Modelling: understanding pandemics and how to control them

Glenn Marion[1,2], Liza Hadley[3], Valerie Isham[4], Denis Mollison[5], Jasmina Panovska-Griffiths[6,7], Lorenzo Pellis[8,9,11], Gianpaolo Scalia Tomba[10], Francesca Scarabel[8,11], Ben Swallow[14,2], Pieter Trapman[15], Daniel Villela[16]

Author affiliations:

[1] Biomathematics and Statistics Scotland, Edinburgh, UK
[2] Scottish COVID-19 Response Consortium, www.gla.ac.uk/scrc/
[3] Disease Dynamics Unit, Department of Veterinary Medicine, University of Cambridge, UK
[4] Department of Statistical Science, University College London, UK
[5] Department of Actuarial Mathematics and Statistics, Heriot-Watt University, UK
[6] The Big Data Institute, Nuffield Department of Mathematics, University of Oxford, Oxford, UK
[7] The Queen's College, Oxford University, UK
[8] Department of Mathematics, University of Manchester, UK
[9] The Alan Turing Institute, London, UK
[10] Department of Mathematics, University of Rome Tor Vergata, Rome, Italy
[11] Joint UNIversities Pandemic and Epidemiological Research, UK
[14] School of Mathematics and Statistics, University of Glasgow, UK
[15] Department of Mathematics, Stockholm University, Stockholm, Sweden
[16] Program of Scientific Computing, Fundação Oswaldo Cruz, Rio de Janeiro, Brazil

Abstract

New disease challenges, societal demands and better or novel types of data drive innovations in the structure, formulation and analysis of epidemic modelling. Innovations in modelling can lead to new insights into epidemic processes and better use of available data, yielding improved control and stimulating collection of better data and new data types. Here we identify key challenges for the structure, formulation, analysis and use of mathematical models of pathogen transmission relevant to current and future pandemics.

Keywords

Introduction

Mathematical and computational modelling of the spread and effects of infection has old roots, going back at least to the works of Daniel Bernoulli in the eighteenth century, with the first significant use of epidemic modelling for successful intervention being perhaps that of Ross and Hudson on malaria, early in the 20th century (Bacaër 2011). Since then, and especially over the last 50 years, the field has expanded hugely (Diekmann *et al*., 2013). Heesterbeek *et al*., (2015) provide a set of historical highlights in the development of modelling methodology and concepts since 1950, demonstrating how developments have been stimulated by the many and diverse infectious disease challenges experienced over this period.

In developing our understanding of pandemics and how best to control them, models are needed that address a wide range of aspects including aetiology, epidemiology, natural history, and individual and societal response. Almost all models exhibit threshold behaviour whereby epidemic spread occurs when a certain combination of parameters that is dependent on model structure, (e.g. the real time reproduction number R_t) exceeds a threshold. Typically, the aim of interventions is to bring epidemic spread below this threshold, and/or to deal with the consequences of being close to it. Interventions will depend on disease impacts and wider societal issues that may also have important implications for modelling.

Given the complex questions and uncertainties that arise during infectious disease emergencies, it is important to note that modelling does not offer exact or binary answers but, rather, provides tools to enable both understanding and quantification of phenomena. Such use of models is discussed in other contributions to this special issue. For example, modelling tools can be used to develop a suite of possible answers for informing and advising policy (Hadley *et al*., 2021) and for designing and assessing interventions (Kretzschmar *et al*., 2021). An important step in this process is the statistical estimation of key quantities that characterise transmission and other aspects of disease dynamics (Swallow *et al*., 2021), which is critically dependent on the availability, accessibility and reliability of data (Shadbolt *et al*., 2021). The specific contribution of the current paper is to focus on the challenges for the structure, formulation, and analysis of models to understand disease dynamics better and to support real time response to future pandemics. We argue that such challenges extend to ensuring model assumptions and results are communicated in an open, transparent manner and that modelling has a critical role to play in informing the intelligent collection of data for future pandemics.

During pandemics such as COVID-19, much of the interest is in difficult questions that cannot be answered empirically or for which insufficient information is available solely from data when decisions need to be made. Modelling allows a technical framework to be developed in which such questions can be explored. During the COVID-19 pandemic modelling has been applied more prominently and widely than ever to inform and advise public health policy. See for example the special issue of the journal *Philosophical*

Transactions B on *Modelling that shaped the early COVID-19 pandemic response in the UK* (Brooks-Pollock *et al*., 2021).

Modellers must grapple with uncertainty driven by spatial, temporal and societal heterogeneities, stochasticity in dynamics and incomplete knowledge of parameters and model structure. Combining outcomes across multiple models that vary in complexity, allows assessment of the robustness of conclusions to a range of assumptions. For example, the relationship between the real time growth rate of the number of cases or infected hosts and the reproduction number R_t is highly sensitive to the distribution of the infectivity profile of an infectious host (Roberts and Heesterbeek, 2007; Wallinga and Lipsitch, 2007). On the other hand, the relationship depends to a much lesser extent on population structure (assuming that structure is characterized by several distinct classes of individuals, a social network structure or households; Trapman *et al.*, 2016). More detailed understanding of transmission mechanisms and the effects of interventions requires more complex models, but inevitably also very many (often unknown) parameters and typically new or larger datasets. The value that statistical inference can extract from data is intimately tied to the structure and formulation of models. As the statistician George Box wrote `all models are wrong, but some are useful' (Box 1979), meaning that models are simplifications of reality, thus false or wrong, but can still retain the essential features of the process, generating understanding and allowing prediction.

Controlling pandemics is a wicked problem (Rittel and Webber 1973) hence characterized by incomplete, contradictory and changing requirements, difficulties in obtaining data relevant for decision-making, and where seemingly well-motivated efforts may lead to unintended and even self-defeating outcomes. Mathematical modelling of pandemics is therefore critical:

- yielding qualitative understanding of how various factors influence spread of infection, e.g. differential within host progression of disease, heterogeneity in population response, infectivity, susceptibility, contact intensity and structures, and allowing *in silico* experiments to determine the effects of possible interventions;
- providing a framework for synthesising multiple sources of information including understanding how best to collect, analyse and interpret observations (data) related to the infection and its spread in the population, and also for determining what further data would most usefully be collected;
- allowing estimation of current and historic trends including non-observed or non-observable quantities as well as short term projections and quantitative (future) scenarios useful for health policy decision making, planning and evaluation of interventions.

Work during the COVID-19 pandemic has revealed not only a lack of information on aspects of the infection itself, but also important gaps in the available modelling tools and theoretical understanding needed. This paper discusses the key challenges in modelling related to pandemics. In doing so, we build upon many of the challenges highlighted in the special issue of the *Epidemics* journal (vol.10) on Challenges in Modelling Infectious Disease Dynamics (Lloyd-Smith *et al*., 2015). That issue included challenges related to global transmission models (Britton *et al*., 2015), meta-population and household approaches (Ball *et al*., 2015), and explicitly spatial (Riley et al., 2015) and detailed network representations (Pellis *et al*., 2015). Although such models are typically built on stochastic processes, deterministic models can also provide critical insights (discussed in Roberts *et al*., 2015). Many of the challenges raised in the 2015 volume, relating for example to pathogen mutation-evolution, multi-strain systems, inferential methods for data on the emerging phase of epidemics in structured populations, computationally efficient methods for calculating thresholds and early exponential growth rates, and designing network-based/spatial interventions, remain important during modelling of the spread of SARS-CoV-2 and its variants. In this paper, we highlight to what extent these challenges have been addressed in the intervening 6 years, which are outstanding, and add new challenges seen from a mid-pandemic rather than `peacetime' perspective. Its overall structure and interconnections are shown in Figure 1. We first discuss challenges in the formulation and analysis of models from the interlinked perspectives of between-host contact processes (Section 1), within-host dynamics (§2) and the characteristics of the pathogen (§3). We then look beyond model formulation and analysis to challenges where models and modellers can play a critical role in improving modelling and the impact of modelling in future pandemics through: better collection of data during outbreaks including design of testing, surveillance and contact tracing (§4); and in improving capabilities for real time decision support and ensuring greater openness, transparency and trust in communicating modelling results to policy makers and the public (§5).

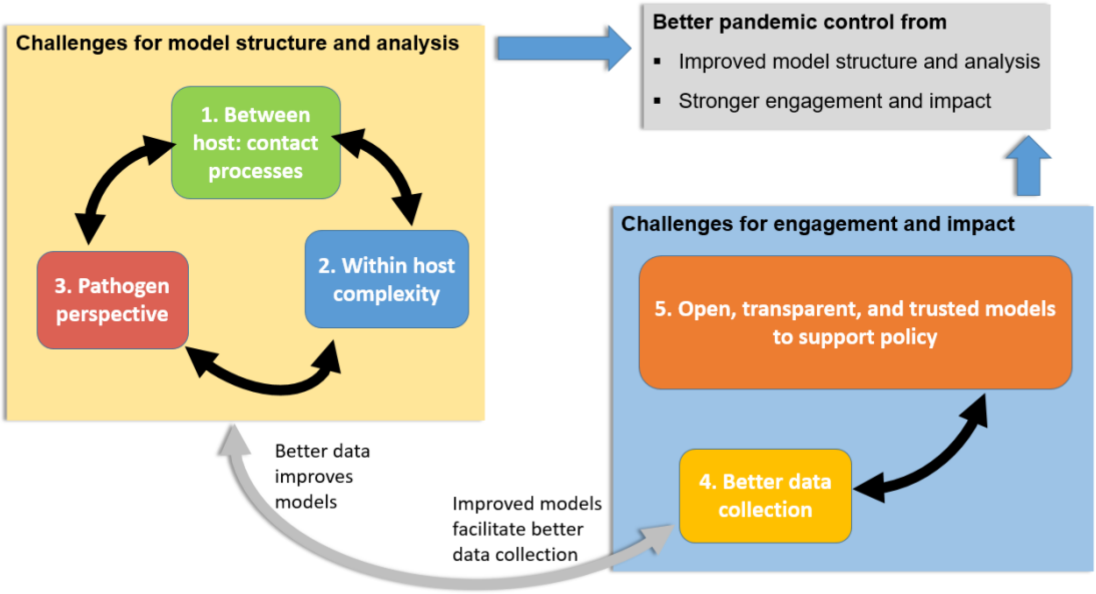

Figure 1. Challenges for pandemic modelling: Enhancements of models and associated analytical tools (see challenges in §1, §2 and §3) and greater engagement with public health stakeholders (§4 and §5) are mutually reinforcing and will ultimately lead to better pandemic control.

1. Between host: Modelling infectious contact processes

The representation of infectious contacts underpins all dynamic transmission modelling and is therefore the central challenge in modelling future pandemics. COVID-19 has revealed significant deficiencies in our ability to model contact processes, but it has also generated huge amounts of data across the world and under different interventions (Chang *et al*., 2021; Vigfusson *et al*., 2021; Jia et al., 2020). A challenge and opportunity is to use these data to inform modelling approaches and in particular to better understand disease contact processes and to allow models to move beyond use of observed pre-outbreak contact patterns (Conlan *et al*., 2021) in future pandemics.

While they appeared less influential a decade ago (Lloyd-Smith *et al*., 2015), agent-based models (ABMs) have proved to be an attractive approach particularly in response to infectious disease outbreaks like Ebola and COVID-19 (see e.g. Kiskowski and Chowell 2016; Kerr *et al*., 2020), allowing investigation of many aspects highlighted here. However, simplified deterministic and stochastic models remain fundamental in providing analytic insights into how key aspects of contact patterns and human behaviour affect epidemic dynamics and outcomes. Analytical approaches can also enable development of simplified models more amenable to formal methods of statistical inference and uncertainty quantification.

Here we identify substantive challenges for the modelling and understanding of contact processes under three broad headings: household models, multi-scale approaches that couple meta-population, spatial or network structures, and the need for more predictive models of behavioural impact on contacts.

Household models and extensions

Households represent a key structure in many human societies. For directly transmissible pathogens, they are a fundamental epidemiological unit because for many infectious diseases the stable and more intimate nature of contacts between household members typically translates into significantly higher probability of transmission than with individuals outside the household. Furthermore, household sizes and compositions are also typically readily available from census data, and many control policies are targeted at households. Some of these considerations can also be extended to other kinds of stable groupings of individuals, such as workplaces, schools, etc., where contacts are closer than average.

Households are also natural observation points for estimating transmission parameters and how individual (e.g. age, sex, occupation, etc.) and household (e.g. socio-economic status, overcrowding) properties affect susceptibility and infectivity. Model structure is intimately connected with estimation. Optimal design of household-stratified data collection has recently attracted increased interest (Kinyanjui *et al*., 2016), and the COVID-19 pandemic has seen an explosion of large-scale data collection (see for example the UK Office for National Statistics (2021)). Observables have traditionally focussed on household final size data (Demiris and O'Neill, 2005), though more recent work has investigated biases arising in a growing epidemic (e.g. Ball and Shaw 2015). However, challenges remain in the context of a pandemic with interventions imposed and lifted in rapid succession or in the presence of imperfect case ascertainment, in particular due to heterogeneous symptoms, test-seeking behaviour and disease outcome. Final size household data can be used by separating the epidemic into suitably distinct phases (House and Pellis, 2021), but methodological developments are urgently needed for using data collected in real-time and on a background of a dynamically varying population prevalence and hence risk of introduction into households, as well as to make full use of test results of household members collected over time or other temporal data. A key challenge is thus to improve analytic understanding of the temporal dynamics of household models under variable conditions and with detailed endpoints.

Important mathematical developments in the context of models with household structure have been made over the past 25 years (e.g. Ball, Mollison and Scalia-Tomba (1997), Ball, Sirl and Trapman (2014), Pellis, Ball and Trapman (2012)). However, the vast majority of theoretical results are confined to time-integrated epidemiological quantities, such as reproduction numbers, final size distributions or the probability of a large outbreak. Such results have been influential in informing the likely impact of specific policies, for example providing intuition on the role of household bubbles in allowing more social interaction while limiting the increase in transmission (Danon *et al*, 2021), or on how school closure and reopening affect the network linking households and thus the household reproduction number. However, other forms of household-based interventions that received significant interest during the COVID-19 pandemic, including in- or out-of-household isolation of single individuals or quarantining of all household members following symptom-based case detection (Overton *et al*., 2020) or contact tracing (Fyles *et al*., 2020), result in responsive changes to transmission parameters during a within-household outbreak. Obtaining theoretical results for models with non-constant parameters, even in simple cases, remains an open problem. Similarly, although analytical results for the real-time growth rate are available (Ball *et al*., 2015, §6), results on real-time non-linear epidemic dynamics in the presence of households would be valuable.

Given these theoretical challenges regarding the temporal dynamics of the epidemic, it is unsurprising that practical questions that require models explicitly accounting for household structure have been mainly tackled with ABMs. Recent examples include, models with both households and communities of households to study the spread of Ebola in West Africa (Kiskowski 2014) or household bubbles in the context of COVID-19 (Leng *et al*., 2021), and models with dynamic household structure, for instance with individuals distributed across multiple dwellings (Chisholm *et al*., 2020) or with explicit demographic change (Geard *et al*., 2015). A general challenge consists in formulating analytically tractable models to tackle these questions, which includes understanding how much complexity is really needed and whether general principles from simple household models can shed light on the behaviour of more complex ABMs.

Extensions to the simplest household models that allow for individual heterogeneity exist (see e.g. Ball, Britton and Sirl, 2011)). Potential extensions include models with multiple types of both households and household members, together with realistic contacts and other disease-dependent parameters between the different types. Such models might allow for better understanding of household infection dynamics, more targeted interventions, as well as a more realistic description of consequences of socio-economic differences between household types (Villela, 2021). Furthermore, such models might incorporate assortative mixing between households based on size or composition, e.g. households with children are more likely to be linked with each other through schools, and are likely larger than average. Effects of interventions such as school closures could be better captured in such models than when household connections are uncorrelated.

Models with further structures beyond single households might also be both of practical interest and amenable to mathematical tractability. Recent advances have been made on models with hierarchical structures (e.g. Gandolfi and Cecconi, 2016). Some work also exists when other small mixing groups overlap with households, e.g. schools or workplaces (e.g. Pellis, Ferguson and Fraser, 2011). However, overlap creates dependencies between units and the techniques (Ball, Sirl and Trapman, 2014) used in these studies, which essentially assume that overlap is restricted to one individual, would profit from generalisation.

From a practical point of view, understanding the role of households on transmission dynamics, epidemic outcome and impact of interventions is crucial. However, the meaning of understanding is highly dependent on context, desired aims and model choice (in turn dictated by data availability). For example, Pellis *et al*., (2020) narrows the question to the context of an emerging epidemic and two age classes, and formulates it as a problem of model comparison. More work is required to study the role of households on transmission in other contexts, starting from the formulation of questions precise enough to allow investigation. Substantial work is also needed on comparison of structurally different models, not only in terms of model fit to data, but also in terms of predictions (model assessment).

Multi-scale models: meta-population, spatial and network

Britton *et al*., (2015) posed the question, "Is the classification into global, network, meta-population and spatial models sufficient for the range of contact structures of interest in understanding infectious disease dynamics?" Here we argue for the need to understand hybrid models that combine or bridge between meta-population, social network and spatial structures. As argued above, meta-population structures of households and workplaces have proved crucial in understanding and controlling COVID-19 as have detailed individual level models (Chang *et al*., 2021) based on networks captured by GPS tracking that reveal local spatial structure. Moreover, at large scales real word disease incidence shows substantial spatial variation. Attempts to address such issues have made use of spatially explicit individual-based models (Lau *et al*., 2017) and ABMs (Kerr *et al*., 2020) including cases described above where household structures are embedded within broader contexts (e.g. Kiskowski & Chowell, 2014).

A significant challenge is to develop approaches to analyse, and simplify such models to enable greater understanding of disease dynamics and control. Methods such as moment-closure and pair approximations are useful in developing simplified representations and analytic results in spatial and network models (Barnard *et al*., 2019). Another promising approach is to develop systematic methods (Khudabukhsh *et al*., 2019) to enable coarse-graining of networked individual-based models to generate more tractable representations. Although it is understood that higher-order network structure impacts epidemic outcomes (Ritchie et al 2014), an important challenge is to understand what properties of real world networks impact the accuracy of analytic results obtained under different assumptions (Silva et al., 2020; Wu and Hadzibeganovic, 2020). The concept of universality classes (Chung *et al*., 2016) may be useful in classifying epidemic dynamics on real world networks.

Analysis of models that bridge the gap between meta-population, network or spatial representation of contact processes are rare. One example of analytical work that attempts to address combined meta-population and network models uses moment-closure to develop simplified representations and analytic results that suggest detailed case-reporting data can be informative of connectivity between metapopulations (Meakin and Keeling, 2019). On the other hand, Haw *et al.*, (2020) develop and analyse the output of hybrid models that combine interactions on social networks with spatial movement showing they can give rise to sub-exponential outbreak dynamics with lower, later epidemic peaks that are hard to explain in more standard models. A potential route to greater analytic understanding of such models might be to exploit links between spatial models (Riley *et al*., 2015) and more general network representations. For example, the spatial structure of human infectious contacts, and how it is affected by interventions, is not well understood. There is an unexplored part of model space between strictly spatial models, where the k-neighbourhood of an individual, the number that can be reached in a chain of k successive contacts, is of order k^2, and homogeneous-mixing models where it is of order e^k. These correspond to networks with typical distance between individuals varying from the square root to the log of population size. For example, initial work in this area shows that networks with local clustering are effectively homogeneous-mixing in their spatial structure (Mollison 2004). Promising work has shown how to develop analytic approaches for models that combine spatial features of scale-free networks with nested community structure (Gandolfi and Cecconi, 2016). More recent work has shown how to embed networks within a spatial structure to better explore the impact of controls like social distancing and travel restrictions (Hulshof *et al*., 2020). Tractable models that better represent mixing at a range of scales e.g. in human populations, would allow better understanding and even prediction of how epidemics spread and how individual and public health responses to them could best limit the impact of outbreaks and reduce persistence of endemic disease.

Behaviour and contact processes

Behavioural responses significantly increase the challenge of assessing the impact of public health interventions (Michie and West, 2020). Throughout the COVID-19 pandemic models have been widely used to predict the impact of non-pharmaceutical interventions, or NPIs (Flaxman *et al*., 2020). NPIs represent measures to reduce ongoing and limit future transmission via social-distancing (including lockdowns) and in the case of respiratory infections the use of face masks. Assessment of NPIs is further complicated by the fact that awareness of disease spread itself, or simply the knowledge that NPIs are being discussed, may alter contact behaviour (Zhou *et al*., 2020). Furthermore, behavioural responses to vaccination and to exogenous factors such as seasonality, societal events e.g. religious and other festivals, economic and natural shocks to society may also prove significant.

Currently the typical approach to assessing the impact of NPIs or other behavioural responses to outbreaks is to modify parameters of epidemic models to capture resulting or anticipated changes e.g. reducing contact rates to represent social distancing measures. This is reliant on expert judgment about the impact of NPIs or rich sources of data on observed responses, ideally under NPIs, to allow robust calibration with uncertainties in predictions quantified. A challenge here is to develop methods that use transmission models and the observed epidemic to robustly quantify changes in contact patterns. Such approaches offer a valuable quantitative framework, but are phenomenological and at best can only explore direct impacts of known and anticipated responses to pandemics and NPIs. For example, they do not account for compensatory behaviours.

Improved understanding and prediction in pandemics and large-scale outbreaks and endemic disease scenarios requires better quantification of behaviour (Funk *et al.*, 2015) and the dynamics of social systems. These dynamics include movement patterns at various scales associated with different activities e.g. work and leisure including how these are impacted by public health intervention, business response and individual behaviour change in direct and indirect response to a pandemic.

Anonymised mobile-phone call detail records (CDRs) have been used to show behavioural change of infected individuals (Vigfusson *et al*., 2021). This poses challenges and opportunities to standard disease transmission modelling to account for such behavioural changes and heterogeneities in such responses across populations due to socio-economic factors e.g. that limit opportunities to self-isolate (Bharti, 2021, Gauvin *et al*., 2021). These complexities are further compounded by large-scale changes in society in response to perception of disease outbreaks and public health interventions. A challenge here is how to make use of available data at different levels of granularity from ‘big data’ on tracking of individuals to aggregate societal data, such as population flows (Jia et al., 2020), transport usage and retail sales, to develop sufficiently predictive models that can, for example, anticipate changes in contact patterns e.g. when pubs, bars and restaurants close (Tang, 2020).

In recent years there has been significant interest in game theoretic approaches to understand and predict behavioural responses to disease threats (Chang *et al*., 2020). Another potential way forward is the development of mechanistic models of dynamic contact processes that account for constraints on individual behaviour using individual propensities for types of contact behaviour. Such models may be amenable to analysis and parameterisation using big data describing historic movement patterns (for moves in this direction see e.g. Knight *et al*., 2021). Treatment of such models as dynamic networks may offer a fruitful approach to develop better understanding of behavioural responses and effect of social distancing measures (Valdano *et al*., 2018; Barnard, et al., 2018; Britton *et al*., 2019).

2. Within host complexity: beyond SEIR models

There is increasing recognition that within host dynamics play a critical role in disease transmission. Ways to account for this range from the use of expanded state spaces in compartmental models beyond standard models such as SEIR e.g. representing separate symptomatic and asymptomatic pathways for COVID-19, to models of individual viral loads and even immune response dynamics. A more complete representation of within host dynamics could be validated by availability of detailed longitudinal data e.g. from repeated testing, but also by better use of existing data sources e.g. by recognising the non-binary nature of diagnostic test results. The modelling challenges here are to develop models that can be parameterised and tested using such data as well as approaches that can be used to analyse them.

Better disease progression models

Although the adoption of discrete states e.g. susceptible, exposed, infectious, recovered, has been extraordinarily successful in epidemiological modelling, these must be adapted to particular conditions or properties of the disease, such as presymptomatic infectivity or degrees of severity of symptoms (see e.g. Anderson *et al*., 2020) or different levels of subsequent immunity. This requires a judicious synthesis of clinical and epidemiological observations, especially in the early phases of a new epidemic. Ideally, modelled states should correspond to distinct and measurable clinical conditions, but this is rarely the case. Parameterisation must then be addressed using inference exploiting longitudinal data with individual-based stochastic models and other proxy information, such as viral load; key challenges are computational complexity of inference (Swallow *et al.*, 2021) and access to longitudinal data (Shadbolt *et al.*, 2021). Recently, describing infectivity progression by an infectivity profile, which is directly coupled to the generation time distribution, has become popular. This kind of description emphasises the continuous nature of disease progression. Adding similar descriptions of symptoms, severity of disease and development of immunity would constitute further improvements. Recent work on inferential questions (e.g. Britton & Scalia Tomba, 2019) has also highlighted the need for describing joint properties of disease states, such as joint distributions of latent, incubation and infectious periods, which lead to related inferential challenges. Coupling within host disease progression models with measurement of viral load (Kissler *et al*., 2021) could provide information on time of infection and infectivity profiles and thus inform the dynamics of outbreaks from cross-sectional data (Hay *et al*., 2020; Rydevik *et al*., 2016). For example, the association between viral load and transmission strength and period across the population could be untangled, enabling better understanding of the difference in development of an early, exponentially growing epidemic and a developed but fragmented epidemic in heterogeneous populations (Lythgoe, Pellis and Fraser, 2013).

Immunity and vaccination

Further work is required to account for the dynamic distribution of immunity across heterogeneous populations resulting from both transmission and vaccination. For example, there is still a challenge in analysing and describing how heterogeneity in the population impacts distribution of immunity after an epidemic or wave of an outbreak, and which heterogeneities should be taken into account (Gomes *et al*., 2020). Some work has been done regarding the impact of multi-type populations (Britton, Trapman and Ball, 2020), where it is shown that immunity caused by an earlier wave of an epidemic is distributed over the population in a substantially more efficient way than if the immunity is obtained through vaccination programmes that do not target those making disproportionally many potentially infectious contacts. A challenge in modelling vaccination is how to differentiate between the waning of immunity following infection and that following vaccination. Disentangling these effects will likely require data on host immune response and information on pathogen exposure. Realistic modelling of waning immunity, and possibly boosting of previously acquired immunity via exposure to infection, remains an active area of research linking within and between host dynamics (Heffernan and Keeling, 2009).

Mathematical models can help in the design of optimal vaccination programmes and to assess their effectiveness. However, work is needed to develop models able to identify optimal distribution strategies and vaccination thresholds when resources e.g. doses, are limited in terms of overall quantities available and rates of supply. Model-based searches for optimal strategies are computationally intensive, especially in the absence of precise estimates or prior knowledge about multiple aspects. Depending on the specific questions asked, several issues must be accounted for: granularity in the population, for instance in terms of age or risk groups and localities; different vaccine modes of action, e.g. transmission blocking vs reducing disease severity (Hodgson *et al*., 2021), especially when transmission and serious disease are distributed differently across the population (i.e. when these are negatively correlated); and different vaccine efficacies or dosing schedules (e.g., requirement for one vs two doses). Additional challenges are related to the role of mathematical models in: enabling vaccine efficacy to be better estimated from surveillance data; exploring interactions between vaccination, disease-induced immunity and NPIs, for instance to identify possible roadmaps towards lifting of restrictions (Panovska-Griffiths *et al*., 2021; Whittles *et al*., 2021); and developing better understanding of the potential for vaccine escape and capturing the effect of variation in vaccine efficacy across variants (Day *et al*., 2020). See Kretzschmar *et al*., (2021) and Madewell *et al*., (2021) for further discussion of vaccination challenges.

3. A pathogen perspective

Multiple pathogen strains: from neutrality to selection

Driven by the increased resolution in our ability to observe pathogens afforded by application of molecular biology tools, an urgent challenge for epidemiological modelling is representation of the dynamics of pathogen heterogeneity e.g. multiple pathogen strains (Lythgoe *et al*., 2021). A standard approximation is to assume that mutations are neutral and therefore that transmission is unaffected by pathogen strain (Frost *et al*., 2015). Under this assumption phylogenetic data on pathogens can inform inference of contact networks and be combined with standard epidemiological observations (Volz and Frost, 2013; Lau *et al*., 2015). However, significant challenges remain in terms of embedding phylodynamics (Grenfell *et al*., 2004) within disease transmission models. In particular, current approaches do not adequately account for within host diversity of pathogen, host immunity and pathogen load or selective pressure amongst competing strains (Lau *et al*., 2019; Metcalf *et al*., 2015; Wikramaratna *et al*., 2015). The latter is particularly urgent given concern over the emergence of novel COVID-19 strains with higher transmissibility.

More work is needed to assess the impact of the evolutionary pressures imposed by vaccination and other control campaigns on pathogens (Read et al., 2015). Currently such problems are tackled using models that focus on the potential for invasion of a variant in the presence of a dominant strain, or model fixed and typically small number of competing strains (Day *et al*., 2020). Greater flexibility is afforded by ABMs that represent multiple strains, but these come with significant computational and analytic challenges. Tools from quantitative genetics may prove useful in developing analytical insight into such problems (Day and Gandon, 2007; Day *et al*., 2020b), and there is increasing recognition of the need to couple evolutionary and ecological dynamics (Lion, 2018) Addressing host genetics e.g. that may affect susceptibility, infectivity and recovery (Pooley *et al*., 2021), would add further complexity (Frost *et al*., 2015) and point toward multi-scale models that represent both within and between host pathogen dynamics (Gog *et al*., 2015).

Environmentally persistent pathogens and indirect transmission

Many pathogens persist in the environment necessitating modelling of indirect, or environmental transmission in addition to, or instead of, direct transmission resulting from contacts between infectious and susceptible individuals. Future pandemics may be caused by pathogens that are more persistent in the environment than SARS-CoV-2, and there are numerous challenges associated with modelling resultant environmental transmission (Hollingsworth *et al*., 2015). However, in simple scenarios, direct transmission models can accurately represent epidemic outbreaks of environmentally transmitted pathogens as long as there is no significant timescale separation between host infectious period and environmental persistence of the pathogen (Benson *et al*., 2021). However, caution is needed, since for pathogens that persist in the environment for long periods of time the behaviour of direct and indirect transmission models will be markedly different e.g. the re-emergence of environmentally transmitted diseases in cases where there are no remaining infectious individuals. Thus, un-accounted for environmental transmission is likely to impact evaluation of control measures. Furthermore, environmental transmission is likely to increase the degree of connectivity compared with an observed direct contact process by broadening the effective contact network. For example, two individuals who visit a given location but at different times may nonetheless have effectively been in contact. A key challenge therefore is to develop modelling approaches that account for such differences implicitly or explicitly and methods that enable integration of environmental pathogen load measurements (Wade et al., 2021) into transmission modelling. Other major challenges include integrating understanding of local environmental transmission in and across a range of settings (Morawska *et al*., 2021; Wang *et al*., 2021) within e.g. city or national scale modelling of disease dynamics. Perhaps a promising place to start is to develop modelling that accounts for variation in relative environmental exposure (Jones *et al*., 2021).

A critical class of strongly environmentally persistence pathogens are vector-borne diseases, such as malaria and infections by arboviruses (DENV, CHIKV, ZIKV), leading to disease endemic levels, marked by seasonality, usually by vector abundance. Modelling of endemicity in these scenarios requires analytical treatment using seasonal oscillations to be considered for instance in interventions (Bacaër and Guernaoui, 2006; Griffin, 2015). Seasonal variations also appear in the cycle of mosquito-borne infections, for instance due to temperature-dependent incubation period, in which parasites remain latent in infected mosquitoes (extrinsic incubation period), with direct impact on the generation time of the disease. The consequences of time-varying generation time have been examined for the estimation of the effective reproduction number (Codeço *et al*., 2018; Siraj *et al*., 2017). However, further characterisation of the seasonal treatment, including modelling of other environment-dependent biological mechanisms, is an open challenge.

4. Better design of testing, surveillance and contact tracing

Public health interventions like symptomatic testing, surveillance and contact tracing and other NPIs, in addition to being vital to control pandemics, are also currently under-exploited pseudo-experiments that provide untapped potential to inform on key parameters and processes. For example, more accurate meta-data on who is tested and why may enable better use of case reports. Modelling is central to demonstrating the benefits of better exploitation of such data sources e.g. more effective intervention, as well as being central to implementing protocols to extract information from data collected during them. Given a lack of knowledge, there may be considerable benefit to trialling and assessing different interventions in different places (Michie and West, 2020). There are clearly ethical and political considerations, but for example where this is already happening (see e.g. Islam *et al*., 2020), it would be advisable to ensure that sufficient data are collected to enable as complete an assessment as possible.

Data collection during outbreaks

An important challenge is in predicting which datatypes may be useful in a future outbreak, many aspects of which are currently unknown. There are data and estimation aspects to this challenge but also a significant modelling challenge in terms of developing appropriate models and scenarios to assess the value of potential data. Prior to the next significant pathogen outbreak, there is therefore considerable work to do in assessing the informative value of various data types and data collection systems (represented in terms of planned protocols, and departures that occur from plans in implementation). This problem can be tackled through a combination of modelling different pandemic scenarios and using inferential tools (Swallow *et al*., 2021) to assess the value of different data collection systems. Even within the same broad scenario, the usefulness of certain data will depend on model structure, so this work must also anticipate that a range of models likely to be used throughout an outbreak. This scenario planning will enable recommendations on what data should be collected at different stages of an outbreak, with the aim of maximising societal ability to respond. Conflict between the primary purpose of such interventions and data gathering can potentially be minimised by developing general guidelines for data collection (Shadbolt *et al*., 2021).

A further challenge is the development of tools for adaptive design to prioritise data collection in real time. These could be built around sensitivity analysis of specific model outputs/value of information studies (Jackson *et al*., 2019) or surveillance of current data streams for changes (Xiang and Swallow, 2021). Stochastic and network models, where sensitivity analysis is significantly more complex, would benefit from further development of tools and software to enable uptake on a wider scale. Sensitivity analysis can also inform prediction of future necessary datasets, giving data collectors the time to implement required protocols.

Contact tracing

Contact tracing represents a particularly difficult challenge for mathematical epidemiology in terms of its ability to inform effective real world intervention and real time data collection. This is due to the complexities of capturing the contact patterns between individuals and the testing and tracing process which propagates over the network of contacts and locally modifies it at the same time. Better understanding of this process is needed to allow interpretation and use of contact-tracing data to inform models and better characterise outbreaks.

Some important challenges in creating mathematically rigorous results for stochastic models with contact tracing remain to be addressed. In particular, contact tracing creates dependencies between durations of infectious periods (and thus cumulative infectivity) for infectors and their infectees. Because of this, the ordinary theory of branching processes does not suffice, even if it is possible to deduce the correct (marginal) distribution of the number of other people an infected person infects (Müller *et al*., 2000, Müller and Hösel 2020). To obtain theoretical results on questions such as "what is the probability that a major outbreak occurs in a population with a functioning contact tracing infrastructure?" new models that allow for dependencies created by contact tracing and that can be analysed in a mathematically rigorous way need to be developed.

Further important challenges are connected with improving the description of the real world. Clusters and superspreading events are known to be decisive for contact tracing effectiveness (House and Keeling 2010, Endo *et al*., 2020). Intermediate settings of society, like households, workplaces, or in more general places of aggregation or at events, are particularly important as policies often act on this scale (Kucharski *et al*., 2020, Fyles *et al*., 2021, Kretzschmar *et al.*, 2020). Another important aspect is the realistic modelling of time and resource constraints, as contact tracing typically requires an extensive infrastructure to identify infected cases and swiftly search and isolate the contacts of a confirmed case. Therefore, theoretical results (e.g., the efficacy of backward versus forward contact tracing, or results on the controllability of the epidemic) that are often derived under assumptions of unlimited tracing capabilities, should be carefully evaluated in the presence of limitations to capacity. In this context, mathematical modelling has proved useful in assessing the effectiveness of new technologies like digital contact tracing compared to manual contact tracing (Kretzschmar *et al*., 2020, Kucharski *et al*., 2020, Ferretti *et al*., 2020). Further work involving resource limitations is critical.

Estimating model parameters from surveillance data are crucial to evaluate the effectiveness of interventions and identify margins for improvement. Some work in the direction of developing deterministic models that are efficient to solve numerically (hence suitable for model fitting) while capturing the essential features of individual contacts with rigorous probabilistic arguments has been done using time-since-infection models (Müller *et al*., 2000, Scarabel *et al*.,2021) or deterministic compartmental models (Sturniolo *et al*., 2021). However, estimation of contact tracing parameters via model fitting remains an open challenge (Swallow *et al*., 2021).

5. Open, transparent, and trusted models to support policy

The challenges addressed here focus on the need to ensure models that support public health policy are timely, better enable decision making and are transparent and openly scrutinised. Using epidemiological models to inform public health policy requires that models be trusted by policy makers and the public alike. Although the road to public trust is complex and subject to a range of forces, greater openness and transparency of model code, data, and underlying assumptions than is currently typical are required to aid reproducibility and engender increased trust, including being clear and honest about the level of uncertainty inherent in each analysis. Such an approach will allow scrutiny within and beyond teams using model outputs which will lead to more robust understanding, through testing of assumptions by a wider scientific community, and ultimately better policy and greater public confidence.

Modelling for real time decision support

Perhaps the key problem for real time decision support is the rapidly evolving landscape in terms of both knowledge of the pathogen and its impact, and hence on the priority questions for public health. Analysis therefore must be timely. This requires faster tools for simulation, statistical estimation, model assessment and uncertainty quantification (Swallow *et al*., 2021), greater understanding of available data (Shadbolt *et al*., 2021) but also analytical understanding of a wider classes of models as discussed above. Another critical issue for real time decision support is the need for effective quantification of uncertainty and good communication with decision makers. For example, what uncertainties to account for should be influenced by the requirements of decision makers e.g. as part of constructing scenarios to assess alternative interventions.

Openness within and beyond the scientific community

Issues with scientific peer review, reproducibility, and ultimately public trust arise when models used to inform policy cannot be readily scrutinised, are not properly documented, or are not immediately and openly available. Improved tools and a greater culture of openness (such as making model software open source as standard) are needed to support open modelling for future pandemics and to lay bare the assumptions and limitations of such approaches.

It is important that modellers communicate the dependence of predictions on such assumptions and make clear their limitations. There need to be transparent and accessible links between models, data and assumptions, and use of best practices in terms of documenting and testing open-source code. The need for such an approach is demonstrated by criticisms of models used to inform public health policies during the COVID-19 pandemic (Rice *et al*., 2020). Shadbolt *et al*., (2021) describe a roadmap for the development of suitable standards and software that would provide traceability and transparency tools that accessibly link model outputs to data and assumptions.

To further increase trust in models used to inform policy, in addition to such robust systems, new and accepted open epidemiology standards are needed to provide documentation of model quality, reproducibility and fitness for purpose. This includes the need for modelling teams to provide evidence of defensible parameters and, ideally, model inference from data or transparent sources for model parameters and model structure. Even more critical is the need to demonstrate evidence of model testing and assessment against simulated or ideally observational or experimental data. For example, is the model able to predict future trends and to what extent? Finally, ideally models and model-based studies used to inform public health policy and intervention should be rigorously peer-reviewed. The Royal Society's Rapid Assistance in Modelling the Pandemic (RAMP) initiative, which crowd-sourced rapid reviews of the burgeoning scientific literature, provides a paradigm for how this could be done. In summary, there is an urgent need to develop and ensure wide adoption of open modelling standards and tools that meet the above requirements. Despite inherent time constraints in pandemic response such tools should facilitate both replication and reproducibility including rigorous testing and adaptation of epidemiological models. Indeed, there have been some moves towards this during the COVID-19 pandemic with, for example, reanalysis of influential model results (Rice *et al*., 2020).

Deeper engagement with policy and decision makers

A further broader challenge is that of communication on the modeller-decision-maker interface. In terms of communication of modelling results, it is critical not to overwhelm decision-makers with too much information. Adoption of suitable visualisation techniques would focus on how best to present information in a way that supports human decision-making abilities (Chen *et al*., 2021). Such communication needs to convey both uncertainty and model assumptions and inherent limitations (Swallow *et al*., 2021; Hadley *et al*., 2021). The development of the open epidemiology standards and data pipeline tools discussed above would provide a framework within which to structure such communications. However, modellers should also give greater consideration to the presentation of results and the art of writing succinct honest, executive summaries and syntheses across multiple sources of evidence. Such communication must take into account the way in which users of the information provided are likely to interpret what is presented.

To deepen trust and increase the effectiveness of interactions across the modeller-decision-maker interface, modellers should also engage in co-construction of models and model-based analysis with policy, decision makers and other stakeholders. Such co-development would focus on formulating policy relevant questions which could be usefully addressed by modelling, and also on the development and parameterisation of models used in answering these questions. We note that the tools of expert elicitation will be invaluable in realising these ambitions (Swallow *et al*., 2021). In line with the points made previously about the need for transparency and trust in models, it is vital to ensure that such elicitation processes are well documented and open to scrutiny. Adoption of these deeply collaborative approaches would likely lead to substantially greater public health benefits derived from epidemiological modelling efforts. Moreover, the time consuming nature of such interactions suggests that the necessary exercises should begin well ahead of future pandemics. This would allow identification and testing of ideas and procedures but perhaps most importantly the development suitable relationships at institutional and individual level.

Discussion

Despite the substantial public health benefits derived to date from epidemiological models of infectious disease there are still many opportunities to enhance modelling to better inform management and control of future pandemics. Here a wide range of challenges for epidemiological modelling that need to be addressed if such benefits are to be realised have been identified. These fall under structural modelling and analytic challenges to better account for contact processes and host and pathogen complexity, and a set of challenges related to improving the impact of modelling in future pandemics (see Figure 1).

The focus on directly transmitted pathogens is justified since future pandemic risks are greatest in the case of direct person to person transmission. However, many current and future high impact disease threats result from vector borne pathogens often affecting groups and populations that lack sufficient resources to respond effectively. Climate change is likely to increase risks from vector borne pathogens as environmental conditions allow vector distributions to spread or shift to either overlap with populations not currently exposed, or to increase exposure in areas already affected. These threats raise a further set of modelling challenges. Nonetheless, many of the issues discussed here remain relevant to reducing wider impacts of infectious diseases and pandemic preparedness.

The key theme identified here is the need to make models that are more effective at providing decision makers with relevant information during pandemics. This will be underpinned by better communication with public health stakeholders at all stages from identifying questions, through model development, to the synthesis of policy advice built on model outputs and expert judgement. Such communication needs to: be built on better understanding of how target audiences perceive presented information; express the limitations of models and uncertainty in model outputs; and critically build a clear understanding of what questions are addressed by specific analyses to enable more effective decision making. An important step in achieving this is the development of tools and standards through which to make models and model outputs used to inform policy more open, transparent and trusted (Shadbolt *et al*., 2021).

In terms of impacting pandemic control, challenges identified include more realistic modelling of time and resource constraints including rates at which interventions including contact tracing and vaccination can be conducted. The modelling of contact tracing is itself identified as a particularly difficult problem due to induced dependencies between disease contacts. A further important opportunity for models to contribute to pandemic control is by using available data better to estimate the underlying characteristics of outbreaks e.g. real time reproduction numbers, transmission rates and disease progression. Furthermore, the largely untapped potential to use models to enhance possibly adaptive collection of data during a pandemic could transform future public health response. The ability to achieve such goals is dependent on continued improvements in the formulation, development and testing of models and theory of infectious disease.

Current critical gaps in the modelling toolkit include the need further to develop models to account for heterogeneity and more complex dynamics of contact structures including host behaviour, hosts and pathogens. Developments in modelling within host dynamics will enhance implementation of interventions including contact tracing and vaccination by improved theoretical understanding and better quantification through more complete use of information from disease diagnostics. Similarly, the explosion of information on pathogen genetics raises significant modelling challenges in terms of representing the diversity and evolution of pathogens, and ultimately how these interact with host genetics. An urgent and overarching requirement is the need to enhance the modelling of contact processes in human society and how these respond to public health messages and interventions including compensatory behaviours. Continued threats from SARS-CoV-2 and risks of future pandemics point to the need for models that more fully couple the dynamics of human societies and pathogens. Such human-disease system models would enable exploration of public health and other policies that would make humanity more resistant to the emergence of zoonotic diseases and less vulnerable to pandemic spread when they inevitably arise.

Acknowledgements

The authors would like to thank the Isaac Newton Institute for Mathematical Sciences, Cambridge, for support during the Infectious Dynamics of Pandemics programme where work on this paper was undertaken. This work was supported by EPSRC grant no. EP/R014604/1.

G.M. is supported by the Scottish Government's Rural and Environment Science and Analytical Services Division (RESAS).

L.P. is funded by the Wellcome Trust and the Royal Society (grant no. 202562/Z/16/Z). L.P. and F.S. are supported by the UKRI through the JUNIPER modelling consortium (grant no. MR/V038613/1). L.P. is also supported by The Alan Turing Institute for Data Science and Artificial Intelligence.

Daniel Villela is a fellow from National Council for Scientific and Technological Development (Ref. 309569/2019-2, 441057/2020-9).

Authors' contributions

All authors took part in discussions and wrote sections of the manuscript. G.M. coordinated discussions throughout and compiled the final version of the manuscript. All authors edited the manuscript and approved the final version for publication.

References

Anderson, R.M., Heesterbeek, H., Klinkenberg, D., Hollingsworth, T.D., 2020. How will country-based mitigation measures influence the course of the COVID-19 epidemic? Lancet 395(10228), P931-934. https://doi.org/10.1016/S0140-6736(20)30567-5

Bacaër, N., 2011. *A short history of mathematical population dynamics*. Springer Science and Business Media.

Bacaër, N., Guernaoui, S., 2006. The epidemic threshold of vector-borne diseases with seasonality. J. Math. Biol. 53, 421–436. https://doi.org/10.1007/s00285-006-0015-0

Ball, F., Britton, T., & Sirl, D. (2011). Household epidemic models with varying infection response. *Journal of mathematical biology*, *63*(2), 309-337. https://doi.org/10.1007/s00285-010-0372-6

Ball, F., Britton, T., House, T., Isham, V., Mollison, D., Pellis, L., Scalia Tomba, G., 2015. Seven challenges for metapopulation models of epidemics, including households models. Epidemics 10, 63–67. https://doi.org/10.1016/j.epidem.2014.08.001

Ball, F., Britton, T., Leung, K.Y., Sirl, D., 2019. A stochastic SIR network epidemic model with preventive dropping of edges. Journal of Mathematical Biology 78, 1875–1951. https://doi.org/10.1007/s00285-019-01329-4

Ball, F., Mollison, D., and Scalia-Tomba, G. (1997). Epidemics with Two Levels of Mixing. The Annals of Applied Probability, **7**(1), 46-89. https://doi.org/10.1214/aoap/1034625252

Ball, F., Neal, P., 2008. Network epidemic models with two levels of mixing. Mathematical Biosciences 212(1), 69-87. https://doi.org/10.1016/j.mbs.2008.01.001

Ball, F., Shaw, L., 2015. Estimating the within-household infection rate in emerging SIR epidemics among a community of households. J. Math. Biol. 71, 1705–1735. https://doi.org/10.1007/s00285-015-0872-5

Ball, F., Sirl, D., 2018. Evaluation of vaccination strategies for SIR epidemics on random networks incorporating household structure. J. Math. Biol. 76, 483–530. https://doi.org/10.1007/s00285-017-1139-0

Ball, F.G., Sirl, D.J. and Trapman, P., 2014. Epidemics on random intersection graphs. Annals of Applied Probability 24(3), 1081-1128. https://doi.org/10.1214/13-AAP942

Ball, F., Sirl, D. and Trapman, P., 2010 Analysis of a stochastic SIR epidemic on a random network incorporating household structure. Mathematical Biosciences 224(2), 53-73. https://doi.org/10.1016/j.mbs.2009.12.003

Barnard, R.C., Kiss, I.Z., Berthouze, L., Miller, J.C., 2018. Edge-Based Compartmental Modelling of an SIR Epidemic on a Dual-Layer Static–Dynamic Multiplex Network with Tunable Clustering. Bull Math Biol 80, 2698–2733. https://doi.org/10.1007/s11538-018-0484-5

Barnard, R.C., Berthouze, L., Simon, P.L. and Kiss, I.Z., 2019 . Epidemic threshold in pairwise models for clustered networks: closures and fast correlations. J. Math. Biol. 79, 823–860. https://doi.org/10.1007/s00285-019-01380-1

Benson, L., Davidson, R., Green, D., Hutchings, M., Marion, G., 2021. When and why direct transmission models can be used for environmentally persistent pathogens. Under review.

Bharti, N., 2021. Linking human behaviors and infectious diseases. Proceedings of the National Academy of Sciences 118 (11) e2101345118. https://doi.org/10.1073/pnas.2101345118

Britton, T., House, T., Lloyd, A.L., Mollison, D., Riley, S., Trapman, P., 2015. Five challenges for stochastic epidemic models involving global transmission. Epidemics 10, 54–57. https://doi.org/10.1016/j.epidem.2014.05.002

Britton, T., Scalia Tomba, G., 2019. Estimation in emerging epidemics: biases and remedies. J.R. Soc. Interface 16, 20180670. https://doi.org/10.1098/rsif.2018.0670

Britton, T., Trapman, P., Ball, F., 2021. The risk for a new COVID-19 wave--and how it depends on R0, the current immunity level and current restrictions." Royal Society Open Science 8(7), 210386. https://doi.org/10.1098/rsos.210386

Brookmeyer, R., Gail M. H., 1986. Minimum size of the acquired immunodeficiency syndrome (AIDS) epidemic in the United States. Lancet, 328(8519), 1320-1322. https://doi.org/10.1016/S0140-6736(86)91444-3

Brookmeyer, R., Gail M. H. (1987) Methods of projecting the AIDS epidemic. Lancet, 330(8550), 99. https://doi.org/10.1016/S0140-6736(87)92761-9

Brooks-Pollock, E., Danon, L., Jombart, T., Pellis, L., 2021. Modelling that shaped the early COVID-19 pandemic response in the UK. Philosophical Transactions of the Royal Society B, 376 (1829), 20210001. https://doi.org/10.1098/rstb.2021.0001

Chang, S., Pierson, E., Koh, P.W., Gerardin, J., Redbird, B., Grusky, D., Leskovec, J., 2021. Mobility network models of COVID-19 explain inequities and inform reopening. *Nature* **589,** 82–87. https://doi.org/10.1038/s41586-020-2923-3

Chang, S.L., Piraveenan, M., Pattison, P., Prokopenko, M., 2020. Game theoretic modelling of infectious disease dynamics and intervention methods: a review. Journal of Biological Dynamics, 14(1), 57-89. https://doi.org/10.1080/17513758.2020.1720322

Chen, M., Abdul-Rahman, A., Archambault, D., Dykes, J., Slingsby, A., Ritsos, P. D., Torsney-Weir, T., Turkay, C., Bach, B., Borgo, R., Brett, A., Fang, H., Jianu, R., Khan, S., Laramee, R.S., Nguyen, P. H., Reeve, R., Roberts, J. C., Vidal, F., Wang, Q., Wood, J., Xu, K., 2021. RAMPVIS: Answering the Challenges of Building Visualisation Capabilities for Large-scale Emergency Responses. *Submitted to* Epidemics. VSI Challenges for Future Pandemics

Chisholm, R.H., Crammond, B., Wu, Y., Bowen, A.C., Campbell, P.T., Tong, S.Y.C., McVernon, J., Geard, N., 2020. A model of population dynamics with complex household structure and mobility: implications for transmission and control of communicable diseases. PeerJ 8, e10203 https://doi.org/10.7717/peerj.10203

Chung, K., Baek, Y., Ha, M., Jeong, H., 2016. Universality classes of the generalized epidemic process on random networks. Phys. Rev. E 93, 052304. https://doi.org/10.1103/PhysRevE.93.052304

Codeço, C.T., Villela, D.A., Coelho, F.C., 2018. Estimating the effective reproduction number of dengue considering temperature-dependent generation intervals. Epidemics 25,101-11. https://doi.org/10.1016/j.epidem.2018.05.011

Conlan, A.J.K., Klepac, P., Kucharski, A.J., Kissler, S., Tang, M.L., Fry, H., Gog, J.R., 2021. Human mobility data from the BBC Pandemic project. medRxiv 2021.02.19.21252079. https://doi.org/10.1101/2021.02.19.21252079

Danon L., Lucas, L., Brooks-Pollock E., 2021. Household bubbles and COVID-19 transmission: insights from percolation theory. Phil. Trans. R. Soc. B376, 20200284.http://doi.org/10.1098/rstb.2020.0284

Day, T. and Gandon, S., 2007. Applying population-genetic models in theoretical evolutionary epidemiology. Ecology Letters 10(10), 876–888. https://doi.org/10.1111/j.1461-0248.2007.01091.x

Day, T., Gandon, S., Lion, S., Otto, S.P., 2020. On the evolutionary epidemiology of SARS-CoV-2. Current Biology *30(15)*, R849–R857. https://doi.org/10.1016/j.cub.2020.06.031

Day et al 2020b Day, T., Parsons, T., Lambert, A., Gandon, S., 2020. The Price equation and evolutionary epidemiology. Phil. Trans. R. Soc. B 375(1797), 20190357. https://doi.org/10.1098/rstb.2019.0357

Demiris, N., O'Neill, P. D. (2005). Bayesian inference for epidemics with two levels of mixing. Scandinavian journal of statistics, 32(2), 265-280. https://doi.org/10.1111/j.1467-9469.2005.00420.x

Diekmann O., Heesterbeek J.A.P., Britton T. (2013) Mathematical Tools for Under-standing Infectious Disease Dynamics. Princeton: Princeton Univ. Press

Endo, A., Centre for the Mathematical Modelling of Infectious Diseases COVID-19 Working Group, Leclerc, Q.J., Knight, G.M., Medley, G.F., Atkins, K.E., Funk,S., Kucharski, A.J., 2021. Implication of backward contact tracing in the presence of overdispersed transmission in COVID-19 outbreaks. Wellcome Open Research, 5, 239. https://doi.org/10.12688/wellcomeopenres.16344.2

Ferretti, L., Wymant, C., Kendall, M., Zhao, L., Nurtay, A., Abeler-Dörner, L., Parker, M., Bonsall, D., Fraser, C., 2020. Quantifying SARS-CoV-2 transmission suggests epidemic control with digital contact tracing, Science, 368(6491), eabb6936. https://doi.org/10.1126/science.abb6936

Flaxman, S., Mishra, S., Gandy, A., Unwin, J.T., Mellan, T.A., Coupland, H., Whittaker, C., Zhu, H., Berah, T., Eaton, J.W., Monod, M., Imperial College COVID-19 Response Team, Ghani, A.C., Donnelly, C.A., Riley, S., Vollmer, M.A.C., Ferguson, N.M., Okell, L.C., Bhatt, S., 2020. Estimating the effects of non-pharmaceutical interventions on COVID-19 in Europe. Nature 584, 257–261. https://doi.org/10.1038/s41586-020-2405-7

Frost, S.D.W., Pybus, O.G., Gog, J.R., Viboud, C., Bonhoeffer, S., Bedford, T. (2015) Eight challenges in phylodynamic inference   Epidemics 10, 88–92. https://doi.org/10.1016/j.epidem.2014.09.001

Fyles, M., Fearon, E., Overton, C., University of Manchester COVID-19 Modelling Group, Wingfield, T., Medley, G.F., Hall, I., Pellis, L., House, T., 2021. Using a household-structured branching process to analyse contact tracing in the SARS-CoV-2 pandemic. Phil.Trans. R. Soc. B 376: 20200267. https://doi.org/10.1098/rstb.2020.0267

Funk, S., Bansal, S., Bauch, C., Eames, K., Edmunds, J., Galvani, A., Klepac, P., 2015. Nine challenges in incorporating the dynamics of behaviour in infectious diseases models Epidemics 10, 21-25. https://doi.org/10.1016/j.epidem.2014.09.005

Gail M.H., Rosenberg P.S. (1992) Perspectives on Using Backcalculation to Estimate HIV Prevalence and Project AIDS Incidence. In: Jewell N.P., Dietz K., Farewell V.T. (eds) AIDS Epidemiology. Birkhäuser, Boston, MA. https://doi.org/10.1007/978-1-4757-1229-2_1

Gandolfi, A., Cecconi, L., 2016. SIR epidemics on a scale-free spatial nested modular network. Advances in Applied Probability 48(1), 137-162. https://doi.org/10.1017/apr.2015.10

Gauvin, L., Bajardi, P., Pepe, E., Lake, B., Privitera, F., Tizzoni, M., 2021. Socioeconomic determinants of mobility responses during the first wave of COVID-19 in Italy: from provinces to neighbourhoods. J. R. Soc. Interface 18(181), 20210092. https://doi.org/10.1098/rsif.2021.0092

Geard, N., Glass, K., McCaw, J.M., McBryde, E.S., Korb, K.B., Keeling, M.J., McVernon, J., 2015. The effects of demographic change on disease transmission and vaccine impact in a household structured population, Epidemics, Volume 13, 56-64. https://doi.org/10.1016/j.epidem.2015.08.002

Gog, J.R., Pellis, L., Wood, J.L.N., McLean, A.R., Arinaminpathy, N., Lloyd-Smith, J.O., 2015. Seven challenges in modeling pathogen dynamics within-host and across scales Epidemics 10, 45–48. https://doi.org/10.1016/j.epidem.2014.09.009

Goldstein, E., Paur, K., Fraser, C., Kenah, E., Wallinga, J., Lipsitch, M., 2009. Reproductive numbers, epidemic spread and control in a community of households. Mathematical biosciences, 221(1), 11-25. https://doi.org/10.1016/j.mbs.2009.06.002

Gomes, M.G.M., Aguas, R., King, J.G., Langwig, K.E., Souto-Maior, C., Carneiro, J., Penha-Gonçalves, C., Gonçalves, G., Ferreira, M.U., 2021. Individual variation in susceptibility or exposure to SARS-CoV-2 lowers the herd immunity threshold. MedRxiv, 2020.04.27.20081893 https://doi.org/10.1101/2020.04.27.20081893

Grenfell, B.T., Pybus, O.G., Gog, J.R., Wood J.L., Daly, J.M., Mumford, J.A., Holmes, E.C., 2004. Unifying the epidemiological and evolutionary dynamics of pathogens. Science, 303(5656), 327–332. https://doi.org/10.1126/science.1090727

Griffin, J.T., 2015. The Interaction between Seasonality and Pulsed Interventions against Malaria in Their Effects on the Reproduction Number.

PLoS Comput Biol 11(1), e1004057. https://doi.org/10.1371/journal.pcbi.1004057

Gyllenberg, M., Hanski, I., 1992. Single-species metapopulation dynamics: A structured model, Theoretical Population Biology, 42(1), 35-61. https://doi.org/10.1016/0040-5809(92)90004-D

Hadley, L., Challenor, P., Dent, C., Isham, V., Mollison, D., Robertson, D.A., Swallow, B., Webb, C.R., 2021. Challenges on the interaction of models and policy for pandemic control. Submitted to Epidemics, VSI Challenges for Future Pandemics.

Hay, J.A., Kennedy-Shaffer, L., Kanjilal, S., Lennon, N.J., Gabriel, S.B., Lipsitch, M., Mina, M.J., 2021. Estimating epidemiologic dynamics from cross-sectional viral load distributions. Science, 373(6552), abh0635. https://doi.org/10.1126/science.abh0635

Haw, D.J., Pung, R., Read, J.M., Riley, S., 2020. Strong spatial embedding of social networks generates nonstandard epidemic dynamics independent of degree distribution and clustering. Proceedings of the National Academy of Sciences 117 (38) 23636-23642. https://doi.org/10.1073/pnas.1910181117

Heesterbeek, H., Anderson, R.M., Andreasen, V., Bansal, S., De Angelis, D., Dye, C., Eames, K.T.D., Edmunds, W.J., Frost, S.D.W., Funk, S., Hollingsworth, T.D., House, T., Isham, V., Klepac, P., Lessler, J., Lloyd-Smith, J.O., Metcalf, C.J.E., Mollison, D., Pellis, L., Pulliam, J.R.C., Roberts, M.G., Viboud, C. and Isaac Newton Institute IDD Collaboration, 2015. Modeling infectious disease dynamics in the complex landscape of global health. Science 347 (6227), 265-270. https://doi.org/10.1126/science.aaa4339

Heffernan J.M., Keeling M.J., 2009. Implications of vaccination and waning immunity. *Proc. R. Soc. B.,* 276 (1664)**,** 2071–2080. http://doi.org/10.1098/rspb.2009.0057

Hodgson, S.H., Mansatta, K., Mallett, G., Victoria Harris, V., Emary, K.R.W., Pollard, A.J., 2021. What defines an efficacious COVID-19 vaccine? A review of the challenges assessing the clinical efficacy of vaccines against SARS-CoV-2. Lancet Infectious Diseases 21(2), E26–E35. https://doi.org/10.1016/S1473-3099(20)30773-8

Hollingsworth, T.D., Pulliam, J.R.C., Funk, S., Truscott. J.E., Isham, V., Lloyd, A.L., 2015. Seven challenges for modelling indirect transmission: vector-borne diseases, macroparasites and neglected tropical diseases. Epidemics, 10, 16-20. https://doi.org/10.1016/j.epidem.2014.08.007

House, T., Keeling, M.J., 2010. The impact of contact tracing in clustered populations. PLOS Computational Biology, 6(3), e1000721. https://doi.org/10.1371/journal.pcbi.1000721

House, T., Pellis, L., Pouwels, K.B., Bacon, S., Eidukas, A., Jahanshahi, K., Eggo, R.M., Walker, A.S., 2021. Inferring Risks of Coronavirus Transmission from Community Household Data. arXiv preprint. https://arxiv.org/abs/2104.04605

Isham, V., 1989. Estimation of the incidence of HIV infection. Phil. Trans. Roy. Soc. Lond., B, 325(1226), 113-121. https://doi.org/10.1098/rstb.1989.0076

Islam, N., Sharp, S.J., Chowell, G., Shabnam, S., Kawachi, I., Lacey B., Massaro, J.M., D’Agostino Sr, R.B., White, M., 2020. Physical distancing interventions and incidence of coronavirus disease 2019: natural experiment in 149 countries. BMJ, 370 :m2743 https://doi.org/10.1136/bmj.m2743

Jackson, C., Presanis, A., Conti, S., De Angelis, D., 2019. Value of Information: Sensitivity Analysis and Research Design in Bayesian Evidence Synthesis. Journal of the American Statistical Association, 114(528), 1436-1449. https://doi.org/10.1080/01621459.2018.1562932

Jia, J.S., Lu, X., Yuan, Y. Xu, G., Jiaand, J., Christakis, N.A., 2020. Population flow drives spatio-temporal distribution of COVID-19 in China. Nature 582, 389–394. https://doi.org/10.1038/s41586-020-2284-y

Jones, B., Sharpe, P., Iddon, C., Hathway, E.A., Noakes, C.J., Fitzgerald, S., 2021. Modelling uncertainty in the relative risk of exposure to the SARS-CoV-2 virus by airborne aerosol transmission in well mixed indoor air, Building and Environment, 191,107617. https://doi.org/10.1016/j.buildenv.2021.107617

Jorritsma, J., Hulshof, T. Komjáthya, J., 2020. Not all interventions are equal for the height of the second peak. Chaos, Solitons & Fractals 139, 109965. https://doi.org/10.1016/j.chaos.2020.109965

Kerr, C.C., Stuart, R.M., Mistry, D., Abeysuriya, R.G., Rosenfeld, K., Hart, G.R., Núñez, R.C., Cohen, J.A., Selvaraj, P., Hagedorn, B., George, L., Jastrzębski, M., Izzo, A., Fowler, G., Palmer, A., Delport, D., Scott, N., Kelly, S., Bennette, C.S., Wagner, B., Chang, S., Oron, A.P., Wenger, E., Panovska-Griffiths, J., Famulare, M., Klein, D.J., 2021. Covasim: an agent-based model of COVID-19 dynamics and interventions. MedRxiv 2020.05.10.20097469. https://doi.org/10.1101/2020.05.10.20097469

KhudaBukhsh, W., Auddy, A., Disser, Y., Koeppl, H., 2019. Approximate lumpability for Markovian agent-based models using local symmetries. Journal of Applied Probability 56(3), 647-671. https://doi.org/10.1017/jpr.2019.44

Kinyanjui, T.M., Pellis, L., House, T., 2016. Information content of household-stratified epidemics, Epidemics 16, 17-26. https://doi.org/10.1016/j.epidem.2016.03.002

Kiskowski. M., Chowell, G., 2016. Modeling household and community transmission of Ebola virus disease: Epidemic growth, spatial dynamics and insights for epidemic control. Virulence 7(2), 163-173. https://doi.org/10.1080/21505594.2015.1076613

Kissler, S., Joseph R. Fauver, J.R., Christina Mack, C., Caroline G. Tai, C.G., Breban, M.I., Watkins, A.E., Samant, R.M., Anderson, D.J., Ho, D.D., Grubaugh, N.D., Grad, Y., 2021. Densely sampled viral trajectories suggest longer duration of acute infection with B.1.1.7 variant relative to non-B.1.1.7 SARS-CoV-2." Preprint, 2021. https://dash.harvard.edu/handle/1/37366884

Knight, M.A., White, P.C.L., Hutchings, M.R., Davidson, R.S., Marion, G., 2021. Generative models of network dynamics provide insight into the effects of trade on endemic livestock disease. Royal Society Open Science 8(3), 201715. https://doi.org/10.1098/rsos.201715

Kretzschmar, M.E., Ashby, B., Fearon, E., Overton, C.E., Panovska-Griffiths, J., Pellis, L., Quaife, M., Rozhnova, G., Scarabel, F., Stage, H.B., Swallow, B., Thompson, R.N., Tildesley, M.J., Daniel Villela, D., 2021. Challenges for modelling interventions for future pandemics. *Submitted to Epidemics, VSI Challenges for Future Pandemics.*

Kretzschmar, M.E., Rozhnova, G., Bootsma, M.C.J., van Boven, M., van de Wijgert, J.H.H.M., Bonten, M.J.M., 2020. Impact of delays on effectiveness of contact tracing strategies for COVID-19: a modelling study, The Lancet Public Health, Volume 5(8), E452-E459. https://doi.org/10.1016/S2468-2667(20)30157-2.

Kucharski, A.J., Klepac, P., Conlan, A.J.K., Kissler, S.M., Tang, M.L., Fry, H., Gog, J.R., Edmunds, W.J., Emery, J.C., Medley, G., Munday, J.D., Russell, T.W., Leclerc, Q.J., Diamond, C., Procter, S.R., Gimma, A., Sun F.Y., Gibbs, H.P., Rosello, A., van Zandvoort, K., Hué, S., Meakin, S.R., Deol, A.K., Knight, G., Jombart, T., Foss, A.M., Bosse, N.I., Atkins, K.E., Quilty, B.J., Lowe, R., Prem, K., Flasche, S., Pearson, C.A.B., Houben, R.M.G.J., Nightingale, E.M., Endo, A., Tully, D.C., Liu, Y., Villabona-Arenas, J., O'Reilly, K., Funk, S., Eggo, R.M., Jit, M., Rees, E.M., Hellewell, J., Clifford, S., Jarvis, C.I., Abbott, S., Auzenbergs, M., Davies, N.G., Simons, D., 2020. Effectiveness of isolation, testing, contact tracing, and physical distancing on reducing transmission of SARS-CoV-2 in different settings: a mathematical modelling study, The Lancet Infectious Diseases 20(10),1151-1160. https://doi.org/10.1016/S1473-3099(20)30457-6.

Lau, M.S.Y., Dalziel, B.D., Funk, S., McClelland, A., Tiffany, A., Riley, S., Metcalf, C.J.E., Grenfell, B.T., 2017. Spatial and temporal dynamics of superspreading events in the 2014–2015 West Africa Ebola epidemic. PNAS **114** (9) 2337-2342. https://doi.org/10.1073/pnas.1614595114

Lau, M.S.Y., Marion, G. Streftaris, G., Gibson, G.J., 2015. A systematic Bayesian integration of epidemiological and genetic data. PLOS Computational Biology, vol. 11, pp. e1004633. https://doi.org/10.1371/journal.pcbi.1004633

Leng, T., White, C., Hilton, J., Kucharski, A., Pellis, L., Stage, H., Davies, N.G., Centre for Mathematical Modelling of Infectious Disease 2019 nCoV Working Group, Keeling, M.J., Flasche, S., 2021. The effectiveness of social bubbles as part of a Covid-19 lockdown exit strategy, a modelling study [version 2; peer review: 2 approved]. Wellcome Open Res 5, 213. (https://doi.org/10.12688/wellcomeopenres.16164.2)

Lion, S., 2018. Theoretical Approaches in Evolutionary Ecology: Environmental Feedback as a Unifying Perspective. The American Naturalist 191(1), 21–44. https://doi.org/10.1086/694865

Longini, I.M., Paul, J.R., Fine, E. M., Thacker, S.B., 1986. Predicting the global spread of new infectious agents. American Journal of Epidemiology, 123(3), 383–391. https://doi.org/10.1093/oxfordjournals.aje.a114253

Lythgoe, K.A., Hall. M., Ferretti, L., de Cesare, M., MacIntyre-Cockett, G., Trebes, A., Andersson, M., Otecko, N., Wise, E.L., Moore, N., Lynch, J., Kidd, S., Cortes, N., Mori, M., Williams, R., Vernet, G., Justice, A., Green, A., Nicholls, S.M., Ansari, M.A., Abeler-Dörner, L., Moore, C.E., Peto, T.E.A., Eyre, D.W., Shaw, R., Simmonds, P., Buck, D., Todd, J.A., on behalf of the Oxford Virus Sequencing Analysis Group (OVSG), Connor, T.R., Ashraf, S., da Silva Filipe. A., Shepherd, J., Thomson, E.C., The COVID-19 Genomics UK (COG-UK) Consortium, Bonsall, D., Fraser, C., Golubchik, T., 2021. SARS-CoV-2 within-host diversity and transmission. Science 372(6539), eabg0821. https://doi.org/10.1126/science.abg0821

Lythgoe, K.A., Pellis, L., Fraser, C., 2013. Is HIV short-sighted? Insights from a multistrain nested model. Evolution 67(10), 2769-2782. https://onlinelibrary.wiley.com/doi/full/10.1111/evo.12166

Madewell, Z.J., Dean, N.E., Berlin, J.A., Coplan, P.M., Davis K.J., Struchiner, C.J., Halloran, M.E., 2021. Challenges of Evaluating and Modeling Vaccination in Emerging Infectious Diseases. Submitted to Epidemics, VSI Challenges for Future Pandemics.

Meakin, S.R., Keeling, M.J., 2019. Correlations between stochastic endemic infection in multiple interacting subpopulations. Journal of Theoretical Biology 483,109991. https://doi.org/10.1016/j.jtbi.2019.109991

Metcalf, C.J.E., Birger, R.B., Funk, S., Kouyos, R.D., Lloyd-Smith, J.O., Jansen, V.A.A., 2015. Five challenges in evolution and infectious diseases. Epidemics 10, 40–44. https://doi.org/10.1016/j.epidem.2014.12.003

Michie, S., West, R., 2020. Behavioural, environmental, social, and systems interventions against COVID-19. BMJ 370, m2982 http://dx.doi.org/10.1136/bmj.m2982

Morawska, L., Allen, J., Bahnfleth, W., Bluyssen, P.M., Boerstra, A., Buonanno, G., Cao, J., Dancer, S.J., Floto, A., Franchimon, F., Greenhalgh, T., Haworth, C., Hogeling, J., Isaxon, C., Jimenez, J.L., Kurnitski, J., Li, Y., Loomans, M., Marks, G., Marr, L.C., Mazzarella, L., Melikov, A.K., Miller, S., Milton, D.K., Nazaroff, W., Nielsen, P.V., Noakes, C., Peccia, J., Prather, K., Querol, X., Sekhar, C., Seppänen, O., Tanabe, S., Tang, J.W., Tellier, R., Tham, K.W., Wargocki, P., Wierzbicka, A., Yao, M., 2021. A paradigm shift to combat indoor respiratory infection. Science, 372(6543), 689-691. https://doi.org/10.1126/science.abg2025

Morris, M., Kretzschmar, M., 1997. Concurrent partnerships and the spread of HIV. AIDS **11**(5), 641-648. ISSN 0269-9370

Müller, J., Hösel, V., 2021. Contact Tracing and Super-Spreaders in the Branching-Process Model. arXiv, 2010.04942. https://arxiv.org/abs/2010.04942v2

Müller, J., Kretzschmar, M., Dietz, K., 2000. Contact tracing in stochastic and deterministic epidemic models, Mathematical Biosciences 164(1), 39-64, https://doi.org/10.1016/S0025-5564(99)00061-9.

Office for National Statistics. Coronavirus (COVID-19) Infection Survey, UK: 9 April 2021. (Accessed on 06/09/2021). https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/conditionsanddiseases/bulletins/coronaviruscovid19infectionsurveypilot/9april2021

Overton, C.E., Stage, H.B., Ahmad, S., Curran-Sebastian, J., Dark, P., Das, R., Fearon, E., Felton, T., Fyles, M., Gent, N., Hall, I., House, T., Lewkowicz, H., Pang, X., Pellis, L., Sawko, R., Ustianowski, A., Vekaria, B., Webb, L., 2020. Using statistics and mathematical modelling to understand infectious disease outbreaks: COVID-19 as an example, Infectious Disease Modelling, 5, 409-441. https://doi.org/10.1016/j.idm.2020.06.008.

Panovska-Griffiths, J,, Stuart, R.M., Kerr, C.C., Rosenfield, K., Mistry, D., Waites, W., Klein, D.J., Bonell, C., Viner, R.M., 2021. Modelling the impact of reopening schools in the UK in early 2021 in the presence of the alpha variant and with roll-out of vaccination against SARS-CoV-2. medRxiv 2021.02.07.21251287. https://doi.org/10.1101/2021.02.07.21251287

Pellis, L., Ball, F., Bansal, S., Eames, K., House, T., Isham, V., Trapman, P., 2015. Eight challenges for network epidemic models. Epidemics 10, 58–62. https://doi.org/10.1016/j.epidem.2014.07.003

Pellis, L., Ball, F., & Trapman, P. (2012). Reproduction numbers for epidemic models with households and other social structures. I. Definition and calculation of R0. Mathematical biosciences, 235(1), 85-97. https://doi.org/10.1016/j.mbs.2011.10.009

Pellis, L., Ball, F., Trapman, P., 2012. Reproduction numbers for epidemic models with households and other social structures. I. Definition and calculation of R0. Mathematical biosciences 235(1), 85-97. https://doi.org/10.1016/j.mbs.2011.10.009

Pellis, L., Cauchemez, S., Ferguson, N. M., & Fraser, C. (2020). Systematic selection between age and household structure for models aimed at emerging epidemic predictions. Nature communications, 11(1), 906. https://doi.org/10.1038/s41467-019-14229-4

Pellis, L., Ferguson, N. M., & Fraser, C. (2011). Epidemic growth rate and household reproduction number in communities of households, schools and workplaces. Journal of mathematical biology, 63(4), 691-734. https://doi.org/10.1007/s00285-010-0386-0

Pooley, C.M., Marion, G., Bishop, S.C., Bailey, R.I., Doeschl-Wilson, A.B., 2020. Estimating individuals' genetic and non-genetic effects underlying infectious disease transmission from temporal epidemic data. PLOS Computational Biology 16(12), e1008447. https://doi.org/10.1371/journal.pcbi.1008447

Read, A.F., Baigent, S.J., Powers C., Kgosana, L.B., Blackwell, L., Smith, L.P., Kennedy, D.A., Walkden-Brown, S.W., Nair, V.K., 2015. Imperfect Vaccination Can Enhance the Transmission of Highly Virulent Pathogens. PLoS Biol 13(7), e1002198. https://doi.org/10.1371/journal.pbio.1002198

Rice, K., Wynne, B., Martin, V., Ackland, G.J., 2020. Effect of school closures on mortality from coronavirus disease 2019: old and new predictions BMJ 371, m3588 https://doi.org/10.1136/bmj.m3588

Riley, S., Eames, K., Isham, V., Mollison, D., Trapman, P., 2015. Five challenges for spatial epidemic models. Epidemics 10, 68-71. https://doi.org/10.1016/j.epidem.2014.07.001

Ritchie, M., Berthouze, L., House, T., Kiss, I.Z., 2014. Higher-order structure and epidemic dynamics in clustered networks. Journal of Theoretical Biology 348, 21-32. https://doi.org/10.1016/j.jtbi.2014.01.025

Roberts, M., Andreasen, V., Lloyd, A., Pellis, L., 2015. Nine challenges for deterministic epidemic models. Epidemics 10, 49–53. https://doi.org/10.1016/j.epidem.2014.09.006

Rydevik, G., Innocent, G.T., Marion, G., Davidson, R.S., White, P.C.L., Billinis, C., Barrow, P., Mertens, P.P.C., Gavier-Widén, D., Hutchings, M.R., 2016. Using Combined Diagnostic Test Results to Hindcast Trends of Infection from Cross-Sectional Data. PLoS Comput Biol 12(7), e1004901. https://doi.org/10.1371/journal.pcbi.1004901

Scarabel, F., Pellis, L., Ogden, N.H., Wu, J., 2021. A renewal equation model to assess roles and limitations of contact tracing for disease outbreak control, *Royal Society Open Science*, 8:202091 http://doi.org/10.1098/rsos.202091

Shadbolt, N., Brett, A., Chen, M., Marion, G., McKendrick, I.J., Panovska-Griffiths, J., Pellis, L., Reeve, R., Swallow, B., 2021. The Challenges of Data in Future Pandemics. Submitted to Epidemics, VSI Challenges for Future Pandemics.

Silva, D.H., Rodrigues, F.A., Ferreira, S.C., 2020. High prevalence regimes in the pair-quenched mean-field theory for the susceptible-infected-susceptible model on networks. Phys. Rev. E 102, 012313. https://doi.org/10.1103/PhysRevE.102.012313

Siraj, A.S., Oidtman, R.J., Huber, J.H., Kraemer, M.U., Brady, O.J., Johansson, M.A., Perkins, T.A., 2017. Temperature modulates dengue virus epidemic growth rates through its effects on reproduction numbers and generation intervals. PLoS neglected tropical diseases 11(7), e0005797. https://doi.org/10.1371/journal.pntd.0005797

Stoner, O., Economou, T., 2020. Multivariate hierarchical frameworks for modeling delayed reporting in count data. Biometrics 76(3), 789-798. https://doi.org/10.1111/biom.13188

Sturniolo, S., Waites, W., Colbourn, T., Manheim, D., Panovska-Griffiths, J., 2021. Testing, tracing and isolation in compartmental models. PLOS Computational Biology 17(3), e1008633. https://doi.org/10.1371/journal.pcbi.1008633

Swallow, B., Birrell, P., Blake, J., Burgman, M., Challenor, P., Coffeng, L.E., Dawid, P., De Angelis, D., Goldstein, M., Hemming, V., Marion, G., McKinley, T.J., Overton, C., Panovska-Griffiths, J., Pellis, L., Probert, W., Shea, K., Villela, D., Vernon, I., 2021. Challenges in estimation,uncertainty quantification and elicitation for pandemic modelling. Submitted to Epidemics, VSI Challenges for Future Pandemics.

Tang, D., 2020. Decentralised, privacy-preserving Bayesian inference for mobile phone contact tracing. arXiv, 2005.05086. https://arxiv.org/abs/2005.05086

Thompson, R.N., Hollingsworth, T.D., Isham, V., Arribas-Bel, D., Ashby, B., Britton, T., Challenor, P., Chappell, LH.K., Clapham, H., Cunniffe, N.J., Dawid, A.P., Donnelly, C.A., Eggo, R.M., Funk, A., Gilbert, N., Glendinning, P., Gog J.R., Hart, W.S., Heesterbeek, H., House, T., Keeling, M., Kiss, I.Z., Kretzschmar. M.E., Lloyd, A.L., McBryde, E.S., McCaw, J.M., McKinley, T.J., Miller, J.C., Morris, M., O'Neill, P.D., Parag, K.V., Pearson, C.A.B., Pellis, L., Pulliam, J.R.C., Ross, J.V., Scalia Tomba, G., Silverman, B.W., Struchiner, C.J., Tildesley, M.J., Trapman, P., Webb, C.R., Mollison, D., Restif, O., 2020. Key questions for modelling COVID-19 exit strategies. Proc. R. Soc. B, **287**(1932), 20201405. https://doi.org/10.1098/rspb.2020.1405

Trapman, P., Ball, F., Dhersin, J-S., Tran, V.C., Wallinga, J., Britton, T., 2016. Inferring R 0 in emerging epidemics—the effect of common population structure is small. Journal of the Royal Society Interface 13(121), 20160288. https://doi.org/10.1098/rsif.2016.0288

Valdano, E., Fiorentin, M.R., Poletto, C., Colizza, V., 2018. Epidemic Threshold in Continuous-Time Evolving Networks. Physical Review Letters 120, 068302. https://doi.org/10.1103/PhysRevLett.120.068302

Vegvari, C., Abbott, S., Ball, F., Brooks-Pollock, E., Challen, R., Collyer, B.S., Dangerfield, C., Gog, J.R., Gostic, K.M., Heffernan, J.M., Hollingsworth, T.D., Isham, V., Kenah, E., Mollison, D., Panovska-Griffiths, J., Pellis, L., Roberts, M.G., Tomba, G.S., Thompson, R.N., Pieter Trapman, P., 2020. Commentary on the use of the reproduction number R during the COVID-19 pandemic. https://www.newton.ac.uk/files/preprints/ni20003.pdf

Verity, R., Okell, L.C., Dorigatti, I., Winskill, P., Whittaker, C., Imai, N., Cuomo-Dannenburg, G., Thompson, H., Walker, P.G.T., Fu, H., Dighe, A., Griffin, J.T., Baguelin, M., Bhatia, S., Boonyasiri, A., Cori, A., Cucunubá, Z., FitzJohn, R., Gaythorpe, K., Green, W., Hamlet, A., Hinsley, W., Laydon, D., Nedjati-Gilani, G., Riley, S., van Elsland, S., Volz, E., Wang, H., Wang, Y., Xi, X., Donnelly, C.A., Ghani, A.C., Ferguson, N.M., 2020. Estimates of the severity of coronavirus disease 2019: a model-based analysis. The Lancet, 20(6), 669-677. https://doi.org/10.1016/S1473-3099(20)30243-7

Vigfusson, Y., Karlsson, T.A., Onken, D., Song, C., Einarsson, A.F., Kishore, N., Mitchell, R.M., Brooks-Pollock, E., Sigmundsdottir, G., Danon, L., 2021. Cell-phone traces reveal infection-associated behavioral change. Proceedings of the National Academy of Sciences, 118(6) e2005241118. https://doi.org/10.1073/pnas.2005241118

Villela D.A,M., 2021. Household crowding hampers mitigating the transmission of SARS-CoV-2. Revista da Sociedade Brasileira de Medicina Tropical, 54, e08212020 https://doi.org/10.1590/0037-8682-0821-2020

Volz, E.M., Frost, S.D.W., 2013. Inferring the Source of Transmission with Phylogenetic Data. PLoS Comput Biol 9(12): e1003397. https://doi.org/10.1371/journal.pcbi.1003397

Wade, M., Jacomo, A.L., Armenise, E., Brown, M., Bunce, J., Cameron, G., Fang, Z., Farkas, K., Gilpin, D., Graham, D., Grimsley, J., Hart, A., Hoffmann, T., Jackson, K., Jones, D., Lilley, C., McGrath, J., McKinley, J., McSparron, C., Nejad, B.F., Morvan, M., Quintela-Baluja, M., Roberts, A., Singer, A., Souque, C., Speight, V., Sweetapple, C., Watts, G., Weightman, A., Kasprzyk-Hordern. B., 2021. Understanding and managing uncertainty and variability for wastewater monitoring beyond the pandemic: Lessons learned from the United Kingdom National COVID-19 Surveillance Programmes. Earth and Space Science Open Archive, 10507606.2. https://doi.org/10.1002/essoar.10507606.2

Wallinga, J., Lipsitch, M., 2007. How generation intervals shape the relationship between growth rates and reproductive numbers. Proceedings of the Royal Society B: Biological Sciences 274(1609), 599-604. https://doi.org/10.1098/rspb.2006.3754

Wang, Z., Galea, E.R., Grandison, A., Ewer, J., Jia, F., 2021. Inflight transmission of COVID-19 based on experimental aerosol dispersion data, Journal of Travel Medicine 28(4), taab023. https://doi.org/10.1093/jtm/taab023

Whittles, L.K, Imai, N., Knock, E.S., Perez-Guzman, P.N., Sonabend, R., Ghani. A., Ferguson, N.M., Baguelin, M., Cori, A., 2021. “Unlocking” Roadmap Scenarios for England v2 https://www.gov.uk/government/publications/imperial-college-london-unlocking-roadmap-scenarios-for-england-18-february-2021

Wikramaratna, P.S., Kucharski, A., Gupta, S., Andreasen, V., McLean, A.R., Gog, J.R., 2015. Five challenges in modelling interacting strain dynamics Epidemics 10, 31–34. https://doi.org/10.1016/j.epidem.2014.07.005

Wu, Q., Hadzibeganovic, T., 2020. An individual-based modeling framework for infectious disease spreading in clustered complex networks. Applied Mathematical Modelling 83, 1-12. https://doi.org/10.1016/j.apm.2020.02.012

Xiang, W., Swallow, B., 2021. Multivariate spatio-temporal analysis of the global COVID-19 pandemic. medRxiv preprint. https://doi.org/10.1101/2021.02.08.21251339

Zhou. W., Wang, A., Xia, F., Xiao, Y., Tang, S., 2020. Effects of media reporting on mitigating spread of COVID-19 in the early phase of the outbreak. Mathematical Biosciences and Engineering 17(3), 2693-2707. https://doi.org/10.3934/mbe.2020147