which incidence

(force of infection) changes with time, due to positive or negative feedbacks resulting from changes in case counts (ie, prevalence of infection), immunity, or differential mortality of high-risk individuals. This contrasts with the static risk of infection incorporated into some economic models of infectious disease control.[24]

Static model
A model that assumes the incidence of infection is independent of the prevalence of infection and, therefore, time. It often refers to Markov (decision tree) models used in medical decision-making and health economics.[24] Use of the term 'static' in this case is differentiated from the dynamic equilibrium achieved within transmission models of endemic diseases.

TYPES OF MODELS
Compartmental model
A model that categorises hosts into key stages (ie, compartments or states) of infection (eg, susceptible, infected, infectious, recovered) experienced at some point in time during the life of an individual (figure 2, Eqn. 1.1—1.5).

S-E-I-R model
Model-specific terminology that divides the natural history of infection into compartments.

▶ Susceptible (S): The state at which the host is not infected but could become infected. Depending on the disease, entry into the susceptible state can occur at birth, onset of sexual maturity (eg, with sexually transmitted diseases), or loss of protective immunity.
▶ Exposed (latent infection) (E): The state that follows infection. The host is harbouring the pathogen (ie, infected) but cannot transmit it yet (ie, is non-contagious), due to low pathogen burden during this early period of multiplication, or due to local immune mechanisms (eg, latent stage tuberculosis). Clinical manifestations may or may not be present.
▶ Infectious (I): The state at which the host is harbouring the pathogen and can transmit it to another host (directly or via a vector). This contagious period can reflect colonisation or disease (clinical manifestation of infection).
▶ Recovered/removed (R): Describes an immune state wherein the host is no longer susceptible to infection for a fixed (and possibly temporary) period of time, or deceased. While the implications of death versus recovery with long-term immunity are very different in clinical practice, these outcomes have an identical impact on disease dynamics.

Different combinations of these states can be harnessed to represent the biology of infection in models (eg, S-I: for lifelong infection without recovery (eg, herpes simplex virus); S-I-S: for transient infections such as gonorrhoea that do not confer natural immunity; S-I-R: for childhood infections that confer natural immunity) and will influence the behaviour of the system, ie, the time trends in the prevalence and incidence of infection (figure 3). In compartmental models, movement between compartments occurs at an average rate (r). The change in the number of hosts (state variable) per compartment over time is translated using differential (deterministic models) or stochastic (stochastic models) equations solved with analytical or numerical methods.

Incubation
The time from infection until the onset of symptoms, rather than infectiousness.

Deterministic model
In compartmental models defined by differential equations every host follows the same average clinical life course. The overall
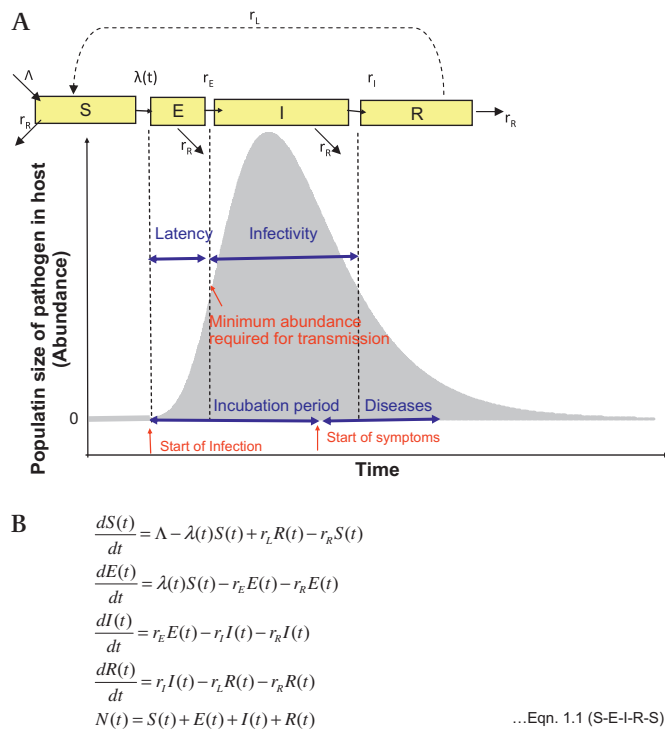
A

B

$$\frac{dS(t)}{dt} = \Lambda - \lambda(t)S(t) + r_L R(t) - r_R S(t)$$

$$\frac{dE(t)}{dt} = \lambda(t)S(t) - r_E E(t) - r_R E(t)$$

$$\frac{dI(t)}{dt} = r_E E(t) - r_I I(t) - r_R I(t)$$

$$\frac{dR(t)}{dt} = r_I I(t) - r_L R(t) - r_R R(t)$$

$$N(t) = S(t) + E(t) + I(t) + R(t)$$

...Eqn. 1.1 (S-E-I-R-S)

Figure 2 (A) Flowchart of a compartmental model illustrating the main stages of infection for an S-E-I-R-S system, following pathogen abundance over time. The transmission process begins when the susceptible host (S) is exposed to the pathogen. Following inoculation, the host is considered infected but does not transmit the pathogen (exposed/latent phase (E)). Exposed individuals become infectious to others if they shed a sufficient quantity of pathogen (I). Over time the pathogen number declines if adequately controlled by the host immune system, successful treatment or natural death. Depending on the agent, the host can therefore recover and become immune (R) for life, or develop immunity for a limited time and become susceptible again (S), or eventually leave the population (death). The host develops symptomatic disease some time after infection depending on the duration of the incubation period (time period between infection and disease). In the flow chart, each compartment represents the number of individuals in the given infection state at time t (state variable). The arrows, quantified by state-specific parameter in Eqn 1 (B), represent the flow of individuals from one state to another. A higher degree of complexity can be included by adding additional compartments (eg, typically three compartments are used to reflect the three main infectivity stages of HIV infectivity/infection: acute/primary, low/asymptomatic, medium/symptomatic). Figure adapted from Aron.[63] (B) Differential equations translating a S-E-I-R-S deterministic model into mathematical terms. Each equation represents the change in the state variables at time t, which depends on the number of individuals in each compartment and the value of the parameters quantifying the flow between compartments. The flow of individuals between stages of infection occurs at an average state-specific per capita rate: r_E=rate of becoming infectious, r_I=rate of developing protective immunity, r_L=rate of loss of immunity, rate of deaths=r_R. The force of infection, $\lambda(t)$, is intrinsically dependant on the number/prevalence of infectious individuals in the population, β (the per contact transmission probability) and c′ (contact rate). The parameter Λ represents the renewal of new susceptible hosts with Λ=0 for closed population and Λ>0 for open population. For a constant population, $\Lambda = r_R N(t)$ (that is, births equal deaths per unit time). Depending on the model complexity, the rate of flow between states (parameter) can be a fixed value, a function of other state variables, or change with time.

course of infection is always the same for all simulations under the predefined model parameters and initial conditions. Therefore, deterministic models reflect the 'average' behaviour of the system. As the state variables are continuous it is possible to have fractions of individuals, and disease is only eradicated asymptotically—making this type of model a poor choice if stochastic effects are of interest.

Stochastic model

Probabilistic model representing stochastic (random) processes. Each transition (movement from one state to another such as infection, recovery, etc) denotes an event that can occur to each individual in a time interval according to a probability that is proportional to the corresponding rate in the deterministic framework. These models can be solved analytically or by simulation. In the later case, each simulation using the same predefined set of constraints results in a different realisation due to chance or stochastic fluctuations. For that reason, models are typically run for a large number of simulations under a given set of constraints in order to estimate the average outcome and the variation around this average. This average outcome generated by the numerous simulations usually provides comparable predictions to deterministic models, especially when the population under study or the force of infection is large.

Stochastic models are used for problems in which random fluctuations are likely to be important: localised outbreaks, small population sizes, rare disease (low force of infection or incidence), and near the start (establishment) or possible end (persistence) of a larger epidemic.[34] Incorporating random fluctuations (noise) enables us to estimate the probability of local extinction (fade-out) and a critical community size at which fade-outs become rare.

State variable

The number of individuals in each compartment at any given time—the value of which varies intrinsically within the system over time (in the absence of any intervention).

Seed or initial conditions

The initial values of the state variable at the start of the simulation and in particular the initial number of infectious individuals (seed) to start a simulated epidemic.

Homogeneous/heterogeneous population

In a homogeneous population, individuals in each disease state are assumed to have the same characteristics. Demographic, biological, or behavioural heterogeneity is introduced by dividing the population into additional subcompartments representing age groups (age-structured model), specific social groups, or behaviour (eg, sexual activity classes). In compartmental models, contact between hosts (or host and vector) is assumed to be instantaneous (ie, negligible duration) although alternative assumptions regarding the length of partnerships (and thereby, concurrency) may be incorporated in individual-based and pair models.

Individual-based (agent-based, microsimulation) model

Individual-based models facilitate the inclusion of different sources of heterogeneity (eg, biological, behavioural) and the representation of intricate contact patterns (eg, duration of relationships, concurrency, within-household contacts, mobility) because each individual host is represented uniquely and can be assigned a specific characteristic. Individual-based models are necessarily stochastic and are solved using simulation techniques.[35–37]
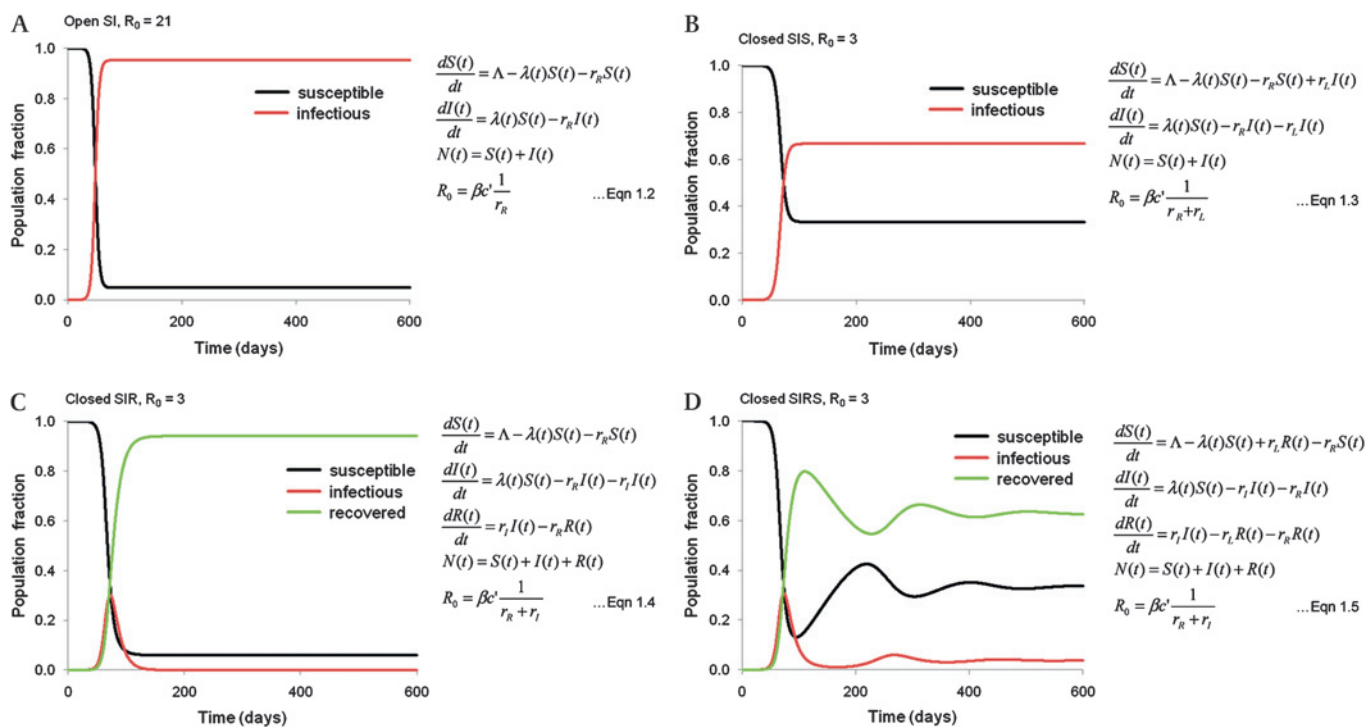
Figure 3 Influence of different model structure on the fraction of susceptible (black), infectious (red), and recovered hosts (green) over time (days) and on R_0 for (A) open S-I (susceptible-infectious), (B) closed S-I-S (susceptible-infectious-susceptible), (C) closed S-I-R (susceptible-infectious-recovered) and (D) closed S-I-R-S (susceptible-infectious-recovered-susceptible) models in a closed homogeneous population. In (A), individuals enter the system as a susceptible host, who upon acquiring infection, remain infected (and infectious) for life. (B) In this model infected individuals recover from infection and become susceptible again. (C) In this model infected and infectious individuals develop lifelong immunity once recovered from infection. (D) In this S-I-R-S model recovered individuals lose their immunity after $1/r_L$ days. Model parameters common to (A–D): $\beta c'=0.3$, $r_I=0.1$. (A) $r_R=1/70$; (B,C,D) $r_{R=0}$. (A,C) $r_L=0$. (B) $r_L=0.1$. (D) $r_L=0.006$.

Network model
Models in which the full contact structure of individuals over a given period of time is explicitly represented and studied (figure 4A). Resulting patterns of disease spread using network theory may be assessed analytically, using concepts from physics, fluid dynamics and population biology.[37–41] Despite considerable advancement in the field, network model complexity remains limited by our empirical understanding of contact networks.[42]

Within-host model
Models that represent a system of cell—pathogen interaction rather than host—pathogen interaction (between-host models).

Pair model
Compartmental models that explicitly define the partnership formation and dissolution process by representing a pair of hosts instead of single hosts. The length of partnerships (rather than instantaneous formation and dissolution) is thus taken into account. The population is subdivided by a single and couples of hosts stratified by disease stages of each host in the pair (eg, infected but single, pair between susceptible and an infected host) (figure 4B).[43 44]

Metapopulation (patch) model
Models that incorporate the within and between interplay of subpopulations (disaggregated in space[45] or social networks).[46] For example, cities (with individuals residing within their confines) may be linked to other cities by a small number of individuals who travel between the two. By separating yet linking these two different groups of hosts, we may obtain better realisations of the transmission dynamics of infection over space and time (figure 4C).

Parameters
Variables that determine the rates of movement (or probability of an event) between states in the model. Parameter values are independent of the state variables but can be subgroup and time specific. Model parameters are estimated from empirical data, literature review, expert opinion, or as a result of model fitting.

Forcing (extrinsic forcing)
The modification of parameter values that reflect behavioural or environmental change over time. For example, oscillatory forcing functions allow models to reproduce periodic surges in incidence that mimic the 'seasonality' of influenza.[47]

KEY CONCEPTS
Force of infection
$\lambda(t)$, given as the per capita incidence rate of infection per susceptible host and the function is central to transmission dynamics models. It summarises the transmission process between infected and susceptible hosts, and depends on the prevalence of infection in the population, $I(t)/N(t)$, the contact rate (c'), and the transmission probability per contact (β). For a compartmental model under an assumption of homogenous mixing, $\lambda(t)$ is given by:

$$\lambda(t) = \beta c' \frac{I(t)}{N(t)} \qquad \ldots \quad \text{Eqn 1}$$

where $I(t)$ and $N(t)$ are the number of infectious and total population of hosts at time t, respectively. Transmission
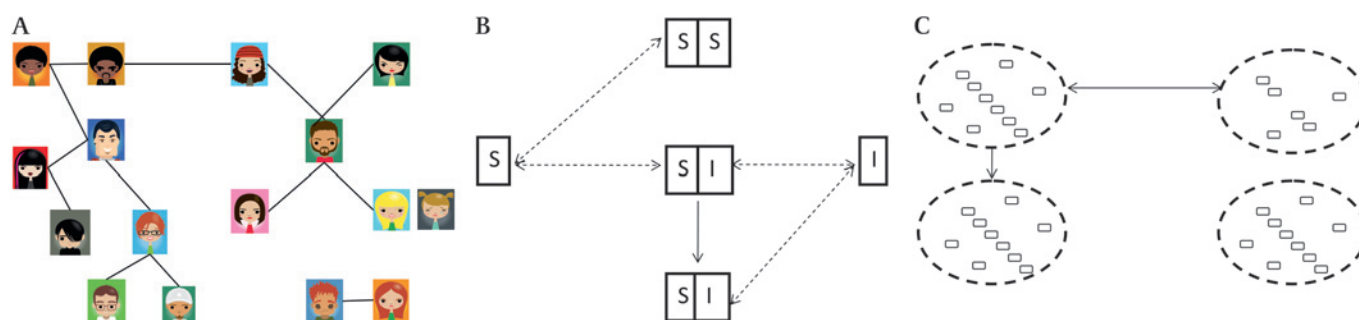
Figure 4 (A) Schematic representation of the contact structure of an individual-based network model of a small heterosexual population of 16 individuals. The contacts or partnerships are represents by the lines/links between individuals. In network models, the links between individuals are tracked over time. In this example, the network has three components of size one, two and 13 individuals. The network has a total of 13 links and an average of 1.625 partners per individual. (B) Schematic representation of a pair model: individuals transition between a single state and a pair state (dashed arrows) in addition to moving through susceptible infectious compartments (given by subscripts S and I, respectively). Transmission results in a discordant pair now containing two infectious persons (black arrow). The formation and dissolution of partnerships are given by the inverse of the duration of time spent in the single and pair states. (C) In the metapopulation model, subgroups are enclosed within the oval regions, and are linked by the connection of one or more individuals between subgroups. Subgroups may be defined by geographical, temporal, or social network demarcations.[43 44 46]

between infecteds and susceptibles will also depend on how the contact structure between the two is expected to change with population size (whether transmission is density dependent or frequency dependent).

► Density-dependent (pseudo mass action) transmission: Contacts are assumed to be proportional to the population size/density (C' = CN ≈ N and Eqn 1 reduces to $\lambda(t) = \beta I(t)$). The number of new infecteds, NI(t), with NI(t) = $\lambda(t)*S(t)$, thus depends on the number of infectious and susceptible, S(t), hosts in the population, assuming random mixing.[1 15 16]

► Frequency-dependent (mass action) transmission: The number of contacts is assumed to be independent of population size (C' = $\lambda(t)$ as Eqn 1).[1]

Incorporating heterogeneities in contacts that arise from age, space, or behavioural differences between individuals can produce more realistic incidence trends (requires modifying Eqn 1).[31]

Transmission probability (β)
The probability that an infectious host will infect a susceptible one following an adequate contact.

Contacts
The type of contacts required for transmission of the infection depends on the mode of transmission (eg, airborne, sexual). For example, physical contact or close conversation for directly transmitted infection (eg, measles, influenza) or physical contact for sexually transmitted infections, insect bites for vector-borne infection.

Contact (mixing) pattern
The contact pattern depends on the frequency (c') and variability of contact per unit time in the host population.

► Homogeneous—random: Each host has the same average rate of contact with another host. Contacts between hosts are made uniformly with equal probability.

► Heterogeneous—non-random: Some hosts have, on average, a higher rate of contact than others due to social, spatial, or behavioural differences, introducing heterogeneity that can be captured in models. Contacts between hosts are non-random, ie, made with unequal probability.

In the presence of heterogeneity, contact between hosts can occur across a spectrum—from assortative to proportional to disassortative—of mixing patterns describing who mixes (makes contact) with whom.[48]

► Assortative mixing: Hosts of similar risk/subgroups are more likely to mix with each other.[48]

► Proportional mixing: Hosts of any risk group have a probability of mixing with a member of another risk/subgroup based on the size and the contact rates of the latter.[48] If all groups are of equal size, hosts with higher contact rates are more likely to be chosen/contacted by any given host.

► Disassortative mixing: Hosts are more likely to interact with members of a risk group other than their own.[48]

The mixing pattern influences the dynamics of infection at the different stages of an epidemic.[32 49]

Epidemic stages
Establishment, spread, equilibrium prevalence and persistence.

Basic reproductive (reproduction) number/ratio (R₀)
The average number of hosts that become infected as a result of the entry of one infectious host into a completely susceptible population in the absence of intervention. By definition, R_0 is an idealised quantity that is virtually impossible to observe directly but can be estimated indirectly.[31] For an infectious disease to establish and propagate in the population, R_0 must be greater than 1. The main components of this ratio include the transmission probability, duration of infectiousness (D, where D=1/ rate of recovery from infection), and contact pattern between hosts.[1] The expression for R_0 depends on the model structure; for a simple S-E-I-R model (in an open population with homogenous mixing, Eqn 1.1), R_0 is given by:

$$R_0 = \beta c'D \qquad \ldots \quad \text{Eqn 2}$$

Infectious disease control programmes are effective because they can reduce any of these components: β can be reduced through personal protective items (eg, N95 masks, condoms); c' can be reduced by quarantine, isolation, or behaviour change; and D can be reduced through treatment (as with tuberculosis). In malaria transmission models, the notation Z_0 is sometimes used to define the average number of infected mosquitoes whose infections are the result of one infected mosquito in a finite human population.[50]

Effective reproductive number/ratio (Rₜ)
The average number of new infections caused by an infection at time t in a population in which not all hosts are susceptible. For

endemic models with homogenous mixing, $R_t=R_0S(t)$, where $S(t)$ is the fraction of hosts who are susceptible at time t. R_t varies with time and equals 1 when the system has reached an endemic equilibrium (steady state).[1 51]

Eradication fraction (herd immunity threshold) (S_h)

Under a compartmental framework with homogenous mixing, the minimum fraction of susceptibles that must be immune (or vaccinated at birth (assuming 100% vaccine efficacy)) to reduce R_t below 1 and eradicate infection; that is, by the removal of susceptible hosts[51−53] ($R_t=R_0S(t),<1$ and $S_h=1−1/R_0$).[31 32 51] When heterogeneity in mixing is introduced, this relationship holds so long as vaccination is applied at random.[31] An infectious process with a lower R_0 requires a lower eradication fraction. Host 'immunity' may be acquired following infection or vaccination.

Elimination/eradication

Theoretically, elimination or eradication can be achieved by reducing R_0 below 1 or maintaining $R_t<1$ by keeping $S(t)<S_h$. Although the term 'eradication' is often used, many models focus on geographically (or socially) distinct populations, and the term 'elimination', in an applied context, is more appropriate. Outbreaks can sometime occur even when R_0 is less than 1 as a result of random fluctuations in the number of new infections generated at any given time point.[32]

Herd immunity

Strictly speaking, a population has herd immunity if the eradication fraction has been achieved. However, it is often used to refer to the indirect population-level effects that vaccinated hosts confer to susceptible hosts as a result of reduced force of infection as the reduced prevalence (among vaccinated persons or those having received the intervention) simultaneously reduces the exposure of still susceptible hosts to infection.[51−54]

Stochastic eradication/elimination or epidemic fade-out

A disease may become locally extinct by chance alone (even if $R_0>1$) in small population or when the disease is rare; a phenomenon simulated within stochastic models but not in deterministic models. In the latter, the basic reproductive number (R_0) is the only criterion for the persistence of an infection in a community.

Doubling time

At (or very near) the start of a growing outbreak, the time for the number of infecteds to increase twofold.

SIMULATION

A technique that uses mathematics to mimic the operations of the real-world process described by our model. The term is commonly used within stochastic frameworks when models are solved using a predefined computational algorithm (such as Monte Carlo simulations) under the same set of parameter values (recall that each simulation can produce a varied outcomes due to random chance). In a deterministic framework, simulations usually refer to model runs under different parameter values (because each simulation provides the same output for a unique set of parameters).

Monte Carlo simulation

A type of computational algorithm used with stochastic (compartmental, network, individual-based) models in which the probability of any given transition is randomly sampled from a predefined probability distribution.

Model fitting

Refers to calibrating model parameters (parameterisation) to the observed epidemiological data (eg, prevalence of infection). It can be performed using different statistical techniques such as least squares, maximum likelihood methods, or using a Bayesian framework. In the latter, parameter sets are sampled from a predefined distribution of plausible values (prior distribution). Only model realisations that are likely (assessed with a fitting criteria) to be compared to the observed outcome are retained to make predictions that reflect uncertainty in parameter assumptions (posterior distribution of model outcomes).

Analytical solution

The procedure of solving equations using mathematical theory to obtain a closed form solution for the relationship between model parameters and an outcome of interest (eg, equilibrium prevalence of infection, R_0). Generally, these methods are used with relatively simple deterministic or stochastic models and recently even with network models.[31 39]

Numerical solution

In the context of compartmental models, the term refers to the solution of a set of differential equations (deterministic or stochastic) obtained by numerical integration using approximation techniques (eg, Euler, Runge Kutta) and mostly with computer programming.[31] In the context of stochastic models (compartmental, network or individual based) solved by Monte Carlo simulations, it refers to the realisations of the probability events at each time point.

Sensitivity analysis

An assessment of the effect of the input parameters on model predictions over the range of parameter values of interest. It helps evaluate the robustness of the results to changes in conditions or external validity of model outcomes outside the simulation setting (ie, different population). Can be conducted as one-way (adjustment of one parameter while all others are held constant) or by multivariate analysis. Note that the range explored can influence model results and conclusions. The results of sensitivity analysis tend to be more general than those of uncertainty analysis because of the wider range explored.

Uncertainty analysis

Although the terms sensitivity and uncertainty analysis are often used interchangeably they are slightly different. Uncertainty analysis is a type of sensitivity analysis in which the input parameter range is narrower than in sensitivity analysis, restricted to parameter values that are realistic for the specific population studied, and preferably obtained after model fitting. It helps elucidate which parameters influence model projections and lead to uncertainty in the outcome of interest in the specific population studied. The results of uncertainty analysis are more context specific than those of sensitivity analysis. Uncertainty and sensitivity analysis can give different results because they cover different parameter ranges.

Latin hypercube sampling

An efficient sampling method for multidimensional parameter distribution often used in multivariate uncertainty analysis to minimise the number of parameter sets explored. When 'P' parameters are explored, each parameter space is divided into 'n' equal probability intervals. For each parameter, one interval is chosen randomly without replacement. This process is repeated N times to produce N multivariate parameter sets.[55 56]

What is already known on this subject

Research aimed at investigating the population-level epidemiology of infectious diseases has increased in recent decades, expanded by the use of mathematical models.

What this study adds

This review highlights terminology common to mathematical models in an effort to facilitate dialogue between modellers, clinicians and public health specialists.

MODEL DEVELOPMENT
Complexity and limits
The degree of complexity required in a model is often a matter of contention and, ultimately, a subjective decision. By definition, a model is a simplification of reality and all models are based on assumptions, even statistical models. The degree of complexity required is intrinsically linked to the research question and limited by the level of computational difficulties introduced. The ideal model should be simple (parsimonious) yet strive to represent reality as adequately as possible by capturing the key demographic, biological and behavioural features necessary to address the question of interest.[57] It has been argued that in many instances compartmental models are simpler to develop and solve, allowing for analytical tractability.[31] As previously discussed, some classes of network models that represent more complex contact structures can also be solved analytically.[31 39] However, when compartmental or network models become too complex, it is easier to use individual-based models.

Generally, simpler models are used to gain a general and intuitive understanding of key principles when little is known about an infection[58] or when studying a novel research question. More complex models are preferred when detailed information is available, accurate projections are required, or new hypotheses necessitate testing (eg, importance of heterogeneity of a given risk factor). However, more complex models (eg, agent-based models) are not necessarily better than simpler ones because greater uncertainty is introduced as the number of assumptions and parameters increases and data to validate models become scarce, making model findings difficult to interpret.

In recent years, due to increased computer power and data availability, the application of statistical methods to 'test and validate' model assumptions and model simplification is an advancing area of study.[59–61] Understanding host behaviour and host–pathogen interactions and collecting empiric data on outcomes strengthen the utility and validation of models.[62]

Model interpretation
Model findings must be interpreted within the limits of their assumptions, parameter values and initial conditions. Good modelling articles ensure that model assumptions and definitions of each component are described in sufficient detail, enabling transparency and reproducibility.

REFERENCES
1. **Anderson RM,** May RM. *Infectious diseases of humans: dynamics and control*. Oxford: Oxford University Press, 1991.
2. **Barreto ML,** Teixeira MG, Carmo EH. Infectious diseases epidemiology. *J Epidemiol Community Health* 2006;**60**:192–5.
3. **Geisecke J.** *Modern infectious diseases epidemiology*, 2 edn. London: Edward Arnold, 2002.
4. **Boily MC,** Lowndes CM, Vickerman P, *et al*. Evaluating large-scale HIV prevention interventions: study design for an integrated mathematical modelling approach. *Sex Transm Infect* 2007;**83**:582–9.
5. **Boily M,** Masse B. Mathematical models of disease transmission: a precious tool for the study of sexually transmitted diseases. *Can J Public Health* 1997;**88**:255–65.
6. **Kermack WO,** McKendrick AG. Contributions to the mathematical theory of epidemics II — the problem of endemicity. *Proceedings of the Royal Society of London* 1932;**138**:55–83.
7. **Kermack WO,** McKendrick AG. Contributions to the mathematical theory of epidemics III — further studies of this problem of endemicity. *Proceedings of the Royal Society of London* 1933;**141**:94–122.
8. **Ross R.** Some quantitative studies in epidemiology. *Nature* 1911;**87**:466.
9. **Bartlett MS.** Deterministic and stochastic models for recurrent epidemics. *Proceedings Third Berkeley Symposium on Mathematical Statistics and Probability* 1956:81–109.
10. **Kermack WO,** McKendrick AG. Contributions to the mathematical theory of epidemics — I. *Bull Math Biol* 1927;**53**:33–55.
11. **Macdonald G.** The analysis of equilibrium in malaria. *Trop Dis Bull* 1952;**49**:813–29.
12. **Heesterbeek JAP.** A brief history of R0 and a recipe for its calcluation. *Acta Biotheor* 2002;**50**:189–204.
13. **Anderson RM.** The Kermack–Mckendrick epidemic threshold theorem — discussion. *Bull Math Biol* 1991;**53**:3–32.
14. **Soper HE.** The interpretation of periodicity in disease prevalence. *J R Stat Soc* 1929;**92**:34–73.
15. **Hamer WH.** The Milroy lectures on epidemic disease in England — the evidence of variability and of persistency of type. *Lancet* 1906;**1**:733–9.
16. **Ross R.** *The prevention of malaria*, 2 edn. London: Murray, 1911.
17. **Hethcote HW,** Thieme HR. Stability of the endemic equilibrium in epidemic models with subpopulations. *Math Biosci* 1985;**75**:205–27.
18. **Edmunds WJ,** Gay NJ, Kretzschmar M, *et al*. The pre-vaccination epidemiology of measles, mumps and rubella in Europe: implications for modeling studies. *Epidemiol Infect* 2000;**125**:635–50.
19. **Lloyd-Smith JO,** Schreiber SJ, Kopp PE, *et al*. Superspreading and the effect of individual variation on disease emergence. *Nature* 2005;**438**:355–9.
20. **Panagiotopoulos T,** Antoniadou I, Valassi-Adam E, *et al*. Increase in congenital rubella occurrence after immunisation in Greece: retrospective survey and systematic review. *BMJ* 1999;**319**:1462–7.
21. **Richens J,** Imrie J, Copas A. Condoms and seat belts: the parallels and the lessons. *Lancet* 2000;**355**:400.
22. **Boily MC,** Bastos FI, Desai K, *et al*. Changes in the transmission dynamics of the HIV epidemic after the wide-scale use of antiretroviral therapy could explain increases in sexually transmitted infections — results from mathematical models. *Sex Transm Dis* 2004;**31**:100–12.
23. **Boily MC,** Godin G, Hogben M, *et al*. The impact of the transmission dynamics of the HIV/AIDS epidemic on sexual behaviour: a new hypothesis to explain recent increases in risk taking-behaviour among men who have sex with men. *Med Hypotheses* 2005;**65**:215–26.
24. **Brisson M,** Edmunds WJ. Economic evaluation of vaccination programs: the impact of herd-immunity. *Med Decis Making* 2003;**23**:76–82.
25. **Penny MA,** Maire N, Studer A, *et al*. What should vaccine developers ask? Simulation of the effectiveness of malaria vaccines. *PLoS One* 2008;**3**:e3193.
26. **Maire N,** Tediosi F, Ross A, *et al*. Predictions of the epidemiologic impact of introducing a pre-erythrocytic vaccine into the expanded program on immunization in sub-Saharan Africa. *Am J Trop Med Hyg* 2006;**75**:111–18.
27. **Smith T,** Killeen GF, Maire N, Ross A, *et al*. Mathematical modeling of the impact of malaria vaccines on the clinical epidemiology and natural history of *Plasmodium falciparum* malaria: overview. *Am J Trop Med Hyg* 2006;**75**:1–10.
28. **Boily MC,** Abu-Raddad L, Desai K, *et al*. Measuring the public-health impact of candidate HIV vaccines as part of the licensing process. *Lancet Infect Dis* 2008;**8**:200–7.
29. **Boily MC,** Anderson RM. Human immunodeficiency virus transmission and the role of other sexually transmitted diseases — measures of association and study design. *Sex Transm Dis* 1996;**23**:312–32.
30. **Desai K,** Boily MC, Masse B, *et al*. Using transmission dynamics models to validate vaccine efficacy measures prior to conducting HIV vaccine efficacy trials. In: Abello J, ed. *Discrete methods in epidemiology*. Providence: American Mathematical Society 2006.
31. **Keeling MJ,** Rohani P. *Modelling infectious diseases in humans and animals*. Princeton: Princeton University Press, 2008.
32. **Grassly NC,** Fraser C. Mathematical models of infectious disease transmission. *Nat Rev Micro* 2008;**6**:477.
33. **Brauer F.** Compartmental models in epidemiology. In: Brauer F, Driessche P, Wu J, eds. *Mathematical epidemiology*. Berlin: Springer-Verlag, 2008.
34. **Britton T,** Lindenstrand D. Epidemic modelling: aspects where stochasticity matters. *Math Biosci* 2009;**222**:109–16.
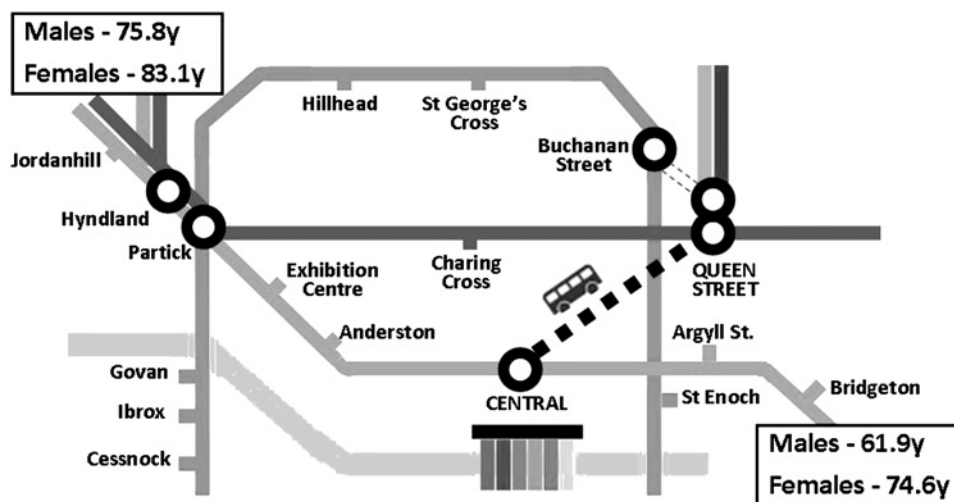
35. **Boily MC,** Poulin R, Masse B. Some methodological issues in the study of sexual networks — from model to data to model. *Sex Transm Dis* 2000;**27**:558—71.
36. **Germann TC,** Kadau K, Longini IM, *et al*. Mitigation strategies for pandemic influenza in the United States. *Proc Natl Acad Sci U S A* 2006;**103**:5935—40.
37. **Eubank S,** Guclu H, Kumar VSA, *et al*. Modelling disease outbreaks in realistic urban social networks. *Nature* 2004;**429**:180—4.
38. **Boily MC,** Asghar Z, Garske T, *et al*. Influence of selected formation rules for finite population networks with fixed macrostructures: implications for individual-based model of infectious diseases. *Math Popul Stud* 2007;**14**:237—67.
39. **Noel PA,** Davoudi B, Brunham RC, *et al*. Time evolution of epidemic disease on finite and infinite networks. *Phys Rev E* 2009;**79**:026101.
40. **Allard A,** Noel PA, Dube LJ, Pourbohloul B. Heterogeneous bond percolation on multitype networks with an application to epidemic dynamics. *Phys Rev E* 2009;**79**:036113.
41. **Keeling MJ,** Danon L, Vernon MC, *et al*. Individual identity and movement networks for disease metapopulations. *Proc Natl Acad Sci U S A* 2010;**107**:8866—70.
42. **Read JM,** Eames KTD, Edmunds WJ. Dynamic social networks and the implications for the spread of infectious disease. *J R Soc Interface* 2008;**5**:1001—7.
43. **Chen MI,** Ghani AC, Edmunds J. Mind the gap: the role of time between sex with two consecutive partners on the transmission dynamics of gonorrhea. *Sex Transm Dis* 2008;**35**:435—44.
44. **Kretzschmar M.** Sexual network structure and sexually transmitted disease prevention — a modeling perspective. *Sex Transm Dis* 2000;**27**:627—35.
45. **Colizza V,** Barrat A, Barthelemy M, *et al*. Predictability and epidemic pathways in global outbreaks of infectious diseases: the SARS case study. *BMC Med* 2007;**5**:34.
46. **Chen MI,** Ghani AC, Edmunds WJ. A metapopulation modelling framework for gonorrhoea and other sexually transmitted infections in heterosexual populations. *J R Soc Interface* 2009;**6**:775—91.
47. **Dushoff J,** Plotkin JB, Simon AL, *et al*. Dynamical resonance can account for seasonality of influenza epidemics. *Proc Natl Acad Sci U S A* 2004;**101**:16915.
48. **Anderson RM,** Gupta S, Ng W. The significance of sexual partner contact networks for the transmission dynamics of HIV. *J Acquir Immune Defic Syndr* 1990;**3**:417—29.
49. **Nagelkerke N,** de Vlas SJ, Jha P, *et al*. Heterogeneity in host HIV susceptibility as a potential contributor to recent HIV prevalence declines in Africa. *AIDS* 2009;**23**:125—30.
50. **Smith DL,** McKenzie FE, Snow RW, *et al*. Revisiting the basic reproductive number for malaria and its implications for malaria control. *PLoS Biol* 2007;**5**:531—42.
51. **Garnett GP.** Role of herd immunity in determining the effect of vaccines against sexually transmitted disease. *J Infect Dis* 2005;**191**(Suppl 1):S97—S106.
52. **Fox JP,** Elveback L, Scott W, *et al*. Herd immunity — basic concept and relevance to public health immunization practices. *Am J Epidemiol* 1971;**94**:179—89.
53. **Fine PE.** Herd immunity—history, theory, practice. *Epidemiol Rev* 1993;**15**:265—302.
54. **Gay NJ,** Hesketh LM, Morgancapner P, *et al*. Interpretation of serological surveillance data for measles using mathematical models — implications for vaccine strategy. *Epidemiol Infect* 1995;**115**:139—56.
55. **Marino S,** Hogue IB, Ray CJ, *et al*. A methodology for performing global uncertainty and sensitivity analysis in systems biology. *J Theor Biol* 2008;**254**:178—96.
56. **Iman RL,** Helton JC. An investigation of uncertainty and sensitivity analysis techniques for computer models. *Risk Analysis* 1988;**8**:71—90.
57. **Koopman JS.** Modeling infection transmission — the pursuit of complexities that matter. *Epidemiology* 2002;**13**:622—4.
58. **Choi BC,** Pak AW. A simple approximate mathematical model to predict the number of severe acute respiratory syndrome cases and deaths. *J Epidemiol Community Health* 2003;**57**:831—5.
59. **Foss AM,** Vickerman PT, Chalabi Z, *et al*. Dynamic modeling of herpes simplex virus type-2 (HSV-2) transmission: issues in structural uncertainty. *Bull Math Biol* 2009;**71**:720—49.
60. **Blower SM,** Dowlatabadi H. Sensitivity and uncertainty analysis of complex models of disease transmission — an HIV model as an example. *Int Stat Rev* 1994;**62**:229—43.
61. **Hoare A,** Regan D, Wilson D. Sampling and sensitivity analyses tools (SaSAT) for computational modelling. *Theor Biol Med Model* 2008;**5**:4.
62. **Garnett GP,** Brunham RC. Magic bullets need accurate guns — syphilis eradication, elimination, and control. *Microbes Infect* 1999;**1**:395—404.
63. **Aron JL.** Mathematical modeling: the dynamics of infection. In: Nelson KE, Williams CM, Graham NM, eds. *Infectious diseases epidemiology: theory and practice*. Gaithersburg: Aspen Publishers, 2001.

Gallery

Illustrating health inequalities in Glasgow

The stark health inequalities in Glasgow are demonstrated by the drop in life expectancy of 2.0 years for men and 1.2 years for women for each station on the railway line between Jordanhill and Bridgeton (more than double the equivalent inequality along London's Jubilee underground line from Westminster to Canning Town) (Figure 1).

G McCartney

Correspondence to Dr Gerry McCartney, NHS Health Scotland, Elphinstone House, 65 West Regent Street, Glasgow, Scotland, UK; gmccartney@nhs.net

Figure 1 Life expectancy differences between areas six railway stations apart in Glasgow.

Life expectancy data refers to 2001-5 and was extracted from the GCPH community health and well-being profiles. Adapted from the SPT travel map by Gerry McCartney.