
REVIEW ARTICLE

Close encounters of the infectious kind: methods to measure social mixing behaviour

J. M. READ^{1*}, W. J. EDMUNDS², S. RILEY³, J. LESSLER⁴ AND
D. A. T. CUMMINGS⁴

¹ *Department of Epidemiology and Population Health, Institute of Infection and Global Health, University of Liverpool, UK*

² *Centre for the Mathematical Modelling of Infectious Diseases, London School of Hygiene and Tropical Medicine, Keppel Street, London, UK*

³ *MRC Centre for Outbreak Analysis and Modelling, Department of Infectious Disease Epidemiology, School of Public Health, Imperial College London, London, UK*

⁴ *Department of Epidemiology, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD, USA*

*Received 7 December 2011; Final revision 8 March 2012; Accepted 10 April 2012;
first published online 12 June 2012*

SUMMARY

A central tenet of close-contact or respiratory infection epidemiology is that infection patterns within human populations are related to underlying patterns of social interaction. Until recently, few researchers had attempted to quantify potentially infectious encounters made between people. Now, however, several studies have quantified social mixing behaviour, using a variety of methods. Here, we review the methodologies employed, suggest other appropriate methods and technologies, and outline future research challenges for this rapidly advancing field of research.

Key words: Contact diary, infectious disease control, mathematical modelling, proximity, social mixing, social network, transmission.

INTRODUCTION

Understanding the spread of respiratory pathogens is a public health priority as many of the greatest threats to human health are spread by direct person-to-person contact. A major challenge is identifying and quantifying the behavioural, social and environmental factors which permit the transmission of such pathogens and generate larger-scale patterns of spread. Accurate predictions of the likely impact of new or re-emerging pathogens and improved targeting of control interventions require a quantitative

understanding of the factors and mechanisms that promote or inhibit the spread of pathogens through populations.

A principal assertion within the epidemiology of respiratory pathogens is that patterns of social mixing behaviour at the individual level contribute to the dynamics of infection seen at the population level. Despite controversy about the precise role different mechanisms may play in infection [1, 2], proximity or social contact is thought to be a major factor in the transmission process for many important infections, including SARS, influenza and tuberculosis [3, 4]. The prevention of social interactions lies at the heart of non-pharmaceutical interventions for these and other respiratory pathogens. However, robust measures of social encounters were, until recently, lacking, making

* Author for correspondence: Dr J. M. Read, Department of Epidemiology and Population Health, Institute of Infection and Global Health, Leahurst Campus, University of Liverpool, CH64 7TE, UK.
(Email: jonread@liverpool.ac.uk)

it difficult to identify the interactions most likely to lead to transmission or the impacts of interventions meant to curtail those interactions. The lack of quantitative measures of infectious contacts makes the parameterization of large-scale transmission models difficult as well, particularly for models which rely explicitly on simulated social contacts to generate infection events.

Currently, there are a number of published studies which have sought to quantify potentially infectious contacts between individuals. This paper seeks to highlight recent and current research in this area and to review the various methodologies and instruments available to quantify social interactions that might lead to transmission. Relevant studies come from many different disciplines and, as such, we have not conducted a systematic review. We discuss the strengths and weaknesses of the methods employed in relevant studies, and suggest promising alternative technologies. Finally, we outline key research challenges currently preventing a better understanding of the role played by specific social contacts in the transmission of respiratory and close-contact infections.

NETWORKS OF CONTACTS

First, we define a potentially infectious contact. Brankston *et al.* [2] identified four main modes of transmission for respiratory infections, each of which defines a contact differently: we refer to these modes as airborne, droplet, direct, and fomite. Airborne transmission involves pathogens expelled from an infectious host through coughing or sneezing within small droplets, which become aerosolized. In this form, pathogens can remain suspended and viable in the air for long periods depending on environmental conditions, may be widely dispersed, and can be inhaled by susceptible hosts causing infection [5–7]. Droplet transmission is where pathogens are expelled from the host in larger, heavier droplets, which do not travel far from the infecting host, quickly settling on surfaces, whether those surfaces are inanimate objects or the mouth or mucosal surfaces of susceptible hosts [6, 8–10]. Direct transmission can occur where the secretions of an infector are transferred directly to the respiratory tract of a susceptible individual, such as through kissing or sharing a cigarette [11–13]. Fomite transmission is where pathogen is deposited by an infector onto an inanimate object (door handles, keyboards, etc.) also handled by susceptible

individuals and ingested [14–16]. It is unlikely that any respiratory pathogens are passed between hosts via a single transmission mode: transmission is more likely to occur to differing degrees via all four modes [17]. However, understanding the relative role of each mode is important, both for control and for defining the type of contact networks to quantify. Significantly, airborne and fomite transmission pathogens do not necessarily require hosts to have close contact to enable infection, while direct and droplet transmission do. Environmental conditions – such as temperature, humidity and ultra-violet light levels – affect the survival rate of pathogens, particularly viruses [8, 18], influencing the exposure risk to individuals and whether it is appropriate to consider them a contact. In a similar way, the use and material of surfaces upon which fomites are deposited influences pathogen viability [7, 19] and who should be considered a contact.

Contacts between individuals are often represented as a network, where nodes represent individuals and links between nodes represent their contacts and opportunities for transmission (Fig. 1). Network theory has provided epidemiology with many insights into how different patterns of contact and their network structures can affect transmission [20]. Visualizing transmission opportunities as networks is difficult, and often fails to capture many of the complex characteristics of the interactions. Sexual partnerships, which may often be more discrete in time and lower in number than social contacts, are difficult to reduce to simple networks based around degree distributions, due to other important characteristics such as sexual behaviour within partnerships which affect transmission risks [21, 22]. Moreover, the dynamical nature of interactions – changes in contact with individuals or changes in the duration or intensity of contact – is well known to be important for the transmission and control of infections, but is difficult to represent in a static network unless weighted links between nodes are used [23–27]. For respiratory infections, where contacts can be classified along a continuum of proximity and duration according to the transmission modes involved, there are likely to be severe limitations by representing interactions as a simple static network. The network approach, however, continues to be a useful way for researchers to think about interactions and transmission between individuals despite its shortcomings. In addition, network theory has shown the importance of different network properties and structures, such as

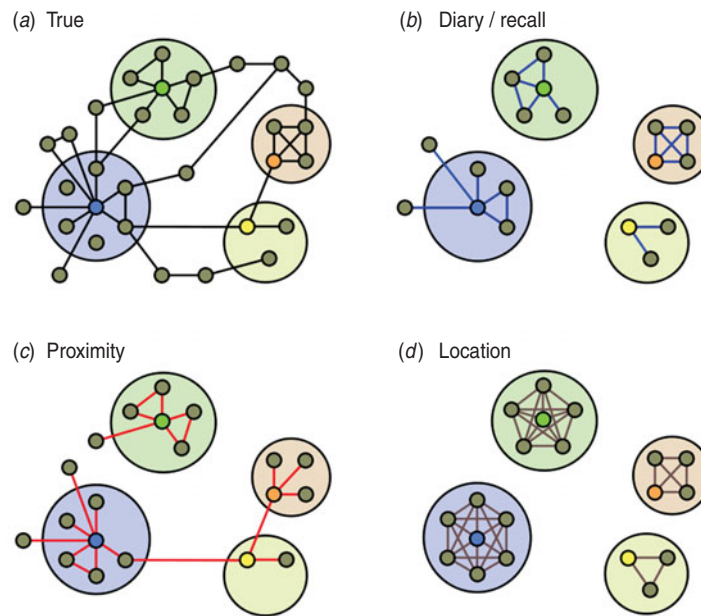


Fig. 1. A stylized illustration of the abilities of different methodologies to capture contact network structures. Individuals are represented as small circles and interactions which permit transmission are links between them. Larger circles represent particular environments or locations in which enclosed individuals are present at the same time. Individuals participating in hypothetical studies are coloured. (a) The true network of contacts between individuals which are viable transmission opportunities for a particular pathogen. (b) The network as measured by anonymous contact diaries, where contacts are not named. Egocentric contact diaries can be used to capture some information on the clustering of contacts, but cannot easily identify links between participants. As the information is subjective it can contain inaccuracies, both missing and erroneous contacts. (c) The network as measured by proximity sensors. Here, clustering may be inferred by the co-location of contacts, and some information on higher-order network structure (longer range links) between participants can be captured. (d) The network inferred by location-based network modelling, where modelled locations are depicted by larger circles. This method is critically dependent on accurate location description, and may lead to overestimation of both the number of contacts made and the local level of clustering.

clustering and centrality, for the spread of infections [20, 28–30].

The network framework permits detailed modelling of different pathogens, for a range of different scenarios, at the individual level. For certain infections where contact tracing is conducted (e.g. SARS or tuberculosis), some network information may be routinely collected in combination with outbreak and case data. For the majority of respiratory infections, however, there is relatively little data available to relate outbreak data to the contact network of cases. Studies of clustered outbreaks involving microbiological sampling may uncover the network structure behind the cases, but are expensive, labour-intensive and difficult. A more practical solution is to measure likely proxies of the true underlying network of potentially infectious contacts, with the presumption that these types of interactions are representative of those that effect transmission. For respiratory infections, these proxy contacts are likely to include spatial and temporal proximity. The methods and

studies described below are all based on this contact-proxy approach.

DIRECT OBSERVATION

The direct observation of individuals and their association in time and space is the foundation of many recent studies of contact between animals and their behavioural ecology [31, 32], but has rarely been employed to study human social interactions from the perspective of disease transmission. There is, however, a history of using video methods in ethnographic studies in the social sciences, for example to study family cohesion [33] or pedestrian interaction dynamics [34].

We know of only two published studies which directly observed and quantified human contact patterns and networks. Polgreen *et al.* [35] observed the number of encounters made by health-care workers during the course of their duties. Villaseñor-Sierra and colleagues [36] asked teachers to observe contacts

between children at daycare centres, where the children were too young to question directly.

Recent advances in software and technology mean that automation of observation studies may now be possible [37–40]; this may help considerably in gathering and analysing interaction data. Observations could be made using video cameras and relevant contacts between participants identified with computer software detection methods (e.g. spatial location and face-recognition technology) [41]. Neck-mounted cameras which automatically take a picture every few seconds have been trialled as a memory aid for patients suffering memory impairment [42, 43], and a similar system could record much of the social interactions an individual might make during a day, as well as detailing some information about the type of interaction and duration. Security camera recordings are a potentially useful data source for interactions in public space, if such data could be accessed in an ethical way: perhaps by using image processing algorithms to anonymize individuals immediately after they are recognized.

However, there are several reasons why this methodology may be unattractive for human subjects' research. Observational studies are likely to be expensive and logistically challenging if they are to capture the behaviour of a large number of people. There are limitations to the ability of current face-recognition and movement-tracking software algorithms to accurately identify and track the motion or interaction of individuals. Observation from multiple angles is often required to assist the software and to reduce occlusion [40], although this is unlikely to be sufficient in high-density public areas. Such studies may also be perceived as too invasive of individuals' privacy, however well the data are anonymized. Additionally, there is the perception that any observation of a subject would need to be exhaustive to capture all the contacts relevant to infection.

Nonetheless, there is great potential for utilizing observations or video recordings of normal interactions to quantify networks of potentially infectious contacts. An important innovation that video-based data provides is that it enables the estimation of several different types of social interaction relevant to transmission. Further, the entire interaction network may be documented (including sharing of objects and fomite pathways). Additionally, there is minimal burden or disruption to participants, although it may require labour-intensive analysis for the researchers. Observational methods may be most suitable to small

populations within closed settings, where all interactions can be observed easily. Household, work- and school-based studies of interactions may benefit greatly from this type of approach. Additionally, this may be the only reliable method to record interactions in locations which are difficult to capture using other methods, for example mass transport environments or crowded public spaces, although, again, there are obvious issues in protecting the anonymity of passive participants.

CONTACT DIARIES

Sociology has a long tradition of measuring and analysing social networks, where typically the emphasis is on relationships between the subject and other individuals (termed the ego and alters in the sociological literature). One method often used is the self-reported contact diary, where study participants are asked to record particular social interactions made within a time period [44–48]. Unfortunately for epidemiologists eager to parameterize models of disease transmission, few of the sociological studies conducted using such instruments are directly relevant. Definitions in these studies of 'interaction' can often include email or telephone communications, or are restricted to a limited subset of all possible encounters, such as those individuals whom participants can name or have a particular relationship with [46, 48].

In attempts to quantify social contact, epidemiological studies which use the contact diary approach have employed tailored definitions of a contact and attempt to only capture infection-relevant encounters. This method has now been applied, from an epidemiological perspective, in a variety of convenience studies [24, 49–52], school-based populations [53–57], healthcare settings [58], and larger studies of the general public [59–62]. The predominant methodology to emerge is for participants to record encounters made with other people that include a face-to-face conversation or physical touch (of skin on skin). Additional characteristics of contacts can also be sought, such as estimates of duration or typical frequency of contact, or the social setting of encounters. An important distinction should be made between 'anonymous' diaries, where subjects only record characteristics of their contacts, and diaries which ask subject to name their contacts – permitting repeat encounters with the same individual to be captured and the full network of participant encounters to be described [24, 63]. This

methodology is clearly dependent on how well individuals can be identified from the information recorded. In addition, it is possible to capture some higher-scale network properties beyond the first degree by asking participants to report which of their contacts also had contact with each other (see Fig. 1*b*). Some studies have assessed different study designs. Mikolajczyk & Kretzschmar [52] compared designs where participants completed diaries prospectively—adding contacts and information as they were made during study days—and retrospectively—reporting contacts the following day. They found, on average, more contacts were reported using the prospective design. In contrast, Beutels *et al.* [50] found no significant difference between prospectively and retrospectively collected contact data in a similarly designed study. McCaw *et al.* [64] compared participants' prospective estimation of the number of contacts with data collected retrospectively, finding participants tended to report fewer contacts when asked prospectively. This study also compared paper diaries with an electronic PDA device which prompted participants to complete contact information following a change in their location; participants reported more contacts when using the paper diary than the PDA device, and they preferred the paper diary, perhaps because it imposed less of a reporting burden.

Retrospective paper-based contact diaries were the recording instrument used by the largest study relevant to respiratory infections conducted so far. The POLYMOD study [60] recorded single-day contact characteristics of 7290 individuals in eight European countries. This study found strong assortativity of contacts by age—contacts tended to be made between people of a similar age. While the study found some differences in average contact rates between countries, there are remarkable similarities between countries, particularly in the age-based contact rates and the distribution of contacts between social settings, duration and intimacy. Indeed, quantifying age-related mixing patterns using the contact diary method has, perhaps, been the most significant advance to date in this field, helping to explain observed serological and infection patterns [59, 65–68].

There are two primary limitations of the contact diary methodology. First, there are difficult-to-quantify potential biases in participant recall and reporting, particularly for participants who have experienced a complicated day with many different contacts in different environments. Studies that

record identifiable contacts can assess these biases to some extent. Read *et al.* [24] found only 30.2% [95% confidence interval (CI) 25.7–35.0] of any identifiable encounters between participants were recorded by both participants. This recording rate improved when restricted to contacts involving touch (90.2%, 95% CI 83.2–95.0) or non-work settings (84.9%, 95% CI 75.5–91.7). Smieszek *et al.* [63] found a higher level of agreement between pairs of participants (65.0%, 95% CI 61.1–68.8) and also found the agreement rate improved when restricted to physical contacts (78.6%, 95% CI 67.1–87.5), and increased with the duration of reported contact, suggesting that short-lived encounters are those most easily forgotten. Collecting longitudinal information on participants' contacts is also subject to study fatigue: several longitudinal studies have found a decrease in reported contacts with time [63, 64, 69].

The third limitation is that defining contacts as those that include conversation or touch potentially limits the reported encounters to a subset of all the social encounters that could permit transmission. For example, there are many circumstances where individuals may be close enough to share airspace and pathogens, such as commuters on a crowded train, yet they may not converse or touch. Further, the conversation and touch definitions fail to capture interactions with shared objects that would permit fomite transmission, although appropriate questions could be included within diaries to attempt to capture such information.

Despite these shortcomings, the contact diary method has important advantages over other methods. The instrument is easy to administer and able to collect information retrospectively; it is able to capture social interactions in a wide range of environments and settings; it does not rely on peer-groups of participants. Additionally, it has been able to capture transmission-relevant mixing sufficiently well to explain age-based patterns of infection of varicella zoster virus and parvovirus B19 [70], mumps [59] and pertussis [65]. The development of the contact diary as an epidemiological instrument is likely to continue.

PROXIMITY SENSORS

The proximity sensing method relies on the use of automated electronic sensors, also called motes or tags, which can measure proximity to or presence within a particular distance of other sensing devices, using radio communication [71, 72]. A network of

spatial proximity of subjects can be derived from participants carrying devices, and contact between subjects inferred. Devices may be designed to only transmit, in which case static receiving base-stations are required to identify the spatial location of subjects [73–76]. Alternatively, devices worn by participants may both transmit and receive signals and therefore may detect the presence of other devices, removing the need for static base-stations [77]. Contact patterns of individuals may be inferred by either of these approaches.

There are two main types of communication signalling: radio-frequency identification (RFID) and Bluetooth. Both operate in a similar manner, transmitting signals that can be picked up by receiving devices. Radio-frequency tagging has been used in animal ecology but until recently has been little used for disease-relevant studies in humans. Advances in technology mean that very small, highly accurate devices with a long battery life are now available for social network studies. Bespoke devices have shown themselves to be ideal for such studies; being lightweight and compact, and able to record a large amount of data. Bluetooth has promise as an epidemiological instrument, because its use is now widespread in many developed societies, functioning in many portable devices such as mobile phones, wireless headsets, car satellite navigation systems or laptop computers, thus providing a unique opportunity for large-scale social observations [78].

A number of studies have now been conducted using proximity sensors worn by subjects in a different settings, including a conference [73], hospital wards [74], schools [75, 77], a public exhibition [73], and several other smaller convenience studies within the academic computer science community [72, 79–82].

Although the objectivity of electronic sensors is appealing compared to contact diaries, a significant disadvantage of proximity sensors is that only interactions between participants can be recorded. This limits their applicability to closed social environments or peer-groups. Another disadvantage of proximity sensors for social encounter studies is that the sensors can potentially record contacts where none may be made in an epidemiological sense (see Fig. 1*c*). This is a particular problem of Bluetooth technology in commercially available devices such as mobile phones. Bluetooth on these devices has a typical range of 5–10 m and can penetrate glass or walls. The range of custom-made RFID devices can be tuned, or the strength of signal recorded, and careful calibration

might overcome the problem seen in mass-produced hardware to some extent. It should also be noted that current sensing technologies are not perfect: reciprocity between electronic sensors, while better than contact diary methods, is not complete in field studies [77]. Nonetheless, automated objective methods of proximity measurement are clearly preferable to self-reported diary studies for particular circumstances, especially in closed environments where participation rates are high and most encounters can be captured.

INFERRING CONTACT PATTERNS FROM SECONDARY DATA ANALYSIS

We leave many traces of our physical presence in the modern world, at individual and collective levels. Examples include: credit card and mobile phone usage; purchasing habits and patterns; healthcare registration and history; vehicle licensing; employment and education history; security camera recordings; patterns of transport and traffic flow; national censuses of households, schools and workplaces. Some, although by no means all, of these types of information have been used to parameterize models of infection by inferring how often and where people interact. The inference of contact patterns from such diverse and disparate information sources is a justified approach given the absence of better estimates of true social mixing patterns. This type of analysis is now conducted routinely on large datasets, in which the movement and interactions of millions of people and are included. The primary objective of these studies is to assess the efficacy of infection control strategies, such as border quarantine controls or travel restrictions.

Socio-demographic information (population demography, employment data, household and workplace size distributions, etc.) have been used in a number of modelling studies to generate detailed simulations of epidemics [83–86]. These studies generate a ‘synthetic’ population, where individuals are typically assigned membership of virtual households and workplaces or schools. The location and size of the households, schools and workplaces is simulated so as to agree in aggregate with available data. Crucially, these modelling studies are forced to make assumptions regarding the transmission probabilities in different social settings and contacts with other household and workplace members, due to the lack of available information from contact studies at the time.

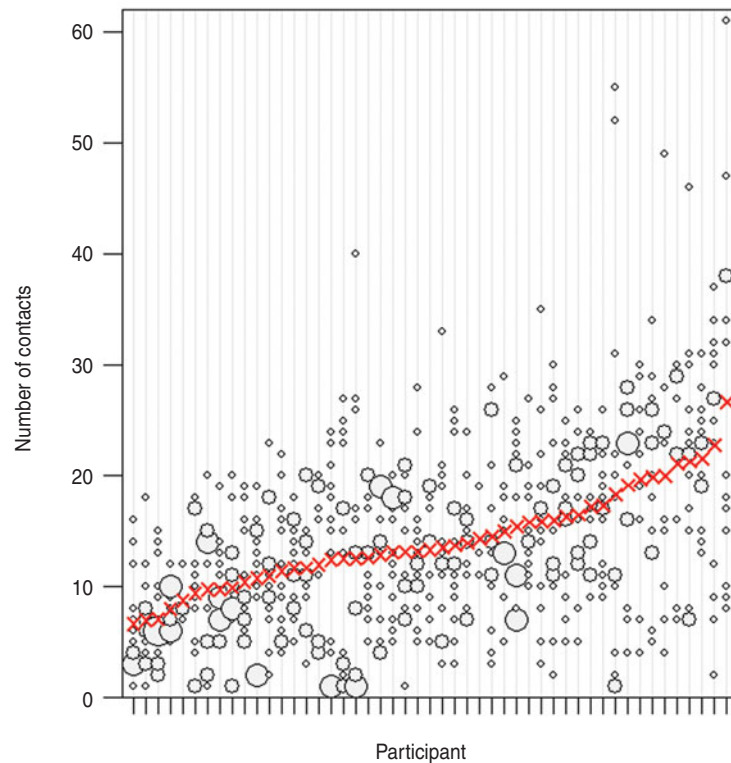


Fig. 2. The daily variation in number of contacts reported by participants in the study by Read *et al.* [24]. Participants are ranked by their average number of contacts (red crosses); the size of circles denotes the number of days where a participant encounters the same number of contacts. There is much day-to-day variation in participants' reported number of contacts.

An alternative approach generates contact networks from location and activity information, most notably the EpiSims model [87]. Here, specific premises (schools, households, business or leisure centres) and their geographical position, as well as their human capacity, are explicitly depicted in the model. Once these locations are integrated with known travel and activity information of residents entire cities are simulated, person by person, hour by hour, journey by journey. The raw simulations are used to generate detailed contact networks [87, 88]. Similar methods have been developed and employed by researchers from the human geography field [89], and by combining with time-use survey information [90, 91]. Two studies have compared contact patterns predicted by synthetic population modelling with observed patterns. The 'Little Italy' study [91] found a good comparison between the model's age-based mixing patterns and Italian contact patterns measured by the POLYMOD study [60]. Yang *et al.* [92] predicted a distribution of daily contacts and clustering metrics for a synthetic population, after calibrating the mean number of contacts to lie close to reported values.

The future of these very detailed simulation studies no doubt lies in the robust incorporation of information from contact studies as well as other data, such as mobile telephone location as a proxy for human location [93].

FUTURE CHALLENGES

The temporal dynamics of contact patterns

It is uncertain how much variation there may be in the daily patterns of social contact of individuals as few of the studies described above have measured the social contact patterns of subjects for longer than a single day. Consequently, it is unknown how long a sampling period is required to capture a representative picture of an individual's contact patterns given day-to-day variation and variability in routine by type of day. This sort of information has only been gathered for small convenience samples [24, 45, 94] (Fig. 2). A related issue is the uncertainty in how contact patterns may change in response to infection or control. It is largely unknown how contact patterns may change due to personal illness, illness of dependants, or

authority imposed school or workplace closures, although some recent studies have attempted to quantify some of these behaviours [56, 57] or inferred them by considering how contact patterns change during holiday periods [58, 69]. Understanding the character and magnitude of behavioural changes has important implications for the way that epidemics should be modelled, and also for estimating the societal impact of the disease itself and the efficacy of non-pharmaceutical interventions to prevent it.

Capturing fomite transmission

Fomite transmission, acquiring infection from shared objects such as door handles or telephones, is thought to be a significant route of infection for many respiratory infections yet it is largely ignored when modellers construct contact networks. There is much uncertainty for many respiratory pathogens as to what proportion of infection occurs via a fomite transmission route: it is not known even for a relatively well-researched pathogen like influenza [15]. Significantly, for the methodologies outlined above, fomite transmission does not require immediate temporal proximity – an individual exposing themselves to a pathogen on a surface or object may do so later than the infecting individual deposited the organisms. Thus, one challenge is how best to deconstruct measured contacts into those where individuals are present at the same time at a location and those who visit the same location but at different times. Explicit temporal and location information of individuals could generate sufficiently detailed information, combined with sampling of objects, surfaces and the environment for viable pathogen and an exposure or infection outcome would elicit valuable information. There is likely to be strong directionality in the fomite contact network, and the significance of this for infection dynamics is unknown.

How big is the small-world of infection?

In the 1950s and 1960s, the concept that society was a ‘small world’ was developed and attempts were made to measure the degrees of separation or steps between individuals in terms of relationship or influence [47, 95]. In the experiments, participants were asked to ensure a letter or package was delivered to a particular target individual, unknown to the starting participants, but they could only do this by passing it onto their friends. Although the experiments and the

reporting of them were flawed or incomplete [96], the ‘result’ that the median chain length was six has entered popular consciousness as a measure of higher-scale social connectivity. While the experimental design is clearly not directly relevant to disease transmission, it serves as an elegant example of large-scale structure within populations. Despite higher-order network structure being thought an important characteristic within social theory and networks of influence [97, 98], and a range of metrics being developed by researchers to quantify such structure [23, 99], it has rarely been quantified in representative populations from the perspective of close-contact or respiratory infection transmission.

The topology of network links beyond the first and second degree greatly influences how fast a respiratory infection can spread across a community or nation, and the most appropriate control [100]. There is a distinction to be made between social networks of encounters – most relevant to transmission – and social networks of influence and information diffusion, which the sociological literature has focused on. There is likely to be considerable overlap between these two types of network and in certain circumstances, such as a major outbreak with a high degree of social awareness, transmission may be a function of both networks [101]. Control efforts may be maximized by considering both networks [102], although this may be premature given neither network type is well quantified.

The closing of schools, workplaces or public gatherings is often suggested as a control measure for epidemics of respiratory infections. These non-pharmaceutical interventions attempt to force extra steps in the path of an infectious wave-front, and so reduce or halt the progression of transmission within a community. However, social control measures may prove futile if there are alternative paths of contacts of similar or even shorter steps between individuals [103]. Epidemiology currently lacks a quantified understanding of how these types of social distancing controls operate, and there is little evidence that they work in the way simple simulation models might predict. Retrospective analysis of social distancing policies in past epidemics can help to determine the strength of effect [104, 105], but sheds little light on the mechanism by which social interaction patterns actually changed. The challenge remains: how to measure relevant higher-order network properties of societies, and how they may change in response to infection threats or intervention.

Embedding contact networks into geographical space

As we move from individual or household scales up to the community or population scales, it seems reasonable to presume an interaction between higher-order network properties and geographical space. People are limited in their capacity to interact by a large extent through social convention and culture, and by restrictions on their movement and travel, whether through architecture, town and transport planning or large-scale geographical limitations. Very few studies have measured epidemiologically relevant social-contacts networks and geographical space. Bates *et al.* [106] found a strong negative correlation between the geographical distance between households and the number of social interactions made between them in rural Ecuador. A study of HIV within the USA [107] measured a social interaction network of sexual encounters and drug use and related this to distance between participants' residences, demonstrating an intriguing relationship and tantalizing clues about how network structure may scale with geography. However, no study has yet provided a reliable estimate on how a social network and its properties important for infection may scale with geographical distance. It may prove possible to infer some higher-scale properties from outbreak data [108]. Moreover, contact-tracing studies may be useful, particularly where genetic distance of infections can help identify transmission pathways.

Linking contact measures with infection risk

The major validation required for all of the methods described above is to directly relate measures of contact to measures of infection. We know of only two studies that attempted to simultaneously measure contact patterns and infection outcome for the same subjects. Group A streptococcus (GAS) infection in children was partly explained by the network structure of observed interactions at daycare [36]. The authors measured social connectivity between a group of 3- and 4-year-olds in a daycare centre by asking teachers to provide a categorical assessment of interaction between the children. Several genetically distinct strains of GAS bacteria were isolated from children. Those carrying identical GAS bacteria were more likely to have extensive contact and high levels of within-group social contact than the mean levels across all children. The study provides a rare simultaneous observation of multiple measures of social contacts and infection. However, analysis was limited

due to small numbers of children in the study and even fewer who became colonized or infected during the study.

Social connectivity measures at the community level, derived from self-reported food-sharing social contacts, were identified as risk factors for diarrhoeal infection [106]. The mean number of reported social contacts and the mean number of food-sharing contacts at the community level were both associated with risk of acquisition of diarrhoeal infection.

There are numerous studies, based around outbreak investigations using contact tracing, which could be used to link contact measures with infection risk. For example, Rea *et al.* [4] retrospectively assigned contact intensities to contacts of SARS patients, and found significant differences in estimated attack rates for contact encountered by patients in different social settings and with different levels of contact. Such analysis is reliant on the investigations reporting total contacts (rather than just those who are positively diagnosed), recording some characteristics of the encounters between contacts and patients (such as duration and setting), and having a good understanding of prior immunity within the study population.

There is, however, strong circumstantial evidence that the contact diary approach, with a definition of contact which includes conversational or physical touch encounters, can capture aspects of social interaction important for the transmission of infection [59, 65, 70]. These studies have investigated the utility of social contact data in explaining patterns of serological prevalence in a population. Wallinga and colleagues estimated age-specific transmission parameters in models with and without social contact data on the age-specific conversational or physical touch contacts to explain patterns of mumps seroprevalence in a cross-sectional survey collected in The Netherlands [59, 109]. Models that included social contact data performed better than models that assumed homogeneous or proportionate mixing between age groups. This result is encouraging for the utility of social contact information obtained by surveys. However, the model using this information was compared to relative simple models of mixing. Comparison to models using other age-specific information would help determine the relative utility of the information in this survey to other approaches. Melegaro *et al.* [70] compared multiple measures of social contact in its ability to predict seroprevalence of parvovirus and varicella in multiple European countries. They found

evidence that more intimate reported contact (i.e. physical contact, longer duration contact, home contact and frequent contacts) were more useful in fitting the seroprevalence pattern of these two diseases. The consistency of this with biological mechanisms of transmission and the consistent pattern across countries provides clear support of the utility of social contact information obtained from surveys. Rohani *et al.* [65] also used the POLYMOD age-mixing data to model pertussis infection in Sweden during subsequent periods of different immunization policies. They found including the contact data provided a reasonable fit to observed changes in incidence without including assumptions of waning immunity, but did not conduct an extensive comparison to a range of alternative models featuring age-specific assumptions.

Many authors have assumed that infection risk is proportional to the number of contacts that occur [60, 65]. However, this is not necessarily the case and methodological challenges remain in linking infection and empirical measures of contact [25]. Multiple advances have been made in recent years in linking infection data to mechanistic models of disease [110–112] and in inferring network structure from egocentric or incomplete network data [113, 114]. These advances will help determine what types of contacts are most likely involved in transmission and will guide future empirical study design and the development of analytical methods.

There are a number of important questions for which a mechanistic understanding of the role of contact networks in disease spread would greatly enhance modelling studies and improve control – from contact tracing to the efficacy of school closure and social distancing. What are the appropriate contacts to inform risk? Does the risk of acquiring respiratory infection depend on an individual's number of contacts, or their centrality within the social network? Does infection risk vary within different parts of the network, or by the social settings or location in which contacts occur?

CONCLUSIONS

A fundamental challenge within infectious disease epidemiology is to understand how individuals 'fit' together, through their interactions in time and space, to form a contact network upon which respiratory infections spread. A plethora of recent studies have quantified social mixing behaviour for a variety of

populations, with a range of different methods including observation, contact diaries and electronic proximity sensors. There is a clear need to continue to quantify social interactions, and to better understand the instruments that capture this information. An improved understanding of social mixing patterns and heterogeneities of behaviour may significantly improve targeted control interventions, and would lead to the development of non-pharmaceutical interventions that minimize transmission while also minimizing social and economic disruption.

Few large-scale studies that can be thought of as representative have been conducted, and little in the way of validation of method against infection risk has been explored. There is still considerable uncertainty as to the most appropriate tool for different research questions and within different social or medical settings. In addition, the diversity of approaches to studying human encounters makes interpreting results difficult. While proximity sensors may relieve the burden of participation that self-reporting diary methods impose, it is not clear how much of the vast amount of information they generate will be useful or how best to interpret such information. To date, no large-scale study has directly compared the social mixing behaviour captured by the range of methods. However, to some extent, a focus on the relative merits of one method of measuring contacts over another is premature. Perhaps the most important challenge is to relate any of the current methods directly to infection risk, for a variety of respiratory pathogens.

ACKNOWLEDGEMENTS

We thank Jon Crowcroft, Ken Eames, Chris Watkins and Eiko Yoneki for stimulating discussion. We also thank three anonymous reviewers for their helpful suggestions. J.M.R. and W.J.E. acknowledge support from the Economic and Social Research Council and the Medical Research Council (grant RES-355-25-0019) for funding workshop meetings. J.M.R., S.R., D.A.T.C. and J.L. acknowledge support from the National Institutes of Health (grant 1 R01 TW008246-01). J.M.R. and D.A.T.C. also acknowledge support from the Centers for Disease Control and Prevention (grant number: 1U01CK000179-01).

DECLARATION OF INTEREST

None.

REFERENCES

1. **Tellier R.** Review of aerosol transmission of influenza a virus. *Emerging Infectious Diseases* 2006; **12**: 1657–1662.
2. **Brankston G, et al.** Transmission of influenza a in human beings. *Lancet Infectious Diseases* 2007; **7**: 257–265.
3. **Musher DM.** How contagious are common respiratory tract infections? *New England Journal of Medicine* 2003; **348**: 1256–1266.
4. **Rea E, et al.** Duration and distance of exposure are important predictors of transmission among community contacts of Ontario SARS cases. *Epidemiology and Infection* 2007; **135**: 914–921.
5. **Yang W, Elankumaran S, Marr LC.** Concentrations and size distributions of airborne influenza a viruses measured indoors at a health centre, a day-care centre and on aeroplanes. *Journal of the Royal Society Interface* 2011; **8**: 1176–1184.
6. **Stilianakis NI, Drossinos Y.** Dynamics of infectious disease transmission by inhalable respiratory droplets. *Journal of the Royal Society Interface* 2010; **7**: 1355–1366.
7. **Weber TP, Stilianakis NI.** Inactivation of influenza a viruses in the environment and modes of transmission: a critical review. *Journal of Infection* 2008; **57**: 361–373.
8. **Tellier R.** Aerosol transmission of influenza a virus: a review of new studies. *Journal of the Royal Society Interface* 2009; **6** (Suppl. 6): S783–790.
9. **Chen SC, et al.** Viral kinetics and exhaled droplet size affect indoor transmission dynamics of influenza infection. *Indoor Air* 2009; **19**: 401–413.
10. **Gralton J, et al.** The role of particle size in aerosolised pathogen transmission: a review. *Journal of Infection* 2011; **62**: 1–13.
11. **Fleisher GR, Pasquariello PS, Warren WS.** Intrafamilial transmission of Epstein-Barr virus infections. *Journal of Pediatrics* 1981; **98**: 16–19.
12. **Odegaard K.** Kissing as a mode of transmission of infectious mononucleosis. *Lancet* 1967; **1**: 1052–1053.
13. **Rhoads MP, Magaret AS, Zerr DM.** Family saliva sharing behaviors and age of human herpesvirus-6b infection. *Journal of Infection* 2007; **54**: 623–626.
14. **Li S, et al.** Dynamics and control of infections transmitted from person to person through the environment. *American Journal of Epidemiology* 2009; **170**: 257–265.
15. **Boone SA, Gerba CP.** Significance of fomites in the spread of respiratory and enteric viral disease. *Applied and Environmental Microbiology* 2007; **73**: 1687–1696.
16. **Winther B, et al.** Environmental contamination with rhinovirus and transfer to fingers of healthy individuals by daily life activity. *Journal of Medical Virology* 2007; **79**: 1606–1610.
17. **Weber TP, Stilianakis NI.** Inactivation of influenza a viruses in the environment and modes of transmission: a critical review. *Journal of Infection* 2008; **57**: 361–373.
18. **Jones RM.** Critical review and uncertainty analysis of factors influencing influenza transmission. *Risk Analysis* 2011; **31**: 1226–1242.
19. **Rheinbaben F, et al.** Transmission of viruses via contact in a household setting: Experiments using bacteriophage straight phix174 as a model virus. *Journal of Hospital Infection* 2000; **46**: 61–66.
20. **Keeling MJ, Eames KT.** Networks and epidemic models. *Journal of the Royal Society Interface* 2005; **2**: 295–307.
21. **Britton T, Nordvik MK, Liljeros F.** Modelling sexually transmitted infections: the effect of partnership activity and number of partners on r_0 . *Theoretical Population Biology* 2007; **72**: 389–399.
22. **Nordvik MK, Liljeros F.** Number of sexual encounters involving intercourse and the transmission of sexually transmitted infections. *Sexually Transmitted Diseases* 2006; **33**: 342–349.
23. **Wasserman S, Faust K.** *Social Network Analysis: Methods and Applications*. Cambridge: Cambridge University Press, 1994.
24. **Read JM, Eames KT, Edmunds WJ.** Dynamic social networks and the implications for the spread of infectious disease. *Journal of the Royal Society Interface* 2008; **5**: 1001–1007.
25. **Smieszek T.** A mechanistic model of infection: why duration and intensity of contacts should be included in models of disease spread. *Theoretical Biology and Medical Modelling* 2009; **6**.
26. **Kretzschmar M, Morris M.** Measures of concurrency in networks and the spread of infectious disease. *Mathematical Biosciences* 1996; **133**: 165–195.
27. **Eames KTD, Keeling MJ.** Monogamous networks and the spread of sexually transmitted diseases. *Mathematical Biosciences* 2004; **189**: 115–130.
28. **Szendroi B, Csányi G.** Polynomial epidemics and clustering in contact networks. *Proceedings of the Royal Society, Series B: Biological Sciences* 2004; **271**: S364–S366.
29. **Eames KTD.** Modelling disease spread through random and regular contacts in clustered populations. *Theoretical Population Biology* 2008; **73**: 104–111.
30. **Smieszek T, Fiebig L, Scholz RW.** Models of epidemics: When contact repetition and clustering should be included. *Theoretical Biology and Medical Modelling* 2009; **6**.
31. **Kasper C, Voelkl B.** A social network analysis of primate groups. *Primates* 2009; **50**: 343–356.
32. **Hamede RK, et al.** Contact networks in a wild tasmanian devil (*Sarcophilus harrisii*) population: Using social network analysis to reveal seasonal variability in social behaviour and its implications for transmission of devil facial tumour disease. *Ecology Letters* 2009; **12**: 1147–1157.
33. **Ochs E, et al.** Video ethnography and ethnoarchaeological tracking. In: Pitt-Catsouphes M, Kossek EE, Sweet S, eds. *The Work and Family Handbook: Multidisciplinary Perspectives and Approaches*. New Jersey: Erlbaum Associates Inc., 2003, pp. 387–409.
34. **Helbing D, et al.** Self-organizing pedestrian movement. *Environment and Planning B: Planning and Design* 2001; **28**: 361–383.

35. **Polgreen PM, et al.** Prioritizing healthcare worker vaccinations on the basis of social network analysis. *Infection Control and Hospital Epidemiology* 2010; **31**: 893–900.
36. **Villaseñor-Sierra A, Quinñonez-Alvarado MG, Caballero-Hoyos JR.** Interpersonal relationships and group a streptococcus spread in a Mexican day-care center. *Salud Publica de Mexico* 2007; **49**: 323–329.
37. **Ge W, Collins RT, Ruback B.** Automatically detecting the small group structure of a crowd. In: *2009 Workshop on Applications of Computer Vision (WACV)*. Snowbird, UT, 2009, pp. 1–8.
38. **Ali S, Shah M.** Floor fields for tracking in high density crowd scenes. In: *10th European Conference on Computer Vision, ECCV 2008*. Marseille, 2008, pp. 1–14.
39. **Moussaid M, et al.** Experimental study of the behavioural mechanisms underlying self-organization in human crowds. *Proceedings of the Royal Society, Series B: Biological Sciences* 2009; **276**: 2755–2762.
40. **Gavrila DM.** The visual analysis of human movement: a survey. *Computer Vision and Image Understanding* 1999; **73**: 82–98.
41. **Cheriyadat AM, Radke RJ.** Detecting dominant motions in dense crowds. *IEEE Journal on Selected Topics in Signal Processing* 2008; **2**: 568–581.
42. **Doherty AR, Moulin CJA, Smeaton AF.** Automatically assisting human memory: a sensecam browser. *Memory* 2011; **19**: 785–795.
43. **Berry E, et al.** The use of a wearable camera, sensecam, as a pictorial diary to improve autobiographical memory in a patient with limbic encephalitis: a preliminary report. *Neuropsychological Rehabilitation* 2007; **17**: 582–601.
44. **Fu Y-c.** Measuring personal networks with daily contacts: a single-item survey question and the contact diary. *Social Networks* 2005; **27**: 169–186.
45. **Fu Yc.** Contact diaries: building archives of actual and comprehensive personal networks. *Field Methods* 2007; **19**: 194–217.
46. **Marsden PV.** Network data and measurement. *Annual Review of Sociology* 1990; **16**: 435–463.
47. **de Sola Pool I, Kochen M.** Contacts and influence. *Social Networks* 1978; **1**: 5–51.
48. **Bernard HR, et al.** Comparing four different methods for measuring personal social networks. *Social Networks* 1990; **12**: 179–215.
49. **Edmunds WJ, O’Callaghan CJ, Nokes DJ.** Who mixes with whom? A method to determine the contact patterns of adults that may lead to the spread of airborne infections. *Proceedings of the Royal Society, Series B: Biological Sciences* 1997; **264**: 949–957.
50. **Beutels P, et al.** Social mixing patterns for transmission models of close contact infections: exploring self-evaluation and diary-based data collection through a web-based interface. *Epidemiology and Infection* 2006; **134**: 1158–1166.
51. **Edmunds WJ, et al.** Mixing patterns and the spread of close-contact infectious diseases. *Emerging Themes in Epidemiology* 2006; **3**: 10.
52. **Mikolajczyk RT, Kretzschmar M.** Collecting social contact data in the context of disease transmission: prospective and retrospective study designs. *Social Networks* 2008; **30**: 127–135.
53. **Glass LM, Glass RJ.** Social contact networks for the spread of pandemic influenza in children and teenagers. *BMC Public Health* 2008; **8**: 61.
54. **Mikolajczyk RT, et al.** Social contacts of school children and the transmission of respiratory-spread pathogens. *Epidemiology and Infection* 2008; **136**: 813–822.
55. **Conlan AJ, et al.** Measuring social networks in British primary schools through scientific engagement. *Proceedings of the Royal Society, Series B: Biological Sciences* 2011; **278**: 1467–1475.
56. **Eames KTD, et al.** The impact of illness and the impact of school closure on social contact patterns. *Health Technology Assessment* 2010; **14**: 267–312.
57. **Eames KT, Tilston NL, Edmunds WJ.** The impact of school holidays on the social mixing patterns of school children. *Epidemics* 2011; **3**: 103–108.
58. **Bernard H, et al.** Nurses’ contacts and potential for infectious disease transmission. *Emerging Infectious Diseases* 2009; **15**: 1438–1444.
59. **Wallinga J, Teunis P, Kretzschmar M.** Using data on social contacts to estimate age-specific transmission parameters for respiratory-spread infectious agents. *American Journal of Epidemiology* 2006; **164**: 936–944.
60. **Mossong J, et al.** Social contacts and mixing patterns relevant to the spread of infectious diseases. *PLoS Medicine* 2008; **5**: 0381–0391.
61. **DeStefano F, et al.** Factors associated with social contacts in four communities during the 2007–2008 influenza season. *Epidemiology and Infection* 2011; **139**: 1181–1190.
62. **Horby P, et al.** Social contact patterns in vietnam and implications for the control of infectious diseases. *PLoS One* 2011; **6**.
63. **Smieszek T, et al.** Collecting close-contact social mixing data with contact diaries: reporting errors and biases. *Epidemiology and Infection* 2012; **140**: 744–752.
64. **McCaw JM, et al.** Comparison of three methods for ascertainment of contact information relevant to respiratory pathogen transmission in encounter networks. *BMC Infectious Diseases* 2010; **10**: 166.
65. **Rohani P, Zhong X, King AA.** Contact network structure explains the changing epidemiology of pertussis. *Science* 2010; **330**: 982–985.
66. **Birrell PJ, et al.** Bayesian modeling to unmask and predict influenza A/H1N1pdm dynamics in London. *Proceedings of the National Academy of Sciences USA* 2011; **108**: 18238–18243.
67. **Goeyvaerts N, et al.** Estimating infectious disease parameters from data on social contacts and serological status. *Journal of the Royal Statistical Society, Series C: Applied Statistics* 2010; **59**: 255–277.
68. **Kretzschmar M, Teunis PFM, Pebody RG.** Incidence and reproduction numbers of pertussis: estimates from serological and social contact data in five European countries. *PLoS Medicine* 2010; **7**: e1000291.

69. **Hens N, et al.** Mining social mixing patterns for infectious disease models based on a two-day population survey in Belgium. *BMC Infectious Diseases* 2009; **9**: 5.
70. **Melegaro A, et al.** What types of contacts are important for the spread of infections?: Using contact survey data to explore European mixing patterns. *Epidemics* 2011; **3**: 143–151.
71. **Weatherall J, Jones A.** Ubiquitous networks and their applications. *IEEE Wireless Communications* 2002; **9**: 18–29.
72. **Hui P, et al.** Pocket switched networks and human mobility in conference environments. In: *ACM SIGCOMM 2005 Workshops: Conference on Computer Communications*. Philadelphia, PA, 2005, pp. 244–251.
73. **Isella L, et al.** What's in a crowd? Analysis of face-to-face behavioral networks. *Journal of Theoretical Biology* 2010.
74. **Isella L, et al.** Close encounters in a pediatric ward: measuring face-to-face proximity and mixing patterns with wearable sensors. *PLoS One* 2011; **6**.
75. **Stehlé J, et al.** High-resolution measurements of face-to-face contact patterns in a primary school. *PLoS One* 2011; **6**.
76. **Cattuto C, et al.** Dynamics of person-to-person interactions from distributed RFID sensor networks. *PLoS One* 2010; **5**.
77. **Salathé M, et al.** A high-resolution human contact network for infectious disease transmission. *Proceedings of the National Academy of Sciences USA* 2010; **107**: 22020–22025.
78. **Eagle N, Pentland A.** Reality mining: sensing complex social systems. *Personal and Ubiquitous Computing* 2006; **10**: 255–268.
79. **Eagle N, Pentland AS, Lazer D.** Inferring friendship network structure by using mobile phone data. *Proceedings of the National Academy of Sciences USA* 2009; **106**: 15274–15278.
80. **Wang Y, Krishnamachari B, Valente TW.** Findings from an empirical study of fine-grained human social contacts. In: *6th International Conference on Wireless On-demand Network Systems and Services, WONS 2009*. Snowbird, UT, 2009, pp. 153–160.
81. **Sastry N, et al.** Data delivery properties of human contact networks. *IEEE Transactions on Mobile Computing* 2011; **10**: 868–880.
82. **McNett M, Voelker GM.** Access and mobility of wireless pda users. *Mobile Computing and Communications Review* 2005; **9**: 40–55.
83. **Longini Jr. IM, et al.** Containing pandemic influenza at the source. *Science* 2005; **309**: 1083–1087.
84. **Ferguson NM, et al.** Strategies for mitigating an influenza pandemic. *Nature* 2006; **442**: 448–452.
85. **Germann TC, et al.** Mitigation strategies for pandemic influenza in the United States. *Proceedings of the National Academy of Sciences USA* 2006; **103**: 5935–5940.
86. **Riley S, Ferguson NM.** Smallpox transmission and control: spatial dynamics in Great Britain. *Proceedings of the National Academy of Sciences USA* 2006; **103**: 12637–12642.
87. **Eubank S, et al.** Modelling disease outbreaks in realistic urban social networks. *Nature* 2004; **429**: 180–184.
88. **Del Valle SY, et al.** Mixing patterns between age groups in social networks. *Social Networks* 2007; **29**: 539–554.
89. **Yang Y, Atkinson PM.** Individual space-time activity-based model: a model for the simulation of airborne infectious-disease transmission by activity-bundle simulation. *Environment and Planning B: Planning and Design* 2008; **35**: 80–99.
90. **Zagheni E, et al.** Using time-use data to parameterize models for the spread of close-contact infectious diseases. *American Journal of Epidemiology* 2008; **168**: 1082–1090.
91. **Iozzi F, et al.** Little Italy: an agent-based approach to the estimation of contact patterns- fitting predicted matrices to serological data. *PLoS Computational Biology* 2010; **6**.
92. **Yang Y, Atkinson P, Etema D.** Individual space-time activity-based modelling of infectious disease transmission within a city. *Journal of the Royal Society Interface* 2008; **5**: 759–772.
93. **Gonzalez MC, Hidalgo CA, Barabasi AL.** Understanding individual human mobility patterns. *Nature* 2008; **453**: 779–782.
94. **Eames KT, Read JM, Edmunds WJ.** Epidemic prediction and control in weighted networks. *Epidemics* 2009; **1**: 70–76.
95. **Travers J, Milgram S.** An experimental study of the small world problem. *Sociometry* 1967; **32**: 425–443.
96. **Kleinfield JS.** The small world problem. *Society* 2002; **39**: 61–66.
97. **Granovetter MS.** The strength of weak ties. *American Journal of Sociology* 1973; **78**: 1360–1380.
98. **Watts DJ, Strogatz SH.** Collective dynamics of 'small-world' networks. *Nature* 1998; **393**: 440–442.
99. **Newman MEJ.** The structure and function of complex networks. *SIAM Review* 2003; **45**: 167–256.
100. **Salathé M, Jones JH.** Dynamics and control of diseases in networks with community structure. *PLoS Computational Biology* 2010; **6**.
101. **Funk S, et al.** The spread of awareness and its impact on epidemic outbreaks. *Proceedings of the National Academy of Sciences USA* 2009; **106**: 6872–6877.
102. **Eames KT.** Networks of influence and infection: parental choices and childhood disease. *Journal of the Royal Society Interface* 2009; **6**: 811–814.
103. **Bansal S, et al.** The dynamic nature of contact networks in infectious disease epidemiology. *Journal of Biological Dynamics* 2010; **4**: 478–489.
104. **Bootsma MC, Ferguson NM.** The effect of public health measures on the 1918 influenza pandemic in U.S. Cities. *Proceedings of the National Academy of Sciences USA* 2007; **104**: 7588–7593.
105. **Hatchett RJ, Mecher CE, Lipsitch M.** Public health interventions and epidemic intensity during the 1918 influenza pandemic. *Proceedings of the National Academy of Sciences USA* 2007; **104**: 7582–7587.

106. **Bates SJ, et al.** Relating diarrheal disease to social networks and the geographic configuration of communities in rural Ecuador. *American Journal of Epidemiology* 2007; **166**: 1088–1095.
107. **Rothenberg R, et al.** Social and geographic distance in HIV risk. *Sexually Transmitted Diseases* 2005; **32**: 506–512.
108. **Caley P, Philp DJ, McCracken K.** Quantifying social distancing arising from pandemic influenza. *Journal of the Royal Society Interface* 2008; **5**: 631–639.
109. **Van Druten JAM, et al.** Mumps, measles and rubella, a longitudinal serological study into degree of protection and risk of infection [in Dutch]. (Eindverslag Praeventiefondsproject no. 28-1348). The Hague, the Netherlands: Praeventiefonds, 1990.
110. **Groendyke C, Welch D, Hunter DR.** Bayesian inference for contact networks given epidemic data. *Scandinavian Journal of Statistics* 2011; **38**: 600–616.
111. **Cauchemez S, et al.** Role of social networks in shaping disease transmission during a community outbreak of 2009 H1N1 pandemic influenza. *Proceedings of the National Academy of Sciences USA* 2011; **108**: 2825–2830.
112. **Ionides EL, Bretó C, King AA.** Inference for non-linear dynamical systems. *Proceedings of the National Academy of Sciences USA* 2006; **103**: 18438–18443.
113. **Rhodes CJ, Jones P.** Inferring missing links in partially observed social networks. *Journal of the Operational Research Society* 2009; **60**: 1373–1383.
114. **Potter GE, et al.** Estimating within household contact networks from egocentric data. *Annals of Applied Statistics* 2011; **5**: 1816–1838.