

Thinking clearly about social aspects of infectious disease transmission

<https://doi.org/10.1038/s41586-021-03694-x>

Caroline Buckee^{1✉}, Abdisalan Noor² & Lisa Sattenspiel³

Received: 2 May 2021

Accepted: 4 June 2021

Published online: 30 June 2021

 Check for updates

Social and cultural forces shape almost every aspect of infectious disease transmission in human populations, as well as our ability to measure, understand, and respond to epidemics. For directly transmitted infections, pathogen transmission relies on human-to-human contact, with kinship, household, and societal structures shaping contact patterns that in turn determine epidemic dynamics. Social, economic, and cultural forces also shape patterns of exposure, health-seeking behaviour, infection outcomes, the likelihood of diagnosis and reporting of cases, and the uptake of interventions. Although these social aspects of epidemiology are hard to quantify and have limited the generalizability of modelling frameworks in a policy context, new sources of data on relevant aspects of human behaviour are increasingly available. Researchers have begun to embrace data from mobile devices and other technologies as useful proxies for behavioural drivers of disease transmission, but there is much work to be done to measure and validate these approaches, particularly for policy-making. Here we discuss how integrating local knowledge in the design of model frameworks and the interpretation of new data streams offers the possibility of policy-relevant models for public health decision-making as well as the development of robust, generalizable theories about human behaviour in relation to infectious diseases.

The ongoing COVID-19 pandemic highlights the continuing importance of global infectious disease threats, and the need to develop rigorous scientific theories to understand, quantify, and forecast the risks that pathogens pose to humanity. One of the most important lessons of the pandemic so far is that the central forces shaping local and global variation in disease burden and dynamics have been social, not biological. Although substantial biological questions remain unanswered, the multiple waves of infection that have been driven by shifting control policies and the heterogeneous public response to them^{1,2}, as well as the disproportionate impact of the disease on poor and marginalized communities around the world^{3–6}, are the defining features of the pandemic's trajectory on local and global scales.

Epidemiological models that describe the spread of infectious diseases through populations have been developed during the pandemic to understand and predict pathogen transmission and to guide public health policies^{7,8}. As a tool for synthesizing current knowledge, identifying key drivers of transmission, and planning public health policy, such models have a long history in research and public health^{9,10}, and are increasingly used to make decisions about health policy and global funding¹¹. Although uncertainties about biological aspects of pathogen transmission may be problematic for modelling, it is the social context—which is important not only in terms of model structure and parameterization but also with respect to the availability and interpretation of epidemiological data—that often presents the biggest challenges for capturing the essential features of disease dynamics^{8,12,13}.

Human societies are structured by cultural forces that define social relations, particularly between kin, and the spread of infection reflects these social structures—starting with the household or family unit, and extending to the structure of workplaces and public spaces, and the physical layouts of villages, towns, cities, and countries. The purpose of a model, whether purely theoretical or fit to data to inform decision-making in a specific context, will determine how detailed these social aspects of transmission need to be, with the adage that a model should be ‘as simple as possible but no simpler’ likewise taking on different meanings depending on the model's intended function. Intrinsic to this decision is a question of scale¹³: capturing population-level dynamics may not require individual-level detail about social interactions, but a model intended to understand local drivers of transmission may.

Data about social relationships that are relevant for modelling pathogen transmission are traditionally collected by censuses and other surveys^{14–16}, but the expansion of access to the internet has started to open up possibilities for a more expansive, real-time, and global approach to the collection of survey data^{17–19} and the development of relevant social science theories about human behaviour. Furthermore, new data streams from mobile devices—for example, via social media—are offering vast, relatively unexplored datasets about human mobility on a global scale^{20,21}. Despite the recent marked increase in the availability of these new datasets—a trend that has accelerated during the COVID-19 pandemic^{22–24}—challenges remain in using them to parameterize transmission models. In particular, the extent to which data from mobile phones provide an accurate proxy for contact rates that spread disease

¹Center for Communicable Disease Dynamics, Department of Epidemiology, Harvard T. H. Chan School of Public Health, Boston, MA, USA. ²World Health Organization, Geneva, Switzerland.

³Department of Anthropology, University of Missouri, Columbia, MO, USA. ✉e-mail: cbuckee@hsph.harvard.edu

Box 1

SEIR models and heterogeneous mixing

Susceptible–exposed–infectious–recovered (SEIR) models provide a mechanistic description of the transmission of a pathogen as it spreads between people in a population. In the simplest of these frameworks, the reproduction number (R_0) of a disease (the average number of secondary cases arising from a single infected individual in an entirely susceptible population) can be defined as $R_0 = bk/r$, where b is the probability of infection given contact, k is the contact rate between people in the population, and r is the rate of recovery. As the contact rate is directly proportional to the reproduction number (R_0) of a disease, the size, speed, spatial heterogeneity, and effect of an epidemic, as well as the interventions needed to prevent and contain it, are all fundamentally linked to the social and cultural processes that generate patterns of exposure. Assumptions must be made about how the contact rate scales with population size, and will depend on the transmission route of the pathogen. Beyond single population models, epidemic models designed to study the spread of infectious diseases in subdivided or geographically separate

populations (also often called metapopulations) also have a long history. Early models of this type considered an arbitrary number of mixing populations^{35,54,96–102}, and several of these efforts were stimulated by the need to develop models that could be used to help control the transmission of gonorrhoea in the USA^{103,104}. These models established the importance of core groups of highly sexually active individuals for the spread of sexually transmitted diseases. Similar early frameworks were developed to describe the importance of non-random biting of mosquito vectors for malaria transmission^{105,106}, and many early models focused on the fundamental question of ‘who mixes with whom’^{28,29,107}. Since then, a substantial body of theoretical work has been developed that examines how heterogeneous local mixing patterns, sometimes formalized as explicit contact network structures, can influence epidemics within a population^{108–110}, often using stylized abstractions of human interactions^{111–113}. The ubiquitous finding of these studies is that heterogeneous contact patterns can be important for both the spread and control of disease.

remains unclear^{25,26}. In fact, it is still difficult to parameterize social aspects of transmission in mechanistic models even in the context of sophisticated approaches to modelling and the addition of powerful new data. Nevertheless, as social scientists embrace and grapple with new data streams, infectious disease modellers have the opportunity to use them in the context of transmission models as “a way of thinking clearly”⁹ about the social drivers of epidemics.

Here we describe key social parameters that modellers must consider to effectively capture the dynamics of pathogen transmission on different scales. We draw a distinction between models in which appropriately disaggregated data about baseline human social dynamics—grounded in local knowledge—can provide mechanistic insights into disease transmission, and the challenges introduced when the quality, resolution, or paucity of epidemiological and behavioural data may constrain predictive power even if the social aspects of a model are well-specified. We also discuss the importance of understanding and predicting deviations from baseline behaviour that result from infection and public health policies, an issue that may need to be addressed using a fundamentally different kind of model structure. Models are increasingly informing target product profiles, global strategies, and investments in global public health programs for many infectious diseases. Social and behavioural aspects of transmission are often ignored in the name of generalizability and parsimony, with authors adopting the language of physics to justify these simplifications while also claiming to provide public health value to specific populations. Too often the integration of these models in decision-making processes at the national level remains weak. We believe that one of the most important challenges for our field is the development of flexible frameworks that integrate social contexts that are relevant for disease, a challenge that requires closer collaboration between social scientists and infectious disease epidemiologists.

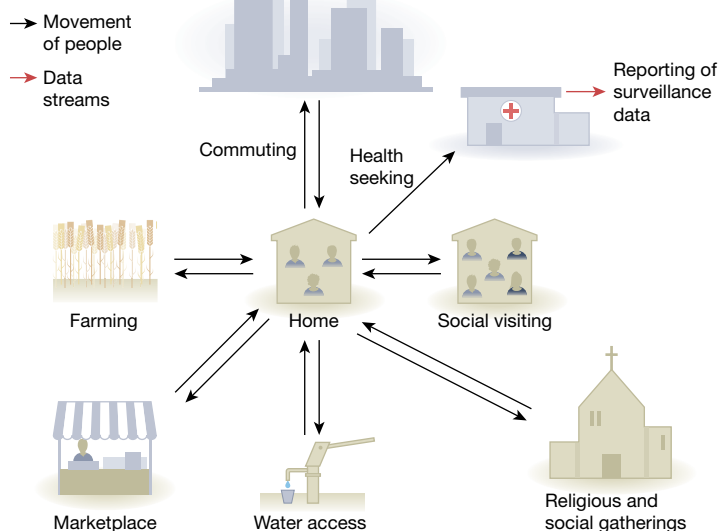
Parameterizing local contact rates

All mechanistic frameworks of infectious disease transmission make assumptions about how frequently people are exposed to disease, for example owing to close physical contact between susceptible and infectious people (the contact rate), and about the probability of infection when exposure occurs. In the well-studied

susceptible–infectious–recovered (SIR) model—first developed by Kermack and McKendrick nearly a century ago²⁷—mixing within a single homogeneous population, and therefore infection risk, was assumed to be random. The conceptual separation of the transmission coefficient into social and biological components was not the norm until the 1980s, as it became increasingly apparent that the population dynamics of HIV were driven disproportionately by transmission within particular demographic groups, which reflected highly non-random patterns of sexual contact^{28–34}. This separation provides conceptual differentiation of the social and cultural forces that drive uneven infection risk within a population from the biology of transmission itself^{30,35,36}, and is an essential feature to include to incorporate the effect of heterogeneous social interactions on transmission (Box 1).

In reality, assumptions of random mixing are always violated, even at the local and within-household scales, and the extent to which models must account for departures from them will depend on the mode of transmission of the pathogen and the purpose of the model. Kin structures are at the heart of all communities (Fig. 1). Comprehensive diary studies have revealed strong, age-structured mixing patterns related to household structures of nuclear families and peer groups shaped by schooling and patterns of employment^{14,16}, but these contact rates can change over time¹⁷ and vary substantially around the world³⁷. Recent technological developments have facilitated bluetooth-based and GPS studies of contact patterns^{37–39}, providing rich, granular data with increasingly large sample sizes, and a foundation for developing general principles of human interaction that could be used in epidemic models. Passively collected, aggregated mobile phone data have also become increasingly available on a near real-time basis and at scale²¹. During the COVID-19 pandemic, many modellers have begun to examine whether local mobility metrics and foot-traffic data are useful proxies for contact rates within populations^{40–46}. This can be effective when changes in mobility that occur on a scale that is measurable using mobile phone data are strongly correlated with contact rates. For example, at the beginning of the COVID-19 pandemic, Badr et al.⁴⁶ used aggregated mobility metrics from mobile phones to show that marked reductions in mobility occurred throughout the USA in mid-March of 2020, regardless of local social distancing policies, and that this was strongly associated with a drop in COVID-19 growth rates across the country. In this case, aggregated mobility data provided a

a Local transmission



b Seasonal and long-distance transmission

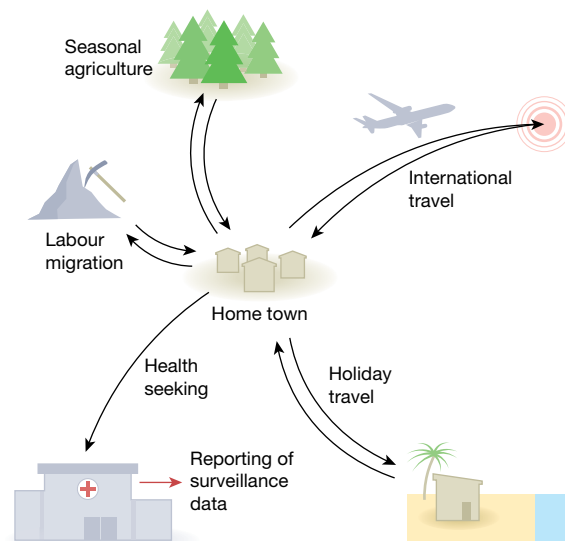


Fig. 1 | The nature of the kin structures in a community strongly influences who mixes with whom. Kin structures can affect patterns of interaction at all geographic scales. For example, at the local level, in some cultures extended families live and interact within the same house or complex. Kin relationships also strongly affect normal patterns of visiting behaviour, as much movement in all human cultures involves visiting kin. In places without running water, it is common for laundry to be done in streams or rivers, and this is often an important social activity, especially for women and children. In some cultures,

women and men have different roles, including differential participation in labour migration and agricultural work, and women may be confined to family compounds during the day. At the regional level, families regularly congregate for larger-scale social events such as weddings and holidays, with longer-distance travel connecting communities. When disease does occur, it is often measured by routine surveillance systems that only identify cases when treatment is sought, and when they are correctly diagnosed and reported.

meaningful proxy for the contact rates that drove changes in transmission on a county level. In general, however, if contact rates are decoupled from mobility patterns measured in this way—which the authors suggest occurred after April 2020 in the USA²⁵—an understanding of local transmission patterns still requires local data collection and/or contextual knowledge.

Careful examination of the interactions between people inside and outside their households that lead to heterogeneous infection risk can produce epidemiological models that yield powerful and generalizable insights, despite their specificity. For example, for the *Aedes* mosquito-borne virus that causes dengue fever, household variation in disease incidence has often been assumed to almost exclusively reflect the spatial distribution of mosquito vectors. However, by closely monitoring people's movements in relation to dengue clusters in Iquitos, Peru, Stoddard et al.^{47,48} showed that patterns of inter-household mobility associated with visiting friends and family were also a major driver of dengue transmission and needed to be considered in addition to spatial variation in mosquito densities. Similarly, by combining detailed survey data and precise location information about an outbreak of another *Aedes* mosquito-borne disease, Chikungunya, in Bangladesh, Salje et al.⁴⁹ reconstructed transmission chains to show that this disease was highly localized to socially connected households within particular communities, and that the 1.5 times higher risk of infection among women coincided with the 1.5 times higher likelihood that they stayed in the home during the day. The power of these studies reflects the combination of social science data and rich epidemiological information, coupled with sophisticated analytics. It is not that models without these details would be wrong per se, but rather that the addition of social science data provides important mechanistic insights into how transmission works on this local scale; behaviours and patterns of household visiting will define the course of any particular outbreak and must be understood when generating context-specific policy.

These concerns in turn emphasize the continuing importance of on-the-ground data collection from surveys, the value of

gender-disaggregated data—which the WHO only made standard practice for global health statistics in 2019⁵⁰—and the role of local knowledge in model and study design. Local knowledge is also key for interpreting epidemiological data used to fit and validate models (Box 2). Despite this, local social phenomena are often left out of disease models¹², sometimes because they are developed in a different context, for use at a different scale, or for academic purposes by researchers who are unaware of local realities, or because the social science data are time-consuming to collect or unavailable. The rich new data sources discussed above raise the question of how much detail should be included in order to understand the mechanisms that drive disease or to capture population-level dynamics at different scales. In all models, a trade-off exists between parsimony and realism that hinges on the scale and purpose of the model: while it is certainly true that “it makes no sense to convey a beguiling sense of ‘reality’ with irrelevant detail, when other equally important factors can only be guessed at,”⁹ it is also the case that a failure to capture the critical deviations from assumptions of random mixing may lead to weak predictions, misspecified estimates of transmission⁵¹, and poor policy decisions. Therefore, matching the model and data structures to the scale of the research or policy question becomes the most important challenge for capturing the social drivers of epidemiological dynamics within a population.

Regional mobility and between-population transmission

Travel outside the community also plays a key role in spreading diseases (Fig. 1), and spatial models of infectious diseases often incorporate travel as a migration rate between populations. Traditionally, simple, theoretically derived gravity and radiation models—both based on the reasonable idea that large populations attract travellers, but not requiring specific data about mobility—have often been used as fixed parameters to describe mobility dynamics in

Box 2

Model fitting and validation

Perhaps one of the least measured or understood uncertainties associated with models of pathogen transmission is in the epidemiological data that are used to parameterize, fit, and validate them. In particular, understanding the social and institutional environments in which case data are generated and reported is essential for any meaningful evaluation of whether a model is fit for purpose. For many pathogens, the data available to fit an epidemiological model are derived from surveillance systems based on the reporting of cases from clinics or hospitals around the country to a central government agency (Fig. 1). Any infection that is not diagnosed within this system will of course not be captured by routine surveillance, and missing cases can occur if infections are asymptomatic, if healthcare is not sought by the infected individual, if cases are misdiagnosed or there is a lack of diagnostic capacity, or if there are reporting delays or failures. These issues are exacerbated among marginalized populations who may not wish to engage in the health system, and in poorly resourced or designed health systems where reporting may be slow, incomplete, and lacking essential metadata. In the context of the COVID-19 pandemic, asymptomatic infections have made it challenging to assess the total number of cases in a community, but additional factors—patchy access

to diagnostic tests throughout the epidemic, variable testing criteria in different contexts, culturally specific treatment-seeking behaviours, and misaligned reporting incentives at different levels of government—have all contributed to wildly unreliable case counts. This has led to challenges in fitting models and estimating the reproduction number^{114,115} during the pandemic. ‘Nowcasting’ approaches have been developed to adjust for reporting delays¹¹⁶, and these have been successfully adapted for COVID-19 in some settings¹¹⁷. In addition, the use of alternative epidemiological data that may be more reliable, or are reported in a different way—for example, using excess mortality¹¹⁸ or even environmental proxies of transmission such as viral concentration in wastewater¹¹⁹—can provide a proxy for cases. Serological studies that measure antibodies in the wake of an infection can also be used to inform models about transmission dynamics^{120,121}. However, few models consider the institutional and social aspects of epidemiological data as more than a ‘reporting rate’ that can be estimated or assumed and is usually opaque with respect to mechanism. Exactly how social aspects of surveillance should be included in modelling frameworks in different settings remains unclear, but is an important consideration for future work.

these metapopulation models^{52–54}. The increasing availability of mobile-phone-derived data on regional mobility is permitting the validation of these frameworks in real-world settings. The results of such studies suggest that gravity models systematically underestimate the volume of long-distance travel in our highly connected world, and may do poorly in rural areas^{55–57}. They also highlight the importance of seasonal patterns of connectivity or asymmetric population shifts, such as holiday travel or displacement due to conflict or natural disasters. A recent comparison of aggregated mobile phone data from three countries showed that seasonal patterns of travel are a general feature of modern societies⁵⁸. For example, in Kenya, this seasonal flux in population density, coinciding with school term times, was shown to be a stronger predictor of the regional patterns of the childhood infection rubella than rainfall and other explanations, explaining Kenya’s unusual three-peak pattern of rubella incidence⁵⁹. Using mobile phone data to measure travel patterns in Bangladesh, Mahmud et al.⁶⁰ observed large travel surges occurring out of the capital city of Dhaka to all parts of Bangladesh during the Eid festivals. In 2017, this holiday coincided with a large Chikungunya outbreak in the city, which spread throughout Bangladesh after the holiday, just as outbreaks of respiratory viruses in the global north at the end of December are often associated with holiday travel^{61,62}.

Mobile phone data therefore provide valuable insights about these relative travel routes of millions of people between different places for the first time, as well as asymmetric movement patterns and large shifts in population density. There are limits to the insights mobile phone data streams can afford, however, primarily related to a gap in social insight; they are spatially coarse relative to contact patterns that spread disease, as previously discussed; they have implicit biases (for example, they do not include children and other people without phones); and they usually do not tell us anything about who is travelling and why. It will be important to measure and quantify bias and representativeness in these datasets^{26,63,64} as they are used more routinely, as well as to engage in meaningful efforts to standardize approaches to both analysis and privacy for this relatively new public health application²⁶.

Analogous to the problems with random mixing assumptions in single-population models, the contact rate between populations often reflects travel by particular subsets of the population; in other words, the probability of travelling is not randomly distributed, but the mobility rates in most models assume that it is, and usually mobile phone data are not disaggregated demographically, in order to preserve the privacy of subscribers. When mobility data are available that are disaggregated, by gender for example, striking differences in mobility may emerge. A study of an urban setting in Latin America⁴¹ illustrates this heterogeneity well—women are both more localized in their movements and visit fewer locations than men, which may be important for infection dynamics in a given setting, depending on the pathogen. Surveys have shown that women with children also travel less to urban centres across sub-Saharan Africa compared to other demographics⁵⁶. In a rural setting in Bangladesh, a survey of patients with malaria showed similar trends, with men travelling much greater distances than women⁶⁵. Given that global gender roles often follow this pattern, it is likely that these findings are general and relevant for building robust disease models, especially when combined with gender-disaggregated health data. Contextual understanding coupled with disaggregated data, often generated using traditional social science approaches, is therefore not something we can do away with in the era of big data. Rather, these new data streams are uncovering dynamics that will be much more powerful when complemented by social science data and analysis.

Some of these gender differences in regional mobility are related to occupational activities, which is another important factor that drives contact rates between populations. Labour migration has long been studied by social scientists, but is challenging to incorporate into epidemic models. The importance of labour migration in particular demographic groups, for example linked to forest and plantation work, agriculture and livestock, or gold mining^{66,67}, has been known to drive regional patterns of malaria transmission for decades⁶⁸. In the Sahel region of Africa, where the malaria burden is intense and highly seasonal, pastoral livestock farming is a key economic activity, and pastoralist communities are highly mobile as they search for pasture

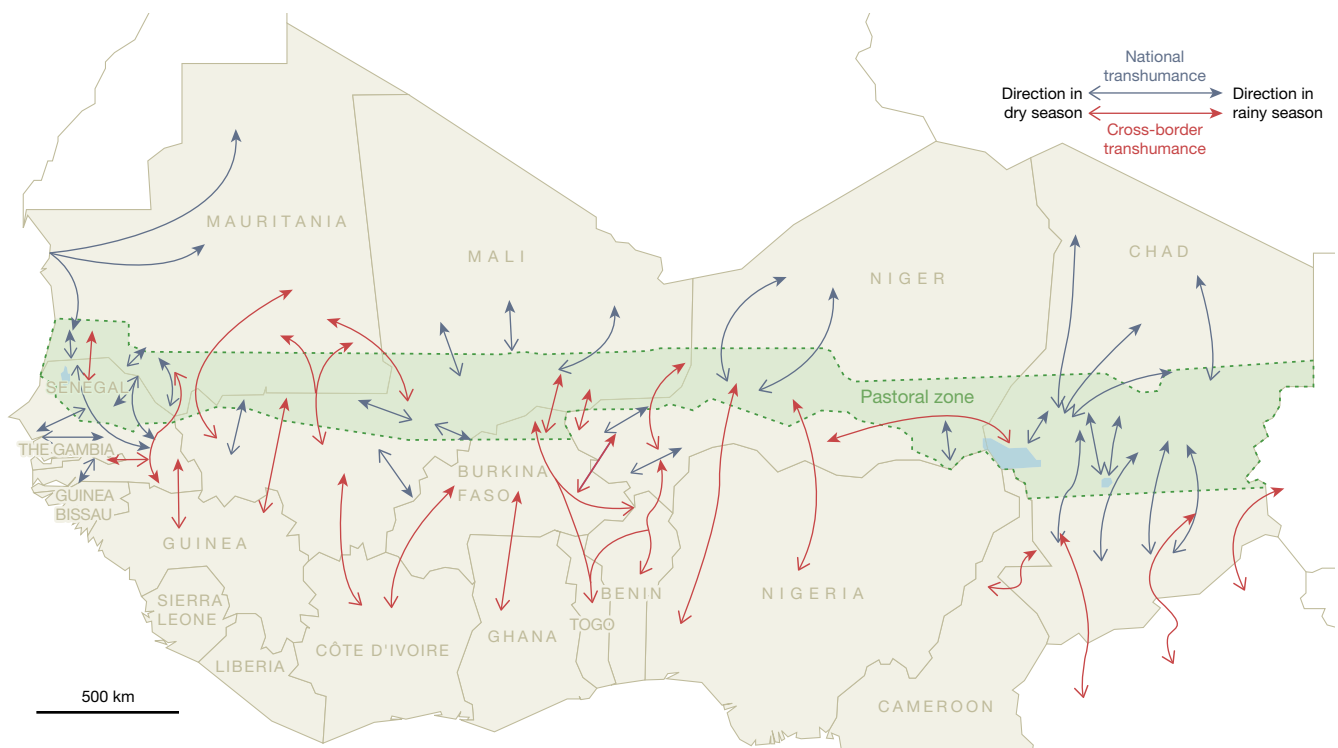


Fig. 2 | Transhumant pastoralists cross through these climatic zones during the course of the year. Transhumance is a type of nomadism in which the seasonal movement of people is driven by the need for livestock pasture. In the rainy season, pastoralists in the Sahel region spread into rich, but short-lived, pastures, while they move further south with the onset of the dry season. After spending the height of the dry season in the more humid south, they move back north before the beginning of the agricultural activities of the rainy season⁹⁴.

These seasonal movements involve both cross-border (red arrows) and national (blue arrows) travel patterns that shape the lifestyles of these populations. The influence of seasonal movements of particular subsets of a population on these different national and international scales is challenging to measure and capture within transmission models of disease. Adapted with permission from ref. ⁹⁵.

and water for their livestock, often in areas that expose them to malaria, and they exhibit seasonal migration within and between countries (Fig. 2). Rapid environmental changes and competition for land have adversely affected pastoralist production systems, which has resulted in conflicts, volatility in mobility patterns, and various other adaptive behaviours that are hard to generalize, but are fundamental to malaria transmission and control in the region.

Mathematical frameworks of malaria, which were among the first epidemiological models to be developed for any infectious disease⁶⁹, struggle to accommodate this kind of mobility. In fact, projections for future scenarios of malaria transmission under various interventions in sub-Saharan Africa that are based in part on mechanistic frameworks may not include any mobility parameters⁷⁰. In the absence of understanding of these social contexts, models may assume that the prevalence of infection reflects local transmission characteristics, rather than imported infections. However, surveillance data from around the world suggest that imported infections actually represent the majority of cases in some settings. In a study in Nairobi, for example, two-thirds of patients with malaria tested in a facility in an informal settlement had a history of travel and nearly 80% of those who had travelled had visited counties with high malaria transmission⁷¹. In settings with frequent importation, therefore, policy targets and funding should focus on managing infections in travellers, not local

mosquito control, and models of malaria transmission that fail to account for mobility will fail to capture the key socioeconomic mechanisms that drive the disease. Although these human aspects of malaria transmission continue to be emphasized as major impediments to elimination⁷², both generalizable and specific mobility frameworks are lacking and are often ignored in malaria transmission models that are used to guide elimination scenario planning, leading to the mistaken general assumption that low incidence regions are straightforward elimination targets⁷³.

New data streams—not only from mobile phones but also from surveys and malaria parasite genetic data, which yield insights into the relatedness of different parasite populations—are allowing more sophisticated modelling approaches to identifying the ‘sources’ and ‘sinks’ of malaria infections, however. Chang et al.⁷⁴ combined mobile phone data with parasite genetic data and surveys to model the spread of malaria in rural Bangladesh, for example. Modelling the expected flow of parasites using these different inputs as mobility parameters showed that there was broad agreement between models; parasites moved east to west as people travelled between the forests and more populous regions, with the survey confirming the importance of labour migration to the forest. In many ways, the relatively sophisticated modelling approach confirmed what the National Malaria Control Program already knew—that people get malaria in the forest—but it provided

useful evidence for this local knowledge, as well as estimates about the volumes of importation and specific routes and hotspots on which to focus interventions. Efforts to combine and validate new data streams, as well as more theoretical models such as gravity and radiation models⁵⁷, with surveys and other social science tools will be an important next step in the development of general mobility frameworks that describe labour migration around the world.

Epidemics are not like the weather

So far we have focused on behaviours that can be included as model parameters, such as contact rates between groups and baseline travel behaviour. In these cases, more, better quality, or different data about contact rates can improve model accuracy. However, human behaviour can also change in response to their awareness of, and information about, a disease. This can create feedback between the real (prevalence-based) or perceived (belief-based) risk of an infection, assessed by individuals in a population based on available information, and the behaviours that in turn drive its transmission, such as contact rate or the use of preventive interventions⁷⁵. Here, human behaviour must be built into epidemiological models, with social parameters mechanistically linked to changes in the disease itself or beliefs about the disease⁷⁶.

For endemic diseases for which prevention requires active participation by affected communities, such as treatment seeking and the use of preventive measures, understanding human behaviour in the context of risk perception and avoidance is essential for the development of dynamic frameworks to predict the impact of public health policies. Treatment seeking and adherence to drug regimens in the context of tuberculosis^{77,78}, the use of condoms in the prevention of HIV transmission^{79,80}, and sleeping under insecticide treated nets to prevent malaria⁸¹ all represent examples of complex human behaviours—particularly adherence to treatment and other interventions—that are challenging to integrate into model frameworks. In fact, these three major infectious disease threats are arguably among the most challenging for which to create robust theoretical frameworks in the context of interventions, for this reason⁸².

Epidemics, and the public alarm they can generate, create particularly strong feedback between behaviour and disease dynamics. For example, Epstein et al.⁸³ modelled two interacting contagion processes that describe the spread of infectious disease and the spread of fear about the epidemic, which leads individuals to effectively remove themselves from the population. These social–epidemiological feedback loops lead to more complex disease dynamics than expected under a model with fixed behaviours; fear of the disease drives behavioural changes in contact rate as people take steps to isolate themselves, leading to flattened epidemic peaks and multiple waves of infection as the perceived and/or real risk of disease fluctuates. This is exactly what we have observed during the COVID-19 pandemic, with the social response to specific interventions such as social distancing driving the variable course of SARS-CoV-2 incidence around the world². The publicly available data from online information sharing on platforms such as Twitter, where information and misinformation spread in parallel with the epidemic itself, will provide a rich source of information for investigating how different societies have reacted to the pandemic⁸⁴. These social media data come with the same caveats discussed above in the context of mobile phone data, and the representativeness of new data streams in different contexts should be reported by data providers, and adequately measured—through social science studies, among others—and accounted for in modelling research and application.

In this context, the COVID-19 pandemic has created an incredible natural experiment on a global scale, with similar policies being enacted around the world in diverse social contexts. Similarities and differences in the trajectories of local epidemics of the same

virus reflect the variable populations involved. Some social dynamics related to the contact rate parameters that we have discussed above do appear to be generalizable over time and in different contexts, and have been repeated around the world during the COVID-19 pandemic. In early 2020, Kissler et al.⁸⁵ measured SARS-CoV-2 prevalence in women who gave birth at different New York City hospitals and found that there were marked differences across the city, ranging from about 10% in Manhattan to as high as 50% in the Bronx. Analysis of mobility data from Facebook users over the same time period showed that these local variations in incidence were strongly associated with continuing commuting behaviour in neighbourhoods with lower socioeconomic status, consistent with the inability of essential workers to lock down. In fact, this inability of lower-income people to reduce mobility as much as those in wealthier neighbourhoods has been associated with differential disease burden and mortality in cities around the world⁸⁶.

Another characteristic response to policies during the pandemic has been an emptying out of urban centres. Throughout history, people have always fled urban centres when an epidemic hits, whether due to an outbreak of cholera in historical London, or due to a perceived but non-existent outbreak of bubonic plague that caused mass panic and the displacement of hundreds of thousands of people from Surat, India in 1995⁷⁶. These shifts in population density and increases in long-distance travel have important and general implications for understanding infectious diseases and designing public health policies. In response to COVID-19 lockdown policies, mobile phone data from around the world have uncovered similar behavioural responses to lockdown policies and travel restrictions, with similar dynamics occurring in urban centres in the USA, France, Spain, India, and Bangladesh⁸⁷. It is likely that the social and demographic factors that drive this urban–rural migration vary greatly in these different settings, with the exodus from Manhattan perhaps representing wealthy people going to country homes and the movement patterns of people in Bangladesh corresponding to the movement of workers in response to the closing and re-opening of garment factories. These differences emphasize the power of coupling large-scale datasets with local context, and highlight the importance of further setting-specific research to untangle the general versus local drivers of these behaviours.

There is currently strong interest in the development of national epidemic modelling and forecasting centres in the USA and elsewhere, with parallels being drawn to the evolution of weather forecasting services. The cases described above bring into question the extent to which disease forecasting efforts can be compared to weather forecasts; however, are shifting behaviours based on local and global information about epidemics (and related policies) ever going to be predictable in the way that physical laws are? If a weatherman forecasts rain and everyone stays at home, it still rains. Not so with infectious diseases. We argue that short-term forecasting efforts that use ensemble or consensus approaches^{88–90} are promising, and for models with the goal of making predictions rather than understanding mechanisms, simple approaches are often adequate, if not more tractable and therefore desirable. This is highlighted by disease forecasting efforts for COVID-19⁸⁸, Ebola⁸⁹, and influenza⁹¹, in which simple models can produce predictions as powerful as those of more complex models over a short timescale. However, a deeper understanding of social and behavioural aspects of risk perception and decision-making is likely to be needed to make medium-term predictions and mechanistic statements about possible future trajectories of epidemics, or to develop models designed to inform interventions in different settings.

Key to these latter models will be more research to understand how people's behaviours will change in response to particular policies; in this case, experimental or quasi-experimental evidence may be required to improve the predictive power of models that include social–epidemiological feedback mechanisms⁹². We argue, therefore,

that rather than creating one large model to forecast disease outbreaks, like a weather forecast, developing distributed research capacity that can respond to specific outbreaks in particular contexts by designing models flexibly and with feedback from local health systems is perhaps a wiser investment.

Thinking clearly about modelling epidemics

If mathematical models are “no more and no less than a way of thinking clearly,”⁹ then it is essential for modellers to think clearly about how modelling decisions about social aspects of transmission reflect the model’s function. In particular, what social phenomena contribute to mechanistic aspects of transmission that are important for population-level dynamics, are these measurable, and, if not, how do they constrain the model’s utility in different contexts? What principles should guide decisions about the scale of a model (for example, individual versus population) given the questions that are being addressed, and what kinds of data are needed to parameterize and validate the resulting model?

There are important distinctions between epidemiological models developed with scientific goals of understanding disease ecology, and epidemiological models developed to inform public health policies, for example. Models with primarily academic or theoretical goals tend to err on the side of abstraction, whereas policy-relevant models may have a more ‘realistic’ depiction of contact rates. Indeed, there is sometimes an implicit assumption in the infectious disease modelling field that integrating setting-specific social interactions in a disease model will inevitably detract from its generalizability, limiting its relevance beyond a particular case study. We argue against this dogma and propose that for models intended to understand mechanisms driving outbreaks—particularly on local scales—“data is the plural of anecdote.”⁹³ It is often by understanding specific social contexts and integrating insights from multiple applications to different contexts that general principles can be drawn. This philosophy requires a distributed generation of knowledge that is unwieldy to integrate into a unified theory, but ultimately can lead to general principles about what can, and what cannot, be ignored about local social contexts for models with different purposes.

In the public health context, substantial investment in modelling capacity is needed at the local and regional levels—not just in the context of dynamical modelling, but also for general statistical and quantitative capacity—to translate the sophisticated and data-rich approaches now available to us into better decision-making. This would ensure that models are being used appropriately in the context of policy. For example, many of the decisions we have discussed, such as whether a simple or a complex model is needed or whether new or different data streams would be helpful, require local literacy in quantitative methods that may be lacking. Partnerships between academic centres or research institutes and public health agencies and governments, as well as better training infrastructure, is therefore needed on an ongoing basis and in the context of endemic pathogens, so that modelling tools can be developed rapidly when a crisis such as COVID-19 occurs.

1. Liu, Y., Morgenstern, C., Kelly, J., Lowe, R. & Jit, M. The impact of non-pharmaceutical interventions on SARS-CoV-2 transmission across 130 countries and territories. *BMC Med.* **19**, 40 (2021).
2. Li, Y. et al. The temporal association of introducing and lifting non-pharmaceutical interventions with the time-varying reproduction number (R) of SARS-CoV-2: a modelling study across 131 countries. *Lancet Infect. Dis.* **21**, 193–202 (2021).
3. Malani, A. et al. Seroprevalence of SARS-CoV-2 in slums versus non-slums in Mumbai, India. *Lancet Glob. Health* **9**, e110–e111 (2021).
In this study, seroprevalence estimates in different parts of Mumbai, India, showed marked differences in SARS-CoV-2 exposure by July 2020, with between 55% and 64% of people in slum regions testing positive for antibodies against SARS-CoV-2, compared with 12–19% of people in non-slum regions.
4. Mackey, K. et al. Racial and ethnic disparities in COVID-19-related infections, hospitalizations, and deaths: a systematic review. *Ann. Intern. Med.* **174**, 362–373 (2021).

5. Mena, G. E. et al. Socioeconomic status determines COVID-19 incidence and related mortality in Santiago, Chile. *Science* **372**, eabg5298 (2021).
This paper showed that in Santiago, Chile, the socioeconomic status of neighbourhoods is strongly associated with COVID-19-associated morbidity and mortality, and linked to mobility patterns and access to healthcare, for example, testing rates.
6. Karmakar, M., Lantz, P. M. & Tipirneni, R. Association of social and demographic factors with COVID-19 incidence and death rates in the US. *JAMA Netw. Open* **4**, e2036462 (2021).
7. Donnelly, C. & Ghani, A. Real-time epidemiology: understanding the spread of SARS. *Significance* **1**, 176–179 (2004).
8. Grassly, N. C. & Fraser, C. Mathematical models of infectious disease transmission. *Nat. Rev. Microbiol.* **6**, 477–487 (2008).
9. May, R. M. Uses and abuses of mathematics in biology. *Science* **303**, 790–793 (2004).
10. Anderson, R. M. & May, R. M. Population biology of infectious diseases: part I. *Nature* **280**, 361–367 (1979).
11. Heesterbeek, H. et al. Modeling infectious disease dynamics in the complex landscape of global health. *Science* **347**, aaa4339 (2015).
12. Ferguson, N. Capturing human behaviour. *Nature* **446**, 733 (2007).
13. Funk, S. et al. Nine challenges in incorporating the dynamics of behaviour in infectious diseases models. *Epidemics* **10**, 21–25 (2015).
14. Mossong, J. et al. Social contacts and mixing patterns relevant to the spread of infectious diseases. *PLoS Med.* **5**, e74 (2008).
In this large diary-based study across countries in Europe, strong age-structured contact patterns were described with variation between different countries; these contact matrices—and others collected in similar ways—are frequently used to parameterize mathematical models of infectious disease transmission.
15. Kretzschmar, M. & Mikolajczyk, R. T. Contact profiles in eight European countries and implications for modelling the spread of airborne infectious diseases. *PLoS One* **4**, e5931 (2009).
16. Prem, K., Cook, A. R. & Jit, M. Projecting social contact matrices in 152 countries using contact surveys and demographic data. *PLOS Comput. Biol.* **13**, e1005697 (2017).
17. Feehan, D. M. & Mahmud, A. S. Quantifying population contact patterns in the United States during the COVID-19 pandemic. *Nat. Commun.* **12**, 893 (2021).
18. Feehan, D. M. & Cobb, C. Using an online sample to estimate the size of an offline population. *Demography* **56**, 2377–2392 (2019).
19. Lazer, D. et al. Computational social science. *Science* **323**, 721–723 (2009).
20. Wesolowski, A., Buckee, C. O., Engø-Monsen, K. & Metcalf, C. J. E. Connecting mobility to infectious diseases: the promise and limits of mobile phone data. *J. Infect. Dis.* **214** (Suppl. 4), S414–S420 (2016).
21. Buckee, C. O. et al. Aggregated mobility data could help fight COVID-19. *Science* **368**, 145–146 (2020).
22. SafeGraph Inc. *SafeGraph Data for Academics* (accessed 27 May 2021); <https://www.safegraph.com/academics>
23. Google. *COVID-19 Community Mobility Reports* (accessed 27 May 2021); <https://www.google.com/covid19/mobility/>
24. COVID-19 Mobility Data Network. *Facebook Data for Good Mobility Dashboard* (accessed 27 May 2021); <https://visualization.covid19mobility.org/?region=WORLD>
25. Badr, H. S. & Gardner, L. M. Limitations of using mobile phone data to model COVID-19 transmission in the USA. *Lancet Infect. Dis.* **21**, e113 (2021).
26. Kishore, N. et al. Measuring mobility to monitor travel and physical distancing interventions: a common framework for mobile phone data analysis. *Lancet Digit. Health* **2**, E622–E628 (2020).
27. Kermack, W. O. & McKendrick, A. G. Contributions to the mathematical theory of epidemics—I. *Bull. Math. Biol.* **53**, 33–55 (1991).
28. May, R. M. & Anderson, R. M. Transmission dynamics of HIV infection. *Nature* **326**, 137–142 (1987).
29. Diekmann, O., Dietz, K. & Heesterbeek, J. A. P. The basic reproduction ratio for sexually transmitted diseases: I. Theoretical considerations. *Math. Biosci.* **107**, 325–339 (1991).
30. Jacquez, J. A., Simon, C. P., Koopman, J., Sattenspiel, L. & Perry, T. Modeling and analyzing HIV transmission: the effect of contact patterns. *Math. Biosci.* **92**, 119–199 (1988).
31. Anderson, R. M., Gupta, S. & Ng, W. The significance of sexual partner contact networks for the transmission dynamics of HIV. *J. Acquir. Immune Defic. Syndr.* **3**, 417–429 (1990).
32. Gupta, S., Anderson, R. M. & May, R. M. Networks of sexual contacts: implications for the pattern of spread of HIV. *AIDS* **3**, 807–818 (1989).
33. Anderson, R. M., Blythe, S. P., Gupta, S. & Konings, E. The transmission dynamics of the human immunodeficiency virus type 1 in the male homosexual community in the United Kingdom: the influence of changes in sexual behaviour. *Phil. Trans. R. Soc. Lond. B* **325**, 45–98 (1989).
34. Diekmann, O., Heesterbeek, J. A. P. & Metz, J. A. J. On the definition and the computation of the basic reproduction ratio R_0 in models for infectious diseases in heterogeneous populations. *J. Math. Biol.* **28**, 365–382 (1990).
35. Nold, A. Heterogeneity in disease-transmission modeling. *Math. Biosci.* **52**, 227–240 (1980).
36. Sattenspiel, L. Population structure and the spread of disease. *Hum. Biol.* **59**, 411–438 (1987).
37. Kit, M. C. et al. Quantifying social contacts in a household setting of rural Kenya using wearable proximity sensors. *EPI Data Sci.* **5**, 21 (2016).
38. Cattuto, C. et al. Dynamics of person-to-person interactions from distributed RFID sensor networks. *PLoS One* **5**, e11596 (2010).
39. Salathé, M. et al. Digital epidemiology. *PLOS Comput. Biol.* **8**, e1002616 (2012).
40. Zhou, Y. et al. Effects of human mobility restrictions on the spread of COVID-19 in Shenzhen, China: a modelling study using mobile phone data. *Lancet Digit. Health* **2**, e417–e424 (2020).

41. Grantz, K. H. et al. The use of mobile phone data to inform analysis of COVID-19 pandemic epidemiology. *Nat. Commun.* **11**, 4961 (2020).
42. Khataee, H., Scheuring, I., Czirok, A. & Neufeld, Z. Effects of social distancing on the spreading of COVID-19 inferred from mobile phone data. *Sci. Rep.* **11**, 1661 (2021).
43. Gao, S. et al. Association of mobile phone location data indications of travel and stay-at-home mandates with COVID-19 infection rates in the US. *JAMA Netw. Open* **3**, e2020485 (2020).
44. Xiong, C., Hu, S., Yang, M., Luo, W. & Zhang, L. Mobile device data reveal the dynamics in a positive relationship between human mobility and COVID-19 infections. *Proc. Natl Acad. Sci. USA* **117**, 27087–27089 (2020).
45. Cot, C., Cacciapaglia, G. & Sannino, F. Mining Google and Apple mobility data: temporal anatomy for COVID-19 social distancing. *Sci. Rep.* **11**, 4150 (2021).
46. Badr, H. S. et al. Association between mobility patterns and COVID-19 transmission in the USA: a mathematical modelling study. *Lancet Infect. Dis.* **20**, 1247–1254 (2020).
In this study, the authors demonstrate a strong association between mobility patterns measured using aggregated mobile phone data across the USA and the transmission of SARS-CoV-2 at the early stages of the pandemic.
47. Stoddard, S. T. et al. House-to-house human movement drives dengue virus transmission. *Proc. Natl Acad. Sci. USA* **110**, 994–999 (2013).
48. Reiner, R. C., Jr, Stoddard, S. T. & Scott, T. W. Socially structured human movement shapes dengue transmission despite the diffusive effect of mosquito dispersal. *Epidemics* **6**, 30–36 (2014).
49. Salje, H. et al. How social structures, space, and behaviors shape the spread of infectious diseases using chikungunya as a case study. *Proc. Natl Acad. Sci. USA* **113**, 13420–13425 (2016).
50. World Health Organization. *Closing Data Gaps in Gender* (accessed 2 May 2021); <https://www.who.int/activities/closing-data-gaps-in-gender>
51. Hébert-Dufresne, L., Althouse, B. M., Scarpino, S. V. & Allard, A. Beyond R_0 : heterogeneity in secondary infections and probabilistic epidemic forecasting. *J. R. Soc. Interface* **17**, 20200393 (2020).
52. Riley, S. Large-scale spatial-transmission models of infectious disease. *Science* **316**, 1298–1301 (2007).
53. Simini, F., González, M. C., Maritan, A. & Barabási, A. L. A universal model for mobility and migration patterns. *Nature* **484**, 96–100 (2012).
54. Xia, Y., Björnstad, O. N. & Grenfell, B. T. Measles metapopulation dynamics: a gravity model for epidemiological coupling and dynamics. *Am. Nat.* **164**, 267–281 (2004).
The authors developed a gravity model formulation that described the mobility between populations in England and Wales, and effectively captured the dynamics of measles in the pre-vaccination era.
55. Wesolowski, A., O'Meara, W. P., Eagle, N., Tatem, A. J. & Buckee, C. O. Evaluating spatial interaction models for regional mobility in sub-Saharan Africa. *PLOS Comput. Biol.* **11**, e1004267 (2015).
56. Marshall, J. M. et al. Key traveller groups of relevance to spatial malaria transmission: a survey of movement patterns in four sub-Saharan African countries. *Malar. J.* **15**, 200 (2016).
57. Marshall, J. M. et al. Mathematical models of human mobility of relevance to malaria transmission in Africa. *Sci. Rep.* **8**, 7713 (2018).
58. Wesolowski, A. et al. Multinational patterns of seasonal asymmetry in human movement influence infectious disease dynamics. *Nat. Commun.* **8**, 2069 (2017).
In this study, the authors compare seasonal travel patterns using aggregated mobile phone data from Namibia, Pakistan, and Kenya, showing strong seasonal, asymmetric movements on a population level in each country.
59. Wesolowski, A. et al. Quantifying seasonal population fluxes driving rubella transmission dynamics using mobile phone data. *Proc. Natl Acad. Sci. USA* **112**, 11114–11119 (2015).
60. Mahmud, A. S. et al. Megacities as drivers of national outbreaks: The 2017 chikungunya outbreak in Dhaka, Bangladesh. *PLoS Negl. Trop. Dis.* **15**, e0009106 (2021).
61. Ewing, A., Lee, E. C., Viboud, C. & Bansal, S. Contact, travel, and transmission: The impact of winter holidays on influenza dynamics in the United States. *J. Infect. Dis.* **215**, 732–739 (2017).
62. Viboud, C. et al. Synchrony, waves, and spatial hierarchies in the spread of influenza. *Science* **312**, 447–451 (2006).
63. Wesolowski, A., Eagle, N., Noor, A. M., Snow, R. W. & Buckee, C. O. The impact of biases in mobile phone ownership on estimates of human mobility. *J. R. Soc. Interface* **10**, 20120986 (2013).
64. Wesolowski, A., Eagle, N., Noor, A. M., Snow, R. W. & Buckee, C. O. Heterogeneous mobile phone ownership and usage patterns in Kenya. *PLoS One* **7**, e35319 (2012).
65. Sinha, I. et al. Mapping the travel patterns of people with malaria in Bangladesh. *BMC Med.* **18**, 45 (2020).
66. Douine, M. et al. Malaria in gold miners in the Guianas and the Amazon: current knowledge and challenges. *Curr. Trop. Med. Rep.* **7**, 37–47 (2020).
67. Yan, S. D. et al. Digging for care-seeking behaviour among gold miners in the Guyana hinterland: a qualitative doer non-doer analysis of social and behavioural motivations for malaria testing and treatment. *Malar. J.* **19**, 235 (2020).
68. Prothero, R. M. Disease and mobility: a neglected factor in epidemiology. *Int. J. Epidemiol.* **6**, 259–267 (1977).
69. Smith, D. L. et al. Ross, Macdonald, and a theory for the dynamics and control of mosquito-transmitted pathogens. *PLoS Pathog.* **8**, e1002588 (2012).
70. Feachem, R. G. A. et al. Malaria eradication within a generation: ambitious, achievable, and necessary. *Lancet* **394**, 1056–1112 (2019).
71. Njuguna, H. N. et al. Malaria parasitemia among febrile patients seeking clinical care at an outpatient health facility in an urban informal settlement area in Nairobi, Kenya. *Am. J. Trop. Med. Hyg.* **94**, 122–127 (2016).
72. Heggenhougen, H. K., Hackethal, V. & Vivek, P. *The Behavioural and Social Aspects of Malaria and its Control. An Introduction and Annotated Bibliography* (TDR, WHO, 2003).
73. World Health Organization. *A Framework for Malaria Elimination* (WHO, 2018).
74. Chang, H. H. et al. Mapping imported malaria in Bangladesh using parasite genetic and human mobility data. *eLife* **8**, e43481 (2019).
In this study, genomic data of the malaria parasite are combined with travel histories and mobile phone data to quantify the routes and volumes of imported cases of malaria in southeast Bangladesh.
75. Funk, S., Salathé, M. & Jansen, V. A. A. Modelling the influence of human behaviour on the spread of infectious diseases: a review. *J. R. Soc. Interface* **7**, 1247–1256 (2010).
76. Edmunds, W. J., Eames, K. & Keogh-Brown, M. in *Modeling the Interplay Between Human Behavior and the Spread of Infectious Diseases* 311–321 (Springer, 2013).
77. Dowdy, D. W., Dye, C. & Cohen, T. Data needs for evidence-based decisions: a tuberculosis modeler's 'wish list'. *Int. J. Tuberc. Lung Dis.* **17**, 866–877 (2013).
78. Houben, R. M. G. J. et al. TIME Impact — a new user-friendly tuberculosis (TB) model to inform TB policy decisions. *BMC Med.* **14**, 56 (2016).
79. Abueleazam, N. N. et al. Can the heterosexual HIV epidemic be eliminated in South Africa using combination prevention? A modeling analysis. *Am. J. Epidemiol.* **184**, 239–248 (2016).
80. Kremer, M. Integrating behavioral choice into epidemiological models of AIDS. *Q. J. Econ.* **111**, 549–573 (1996).
81. Chitnis, N., Schapira, A., Smith, T. & Steketee, R. Comparing the effectiveness of malaria vector-control interventions through a mathematical model. *Am. J. Trop. Med. Hyg.* **83**, 230–240 (2010).
82. Childs, L. M. et al. Modelling challenges in context: lessons from malaria, HIV, and tuberculosis. *Epidemics* **10**, 102–107 (2015).
83. Epstein, J. M., Parker, J., Cummings, D. & Hammond, R. A. Coupled contagion dynamics of fear and disease: mathematical and computational explorations. *PLoS One* **3**, e3955 (2008).
84. Gallotti, R., Valle, F., Castaldo, N., Sacco, P. & De Domenico, M. Assessing the risks of 'infodemics' in response to COVID-19 epidemics. *Nat. Hum. Behav.* **4**, 1285–1293 (2020).
85. Kissler, S. M. et al. Reductions in commuting mobility correlate with geographic differences in SARS-CoV-2 prevalence in New York City. *Nat. Commun.* **11**, 4674 (2020).
86. Scannell Bryan, M. et al. Coronavirus disease 2019 (COVID-19) mortality and neighborhood characteristics in Chicago. *Ann. Epidemiol.* **56**, 47–54.e5 (2021).
87. Kishore, N. et al. Lockdowns result in changes in human mobility which may impact the epidemiologic dynamics of SARS-CoV-2. *Sci. Rep.* **11**, 6995 (2021).
88. Cramer, E. Y. et al. Evaluation of individual and ensemble probabilistic forecasts of COVID-19 mortality in the US. Preprint at <https://doi.org/10.1101/2021.02.03.21250974> (2021).
The authors compared multiple different forecasts of COVID-19 in the USA to evaluate their accuracy, and found that in general, predictions were only accurate on relatively short timescales, and that simple models were often just as accurate as more complex frameworks.
89. Viboud, C. et al. The RAPIDD ebola forecasting challenge: synthesis and lessons learnt. *Epidemics* **22**, 13–21 (2018).
90. Borchering, R. K. et al. Modeling of future COVID-19 cases, hospitalizations, and deaths, by vaccination rates and nonpharmaceutical intervention scenarios — United States, April–September 2021. *MMWR Morb. Mortal. Wkly. Rep.* **70**, 719–724 (2021).
91. Lutz, C. S. et al. Applying infectious disease forecasting to public health: a path forward using influenza forecasting examples. *BMC Public Health* **19**, 1659 (2019).
92. Haushofer, J. & Metcalf, C. J. Which interventions work best in a pandemic? *Science* **368**, 1063–1065 (2020).
93. *LISTSERV 14.4* (accessed 2 May 2021); <https://web.archive.org/web/20080523225000/http://listserv.linguistlist.org/cgi-bin/wa?A2=ind0407a&L=ads-l&P=8874>
94. Migration, U. N. *Regional Policies and Response to Manage Pastoral Movements within the ECOWAS Region* (IOM, 2019).
95. OECD/SWAC. *An Atlas of the Sahara-Sahel: Geography, Economics and Security* (OECD Publishing, 2014).
96. Post, W. M., DeAngelis, D. L. & Travis, C. C. Endemic disease in environments with spatially heterogeneous host populations. *Math. Biosci.* **63**, 289–302 (1983).
97. Watson, R. K. On an epidemic in a stratified population. *J. Appl. Probab.* **9**, 659–666 (1972).
98. Rushton, S. & Mautner, A. J. The deterministic model of a simple epidemic for more than one community. *Biometrika* **42**, 126 (1955).
99. Etienne, R. S. *Mathematical Models & Methods Meet Metapopulation Management* Thesis, Wageningen University (2002).
100. Hanski, I. & Simberloff, D. in *Metapopulation Biology* (eds Hanski, I. & Gilpin, M. E.) 5–26 (Academic, 1997).
101. Hethcote, H. W. An immunization model for a heterogeneous population. *Theor. Popul. Biol.* **14**, 338–349 (1978).
102. Anderson, R. M. & May, R. M. Spatial, temporal, and genetic heterogeneity in host populations and the design of immunization programmes. *IMA J. Math. Appl. Med. Biol.* **1**, 233–266 (1984).
103. Pinsky, P. & Shonkwiler, R. A gonorrhea model treating sensitive and resistant strains in a multigroup population. *Math. Biosci.* **98**, 103–126 (1990).
104. Yorke, J. A., Hethcote, H. W. & Nold, A. Dynamics and control of the transmission of gonorrhea. *Sex. Transm. Dis.* **5**, 51–56 (1978).
105. Hasibeder, G. & Dye, C. Population dynamics of mosquito-borne disease: persistence in a completely heterogeneous environment. *Theor. Popul. Biol.* **33**, 31–53 (1988).
106. Dye, C. & Hasibeder, G. Population dynamics of mosquito-borne disease: effects of flies which bite some people more frequently than others. *Trans. R. Soc. Trop. Med. Hyg.* **80**, 69–77 (1986).
107. Hethcote, H. W., Van Ark, J. W. & Karon, J. M. A simulation model of AIDS in San Francisco: II. Simulations, therapy, and sensitivity analysis. *Math. Biosci.* **106**, 223–247 (1991).
108. Lloyd-Smith, J. O., Schreiber, S. J., Kopp, P. E. & Getz, W. M. Superspreading and the effect of individual variation on disease emergence. *Nature* **438**, 355–359 (2005).
109. Bansal, S., Grenfell, B. T. & Meyers, L. A. When individual behaviour matters: homogeneous and network models in epidemiology. *J. R. Soc. Interface* **4**, 879–891 (2007).
110. Newman, M. E. J. Spread of epidemic disease on networks. *Phys. Rev. E* **66**, 016128 (2002).
111. Keeling, M. The implications of network structure for epidemic dynamics. *Theor. Popul. Biol.* **67**, 1–8 (2005).

112. Keeling, M. J. & Eames, K. T. D. Networks and epidemic models. *J. R. Soc. Interface* **2**, 295–307 (2005).
113. Meyers, L. A., Pourbohloul, B., Newman, M. E. J., Skowronski, D. M. & Brunham, R. C. Network theory and SARS: predicting outbreak diversity. *J. Theor. Biol.* **232**, 71–81 (2005).
114. Pitzer, V. E. et al. The impact of changes in diagnostic testing practices on estimates of COVID-19 transmission in the United States. *Am. J. Epidemiol.* kwab089 (2021).
115. Gostic, K. M. et al. Practical considerations for measuring the effective reproductive number R_t . *PLOS Comput. Biol.* **16**, e1008409 (2020).
116. McGough, S. F., Johansson, M. A., Lipsitch, M. & Menzies, N. A. Nowcasting by Bayesian Smoothing: A flexible, generalizable model for real-time epidemic tracking. *PLOS Comput. Biol.* **16**, e1007735 (2020).
117. Greene, S. K. et al. Nowcasting for real-time COVID-19 tracking in New York City: an evaluation using reportable disease data from early in the pandemic. *JMIR Public Health Surveill.* **7**, e25538 (2021).
118. Woolf, S. H. et al. Excess deaths from COVID-19 and other causes, March–July 2020. *J. Am. Med. Assoc.* **324**, 1562–1564 (2020).
119. Peccia, J. et al. Measurement of SARS-CoV-2 RNA in wastewater tracks community infection dynamics. *Nat. Biotechnol.* **38**, 1164–1167 (2020).
120. Clapham, H. et al. Seroepidemiologic study designs for determining SARS-COV-2 transmission and immunity. *Emerg. Infect. Dis.* **26**, 1978–1986 (2020).
121. Metcalf, C. J. E., Viboud, C., Spiro, D. J. & Grenfell, B. T. Using serology with models to clarify the trajectory of the SARS-CoV-2 emerging outbreak. *Trends Immunol.* **41**, 849–851 (2020).

Author contributions C.O.B., A.N. and L.S. all contributed to the planning and writing of this manuscript.

Competing interests The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to C.B.

Peer review information Nature thanks the anonymous reviewers for their contribution to the peer review of this work.

Reprints and permissions information is available at <http://www.nature.com/reprints>.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© Springer Nature Limited 2021