Remodelling selection to optimise disease forecasts and policies

Authors: M. Gabriela M. Gomes,[1,2*] Andrew M. Blagborough,[3] Kate E. Langwig,[5] Beate Ringwald.[4]

[1] Department of Mathematics and Statistics, University of Strathclyde, Glasgow, United Kingdom.

[2] NOVA School of Science and Technology, Centre for Mathematics and Applications (NOVA MATH), Caparica, Portugal.

[3] Department of Pathology, University of Cambridge, Cambridge, United Kingdom.

[4] Liverpool School of Tropical Medicine, Liverpool, United Kingdom.

[5] Department of Biological Sciences, Virginia Tech, Blacksburg, VA, USA.

[*] Correspondence to: gabriela.gomes@strath.ac.uk.

Abstract: Mathematical models are increasingly adopted for setting disease prevention and control targets. As model-informed policies are implemented, however, the inaccuracies of some forecasts become apparent, for example overprediction of infection burdens and intervention impacts. Here, we attribute these discrepancies to methodological limitations in capturing the heterogeneities of real-world systems. The mechanisms underpinning risk factors of infection and their interactions determine individual propensities to acquire disease. These factors are potentially so numerous and complex that to attain a full mechanistic description is likely unfeasible. To contribute constructively to the development of health policies, model developers either leave factors out (reductionism) or adopt a broader but coarse description (holism). In our view, predictive capacity requires holistic descriptions of heterogeneity which are currently underutilised in infectious disease epidemiology, in comparison to other population disciplines, such as non-communicable disease epidemiology, demography, ecology and evolution.

1. Introduction

Setting realistic targets and developing feasible strategies for disease prevention and control depends on representative models. These can be conceptual, experimental, or mathematical. Mathematical modelling was established in infectious diseases over a century ago [Ross et al 1916; Ross and Hudson 1917; Kermack and McKendrick 1927]. Propelled by the discovery of aetiological agents for infectious diseases, and Koch's postulates, models have focused on the complexities of pathogen transmission and evolution to understand and predict disease trends in greater depth [Heesterbeck et al 2015]. This has led to their adoption by decision makers to inform national and international policy. However, as model-informed policies are being implemented, systematic errors in forecasts become increasingly apparent, most notably their tendency to overpredict infection burdens and overestimate the impact of control measures [Gaolathe et al 2016; Karim 2016; UNAIDS 2017; Frescura et al 2022; Specht et al 2019; Flaxman et al 2020; Gomes et al 2022]. Here, we discuss how these discrepancies could be explained by methodological limitations in capturing the effects of individual variation in real-world systems. We suggest improvements that derive from early theory in the analysis of hazards [Greenwood and Yule 1920].

When a physical, chemical, or biological hazard invades a population, it typically encounters a set of individuals that can vary dramatically in their susceptibility or exposure to the threat. As a result, more susceptible (or exposed) individuals tend to be affected first while the mean susceptibility among those remaining unaffected decreases due to the selective depletion of the most susceptible. This process effectively decelerates growth in the number of disease cases when compared to a scenario of equally susceptible individuals exposed to the same mean hazard (Figure 1). Hence when homogeneous (or insufficiently heterogeneous) models fitted to the early phase of an epidemic are used to project the future, cases tend to be overpredicted. Conversely, if too much individual variation is built into the model, then cases may be underpredicted. Deviations in the quantification of variation that is under selection tend to induce large biases and, therefore, their quantification should play an essential part in the construction of predictive models for infectious as well as non-communicable diseases.

The selective depletion bias just described is pervasive in population studies and has been discovered many times and given many names, such as survivorship bias [Wald 1943], frailty variation [Vaupel et al 1979], phenotypic selection [Haldane 1954; Lande and Arnold 1983], or selective (dis)appearance [Forslund and Pärt 1995, van de Pol and Verhulst 2006]. It has been recognised to affect diverse phenomena. It can create spurious trends in measured rates of mortality [Keyfitz and Littman 1979; Vaupel et al 1979], leading to paradoxical risk

associations [Vaupel and Yashin 1985; Strandberg et al 2013] and conflicting evidence on theories of ageing [Nussey et al 2006]. It may induce misleading expectations for the survival of endangered species [Kendall and Fox 2002; Jenouvrier et al 2018]. It may affect the scope of neutral theories of biodiversity and molecular evolution [Steiner and Tuljapurkar 2012; Gomes et al 2019a]. It may bias estimates of risks of diseases, whether non-communicable [Aalen et al 2015; Stensrud and Valberg 2017] or infectious [Anderson et al 1986; Colgate et al 1988; Dwyer et al 1997; Smith et al 2005; Bellan et al 2015; Gomes et al 2019b; Corder et al 2020; Britton et al 2020; Gomes et al 2022], and efficacy of interventions, such as vaccines [Halloran et al 1996; O´Hagan et al 2012; Gomes et al 2014; Gomes et al 2016; Langwig et al 2017] or symbionts [Pessoa et al 2016; King et al 2018]. Some of these insights gave rise to new research priorities in evolutionary biology [Metcalf and Pavard 2007] while this paper presents a case for an equivalent impetus in infectious disease epidemiology.

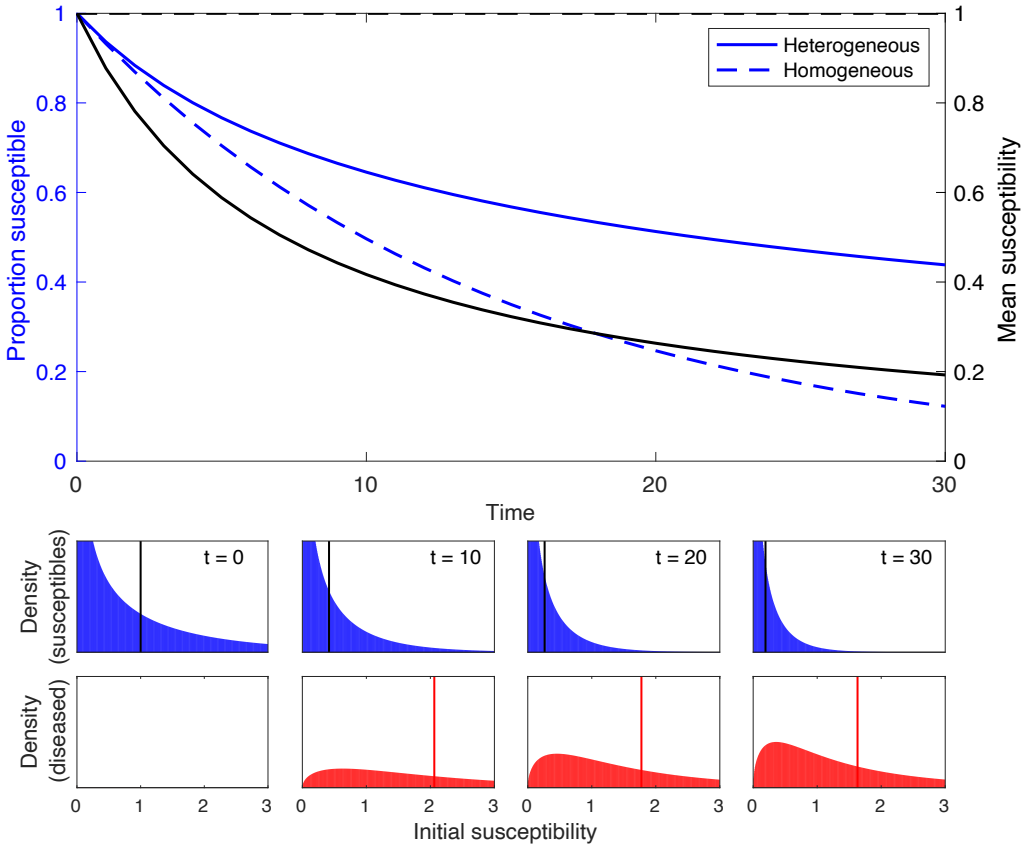

Figure 1: **Depletion of susceptibles in homogeneous and heterogeneous populations.** (Top) Proportion susceptible (blue) and mean susceptibility (black) to a non-communicable disease formulated as a Susceptible-Diseased model with constant exposure to a disease-causing agent [$\lambda = 0.07$] in two scenarios: (homogeneous susceptibility) $dS/dt = -\lambda S$, $dD/dt = \lambda S$ (dashed curves); and (gamma distributed susceptibility [x] with variance 2): $dS(x)/dt = -\lambda x S(x)$, $dD(x)/dt = \lambda x S(x)$ (solid curves). (Botton) Density of susceptible (blue) and diseased (red) individuals over the susceptibility domain at four different time snapshots of the epidemic. Mean susceptibility decreases over time due to the disproportionate depletion of individuals with high susceptibility. The vertical lines mark the mean baseline susceptibility in each context.

In this topical review, we illustrate how unmeasured heterogeneity can have a wide expression in infectious disease dynamics and formulate a pragmatic approach to estimate the most impactful forms that need to be incorporated in mathematical models to eliminate common biases.

2. Heterogeneity affects the accuracy of model forecasts

We use the examples of acquired immunodeficiency syndrome (AIDS) and coronavirus disease 2019 (COVID-19) to illustrate the effects that individual variation in susceptibility

and exposure to infection can have on the performance of mathematical models for the dynamics of endemic and epidemic diseases.

(a) *Endemic infectious diseases*

Since the detection of AIDS in the early 1980s, it has been evident that heterogeneity in individual sexual behaviours needed to be considered in mathematical models for the transmission of the causative agent – the Human Immunodeficiency Virus (HIV) [Anderson et al 1986; Colgate et al 1988]. Much research has been devoted to measuring contact networks in diverse settings and by different methods, to attempt to reproduce transmission dynamics accurately [Woolhouse et al 1997; Keeling and Eames 2005; Leigh Brown et al 2011]. However, other equally important sources of inter-individual variation may have been overlooked. For example, models that omit heterogeneity in infectiousness and susceptibility lead to substantial overestimates of HIV acute phase infectivity, resulting in an overemphasis of the early stage of infection as a driver of new infections as shown by Bellan et al [2015]. By accounting for such heterogeneities, the authors concluded that elevated acute phase infectivity was less likely to compromise "treatment as prevention" measures.

The problem of unaccounted for heterogeneity in models forecasting an infectious disease can be illustrated with the simplest mathematical description of pathogen transmission in a host population. Figure 2 shows the prevalence of infection over time under three alternative scenarios: all individuals are at equal risk of acquiring infection (black trajectories); individual risk is affected by a factor that modifies either their susceptibility to infection (blue); or exposure through connectivity with other individuals (green). Homogeneous models assign every individual a risk factor of 1 (black frequency plot), whereas heterogeneous risk derives from a distribution with mean one (blue and green density plots). As the virus spreads within the population, individuals at higher risk are predominantly infected as indicated at endemic equilibrium (Figure 2 A, B, C, density plots on the right, coloured red) and after 100 years of control (Figure 2 D, E, F). The control strategy applied to endemic equilibrium in the figure is the 90-90-90 treatment as prevention target advocated post-2015 by the Joint United Nations Programme on HIV/AIDS (UNAIDS) whereby 90% of HIV-infected individuals should be detected, with 90% of these receiving antiretroviral therapy, and 90% of these should achieve viral suppression (becoming effectively non-infectious).

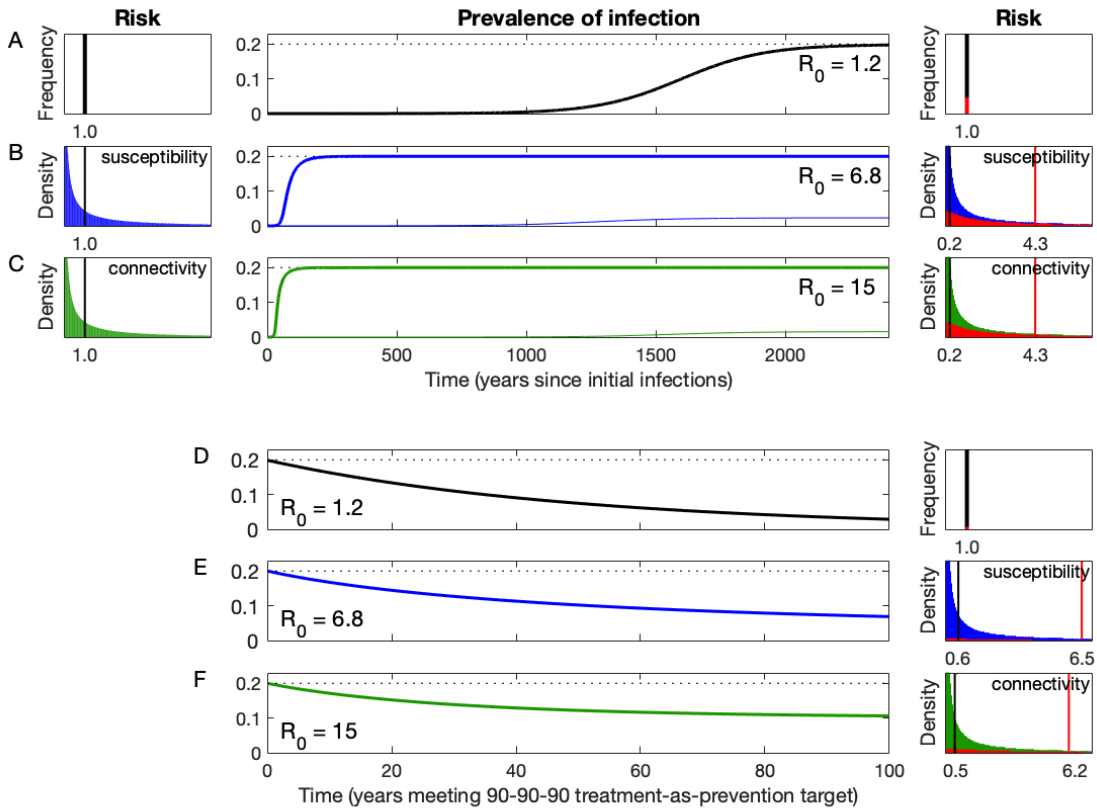

Figure 2: Prevalence trajectories under homogeneous and heterogeneous models. Risk distributions are simulated in three scenarios: homogeneous (A, D) [notice the unrealistic time scale in A]; distributed susceptibility to infection with variance 10 (B, E); distributed connectivity with variance 10 (C, F). In disease-free equilibrium, individuals differ in potential risk in scenarios B and C, but not in scenario A (risk panels on the left). The vertical lines mark the mean risk values (1 in all cases). At endemic equilibrium, individuals with higher risk are predominantly infected (risk panels on the right, where red vertical lines mark mean baseline risk among individuals who eventually became infected), resulting in reduced mean risk among those who remain uninfected (black vertical lines). To compensate for this selection effect, heterogeneous models require a higher R_0 to attain the same endemic prevalence (A, B, C). Interventions that reduce infection also reduce selection pressure, which unintendedly increases mean risk in the uninfected subpopulation and undesirably reduces intervention impact (D, E, F). Models: homogeneous (A, D) $dS/dt = \mu - \beta IS - \mu S$, $dI/dt = \beta IS - \mu I$, and $R_0 = \beta/\mu$; heterogeneous susceptibility (B, E) $dS(x)/dt = q(x)\mu - \beta \int I(u)du\, xS(x) - \mu S(x)$, $dI(x)/dt = \beta \int I(u)du\, xS(x) - \mu I(x)$, and $R_0 = \beta/\mu$; heterogeneous connectivity (C, F) $dS(x)/dt = q(x)\mu - \beta \int uI(u)du\, xS(x) - \mu S(x)$, $dI(x)/dt = \beta \int uI(u)du\, xS(x) - \mu I(x)$, and $R_0 = \int u^2 q(u)du\, \beta/\mu$. In heterogeneous models, $q(x)$ is a probability density function with mean 1 and variance 10, and initial conditions are of the form $S(x,t) = (1-\varepsilon)q(x)$ and $I(x,t) = \varepsilon q(x)$, for some infectious seed $0 < \varepsilon \ll 1$. Gamma distributions were used for concreteness.

Figure 2 shows that heterogeneous models that account for wide biological and social variation require higher basic reproduction numbers (R_0) to reach a given endemic level and predict less impact for control efforts when compared with the homogeneous counterpart model. This holds true regardless of whether heterogeneity affects susceptibility or connectivity and is generalisable to realistic combinations of the two traits. At endemic equilibrium, individuals at higher risk are predominantly infected (red distributions have mean greater than one as marked by the red vertical lines), and hence those who remain uninfected are individuals with lower risk (blue and green distributions have mean lower than one as marked by the black vertical lines). Thus, the mean risk in the uninfected but susceptible subpopulation decreases, and the epidemic decelerates (thin blue and green curves); higher values of R_0 are consequently required if the heterogeneous models are to attain the same endemic level as the homogeneous formulation (heavy blue and green

curves). Finally, interventions are less impactful under heterogeneity because any decrease in transmission collaterally increases the mean risk factor of the uninfected subpopulation (Figure 2, risk panels on the right) offering extra resistance to control. In concrete, these biases could help explain trends in HIV incidence data which lag substantially behind targets informed by model predictions [Granich et al 2009], even in settings that reached the 90-90-90 implementation targets [Gaolathe et al 2016; Karim 2016; UNAIDS 2017; Frescura et al 2022], meanwhile raised to 95-95-95 [UNAIDS 2023].

We emphasise that these results do not oppose previous research showing that antiretroviral treatments can not only delay disease, but also prevent transmission. The 90-90-90 treatment-as-prevention target helped improve access to antiretroviral medicines and save lives globally. The question is how these benefits translate from individual to population level. In our perspective, complementary measures are needed to reduce the susceptibility and exposure of uninfected individuals, especially those most vulnerable of acquiring HIV. In later sections we outline a procedure that seeks to account for effects of the entire heterogeneity of real-world systems.

(b) *Epidemic infectious diseases*

At the end of 2019, a novel severe acute respiratory syndrome coronavirus (SARS-CoV-2) isolated from a patient in China began to spread worldwide causing the COVID-19 pandemic. Countrywide epidemics have been extensively analysed and modelled throughout the world. Early studies projected first waves of infection with attack rates of around 90% if transmission had been left unmitigated [Davies et al 2020; Flaxman et al 2020], while subsequent reports noted that individual variation in susceptibility or exposure might flatten epidemic curves and reduce these estimates substantially [Britton et al 2020; Neipel et al 2020; Rose et al 2021; Tkachenko et al 2021; Montalbán et al 2022; Gomes et al 2022], as shown in Figure 3 (compare the blue [heterogeneous susceptibility] and green [heterogeneous connectivity] curves with the black [homogeneous]). Moreover, these types of variation that are subject to selection through natural infection tend to affect population measures of risk ratios leading to biased interpretations if realistic heterogeneity is not accounted for. For example, the bottom panel in Figure 3 illustrates how reinfection risk is likely to be overestimated when heterogeneity is neglected (black horizontal line represents individual risk ratio while blue and green curves depict time-dependent population risk ratios under heterogeneous susceptibility and connectivity, respectively).

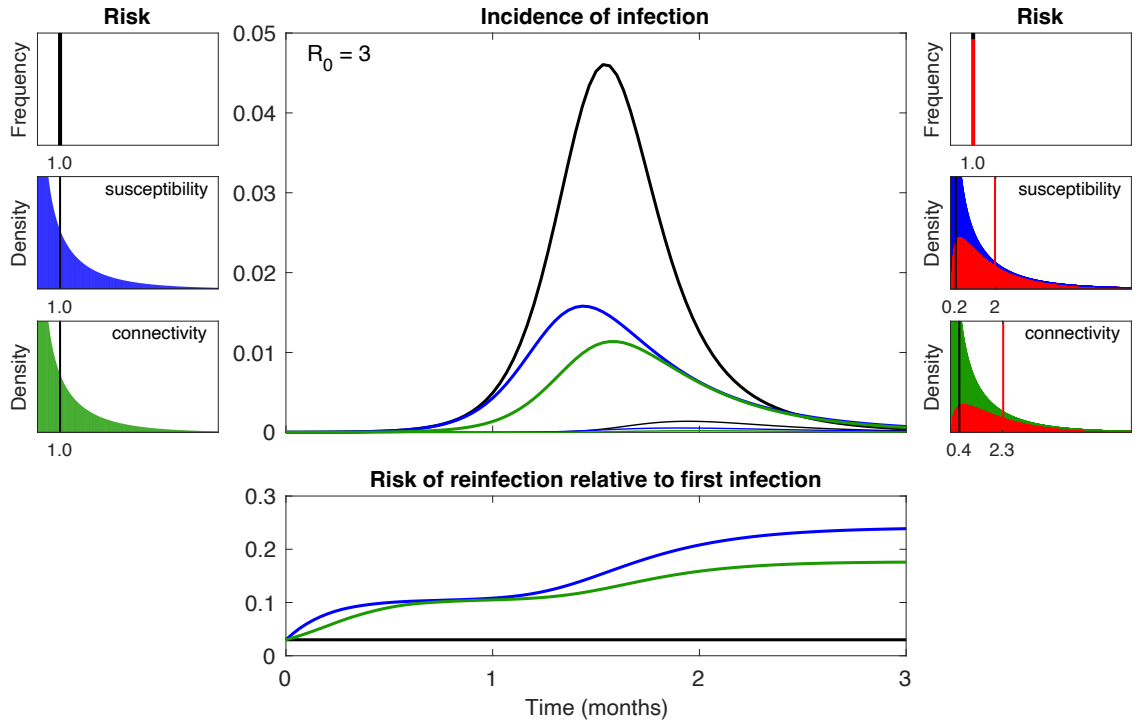

Figure 3: Incidence trajectories under homogeneous and heterogeneous models. Risk distributions are simulated in three scenarios: homogeneous (black); distributed susceptibility to infection with variance 2.5 (blue); distributed connectivity with variance 2.5 (green). On the main panel, heavy lines represent first infection and thin lines are reinfection. Left panels represent distributions of potential individual risk prior to the outbreak, with vertical lines marking mean risk values (1 in all cases). As the epidemic progresses, individuals with higher risk are predominantly infected, depleting the susceptible pool in a selective manner and decelerating epidemic growth. Right panels show in red the risk distributions among individuals who have been infected over 3 months of epidemic spread (mean greater than one when risk is heterogeneous, as marked by red vertical lines) and the reduced mean risk among those who have not been affected (black vertical lines). Models: homogeneous (black) $dS/dt = -\beta IS$, $dI/dt = \beta I(S + \sigma R) - \gamma I$, $dR/dt = \gamma I - \sigma\beta IR$, and $R_0 = \beta/\gamma$; heterogeneous susceptibility (blue) $dS(x)/dt = -\beta \int I(u)du\, xS(x)$, $dI(x)/dt = \beta \int I(u)du\, x[S(x) + \sigma R(x)] - \gamma I(x)$, $dR(x)/dt = \gamma I(x) - \sigma\beta \int I(u)du\, xR(x)$, and $R_0 = \beta/\gamma$; heterogeneous connectivity (green) $dS(x)/dt = -\beta \int uI(u)du\, xS(x)$, $dI(x)/dt = \beta \int uI(u)du\, x[S(x) + \sigma R(x)] - \gamma I(x)$, $dR(x)/dt = \gamma I(x) - \sigma\beta \int uI(u)du\, xR(x)$ and $R_0 = \int u^2 q(u)du\, \beta/\gamma$. In heterogeneous models, $q(x)$ is a probability density function with mean 1 and variance 2.5, and initial conditions are of the form $S(x,t) = (1-\varepsilon)q(x)$, $I(x,t) = \varepsilon q(x)$, and $R(x,t) = 0$, for some infectious seed $0 < \varepsilon \ll 1$. Gamma distributions were used for concreteness. Parameter σ represents the risk of reinfection of each individual relative to their own risk of first infection, here assumed $\sigma = 0.03$. The bottom panel depicts the average risk of reinfection (over the subpopulation at risk of reinfection) relative to the average risk of first infection (over the subpopulation at risk of first infection).

Representing individual variation is necessary to forecast infectious disease dynamics and inform policy. Epidemic curves for COVID-19 are widely available, and it is possible to construct models with inbuilt risk distributions. Their shapes can be inferred by assessing the ability of models to fit simulated trajectories to observed epidemics, while accounting for realistic social and biomedical interventions [Gomes et al 2022]. It has also been highlighted that the interplay between social dynamics and spread of infection may reduce the effects described herein [Tkachenko 2021]. If socioeconomic gradients (main drivers of risk heterogeneity in infectious diseases [Millett et al; Mena et al 2021; Xia et al 2022]) changed over time in such a way that individuals with low susceptibility/exposure early in the epidemic became high susceptibility/exposure in later stages, and vice versa, this could compromise the utility of coefficients of variation estimated early on. Inverting socioeconomic gradients and their health impacts, however, would require a much longer time scale than that of an acute infectious disease pandemic [Braveman, Gottlieb 2014].

There is mounting evidence that, on the contrary, disadvantaged social groups suffer more from both disease and containment measures, exacerbating preexisting risk inequalities [Okonkwo et al 2021]. Gomes et al [2022] estimated similar coefficients of variation by fitting time series encompassing either one or two epidemic waves of COVID-19 in England and Scotland, suggesting long-lasting heterogeneity.

A contrasting and more common approach to incorporate heterogeneity in COVID-19 transmission models has been to focus on specific sources of heterogeneity, such as age structure, households, schools, workplaces, and implement these according to available data (see, for example, [Moore et al 2021; Hilton et al 2022] for differential equation formulations and [Kerr et al 2021 for agent-based models). A strength of this reductionism is to base the implementation of specific heterogeneities on explicit data. A weakness is that it does not usually capture the entire heterogeneity of the real system due to limits in data availability and capacity to process so much complexity, although it is conceivable that this may be overcome in the future. Meanwhile, a holistic compromise can be reached by formulating heterogeneity unspecifically into otherwise homogeneous (or incompletely heterogeneous) models and inferring its magnitude by fitting to trends measured in suitable population studies as outlines in the following sections. Once the biases due to unmodelled heterogeneity are understood it should be unacceptable to base policy on model projections that are not accompanied by a thorough quantitative investigation of the subject, either by directly incorporating informative data into the model, by conducting sensitivity analyses, by aiming to infer heterogeneity as we outline in the following sections, or some combination of these schemes.

3. Heterogeneity affects vaccine efficacy estimation over time and across settings

The need to account for heterogeneity in risk of acquiring infections is generally applicable not only across all models of infectious disease epidemiology, but also in methods intended to evaluate the efficacy of interventions from experimental studies, whether lab-based controlled experiments or field-based randomised controlled trials.

Individual variation in susceptibility or exposure to infection induces biases in cohort studies and clinical trials. Vaccine efficacy trials offer a useful illustration of the problem and expose a pragmatic approach to its solution. In a vaccine trial, two groups of individuals are randomised to receive a vaccine or placebo and disease occurrences are recorded in each group. As disease affects predominantly higher-risk individuals, the mean risk among those who remain unaffected decreases and disease incidence declines. In the vaccine group the same trend occurs at a slower pace (presuming that the vaccine protects to some degree). As a result, the two randomised groups become different over time with more highly susceptible individuals remaining in the vaccine group. The vaccine efficacy, described as $1 - RR$, where RR is the ratio of cases in vaccinated over control, therefore appears to wane [Halloran et al 1996; O'Hagan et al 2012]. This effect will be stronger in settings where transmission intensity is higher, inducing a trend of seemingly declining efficacy with disease burden [Gomes et al 2016]. These concepts are illustrated in Figure 4 by simulating a vaccine trial with heterogeneous and homogeneous models analogous to those utilised in Figures 1-3.

Selection on individual variation in disease susceptibility thus offers an explanation for vaccine efficacy trends that is entirely based on population level heterogeneity, in contrast with individual waning of vaccine-induced immunity [Olotu et al 2016: Bell et al 2022]. It is important to disentangle their roles, as both may occur concurrently in a trial and lead to different interpretations of the same data. To capture this in a timely manner requires multicentre trial designs with sites carefully chosen over a gradient of transmission intensities (e.g., optimally spaced along the incidence axis in Figure 4 C, F), and analyses performed by

fitting curves generated by models that incorporate individual variation. An alternative and more tightly controlled approach would be to use experimental designs in human infection challenge studies, where these are available [Darton et al 2015; Roestenberg et al 2018], to generate dose-response curves and apply similar models [Gomes et al 2014]. These approaches have been successfully applied to animal systems [Langwig et al 2017; Pessoa et al 2016; King et al 2018].

The essential purpose of suggesting these study designs (multicentre trials over a gradient of transmission intensities, or dose-response infection challenges) is to enable selection on individual infection risks to be remodelled (empirically and mathematically) along force of infection (selection) gradients, in such a way that variation and selection can be inferred from observed infection trends.

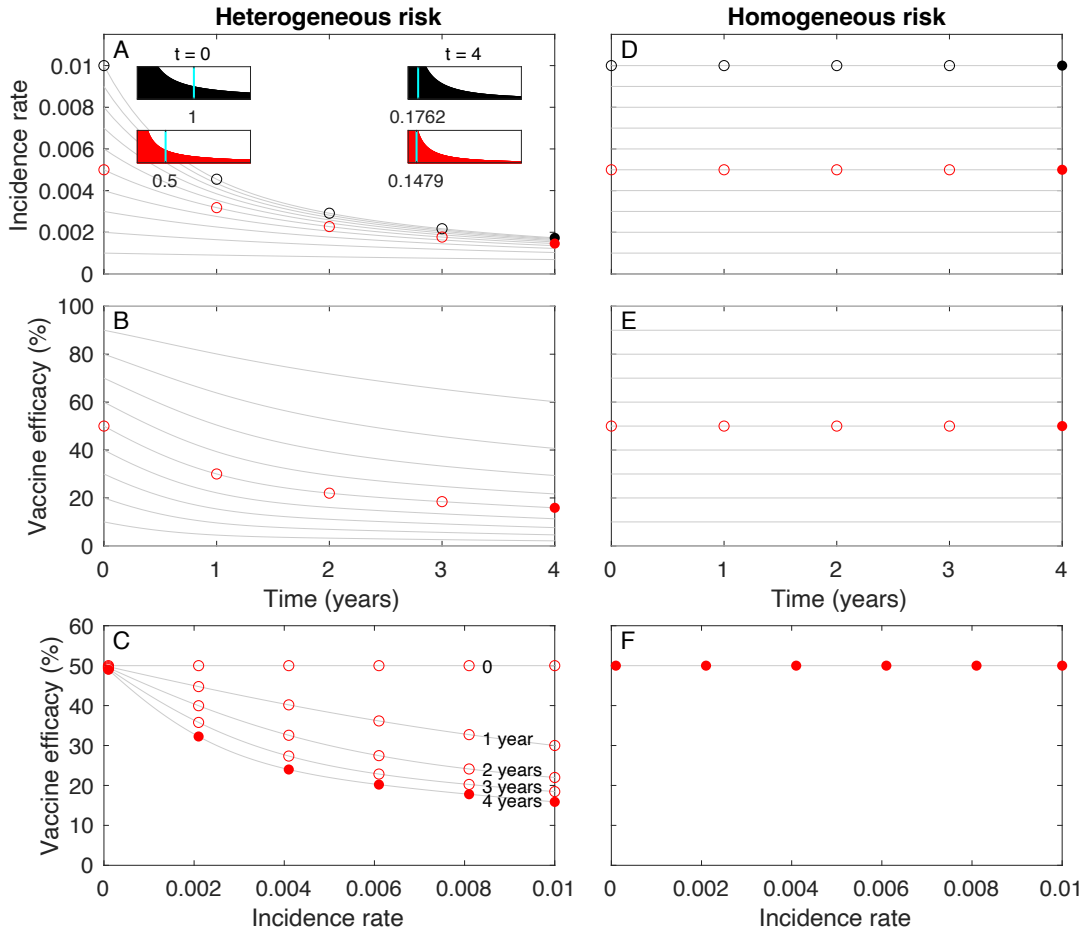

Figure 4: Vaccine efficacy trajectories under homogeneous and heterogeneous models. A, B, C, Heterogeneous susceptibility or exposure (with mean 1 and variance 10) with insets in A depicting susceptibility distributions in control and vaccine groups at the beginning and end of the trial (cyan line is the mean); D, E, F, Homogeneous model. Models: (homogeneous) $dS_c/dt = -\lambda S_c$, $dI_c/dt = \lambda S_c$, and $dS_v/dt = -\sigma\lambda S_v$, $dI_v/dt = \sigma\lambda S_v$; (heterogeneous) $dS_c(x)/dt = -\lambda x S_c(x)$, $dI_c(x)/dt = \lambda x S_c(x)$, and $dS_v(x)/dt = -\sigma\lambda x S_v(x)$, $dI_v(x)/dt = \sigma\lambda x S_v(x)$. Vaccine efficacy is calculated as $[1 - r_v(t)/r_c(t)] \times 100$, where r_v and r_c represent the incidences in vaccinated (v) and control (c) groups, respectively: (homogeneous) $r_c(t) = \lambda$; $r_v(t) = \sigma\lambda$; (heterogeneous) $r_c(t) = \lambda \int x S_c(x,t)dx / \int S_c(x,t)dx$; $r_v(t) = \sigma\lambda \int x S_v(x,t)dx / \int S_v(x,t)dx$. Gamma distributions were used in heterogeneous models for concreteness.

4. Inferring heterogeneities by remodelling selection

Heterogeneities in predisposition to infection depend on the mode of transmission. In respiratory infections, heterogeneity may arise from variation in exposure of the susceptible

host to the pathogen, or the competence of host immune systems to control it. These two processes have multiple component factors. Some of the most studied are age, patterns of inter-personal contacts, exposure to smoke, nutritional status, pre-existing respiratory illness such as asthma or chronic obstructive pulmonary disease, and the presence of other concomitant diseases such as diabetes and HIV. Enteric diseases have other heterogeneities determined by the source and dose of contaminated sources. Vector-borne pathogens may be transmitted by mosquitoes, ticks, snails, and other intermediate hosts, where the risk of onward transmission is affected by heterogeneities in exposure and susceptibility across a complex range of host, demographic, social, geographical, and environmental (including climatic) factors. For example, malaria endemicity is typically measured using the entomological inoculation rate (EIR), determined by multiplying the sporozoite rate (the proportion of mosquitoes that contain infectious sporozoites) by the host biting rate (average number of bites per person per unit time). Global (or even national) EIRs average over substantial individual variability in pathogen exposure and requirements for efficacious interventions [Smith et al 2005]. As for sexually transmitted diseases specific factors include behaviour, age, gender, and sexual orientation.

The mechanisms underpinning single factors for infection and their interactions determine individual propensities to acquire disease. These factors are potentially so numerous and interlinked that to attain a full mechanistic description is usually unfeasible. Even if lists of all putative factors were available, the measurement of effect sizes might be subject to selective depletion bias resulting in underestimated variances [Aalen et al 2015]. To contribute constructively to the development of health policies, model building involves compromises between leaving factors out (reductionism) or adopting a broader but coarse description (holism). Holistic descriptions of heterogeneity are uncommon in the study of disease dynamics.

The awareness that heterogeneities matter in infectious disease analyses has a long history since, already in the 1920s and 1930s, the pioneering work of Kermack and McKendrick [1927] and McKendrick [1939] circumvented the lack of explicit heterogeneity in early models by assuming that only a fraction of the population was accessible to infection in order to fit observed incidences. In 1968, Gart [1968] admitted that “it is difficult to define exactly the size of the population of susceptible hosts” due to the “heterogeneous nature of the population” and, in 1971, the same author formulated a model with several susceptibility groups [Gart 1971] which, in 1985, Ball [1985] compared to the homogeneous version and described how homogeneity assumptions increase the size of epidemics. In 2001, Pastor-Satorras and Vespignani [2001] developed related formalisms to describe epidemics on contact networks. Unfortunately, despite the long-standing recognition that heterogeneity is required for models to fit data and the availability of adequate mathematical models for the effect, there is a widespread belief that unobserved heterogeneity cannot be estimated.

However, unmeasured heterogeneities that respond to selection, can be built into dynamic models and estimated by fitting model outputs to population data, in a similar vein to the 2000 Nobel Memorial Prize in Economic Sciences winning work conducted by James Heckmann (see [Heckmann 1979]). Dynamic models describing state transitions in an infectious or non-communicable disease (or behavioural phenomena in the social sciences) become motors of selection on the inbuilt heterogeneity. It is then the interplay between selection and the baseline heterogeneity that affect model outputs. Hence, taking population measurements along a selection (such as exposure to a hazard) gradient and fitting a model-generated curve to the resulting data can enable the inference of baseline distributions in a holistic manner. While this procedure is established in microbial risk assessment [Haas 1999] and survival or event history analysis [Hougaard 1986; Aalen 2008 and references therein],

its application in the modelling of disease dynamics has been less widespread [Smith et al 2005; Bellan et al 2015; Gomes et al 2019; Corder et al 2020; Dwyer et al 1997; Gomes et al 2022; Stensrud and Valberg 2017]. The intent of this review is to convey the generality of the approach, and its feasibility and importance for model predictability. We introduce the term *remodelling selection* to refer to the body of theory and methods unified across disciplines whereby variation and selection are essentially remodelled, mathematically and empirically, in a way that enables their statistical inference (e.g., [Furumoto et al 1967; Heckmann 1979; Dwyer et al 1997; Hougaard 1986; Haas 1999; Smith et al 2005; Ben-Ami et al 2008; Zwart et al 2011; Gomes et al 2014; Pessoa et al 2016; Stensrud and Valberg 2017; Langwig et al 2017; King et al 2018; Gomes et al 2019; Corder et al 2020; Gomes et al 2022]).

In the case of infectious diseases, selection is exerted primarily by the infectious agent, so the analyst will be fitting model-generated curves to a collection of incidence measurements taken in multiple conditions spanning a range of exposure intensities. When controlled infection experiments can be performed [Darton et 2015; Roestenberg et al 2018], dose-response designs should be adopted. Intuitively, the lowest challenge doses infect mostly highly susceptible individuals while as dose increases more of the less susceptible are also infected. Therefore, dose-response curves are closely related to cumulative distributions of susceptibility, which can be inferred by fitting appropriate models [Furumoto et al 1967; Haas 1999; Ben-Ami et al 2008; Zwart et al 2011; Gomes et al 2014; Pessoa et al 2016; Langwig et al 2017; King et al 2018]. When infection is by natural exposure a similar tactic can be devised. Incidence measurements should be collected from multiple settings, ideally spanning a wide range of exposure intensities. Model-generated curves will then be fitted to the entire dataset, conditioned on individual variation being similar across settings (unless additional prior information is available) [Smith et al 2015; Gomes et al 2019; Gomes et al 2022]. When disease episodes are so frequent that individuals can be characterised by how many occurrences they experienced over a feasible study period, such as with seasonal respiratory viruses or malaria in endemic regions, then heterogeneity may be inferable from a single setting [Corder et al 2020]. In non-communicable diseases, such as cancer, it may be feasible to consider predisposing genes or household characteristics as disease agents, and hence exposure intensities can be structured by familial relatedness [Aalen et al 2015; Stensrud and Valberg 2017]. The commonality is to employ models that have individual variation represented explicitly to enable response to changes in exposure (selection) intensity (Figures 2, 3) should these occur naturally or through interventions. Free from the selection biases exposed in this review, this modelling approach will automatically enable more accurate forecasts to inform policies.

5. Conclusion

There is compelling evidence for the utility of holistic descriptions of individual variation in disease risk, admitting that heterogeneity is so vast in real-world systems that complete mechanistic reconstructions may be currently unachievable. Inspired by other population disciplines and supported by successful applications in both infectious and non-communicable diseases, we describe methods of study design and analyses that enable inferences of heterogeneity by estimating how much selection occurs as susceptible populations are depleted through infection and/or disease. These methods rely on *remodelling selection* along gradients which may result naturally from trends of exposure to a hazard across population strata, in the case of observational studies, or be created by design, in the case of controlled experiments. We advocate for the wide adoption of these approaches in epidemiology to enable accurate disease forecast models.

Contributions

MGMG conceived the idea and drafted the article. All authors contributed to the final writing of this article.

Declaration of interests

We declare no competing interests.

Acknowledgements

This paper benefited from supportive discussions with numerous colleagues, especially Mauricio Barreto, Maxine Caws, Andrea Doeschl-Wilson, Nicholas Feasey, Marcelo Ferreira, Philippe Glaziou, Stephen Gordon, Jessica King, James LaCourse, Christian Lienhardt, Penelope Phillips-Howard, Lisa Reimer, Meta Roestenberg, Jamie Rylance, Bertel Squire, Russell Stothard, Miriam Taegtmeyer, Dianne Terlouw, Rachel Tolhurst, Tom Wingfield. MGMG has received funding from the Innovative Medicines Initiative 2 Joint Undertaking under grant agreement No 101007799 (Inno4Vac). This Joint Undertaking receives support from the European Union's Horizon 2020 research and innovation programme and EFPIA.

References

1. Ross R. 1916 An application of the theory of probabilities to the study of *a priori* pathometry, Part I. *Philos. Trans. R. Soc. Lond. A* **92**, 204–230.
2. Ross R, Hudson,HH. 1917 An application of the theory of probabilities to the study of *a priori* pathometry, Part II. *Philos. Trans. R. Soc. Lond. A* **93**, 212–225.
3. Kermack WO, McKendrick AG. 1927 A contribution to the mathematical theory of epidemics. *Proc. R. Soc. Lond. A* **115**, 700-721.
4. Heesterbeck H *et al.* 2015 Modeling infectious disease dynamics in the complex landscape of global health. *Science* **347**, aaa4339.
5. Gaolathe T *et al.* 2016 Botswana's progress toward achieving the 2020 UNAIDS 90-90-90 antiretroviral therapy and virological suppression goals: a population-based survey. Lancet HIV 3, e221-e230.
6. Karim SA. 2016 Is the UNAIDS target sufficient for HIV control in Botswana? *Lancet HIV* **3**, e195-6.
7. UNAIDS. 2017 Joint United Nations Programme on HIV/AIDS. Ending AIDS: Progress towards the 90-90-90 targets.
8. Frescura L, Godfrey-Faussett P, Feizzadeh AA, El-Sadr W, Syarif O, Ghys PD, on behalf of the 2025 testing treatment target Working Group. 2022 Achieving the 95-95-95 targets for all: A pathway to ending AIDS. *PLOS One* **17**, e0272405.
9. Specht S, Suma TK, Pedrique B, Hoerauf A. 2019 Elimination of lymphatic filariasis in South East Asia. *Br. Med. J.* **364**, k5198.
10. Flaxman S *et al.* 2020 Estimating the effects of non-pharmaceutical interventions on COVID-19 in Europe. *Nature* **584**, 257-261.
11. Gomes MGM *et al.* 2022 Individual variation in susceptibility or exposure to SARS-CoV-2 lowers the herd immunity threshold. *J. Theor. Biol.* **540**, 111063.
12. Greenwood M, Yule GU. 1920 An inquiry into the nature of frequency distributions representative of multiple happenings with particular reference to the occurrence of multiple attacks of disease or of repeated accidents. *J. R. Stat. Soc.* **83**, 255-279.
13. Wald A. 1943 *A method of estimating plane vulnerability based on damage of survivors*. Statistical Research Group, Columbia University. CRC 432. Reprinted in July 1980,

Center for Naval Analyses, Alexandria, VA, Operations Evaluation Group: https://apps.dtic.mil/sti/citations/ADA091073.

14. Vaupel JW, Manton KG, Stallard E. 1979 Impact of heterogeneity in individual frailty on the dynamics of mortality. *Demography* **16**, 439-454.
15. Haldane JBS. 1954 The measurement of natural selection. *Proc. IX Intl. Cong. Genet.* **1**, 480-487.
16. Lande R, Arnold SJ. 1983 The measurement of selection on correlated characters. *Evolution* **37**, 1210-1226.
17. Forslund P, Pärt T. 1995 Age and reproduction in birds – hypotheses and tests. *Tends Ecol. Evol.* **10**, 374-378.
18. Van De Pol M, Verhulst S. 2006 Age-dependent traits: a new statistical model to separate within- and between-individual effects. *Am. Nat.* **167**, 766– 773.
19. Keyfitz N, Littman G. 1979 Mortality in a heterogeneous population. *Popul. Stud.* **33**, 333-342.
20. Vaupel J, Yashin A. 1985 Heterogeneity ruses – some surprising effects of selection on population dynamics. *Am. Stat.* **39**, 176-185.
21. Strandberg TE, Stenholm S, Strandberg AY, Salomaa VV, Pitkälä KH, Tilvis RS. 2013 The "Obesity Paradox", frailty, disability, and mortality in older men: A prospective, longitudinal cohort study. *Am. J. Epidemiol.* **178**, 1452-1460.
22. Nussey DH, Kruuk LEB, Donald A, Fowlie M, Clutton-Brock TH. 2006 The rate of senescence in maternal performance increases with early-life fecundity in red deer. *Ecol. Lett.* **9**, 1342-1350.
23. Kendall BE, Fox GA. 2002 Variation among individuals and reduced demographic stochasticity. *Conserv. Biol.* **16**, 109-116.
24. Jenouvrier S, Aubry LM, Barbraud C, Weimerskirch H, Caswell H. 2018 Interacting effects of unobserved heterogeneity and individual stochasticity in the life history of the southern fulmar. *J. Anim. Ecol.* **87**, 212-222.
25. Steiner UK, Tuljapurkar S. 2012 Neutral theory for life histories and individual variability in fitness components. *Proc. Natl. Acad. Sci U. S. A.* **109**, 4684-4689.
26. Gomes MGM, King JG, Nunes A, Colegrave N, Hoffmann A. 2019a The effects of individual nonheritable variation on fitness estimation and coexistence. *Ecol. Evol.* **16**, 8995-9004.
27. Aalen OO, Valberg M, Grotmol T, Tretli S. 2015 Understanding variation in disease risk: the elusive concept of frailty. *Int. J. Epidemiol.* **4**:1408-1421.
28. Stensrud MJ, Valberg M. 2017 Inequality in genetic cancer risk suggests bad genes rather than bad luck. *Nat. Commun.* **8**: 1165.
29. Anderson RM, Medley GF, May RM, Johnson AM. 1986 A preliminary study of the transmission dynamics of the human immunodeficiency virus (HIV), the causative agent of AIDS. *IMA J. Math. Appl. Med. Biol.* **3**, 229-263.
30. Colgate SA, Stanley AE, Hyman JM, Layne SP, Qualls C. 1988 Risk behavior-based model of the cubic growth of acquired immunodeficiency syndrome in the United States. *Proc. Natl. Acad. Sci. U. S. A.* **86**, 4793-4797.
31. Dwyer G, Elkinton JS, Buonaccorsi J P. 1997 Host heterogeneity in susceptibility and disease dynamics: Tests of a mathematical model. *Am. Nat.* **150**, 685-707.
32. Smith DL, Dushoff J, Snow RW, Hay SI. 2005 The entomological inoculation rate and *Plasmodium falciparum* infection in African children. *Nature* **438**, 492-495.

33. Bellan SE, Dushoff J, Galvani AP, Meyers LA. 2015 Reassessment of HIV-1 acute phase infectivity: accounting for heterogeneity and study design with simulated cohorts. *PLOS Med* **12**, e1001801.
34. Gomes MGM *et al.* 2019b Introducing risk inequality metrics in tuberculosis policy development. *Nat. Commun.* **10**, 2480.
35. Corder RM, Ferreira MU, Gomes MGM. 2020 Modelling the epidemiology of residual Plasmodium vivax malaria in a heterogeneous host population: a case study in the Amazon Basin. *PLOS Comput. Biol.* **16**, e1007377.
36. Britton T, Ball F, Trapman P. 2020 A mathematical model reveals the influence of population heterogeneity on herd immunity to SARS-CoV-2. *Science* **369**, 846-849.
37. Halloran ME, Longini IM Jr., Struchiner CJ. 1996 Estimability and interpretability of vaccine efficacy using frailty mixing models. *Am. J. Epidemiol.* **144**, 83-97.
38. O'Hagan JJ, Hernán MA, Walensky RP, Lipsitch M. 2012 Apparent declining efficacy in randomized trials: Examples of the Thai RV144 HIV vaccine and CAPRISA 004 microbicide trials. *AIDS* **26**, 123-126.
39. Gomes MGM, Lipsitch M, Wargo AR, Kurath G, Rebelo C, Medley GM, Coutinho A. 2014 A missing dimension in measures of vaccination impacts. *PLOS Pathog.* **10**, e1003849.
40. Gomes MGM, Gordon SB, Lalloo DG. 2016 Clinical trials: the mathematics of falling vaccine efficacy with rising disease incidence. *Vaccine* **34**, 3007-3009.
41. Langwig KE *et al.* 2017 Vaccine effects on heterogeneity in susceptibility and implications for population health management. *mBio* **8**, e00796-17.
42. Ben-Ami F, Regoes RR, Ebert D (2008) A quantitative test of the relationship between parasite dose and infection probability across different host-parasite combinations. Proc R Soc B 275: 853–859.
43. Zwart MP, Hemerik L, Cory JS, de Visser JAGM, Bianchi FJJA, et al. (2011) An experimental text of the independent action hypothesis in virus-insect pathosystems. Proc R Soc B 276: 2233–2242.
44. Pessoa D, Souto-Maior C, Gjini E, Lopes JS, Ceña B, Codeço CT, Gomes MGM. 2016 Unveiling time in dose-response models to infer host susceptibility to pathogens. *PLOS Comput. Biol.* **10**, e1003773.
45. King JG, Souto-Maior C, Sartori LM, Maciel-de-Freitas R, Gomes MGM. 2018 Variation in Wolbachia effects on Aedes mosquitoes as a determinant of invasiveness and vectorial capacity. *Nat. Commun.* **9**, 1-8.
46. Metcalf CJE, Pacard S. 2007 Why evolutionary biologists should be demographers. *Tends Ecol. Evol.* **22**, 205-212.
47. Woolhouse MEJ *et al.* 1997 Heterogeneities in the transmission of infectious agents: implications for the design of controls programs. *Proc. Natl. Acad. Sci. U. S. A.* **94**, 338-142.
48. Keeling MJ, Eames KTD. 2005 Networks and epidemic models. *J. R. Soc. Interface* **2**, 295-307.
49. Leigh Brown AJ, Lycett SJ, Weinert L, Hughes GJ, Fearnhill E, Dunn DT. 2011 Transmission network parameters estimated from HIV sequences for a nationwide epidemic. *J. Infect. Dis.* **204**, 1463-1469.

50. Granich RM, Gilks CF, Dye C, De Cock KM, Williams BG. 2009 Universal voluntary HIV testing with immediate antiretroviral therapy as a target for elimination of HIV transmission: a mathematical model. *Lancet* **373**, 48-57.
51. UNAIDS. 2023 Joint United Nations Programme on HIV/AIDS. The path that ends AIDS.
52. Davies NG, Kucharski AJ, Eggo RM, Gimma A, Edmunds WJ, Centre for the Mathematical Modelling of Infectious Diseases COVID-19 working group. 2020 Effects of non-pharmaceutical interventions on COVID-19 cases, deaths, and demand for hospital services in the UK: A modelling study. *Lancet Public Health* **5**, e375–e385.
53. Neipel J, Bauermann J, Bo S, Harmon T, Jülicher F. 2020 Power-Law population heterogeneity governs epidemic waves. *PLOS One* **15**, e0239678.
54. Rose C, Medford AJ, Goldsmith CF, Vegge T, Weitz JS, Peterson AA. 2021 Heterogeneity in susceptibility dictates the order of epidemic models. *J. Theor. Biol.* **528**, 110839.
55. Tkachenko AV, Maslov S, Elbanna A, Wong GN, Weiner ZJ, Goldenfeld N. 2021 Time-dependent heterogeneity leads to transiente suppression of the COVID-19 epidemic, not herd immunity. *Proc. Natl. Acad. Sci. U. S. A.* **118**, e2015972118.
56. Montalbán A, Corder RM, Gomes MGM. 2020 Herd immunity under individual variation and reinfection. *J. Math. Biol.* **85**, 2.
57. Millett GA *et al.* 2020 Assessing differential impacts of COVID-19 on black communities. *Ann. Epidemiol.* **47**, 37-44.
58. Mena G, Martimez PP, Mahmud AS, Marquet PA, Buckee CO, Santillana M, 2021 Socioeconomic status determines COVID-19 incidence and related mortality in Santiago, Chile. *Science* **372**, abg5298.
59. Xia Y *et al.* 2022 Geographic concentration of SARS-CoV-2 cases by social determinants of health in metropolitan areas in Canada: a cross-sectional study. *CMAJ* **194**, E195–E204.
60. Braveman P, Gottlieb L. 2014 The social determinants of health: It's time to consider the causes of the causes. *Public Health Rep.* **129**, 19-31.
61. Okonkwo NE *et al.* 2021 COVID-19 and the US response: accelerating health inequalities. *BMJ Evid. Based Med.* **26**, 176-179.
62. Moore S, Hill EM, Tildesley MJ, Dyson L, Keeling MJ. 2021 Vaccination and non-pharmaceutical interventions for COVID-19: a mathematical modelling study. *Lancet Infect. Dis.* **21**, 793-802.
63. Hilton J *et al.* 2022 A computational framework for modelling infectious disease policy based on age and household structure with applications to the COVID-19 pandemic. *PLOS Comput. Biol.* **18**, e1010390.
64. Kerr CC *et al.* 2021 Covarism: An agent-based model of COVID-19 dynamics and interventions. *PLOS Comput. Biol.* **17**, e1009149.
65. Olotu A *et al.* 2016 Seven-year efficacy of RTS,S/AS01 malaria vaccine among young African children. *N. Engl. J. Med.* **374**, 2519-2529.
66. Bell GJ *et al.* 2022 Malaria transmission intensity likely modifies RTS,S/AS01 efficacy due to a rebound effect in Ghana, Malawi, and Gabon. *J. Infect. Dis.* **226**, 1646-1656.
67. Darton TC *et al.* 2015 Design, recruitment, and microbiological considerations in human challenge studies. *Lancet Infect. Dis.* **15**, 840-851.

68. Roestenberg M, *et al.* 2018 Experimental infection of human volunteers. *Lancet Infect. Dis.* **18**, E312-E322.
69. McKendrick AG. 1939 The dynamics of crowd infection. *Edinb. Med. J.* **47**, 117-136.
70. Gart JJ. 1968 The mathematical analysis of an epidemic with two kinds of susceptibles. *Biometrics* **24**, 557-566.
71. Gart JJ. 1971 The statistical analysis of chain-binomial epidemic models with several kinds of susceptibles. *Biometrics* **28**, 921-930.
72. Ball F. 1985 Deterministic and stochastic epidemic models with several kinds of susceptibles. *Adv. Appl. Probab.* **17**, 1-22.
73. Pastor-Satorras R, Vespignani A. 2001 Epidemic dynamics and endemic states in complex networks. *Phys. Rev. E* **63**, 066117.
74. Heckman JJ. 1979 Sample selection bias as a specification error. *Econometrica* **47**, 153-61.
75. Haas CN, *et al.* 1999 Quantitative Microbial Risk Assessment. John Wiley & Sons, Inc, New York.
76. Hougaard R. 1986 Survival models for heterogeneous populations derived from stable distributions. *Biometrika* **73**, 387-96.
77. Aalen OO, Borgan O, Gjessing HK. 2008 *Survival and Event History Analysis: a process point of view*. Statistics for Biology and Health. Springer, New York.