



The Interval between Successive Cases of an Infectious Disease

Paul E. M. Fine

From the Department of Infectious and Tropical Diseases, London School of Hygiene and Tropical Medicine, London, United Kingdom.

Received for publication March 17, 2003; accepted for publication May 30, 2003.

The interval between successive cases of an infectious disease is determined by the time from infection to infectiousness, the duration of infectiousness, the time from infection to disease onset (incubation period), the duration of any extra-human phase of the infectious agent, and the proportion clinically affected among infected individuals. The interval is important in the interpretation of infectious disease surveillance and trend data, in the identification of outbreaks, and in the optimization of quarantine and contact tracing. This paper discusses the properties of these intervals, as measured between transmission events or between clinical onsets of successive infected individuals, noting the determinants of their ranges and frequency distributions, the circumstances under which secondary cases may arise before primaries, and under which the infection transmission interval will be different from the interval between clinical onsets of successive cases. It discusses the derivation of interval distribution statistics from descriptive data given in standard textbooks, with illustrations from published data on outbreaks, households, and epidemiologic tracing. Finally, it discusses the implications of such measures for studies of secondary attack rates, for the persistence of infection in human communities, for outbreak response, and for elimination or eradication programs.

communicable diseases; disease outbreaks

Abbreviation: AIDS, acquired immunodeficiency syndrome.

The time and place of clinical onsets, as well as the social connections between cases, are among the key observations upon which infectious disease research and control decisions depend. The relations between them carry important information, such as whether there is a single point source outbreak, or multiple sources, or evidence of propagation in time and space, and reflect risk factors relevant in that particular circumstance. The interpretation of such data requires consideration of the time between successive cases in a chain of transmission. Often called the “serial interval” (1), this time period depends upon the temporal relation between infectiousness and clinical onset of a source case and the incubation period of the receiving case. The correct interpretation of such data is particularly important in the context of infectious disease emergency situations. This paper sets out the basic properties of these measures and outlines their interpretation and implications.

THE COURSE OF A SINGLE INFECTION

Regardless of the infectious agent involved and the pathogenesis, three time periods determine the course of individual cases and their successors: 1) the time from infection to onset of infectiousness, 2) the duration of infectiousness, and 3) the time from infection to clinical onset (the incubation period). A fourth interval—the period of time between an infectious agent’s leaving one human and entering another—may also be important if transmission involves a stage in either another host (e.g., insect-borne infections) or the environment (e.g., enteric agents in a water or food supply).

Not all infections necessarily lead to clinical disease. The proportion clinical among infections varies greatly, from below 1 percent for polio to over 90 percent for measles and smallpox. Insofar as subclinical infections may still be infectious, the ability to identify chains of transmission from clinical cases is strongly influenced by this factor and will be very difficult if an appreciable proportion of infections are

Correspondence to Department of Infectious and Tropical Diseases, London School of Hygiene and Tropical Medicine, Keppel Street, London WC1E 7HT, United Kingdom (e-mail: paul.fine@lshtm.ac.uk).

not recognized (as for polio). Whether or not transmission occurs in the absence of clinical manifestations, the interval between successive *infections* is analogous, but not identical, to the serial interval between clinical onsets of successive cases and is called here the “transmission interval.”

Time from infection to infectiousness

Some authors have described this as the “latent period,” although this may not be ideal given that the terms “latency” and “latent period” itself are used in a variety of other microbiologic and epidemiologic contexts (1). The interval is difficult to measure, as we rarely observe infection transmission events, except by inference from detailed contact histories. The start of infectiousness may in theory also be inferred from data on the shedding of infectious agents, as in respiratory secretions or in stools, or the appearance of infectious stages in the blood (e.g., *Plasmodium* gametocytes). Such data are not easily or often collected, however, and thus there are fewer data on this interval than for the incubation period, for example. We do know that, for some infections, individuals may become infectious appreciably before onset of disease (human immunodeficiency virus/acquired immunodeficiency syndrome (AIDS) and hepatitis A being examples), and for others the opposite occurs (e.g., the infectious gametocytes of *falciparum* malaria appear several days after onset of illness) (2).

Duration of infectiousness

This period is also difficult to measure, requiring the accumulation of detailed contact histories of linked cases or else data on shedding of the infectious agent. Beyond this difficulty, the *level* of infectiousness can vary considerably during the course of the infectious period for several reasons. This will reflect the time course of the concentration of agents shed into the environment, as well as of the clinical manifestations of a disease (e.g., coughing), and their implications for individual and social behavior (including restricted activity or hospitalization) and hence for contact. The actual numbers of transmissions of infection from individual cases and their distribution over time, which will ultimately be manifested in the pattern of subsequent disease in the community, will reflect all of these factors.

Incubation period

Although the basic definition of an incubation period (time from infection to clinical onset) is widely accepted, its measurement also has difficulties. In addition to the timing of the infection event, the timing of disease onset may be difficult (it may be insidious, or people may just not remember precisely), and data are sometimes presented with reference to the onset or maximum of some clearly defined symptom, such as a rash. Some descriptions fail to clarify just how the onset was defined, making interpretation difficult. Despite such difficulties, incubation period estimates, typically expressed as ranges, are available for most infectious diseases (2).

An important paper by Sartwell (3) demonstrated that frequency distributions of incubation periods are typically “right skew” or “log normal,” meaning that they have a long right-hand tail but look symmetric if the time axis is plotted on a logarithmic scale. For this reason, some authors have preferred to summarize incubation periods with a median or geometric mean, rather than an arithmetic mean (the geometric mean is the antilog of the arithmetic mean of the logarithms). The actual distributions reflect variations in infectious dose, in replication times of the pathogen, and in levels of susceptibility among members of the host population. There is good evidence for an inverse relation between incubation period and infectious dose (higher dose, shorter incubation) for some infections (typhoid is a classic example) but not others (2). Most of the literature on incubation periods relates to diseases with relatively short periods, with comparatively little discussion on the more difficult problem of defining incubation periods when disease occurs many years after initial infection, as with tuberculosis (4). The important exception to this generalization is AIDS, the incubation period of which has attracted much attention because of its importance in predictions of future disease burden (5).

THE INTERVAL BETWEEN SUCCESSIVE INFECTIONS (TRANSMISSION INTERVAL)

The interval between infection events of successive individuals in a chain of transmission is determined by the time from onset of infection to onset of infectiousness and by the duration and pattern of infectiousness. Its minimum will be just the shortest possible time from infection to infectiousness, and its maximum will be the sum of the maximum interval from infection to infectiousness plus the maximum duration of infectiousness (which may be lifelong for some infections). The frequency distribution of transmission intervals will be determined by the frequency distributions of its two components, which will in turn be a function of infectious dose, level of immunity, and contact patterns. The *average transmission interval*, which is important as it influences the rapidity of infection spread in a community, is the sum of the averages of its two component distributions.

THE INTERVAL BETWEEN SUCCESSIVE CLINICAL CASES (CLINICAL ONSET SERIAL INTERVAL)

Person A transmits infection to person B, and they both become clinically ill (figure 1). What we observe is not the infection transmission but the clinical onsets, whose separation in time is a reflection of two things: 1) the timing of infection transmission with respect to person A’s clinical onset (T_A , defined here as the time of infection transmission from person A *minus* the time of person A’s clinical onset), and 2) the incubation period of case B (I_B , being the time of person B’s clinical onset minus the time of person B’s infection). According to this convention, T_A is positive if transmission occurs *after* person A’s illness onset, but it is negative if transmission occurs prior to person A’s illness. Both T_A and I_B will be determined by several factors in any particular setting. Defined in this manner, the interval

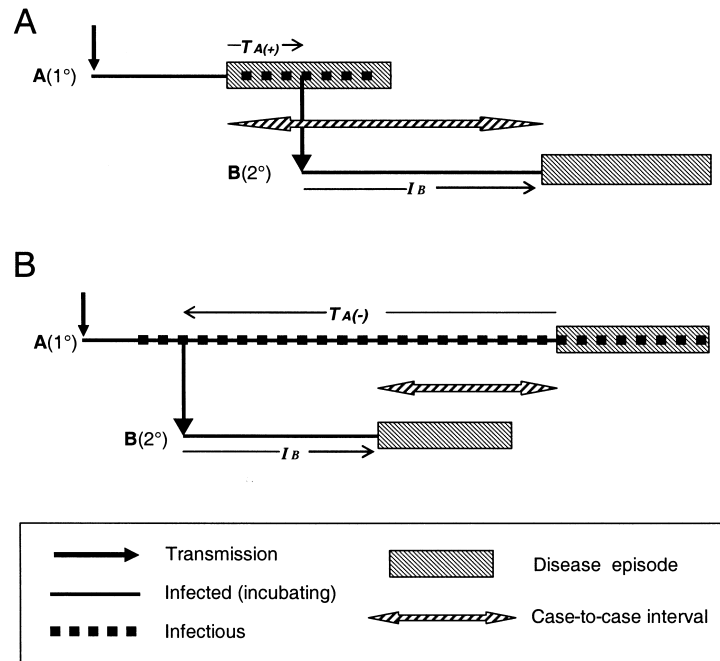


FIGURE 1. Relation between successive cases in a chain of transmission. In part A, transmission from case A occurs relatively late in the course of A's infection ($T_A > 0$), and case B's clinical onset occurs ($T_A + I_B$ time units) after that of case A. In part B, the combination of early transmission, prior to case A's clinical onset ($T_A < 0$), and a short incubation period for case B ($I_B < -T_A$) means that case B experiences clinical onset prior to case A.

between clinical onsets of persons A and B is simply ($T_A + I_B$). Figure 1, part A, illustrates this diagrammatically.

This interval will be *minimum* (called here S_n) if the transmission occurs as early as possible in relation to person A's clinical onset (i.e., T_A is small or negative) and if person B's incubation period is a minimum (I_B is small). Importantly, if in some circumstance the transmission from person A occurs appreciably prior to clinical onset ($T_A < 0$), such that $(-T_A) > I_B$, then the clinical onset of person B will occur *prior to* that of person A (figure 1, part B). This is most likely to occur in circumstances in which the variation in time to infectious onsets, in infectious periods, and in incubation periods is great. An obvious example in which this may happen is with human immunodeficiency virus infection and AIDS, and it is probable that an appreciable number of individual AIDS cases have had clinical onset prior to that of the individual from whom they contracted the infection (6). What *proportion* of cases will have clinical onset prior to that of their source cases is a difficult question, requiring detailed data on courses of infections and upon behavior (number of contacts) over time.

The *maximum* clinical onset interval (called here S_x) will occur if transmission occurs as late as possible relative to person A's clinical onset (thus, T_A is a large positive number) and with a maximum incubation period for person B. The range can be very great for infections with extended periods of infectiousness, certainly many decades and perhaps as long as a century. Consider a child (person A) who contracts

varicella virus infection and associated chickenpox in the first year or two of life and then many decades later has an episode of zoster, thereby infecting his or her great-grandchild (person B), who then has chickenpox as a result. No doubt this has occurred! Other herpesvirus diseases, mycobacterial diseases such as leprosy and tuberculosis, and perhaps also prion diseases also have this potential, because of the propensity for their infections to be lifelong.

The actual clinical onset serial intervals for any disease will follow a frequency distribution between these minimum and maximum extremes. The cleanest observed data on such distributions are seen in records of outbreaks with a single primary case and associated secondary cases (with the proviso that the single source case may not be typical or representative). Figure 2 shows examples of such data on measles and smallpox, showing clear separation of secondary cases from single primaries, as well as also subsequent 3° and 4° cases. (Although the definition of primary, secondary, tertiary [1°, 2°, 3°, ...] cases has little ambiguity, some authors speak of secondaries as the "first generation." There is disagreement on this latter terminology within and between disciplines, with different applications in the context of epidemiology, genetics, migrant studies, and microbiology [e.g., relating to the serial passage of infectious agents in laboratories], the confusion arising because some workers number as generations the successive groups of individuals or cases or organisms, whereas others number the successive transmissions between them. Because of this

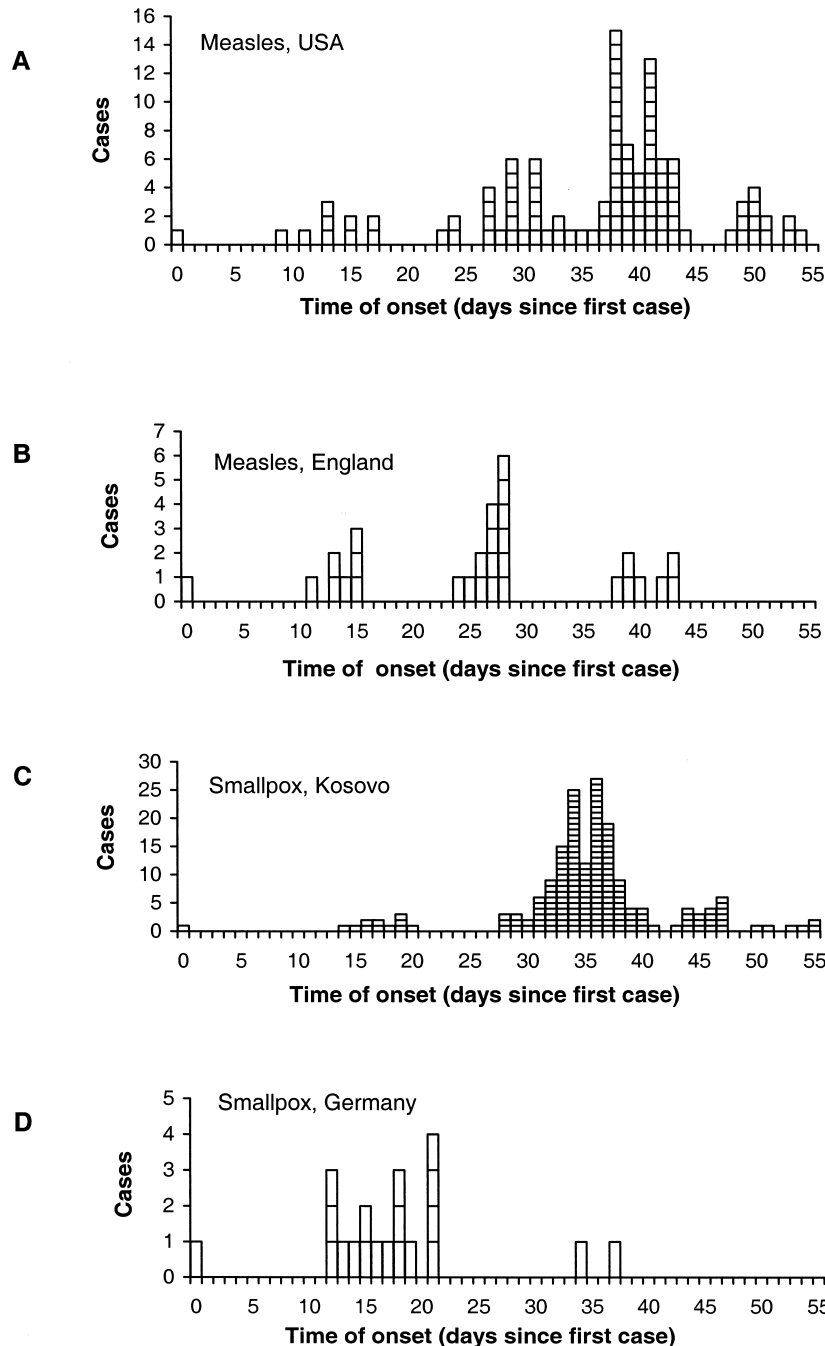


FIGURE 2. Propagative epidemics of smallpox and measles, each initiated by a single case: A, measles in a boarding school in the northeastern United States in 1934 (7); B, measles in an English village in 1932 (8); C, smallpox in Kosovo in 1972 (9); D, smallpox in Meschede, Germany, in 1970 (10). For none of these examples does the original publication state the criterion used for establishing the date of onset.

we will avoid numbering “generations” here.) In many data sets of various diseases, the 3^o and 4^o case groups overlap in time.

According to the above terminology for minimum and maximum intervals, the secondary cases should have onset between S_n and S_x days (for example) “after” the primary, extending over a period of $S_x - S_n$ days (recognizing that S_n

may be a negative number). Similarly, the tertiary cases should in theory occur between $2S_n$ and $2S_x$ days after the primary, extending over a period of $2(S_x - S_n)$ days. Thus, the “wave” of tertiaries should be twice as “wide” as that of the secondaries. This will not always be the case in a real outbreak, however, given that the onsets of any tertiary cases will be contingent upon the timing of the actual secondaries

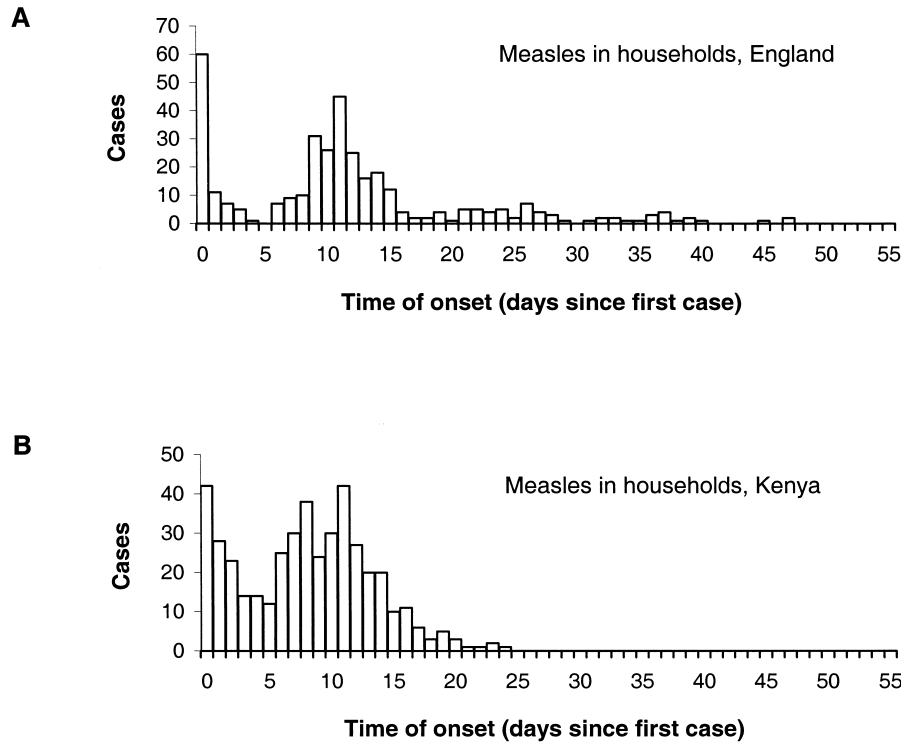


FIGURE 3. Frequency distributions of the interval from primary case onset to the onset of subsequent cases in household outbreaks of measles: A, data from rural England, based upon the date of “fullest flowering of the rash” (11, p. 552); B, data from rural Kenya, based primarily upon the date of rash onset (12).

in that particular episode, which are themselves a sample from the distribution of potential intervals. So, if the onset times of the first and last secondaries are called B_f and B_l , respectively, the tertiary cases can arise from $(B_f + S_n)$ to $(B_l + S_x)$ days, thus over the period $(B_l + S_x) - (B_f + S_n)$ days. The logic would be identical for subsequent generations, but in practice these distributions typically become unclear because of the stochastic nature of the individual intervals and the overlap of generations. If $S_x \geq 2S_n$, overlap can occur between 2° and 3° cases, whereas if $S_x < 2S_n$, but $2S_x \geq 3S_n$, overlap can start only between the 3° and 4° cases, etc. Because of such overlaps, it may be difficult to identify cases by generation in the absence of detailed contact histories.

The observed intervals between clinical onsets may also be affected by control measures. If infected individuals are recognized early, their infectious periods may be curtailed by treatment or isolation, in which circumstance we would expect case-to-case intervals to shorten on average over time, during the course of an epidemic. For this reason, one might expect transmission from the primary to secondary cases in outbreaks to provide the fullest distribution of potential clinical onset intervals for the infection concerned.

HOUSEHOLD DATA

Households provide a special setting for studies of serial intervals. The literature contains several sets of data on

household outbreaks plotted such that the horizontal axis is in effect the serial interval with respect to the initial “primary” case, which is assumed to have introduced the infection into the household. Examples are shown in figure 3. These distributions typically indicate a group of “coprimary” cases, occurring too close in time to the primary to be considered secondaries and thus apparently also contracted outside the household. This is followed by a trough and then a “wave” interpretable as secondary cases, followed by 3° or 4° cases, often with some overlap such that their order assignment may not be entirely clear.

Household outbreaks imply infection transmission under particularly intense contact conditions of close and prolonged physical proximity. Given that close contact should facilitate transmission during the early phase of an infectious period and that incubation periods of some infections may be negatively associated with infectious dose, one might expect case-to-case intervals in households to be shorter than those observed in extradomestic circumstances.

EPIDEMIOLOGIC TRACING

Case-to-case intervals may also be evident through detailed contact tracing, now sometimes supplemented by molecular typing of isolates, which reveals chains of transmission and may even pinpoint exactly when and where transmission must have occurred. Two examples are shown

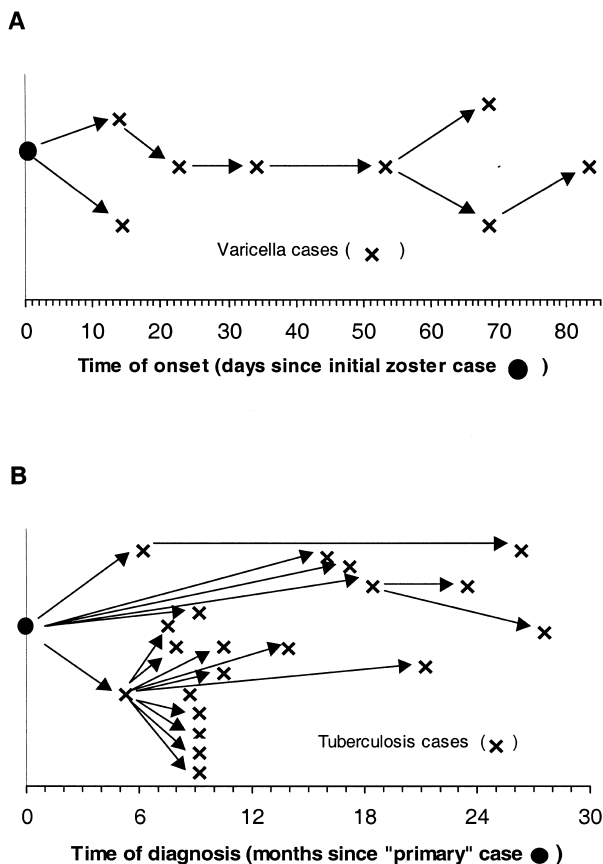


FIGURE 4. Chains of transmission: A, zoster (primary case) and chickenpox cases in the Shetland Islands (13); B, tuberculosis cases in San Francisco, California, linked by molecular fingerprinting and epidemiologic contact tracing (14).

in figure 4. The chickenpox data are from an investigation that provided epidemiologic evidence for the identity of the zoster and varicella viruses. The tuberculosis data are based upon a combination of molecular fingerprinting of isolates and epidemiologic investigations to confirm linkages between cases in San Francisco. Although data such as these show explicit serial intervals, their implied frequency distributions reflect the circumstances of the particular investigations and are likely to favor shorter intervals.

ANALYSIS

We here examine the quantitative implications of standard textbook descriptions with reference to the data and diseases illustrated in figures 2–4.

Measles

According to a widely used reference, the incubation period of measles is “About 10 days, but may be 7 to 18 days from exposure to onset of fever, usually 14 days until rash appears; rarely as long as 19–21 days” (2, p. 331), the “rash

appears on the third to the seventh day” after the onset of prodromal fever, and the infectious period is “from 1 day before the beginning of the prodromal period (usually about 4 days before rash onset) to 4 days after appearance of the rash; minimal after the second day of rash” (2, p. 330). On the basis of this description, we would expect the rash-to-rash serial interval to fall between a minimum of 6 days (i.e., $T_A = -4$; $I_B = 7 + 3$) and a maximum of 25 days ($T_A = 4$; $I_B = 21$). (An extreme interpretation of the text could suggest minimal $T_A = -8$ if a case were infectious 1 day prior to prodrome and if there were 7 days from prodrome to rash, thus giving a minimum rash-to-rash interval of 2 days ($T_A = -8$; $I_B = 7 + 3$). Similarly, the text might be interpreted as saying that I_B could be as long as 18 days to prodrome plus 7 more days to rash, thus making a maximum rash-to-rash interval of 29 days ($T_A = 4$; $I_B = 18 + 7$). The implied range between onsets of prodromes is from 6 ($T_A = -1$; $I_B = 7$) to 29 days ($T_A = 7 + 4$; $I_B = 18$).

The two community measles outbreaks illustrated in figure 2 show very similar temporal patterns. The larger outbreak, which might on account of its size be less influenced by chance than the other smaller example, suggests a range of serial intervals from 9 to 17 days, and all the data in both outbreaks are consistent with this. In contrast, the household data sets illustrated in figure 3 show patterns that are similar between themselves, despite coming from very different settings, but that differ from the community outbreaks. The serial intervals appear to be shorter (minimum, 6 days; mode, 11 days) in the household context than in the community, consistent with transmission’s occurring earlier and perhaps implying a higher dose and a shorter incubation period in the household context.

Smallpox

According to the same reference, the incubation period of smallpox is “From 7–19 days; commonly 10–14 days to onset of illness and 2–4 more days to onset of rash”; and cases are infectious “From the time of development of the earliest lesions to disappearance of all scabs; about 3 weeks. The patient is most contagious during the pre-eruptive period by aerosol droplets from oropharyngeal lesions” (2, p. 457). Interpreting this to mean that cases can be infectious from 4 days prior to 21 days after rash onset, the full range of rash-to-rash intervals would extend from 5 days minimum ($T_A = -4$; $I_B = 7 + 2$) to 44 days maximum ($T_A = 21$; $I_B = 19 + 4$). The interval between illness onsets should be from 3 ($T_A = -4$; $I_B = 7$) to 44 ($T_A = 4 + 21$; $I_B = 19$) days.

The two smallpox outbreaks illustrated here show similar distributions for the successive generations. The Kosovo outbreak (figure 2, part C) was sufficiently large to provide a reasonable distribution, with intervals from primary to secondary cases ranging from 14 to 20 days. This is considerably shorter than the theoretical range, probably reflecting that transmission occurred during a narrow time window around the time of clinical onset of the primary case. Interestingly, the earliest and latest tertiary cases occur at exactly double the times to the earliest and latest secondary cases, as predicted, implying narrow transmission periods from the secondary cases as well. The quaternary case generation is

right truncated, indicating that control (vaccination and quarantine) was successful in breaking transmission entirely from the later tertiary cases. The Meschede outbreak (figure 2, part D) took place in a hospital setting, the secondary cases being attributable to aerosol transmission, prior to the primary case's being moved to another hospital.

Chickenpox

The incubation period is "from 2 to 3 weeks; commonly 14–16 days; may be prolonged after passive immunization ... and in the immunodeficient" (2, p. 93). The period of communicability starts "as long as 5 but usually 1 to 2 days before onset of rash, and continuing until all lesions are crusted (usually about 5 days) ..." (2, p. 94). Assuming that the incubation period is measured to rash onset (the rash occurs early in chickenpox) implies that the interval between case (rash) onsets should range between 9 ($T_A = -5; I_B = 14$) and 26 ($T_A = 5; I_B = 21$) days. The intervals observed in the Shetland outbreak (range, 9–19 days; mean, 14 days) (figure 4, part A) fall within this range. The fact that the intervals following the zoster and varicella cases were similar was emphasized in the original paper as evidence supporting the identity of the agents (13). The primary zoster case was the father of the two secondary cases and, thus, the interval backwards to his source case may have been several decades.

Tuberculosis

For some infections, the definitions become even less clear. For example, "Incubation period—from infection to demonstrable primary lesion or significant tuberculin reaction (note: this is not a clinical endpoint), about 2–10 weeks. While the subsequent risk of progressive pulmonary or extrapulmonary TB [tuberculosis] is greatest within the first year or two after infection, latent infection may persist for a lifetime," and "Period of communicability—Theoretically as long as viable tubercle bacilli are being discharged in the sputum" (2, p. 524). Data such as those in figure 4, part B, are limited to a relatively narrow (for tuberculosis) time window, in this case excluding intervals greater than 2.5 years, many of which must occur as a consequence of the reactivation of long-term latent infections (14). Inferences about the full distribution of case-to-case intervals of tuberculosis are complicated by uncertainties over the proportion of cases among older individuals that represent recent primary or reinfections versus reactivations of old infection. Any attempt to describe the full distribution must thus be based upon assumptions concerning pathogenesis, and it will also be influenced by time trends in infection risk and by the demography of the host population, insofar as the frequency of long incubation and serial intervals will be related to life expectancy (4).

DISCUSSION

Although there is much published discussion of incubation periods, there is relatively little literature on transmission and clinical onset serial intervals. Incubation periods are simpler in concept, and they are of obvious importance in

common source outbreaks (e.g., for identifying the exposure responsible for food-poisoning outbreaks) and in stipulating periods of disease risk or quarantine after known exposure. The interval between clinical onsets, however, is what is typically observed, and it is the more important statistic in several contexts.

Panum (15), in his classic description of measles on the Faroe Islands in 1846, noted that one can infer when patients are infectious by relating data on incubation periods to the case-to-case interval (these intervals are similar in measles, indicating that patients are most infectious at the time of clinical onset). Hope Simpson (16), some of whose data are illustrated in figure 3, part A, and figure 4, part A, used similar reasoning in discussing infectious periods.

Case onset serial intervals are functions of several variables: of times to infectiousness, duration of infectiousness, and time to disease (incubation period). They are also influenced by the pattern of numbers of contacts over time (itself influenced by individual behavioral and social factors), by control interventions, and by stochastic effects. They may be complicated further by dependence between these variables; that is, the time to onset and the duration of infectiousness are unlikely to be independent of each other, let alone independent of the incubation period or of the implementation of control interventions. Examples such as presented here are illustrative, and they will have been influenced by all these factors. Although the range of possible serial intervals can be estimated from descriptions of infectious and incubation ranges, as shown above, the more severe the condition, the more likely it is that actual serial intervals will fall in the shorter part of the range, given that illness, isolation, and control measures will tend to curtail transmission. Serial intervals in households or other close contact settings will tend to be shorter than in community settings, as illustrated in figure 3, given that close contact predisposes to early infection and also that there may be an inverse association between dose and incubation period.

The parameter definitions quoted in this paper are from a reference that is widely used in infectious disease control (2), but it should be noted that such definitions differ between texts. For example, with reference to measles, one may also read the following: "The incubation period (to onset of fever) is 8–13 days, usually about 10 days. The period of communicability starts just before the onset of the prodrome and lasts until four days after the rash appears" (17, p. 140), or "The first symptoms ... occur after a 10–12 day incubation period.... The prodromal stage ... lasts 2 to 4 days ... generally considered to be infectious 2 to 4 days before to 4 days after rash onset" (18, p. 223). Each of these (and several other published) definitions differs slightly from the one used above. None of these references cites the data upon which the estimates are based. None of the published definitions is written in a manner that is easily translated into serial intervals.

In the context of an infectious disease outbreak, a public health officer may be presented with one or more (index) cases of a disease, and among the first questions are those concerning the source (if the source was a clinical case, then it had onset within the range of one serial interval before the index cases) and the extent of propagation (if index cases are

secondary to a single source case, then they should appear over an interval of time equal to $(S_x - S_n)$. Questions of this sort would be of particular importance in the response to a new or emerging infection or to a bioterrorist incident.

The *average* clinical onset serial interval will be the same as the average infection transmission interval (the former being far easier to observe than the latter) only if the distribution of infectiousness periods, relative to time of infection onset, is independent of whether or not individuals manifest signs of clinical disease. This is probably not true for many infections. Even if the average clinical onset and transmission intervals are similar, the frequency distributions of these two statistics will generally be different. This is because the variation in incubation period affects only the former, the interval between clinical onsets, and has no direct influence upon the transmission interval. Although a source case must by definition become infected prior to its secondary, it is possible, as noted above and in figure 1, part B, for a secondary case to have clinical onset prior to that of its source.

In the estimation of time trends of infection, a key statistic is the reproduction number (defined as the average number of transmissions per case) and whether this is greater or less than unity—implying that each case is responsible, on average, for more or less than one successful transmission (19). The statistic is then the factor of increase or decrease in infection incidence over a period equal to the mean transmission interval.

Another context in which the serial interval is important is that of calculating secondary attack rates, themselves the traditional measure of infectiousness or transmissibility. Calculation of such statistics requires a denominator of the number of individuals exposed to a primary case and a numerator of the number of cases attributable to that exposure, which should be defined as those arising within one case onset serial interval (distribution) from the primary case (20).

A particularly long interval between clinical cases may reflect either a long infectious period of the source or a long incubation period of the recipient case. Alternatively, it may reflect that transmission is not by direct contagion, but was indirect, with the agent sequestered in the environment within a vector or in a zoonotic cycle. In other circumstances, a long interval between identified cases may reflect that transmission went unobserved through subclinically infected individuals. The latter is an especially important issue in the context of the current global polio eradication program, whose success requires the eradication of infection but whose surveillance is dependent largely upon identifying clinical cases, although less than 1 percent of infections are associated with clinical disease.

The very long transmission interval for some infections, such as herpesvirus and mycobacterial infections, has important implications for the ability of such agents to persist in small human populations. It was recognized by Black et al. (21) that antibodies for herpesviruses were found consistently in small and isolated populations, indicating that these infections have a small “critical population size” (minimum total human population size necessary for persistence of the infection). This reflects the ability of these agents to persist

for long periods within individuals, ultimately to cause infections in new generations of susceptible individuals that have appeared in the interim.

Yet another problem complicates the measurement of long incubation and serial intervals, as occur with mycobacterial and prion diseases. Insofar as an older individual cannot experience a very long *subsequent* incubation period if s/he becomes infected, but clinical onset in such an individual could reflect a very long *prior* incubation period, it is clear that these intervals will be a function of age and whether they are examined prospectively from a primary case or retrospectively from secondary cases. This issue has been recognized in some analyses of incubation periods of prion diseases (22) and has been examined explicitly in the context of describing the full serial interval distribution of tuberculosis (4).

The propensity for persistent infections and/or long incubation periods, and, hence, long serial intervals, poses a particular problem in the context of eradication or elimination programs, as exemplified by the current World Health Organization efforts with regard to polio and leprosy. It is now known that certain classes of immunodeficient individuals can continue to pass vaccine-derived polioviruses in their feces for several years (23), and, despite the very best control efforts, it is certain that new and infectious cases of leprosy will arise many decades from now, after very long incubation periods. The circumstances surrounding these individuals will influence the long-term trends of these diseases and must be considered in the planning of control strategies.

In addition to data on incubation periods and duration of infectiousness, the accumulation of observations on transmission and clinical onset serial intervals, with explicit stipulation of the onset criterion employed, and their inclusion in standard texts and control manuals, would be a useful contribution to guide public health interpretation and action.

ACKNOWLEDGMENTS

The author thanks Alessia Melegaro for helpful discussion and comments.

REFERENCES

1. Last JM, ed. A dictionary of epidemiology. 4th ed. Oxford, England: Oxford University Press, 2001.
2. Chin J, ed. Control of communicable diseases manual. 17th ed. Washington, DC: American Public Health Association, 2000.
3. Sartwell PE. The distribution of incubation periods of infectious disease. *Am J Hyg* 1950;51:310–18.
4. Vynnycky E, Fine PEM. Lifetime risks, incubation period, and serial interval of tuberculosis. *Am J Epidemiol* 2000;152:247–63.
5. Munoz A, Sabin CA, Phillips AN. The incubation period of AIDS. *AIDS* 1997;11(suppl A):S69–76.
6. Learmont JC, Geczy AF, Mills J, et al. Immunologic and virologic status after 14 to 18 years of infection with an attenuated strain of HIV-1. *N Engl J Med* 1999;340:1715–22.

7. Aycock WL. Immunity to poliomyelitis. Heterologous strains and discrepant neutralisation tests. *Am J Med Sci* 1942;204:455–67.
8. Pickles WN. *Epidemiology in country practice*. Bristol, England: John Wright & Sons, Ltd, 1939.
9. Fenner F, Henderson DA, Arita I, et al. *Smallpox and its eradication*. Geneva, Switzerland: World Health Organization, 1988.
10. Wehrle PF, Posch J, Richter KH, et al. An airborne outbreak of smallpox in a German hospital and its significance with respect to other recent outbreaks in Europe. *Bull World Health Organ* 1970;43:669–79.
11. Hope Simpson RE. Infectiousness of communicable diseases in the household (measles, chickenpox and mumps). *Lancet* 1952;2:549–54.
12. Aaby P, Leeuwenburg J. Patterns of transmission and severity of measles infection: a reanalysis of data from the Machakos area, Kenya. *J Infect Dis* 1990;161:171–4.
13. Hope Simpson RE. Studies on shingles. Is the virus ordinary chickenpox virus? *Lancet* 1954;2:1299–302.
14. Small PM, Hopewell PV, Singh SP, et al. The epidemiology of tuberculosis in San Francisco. *N Engl J Med* 1994;330:1703–9.
15. Panum PL. *Observations made during the epidemic of measles on the Faroe Islands in the year 1846*. New York, NY: Delta Omega Society, 1940.
16. Hope Simpson RE. The period of transmission in certain epidemic diseases: an observational method for its discovery. *Lancet* 1948;2:756–60.
17. Hawker J, Begg N, Blair I, et al. *Communicable disease control handbook*. Oxford, England: Blackwell Science Publishers, 2001.
18. Redd SC, Markowitz LE, Katz SL. Measles vaccine. In: Plotkin SA, Orenstein WA, eds. *Vaccines*. Philadelphia, PA: WB Saunders, 1999:222–66.
19. Anderson RM, May RM. *Infectious diseases of humans: dynamics and control*. Oxford, England: Oxford University Press, 1992.
20. Fine PEM, Jezek Z, Grab B, et al. The transmission potential of monkeypox virus in human populations. *Int J Epidemiol* 1988;17:643–50.
21. Black FL, Hierholzer WJ, Pinheiro F, et al. Evidence for persistence of infectious agents in isolated human populations. *Am J Epidemiol* 1974;100:239–50.
22. Huillard d’Aignaux JN, Cousens SN, Maccario J, et al. The incubation period of Kuru. *Epidemiology* 2002;13:402–8.
23. Fine PEM, Carneiro IAM. Transmissibility and persistence of oral polio vaccine viruses: implications for the global poliomyelitis eradication initiative. *Am J Epidemiol* 1999;150:1001–21.