EPIDEMIC MODELS

Their Structure and Relation to Data

edited by

Denis Mollison

Heriot-Watt University

CAMBRIDGE
UNIVERSITY PRESS

CONTENTS

Part 4: Heterogeneity in human diseases

Part 5: Data analysis: estimation and prediction

Introduction

Denis Mollison

Infectious diseases have played a major role in evolution, of which their influence on human history is just the most recent and best documented example (McNeill 1976, Crosby 1986). The spectacular success of humans in dominating the world's ecology has meant that they — and their domestic animals and crops — provide an unprecedentedly rich resource for parasites. Not surprisingly, parasites have evolved, and continue to evolve, to exploit this resource. Meanwhile the earth's remaining natural ecosystems, often weakened by human influence, are in relative terms even more threatened by the spread of infectious diseases.

Until fairly recently it was possible to look back on over 100 years of fairly steady progress in the control of disease — including the introduction of antibiotics, and vaccination programmes leading to great reductions of the main 'childhood diseases' in developed countries, and to the worldwide eradication of smallpox — and to extrapolate to a future in which infectious disease had been conquered. The rise of AIDS and of vaccine-resistant strains of a number of the diseases thought to be no longer a threat has demonstrated that such simple optimism is unjustified. The war against disease is going to be a long one, in which lasting success can only be won through better understanding.

Epidemics involve processes at all scales, from the global population, through the individual level, right down to the behaviour of the immune system. The resulting dynamical systems are characteristically highly nonlinear, stochastic and subject to natural selection. The corresponding mathematical problems concern both the structure of models that are needed to describe this dynamical diversity, and also the modelling and statistical methods required to deal with heterogeneity, in space, time, social contact and so on. A third, vital and complementary area of work is the adaptation of developments in these areas to particular applied problems.

Much current work is driven by the great concern centred on the AIDS epidemic. This has raised many stimulating technical questions: AIDS is particularly difficult to model because of the complex heterogeneity of the relevant contact structures, and difficulties in collecting and analysing data which are enhanced by its long incubation period and by the social sensitivity associated with sexually transmitted diseases. Nevertheless, similar problems arise for all diseases, and any solutions or improvements in methodology will be of wider applicability.

As well as AIDS, the areas of application include: epidemics of other sexually transmitted diseases; childhood diseases like measles or polio and the

question of appropriate vaccination policies; strategies for disease control in animals; the problems involved in the spatial spread of disease (e.g. rabies); and so on.

Improved computer technology has provided better data bases, and opened up new possibilities for the use of computationally-intensive methods in the analysis of data. It has also allowed the simulation of more realistic models; however this increases the need to improve our understanding of how a relatively few key components can drive the dynamics of such models.

1   The conceptual framework

The present volume opens with a review by Dietz of some of the basic concepts and problems of epidemic modelling; and of a wide range of questions related to vaccination strategies which should in principle be amenable to mathematical analysis.

This is followed by a chapter on the structure of epidemic models. We cannot begin to make use of models until we understand how their dynamics depend on basic components, and how sensitive they are to the way in which they are incorporated into the model; this is especially important where we extrapolate on the basis of a model, for instance for prediction and control.

Stochastic epidemic models, although generally more realistic than deterministic ones, have often been seen as too complex for analysis because of the level of detail they require, for instance relating to the probability that an individual with the disease will infect a particular subset of the population. However, in recent years it has been realised that this micro-structure can be turned to considerable advantage, especially when comparing one model with another, by the simple and elegant technique of *coupling*. Ball reviews this technique in the epidemic context, while Lefèvre and Picard use similar stochastic methods to generalise the classic Reed-Frost model for an epidemic in a closed population.

Nåsell discusses how the concept of an epidemic threshold should be extended to stochastic models, where the question of whether an outbreak can be regarded as 'large' is not as clearcut as in the simplified deterministic case. De Jong, Diekmann and Heesterbeek consider the transmission rate of infection, arguing from both theory and data in favour of the 'true mass-action' assumption, namely that the rate should rise with population *density* but not depend on the overall population size.

Part 1 concludes with a wide-ranging survey by Diekmann, Heesterbeek and Metz of the various possible types of disease dynamics, with special emphasis on the influence of heterogeneity and on how the appropriate modelling approach depends on the population and time scales of interest.

2   Spatial models

Recent developments in spatial stochastic processes and differential equations have applications to animal and plant diseases, where key questions concern long-distance contacts, endemic patterns, and the effectiveness of control zones.

Cliff reviews mapping methods for identifying the spatial corridors used by diseases, and the importance of different components in their spread. This is illustrated by analysis of the spread of measles and influenza over a wide range of population densities and spatial scales. Metz and van den Bosch present a very general framework in which the velocity of spread of epidemics and populations can be calculated from information at the individual level, and illustrate this with applications to a wide variety of cases, mainly of plant and animal diseases.

The breadth of applicability of Metz's approach is achieved at the expense of some loss of accuracy, particularly in neglecting stochastic effects. Durrett shows how significant progress is now being made in analysing the more accurate but relatively intractable stochastic models for the spatial spread of epidemics.

This part concludes with two more specialised papers. Daniels shows, for nonlinear deterministic models, how an elegant perturbation approach can be used to calculate the detailed shape of epidemic wavefronts. Billard *et al.* generalise classical logistic models for the rate of increase of infected tissue in epidemic plant diseases, so as to incorporate randomness and additional infections from the surrounding environment, and show how they can be fitted to experimental data on the spread of *anthracnose* in *Stylosanthes scabra*.

3   Nonlinear time and space-time dynamics

Data on recurrent epidemics of human 'childhood' diseases, especially measles, have played a key part in discussions of the role of nonlinearity and chaos in biological population dynamics. Ellner *et al.* review methods for the detection of nonlinearity and chaos, and apply them to a collection of data sets for four different diseases. Their broad conclusion is that evidence for chaos is generally lacking in these data sets, though measles may be close to it.

Grenfell *et al.* look at measles modelling in more detail: seasonal forcing has a destabilising effect on the dynamics, while the introduction of a realistic age structure to the model promotes stability. They also discuss the more problematic effects of spatial heterogeneity and of variations in the birth rate, in the light of data from a number of major cities of the developed world.

4   Heterogeneity in human diseases

It is vital for successful modelling to take appropriate account of both *population heterogeneity*, i.e. variations between individuals in parameters such as their contact rate or susceptibility (especially important for AIDS and other STDs); and *heterogeneity of mixing*, i.e. how the pattern of contacts depends on spatial location or on the connectivity of social networks. The development of satisfactory models for the contact structures of human diseases offers a challenging opportunity for collaborative research between workers with expertise ranging from random graphs to sociology.

Here, Levin gives an elegant analysis of the general mixing problem which is basic to the modelling of pairing in a heterogeneous population, and discusses the dynamics of formation and dissolution of pairs. Jacquez *et al.* analyse the relation between the epidemic threshold for a population composed of differing subgroups and the threshold of the individual subgroups.

Morris offers a very different approach to the same problem area, showing how log-linear models can provide a framework for estimation of mixing parameters from available sociological data, and for inference on the resulting patterns of transmission. This is illustrated from sexual preference data, and the implications for the spread of sexually transmitted diseases are discussed.

Gupta *et al.* consider a quite different kind of heterogeneity, in the disease agent, showing how this can make a substantial difference to estimates of the basic reproductive ratio in the case of malaria.

5   Data analysis: estimation and prediction

Even for relatively simple epidemic situations, statistical inference faces difficulties including identifiability problems, very complicated likelihood functions, multiple sources of errors, difficulties with the interpretation of 'parameters', and the need to make predictions for the epidemic from which the data are gathered. Becker reviews and illustrates methods to overcome such difficulties, including the use of effective simplifying assumptions, addressing estimation problems in a more specific context, and the use of martingale methods.

For many diseases, complex models are required to aid our understanding of the mechanics of transmission and evaluation of potential control strategies, but the crudeness of data makes the use of relatively simple models the only sensible approach to model fitting and projection. Cairns discusses how these two conflicting requirements may be reconciled through the identification of a minimal set of Primary Components, that is, functions of the basic parameters which dictate epidemic dynamics; illustrating the technique through reference to some simple models for the spread of AIDS.

The World Health Organisation's Onchocerciasis Control Programme in West Africa provides a large scale example of the use of epidemic modelling in disease control. Onchocerciasis (river blindness) is a complex disease – its dynamics depend on the interaction of three populations, the parasites, the human hosts and the vectors – which affects an estimated 18 million people, mainly in west and central Africa. Remme *et al.* describe how epidemic models have been used in the planning, implementation, evaluation and timely adjustment of control strategies.

A major problem in disease control is that vaccines may fail to confer immunity, or may only confer partial immunity. Longini *et al.* review the use of statistical methods to estimate vaccine efficiency both for single outbreaks and endemic conditions.

Norman Bailey, noted for his pioneering work in epidemic modelling (Bailey 1957, 1975), rounds off this final section of the book with a discussion of the vital practical issues involved in integrating modelling with public health decision-making and planning, illustrated from his recent work on the spread and control of HIV/AIDS.

Future work

During the NATO workshop on *Epidemic Models* there were discussions to identify key problem areas: the conclusions of these sessions are given as an Appendix, under the headings 'Model structure', 'Heterogeneity' and 'Data analysis and prediction'. A further review, in part based on these discussions, is given in Mollison, Isham and Grenfell (1994).

Many theoretical challenges remain, for instance to expand the scope of epidemic modelling to deal with the co-evolution of diseases and the species they affect (see e.g. Hamilton and Howard 1994). On the practical side, we need to use our understanding to achieve greater success in the war against disease. It seems ironic, when there is rightly much concern over the number of species disappearing from the earth as an incidental consequence of human activities, that it should be so hard to eliminate a species deliberately, with the smallpox virus still our only complete success.

References for Preface and Introduction

Bailey, N.T.J. (1957) *The Mathematical Theory of Epidemics*, Griffin, London.

Bailey, N.T.J. (1975) *The Mathematical Theory of Infectious Diseases and its Applications*, Griffin, London.

Crosby, A.W. (1986) *Ecological Imperialism: the biological expansion of Europe, 900–1900*, Cambridge University Press, Cambridge.

Grenfell, B.T. and Dobson, A. (eds.) (1995) *Ecology of Infectious Diseases in Natural Populations*, Cambridge University Press, Cambridge.

Hamilton, W.D. and Howard, J.C. (eds.) (1994) *Infection, Polymorphism and Evolution, Phil. Trans. R. Soc. Lond. B*, to appear.

Isham, V. and Medley, G. (eds.) (1995) *Models for Infectious Human Diseases: their Structure and Relation to Data*, Cambridge University Press, Cambridge.

Mollison, D., Isham, V. and Grenfell, B.T. (1994) 'Epidemics: models and data' (with Discussion), *J. R. Statist. Soc. A* **157**, 115–149.

McNeill, William H. (1976) *Plagues and Peoples*, Doubleday, New York.