Epidemiol. Infect. (2005), **133** (Suppl. 1), S45–S47. © 2005 Cambridge University Press
doi:10.1017/S0950268805004334

13. The application of reproduction number concepts to tuberculosis

Vynnycky E, Fine PEM. Epidemiol Infect 1998; **121**: 309–324

AN APPRECIATION BY EMILIA VYNNYCKY

Health Protection Agency, London, UK
(emilia.vynnycky@HPA.org.uk)

Introduction

Epidemiology & Infection probably attracts more papers on mathematical modelling of infectious diseases than does any other epidemiology journal. The most important modelling papers published in the journal were probably those of Anderson and May during the 1980s, which laid the foundations for much of the subsequent modelling work carried out by themselves and their colleagues. Since the start of their partnership, they authored 17 articles between them in the journal, including work quantifying the effect of different vaccination strategies against measles and rubella [1, 2], on the epidemiology of rubella in the United Kingdom [3], and on the effect of age-dependent contact between individuals on the critical level of vaccination coverage required for control [4]. The latter work, published in 1985, was particularly important, since it described methods for incorporating realistic assumptions about (heterogeneous) mixing between individuals into models, an issue which was beginning to be addressed in the mathematical literature but which had not yet reached many epidemiological journals. Other important modelling work published in *Epidemiology and Infection* includes that of McLean et al. (reproduced in this edition) on the control of measles in developing countries [5, 6], and by Garnett and Grenfell on the epidemiology of varicella zoster in developed countries [7, 8].

A major theme of these earlier modelling papers was that of 'reproduction numbers', and most researchers estimated this statistic for acute immunizing infections such as measles, mumps and rubella. Of the two modelling papers selected for publication in the centenary edition of *Epidemiology & Infection*, one [9] focuses directly on this theme and discusses the application of reproduction number concepts to the epidemiology of tuberculosis in England and Wales during the last century.

Historical context of reproduction number measures

The basic and net reproduction numbers are among the most useful theoretical concepts in infectious disease epidemiology. Adapted from demographic theory, they were first applied to an infectious disease (malaria) by Macdonald in 1952 [10, 11], who defined a reproduction 'rate' (now commonly known as the 'net reproduction number') as the average number of secondary infectious cases resulting from each (infectious) case in a given population. Increases or decreases in the incidence of infection or disease thus reflect the magnitude of the 'net reproduction number', i.e. whether or not it exceeds 1. The limiting value of the net reproduction number, defined formally as the number of secondary infectious cases to result from an infectious case in a 'totally susceptible population' [12–14] (the 'basic reproduction number') provides a measure of the 'transmission potential' of an infection under 'ideal' conditions.

Estimates of the basic reproduction number (R_0) for acute immunizing infections are particularly useful since they can be used to calculate the 'herd immunity' threshold, or the critical proportion of the population which needs to be immunized in order to control transmission, using the expression $1-1/R_0$. Estimation of these statistics was facilitated greatly by the work during the 1970s of Dietz and others, who demonstrated that the basic reproduction number for acute immunizing infections can be derived simply

from the (average) life expectancy (L) and the average age at infection (A) using the expressions $R_0 = L/A$ (developed countries) and $R_0 = 1 + L/A$ (developing countries) [15]. These simple expressions assume that individuals mix randomly in the population. The literature and theory on reproduction number concepts has since expanded, dealing with estimation in heterogeneously mixing populations by Heesterbeek and Diekmann [13, 14] and, more recently, by Farrington [16, 17]. The incorporation of age in these methods typically traces back to the significant work by Anderson and May in the journal [4].

Application of reproduction number concepts to tuberculosis

After the early application of reproduction number concepts for malaria, the theory developed subsequently was applied largely to the acute childhood immunizing infections, such as measles, mumps and rubella, much of which was published in this journal [1, 2, 4–6, 18, 19]. This reflected the availability of both good data and good vaccines for these infections, and meant that theoretically derived herd-immunity thresholds might be tested empirically.

It was only later that the ideas were applied to another important airborne infection, tuberculosis, for which a vaccine is widely available [9, 20]. As illustrated in the reproduced article, several properties of the natural history of tuberculosis made this complicated, in particular the potential for reinfection to occur and the very long time period between successive cases in a chain of transmission (the serial interval). A further complication arises because the number of individuals 'effectively contacted' by each case (defined as contact sufficient to lead to infection if the contacted individual has never been infected) appears to have declined over time, e.g. from approximately 22 individuals effectively contacted by a case in 1900 to approximately 1 by 1990 [21].

The fact that immunity after infection with *M. tuberculosis* is not solid means that both uninfected and infected individuals are susceptible to infection (although perhaps to different degrees). Thus, the net reproduction number at a given time depends on the number of individuals infected or reinfected by each case. The derivation of the net reproduction number for tuberculosis is made yet more complicated by the fact that the relative contribution of bacilli from a particular infection event to a given disease episode is unknown: a case can be defined as the secondary case of just one source case if bacilli from one (e.g. the most recent) infection event contributes to each disease episode, but could be a secondary case of multiple prior cases if bacilli from several reinfections contribute to each episode. This uncertainty will particularly affect the interpretability of reproduction number estimates for tuberculosis for high infection risk settings, where a large proportion of tuberculosis morbidity may be attributable to reinfection [22].

Another complication results from the fact that the basic reproduction number is interpreted conventionally as the true transmission potential of an infection, but is typically derived with an implicit assumption that no changes in epidemiological or environmental conditions occur over the time period considered. The decline in the effective contact number, which appears to have occurred, for example in England and Wales during the last century [21], means that the proportion of individuals infected at a particular time who were later reinfected and developed disease attributable to the initial or to a later infection event must have also changed over time, and must have differed between those infected in 1900 and, say, in 1950. As a result, the true transmission potential of a tuberculosis case in a given year may be best described by a variant of the traditional reproduction number concept, defined in the reproduced article [9] (the 'ultimate' basic reproduction number), which takes account of changes in epidemiological conditions over time.

A further complexity is the potentially long time interval between successive cases in a chain of transmission, which means that the *net reproduction number for tuberculosis for a given year may not reflect the trend in disease incidence for that year*. For example, the net reproduction number was estimated to have been close to 1 in England and Wales during the period 1900–1950, even though the mortality rate (which reflected the disease incidence before treatment became available in 1950) declined dramatically throughout this time. This paradox follows from the fact that, of the individuals infected, say, in 1900, some did not develop disease until many years thereafter, for example in 1980. Thus, these individuals would have contributed to estimates of the net reproduction number in 1900, but to estimates of the trend in disease incidence only in 1980. Therefore, the net reproduction number at a given time accurately reflects the current trend in incidence only for diseases, such as the acute immunizing infections, for which the time interval between successive cases in a chain of transmission

varies little between individuals and for which epidemiological conditions do not change over this time interval.

Because of these issues, the simple reproduction number concepts developed with reference to acute immunizing infections, such as measles and rubella, are difficult to interpret when applied to tuberculosis. The complications also mean that reproduction number estimates for tuberculosis cannot be used to calculate a simple herd-immunity threshold for the disease. It is even debatable whether the concept of herd immunity is relevant for tuberculosis at all, given the variable protection imparted by BCG and the fact that infection with *M. tuberculosis* does not provide solid immunity against reinfection. Some of the problems with reproduction number measures for tuberculosis may well apply to other diseases for which the time interval between successive cases in a chain of transmission is very long (e.g. HIV, CJD) and/or for which reinfection can occur (e.g. varicella zoster). Given the current interest and concerns for tuberculosis and for these infections, an understanding of the complications of reproduction number measures and the implications for identifying control strategies is both important and timely.

References

1. **Anderson RM, May RM.** Vaccination against rubella and measles: quantitative investigations of different policies. J Hyg 1983; **90**: 259–325.
2. **Anderson RM, Grenfell BT.** Quantitative investigations of different vaccination policies for the control of congenital rubella syndrome (CRS) in the United Kingdom. J Hyg 1986; **96**: 305–333.
3. **Nokes DJ, Anderson RM, Anderson MJ.** Rubella epidemiology in South East England. J Hyg 1986; **96**: 291–304.
4. **Anderson RM, May RM.** Age-related changes in the rate of disease transmission: implications for the design of vaccination programmes. J Hyg 1985; **94**: 365–436.
5. **McLean AR, Anderson RM.** Measles in developing countries. Part I. Epidemiological parameters and patterns. Epidemiol Infect 1988; **100**:111–133.
6. **McLean AR, Anderson RM.** Measles in developing countries. Part II. The predicted impact of mass vaccination. Epidemiol Infect 1988; **100**: 419–442.
7. **Garnett GP, Grenfell BT.** The epidemiology of varicella-zoster virus infections: the influence of varicella on the prevalence of herpes zoster. Epidemiol Infect 1992; **108**: 513–528.
8. **Garnett GP, Grenfell BT.** The epidemiology of varicella-zoster virus infections: a mathematical model. Epidemiol Infect 1992; **108**: 495–511.
9. **Vynnycky E, Fine PEM.** The long-term dynamics of tuberculosis and other diseases with long serial intervals: implications of and for changing reproduction numbers. Epidemiol Infect 1998; **121**: 309–324.
10. **Macdonald G.** The analysis of equilibrium in malaria. Trop Dis Bull 1952; **49**: 813–829.
11. **Macdonald G.** The epidemiology and control of malaria. London: Oxford University Press, 1956.
12. **Anderson RM, May RM.** Infectious diseases of humans – dynamics and control. Oxford: Oxford University Press, 1992.
13. **Heesterbeek H.** R_0. Amsterdam: Centrum voor Wiskunde en Informatica, 1992.
14. **Diekmann O, Heesterbeek JAP, Metz JAJ.** On the definition and the computation of the basic reproduction ratio R_0 in models for infectious diseases in heterogeneous populations. J Math Biol 1990; **28**: 365–382.
15. **Dietz K.** Transmission and control of arboviruses. In: Ludwig D, Cooke KL, eds. Epidemiology. Philadelphia: SIAM, 1975: 104–121.
16. **Farrington CP, Kanaan MN, Gay NJ.** Estimation of the basic reproduction number for infectious diseases from age-stratified serological survey data. J Roy Statist Soc Series C (Applied Statistics) 2001; **50**: 251–292.
17. **Farrington CP, Whitaker HJ.** Estimation of effective reproduction numbers for infectious diseases using serological survey data. Biostatistics 2003; **4**: 621–632.
18. **Anderson RM, Crombie JA, Grenfell BT.** The epidemiology of mumps in the UK: a preliminary study of virus transmission, herd immunity and the potential impact of immunization. Epidemiol Infect 1987; **99**: 65–84.
19. **Grenfell BT, Anderson RM.** The estimation of age-related rates of infection from case notifications and serological data. J Hyg 1985; **95**: 419–436.
20. **Blower SM, McLean AR, Porco TC, et al.** The intrinsic transmission dynamics of tuberculosis epidemics. Nature Med 1995; **1**: 815–821.
21. **Vynnycky E, Fine PEM.** Interpreting the decline in tuberculosis during the last century – the roles of secular trends in effective contact. Int J Epidemiol 1999; **28**: 327–334.
22. **van Rie A, Warren R, Richardson M, et al.** Exogenous reinfection as a cause of recurrent tuberculosis after curative treatment. N Engl J Med 1999; **341**: 1174–1179.