

Epidemics from the population perspective*

Jonathan Fuller¹

*forthcoming in *Philosophy of Science*

¹ Department of History and Philosophy of Science, University of Pittsburgh, Pittsburgh, PA,
USA

Email: JPF53@pitt.edu

Abstract

Many epidemics consist in individuals spreading infection to others. From the population perspective, they also have population characteristics important in modeling, explaining and intervening in epidemics. I analyze epidemiology's contemporary population perspective through the example of epidemics by examining two central principles attributed to Geoffrey Rose: a distinction between the causes of cases and the causes of incidence, and between 'high-risk' and 'population' strategies of prevention. Both principles require revision or clarification to capture the sense in which they describe distinct perspectives on the same phenomenon (such as an epidemic), each perspective capturing a different level of contrastive analysis.

Acknowledgments:

Thanks to Colin Allen, Alex Broadbent, Michael Dietrich, Marina DiMarco, Sandro Galea, Stephen John, Neil Pearce, Jim Woodward and anonymous reviewers, as well as audiences at the University of Pittsburgh and Aarhus University for helpful feedback and discussion.

1. Introduction

Epidemics – widespread outbreaks of infectious disease – have visited human societies throughout their history. As long as they have occurred, people have sought to make sense of them. Such understandings often adopted a *population perspective*, understanding epidemics as afflictions of civilizations due to factors affecting the population as a whole, be it cosmos or climate (Rosenberg 1992). Epidemics can seem to have a collective life of their own, above and beyond the individual cases of illness of which they are constituted – a pathology of the population.

Epidemiology as a discipline rose in the nineteenth and early twentieth centuries largely to study epidemics with new numerical methods. Epidemiology is typically defined as a science of populations (Krieger 2011), thus reinforcing the idea that epidemics – historically, the epidemiologist’s bread and butter – are best viewed through a population lens. However, from the second half of the nineteenth century, the population perspective on epidemics has been in tension with the now dominant understanding of infectious diseases as illnesses caused by germs (Rosenberg 1992), small pathogens infecting individual bodies. Moreover, according to some epidemiologists (e.g. Pearce 1996; Krieger 2011), since the middle of the twentieth century the epidemiological gaze generally has become fixed on individuals and the causes of their disease, especially with the growth of noncommunicable disease epidemiology. Over the past several decades, dissatisfaction with this conception of epidemiology has spurred a revival of interest in the population perspective, driven in part by population health scientists who study health and disease within and between populations (Keyes and Galea 2016; Valles 2018).

Some of the conceptual groundwork for the revived population perspective was laid by the epidemiologist Geoffrey Rose (1985, 1992). In the present paper, I will examine two of Rose's influential principles. The first principle is that the causes of individual cases of disease are different than the causes of the incidence of disease in a population. The second is that a 'population strategy' in which one intervenes at a population level is often more effective in prevention than a 'high-risk strategy' in which one intervenes to prevent disease in highly susceptible individuals. These two ideas are central to Rose's population perspective. I will analyze these claims as applied to an epidemic, using the example of the coronavirus disease 2019 (COVID-19) pandemic. Part of the difficulty in making sense of Rose's principles is that a population is composed of individuals and an epidemic is made up of infections. Therefore, in what sense (if any) are the causes of infectious disease distinct from the causes of the incidence of infectious disease in an epidemic? And can we sharply distinguish between a strategy of intervening in an epidemic and one of intervening in instances of infectious disease?

Rose's principles require clarification or revision. In section 2, I will introduce Rose's population perspective, focusing on his two central principles. In section 3, I will provide some scientific background on epidemics. Then in section 4, I will argue that Rose's distinction between the causes of cases and the causes of incidence should be understood in terms of different levels of contrastive causal explanation; and in section 5, I will argue that the population strategy should be given a corresponding contrastive interpretation. As its name would suggest, the population perspective offers a particular perspective on epidemics, viewing the etiology and prevention of infectious disease through a coarse-grained lens focused on population characteristics.

2. The population perspective

Epidemiology is often defined as the study of the distribution and determinants of health and disease in *populations* (Krieger 2011). It may seem odd to suggest that epidemiological research could ever lack a population perspective. However, over the past few decades, several epidemiologists have articulated concerns about an individualistic approach in mainstream epidemiology and have advocated for a more population-oriented approach.

For example, Neil Pearce (1996) argues that in contrast to the traditional approach to epidemiology of the nineteenth century that investigated the social and environmental conditions leading to poor health in populations, modern epidemiology has embraced a “reductionist approach [that] focuses on the individual” (678) and he implores epidemiologists to “rediscover the population perspective” (682) lost in recent decades by investigating the causes of disease in populations. Similarly, Nancy Krieger (2011) argues that the dominant individualistic approach to epidemiological theorizing since the mid-twentieth century, the ‘biomedical model’, “views populations simply as the sum of individuals and population patterns of disease as simply reflective of individual cases” (137).

In recent decades, the field of ‘population health science’ has arisen to reorient public health by studying patterns of health within and between populations (Valles 2018). In their textbook *Population Health Science*, Katherine Keyes and Sandro Galea write: “Population health science is the study of the conditions that shape distributions of health within and across populations, and of the mechanisms through which these conditions manifest in the health of individuals” (2016, 1). Keyes and Galea distinguish conditions affecting health at the *population* level from the mechanisms in which these conditions manifest in *individuals*.

Much of the conceptual grounding for the resurgent population perspective in epidemiology was set down by Rose. In a 1985 paper titled “Sick Individuals and Sick Populations”, Rose proposed several radical ideas, later extended in his 1992 book *The Strategy of Preventive Medicine*. Rather than examining the population perspective directly, I will focus on two of Rose’s principles, which are central to his population perspective. The foundations of the population perspective have been further developed by Rose’s contemporaries and successors (Keyes and Galea (2016) consider Rose to be the founder of population health science). I will introduce Rose’s two ideas briefly here. In sections 4 and 5, I will analyze them by applying them to the epidemiology of epidemics.

Rose’s first principle is that the causes of individual cases of disease differ from the causes of the incidence of disease (the number of new cases per unit time¹) or the causes of the mean value of a risk factor in a population. Rose: “I find it increasingly helpful to distinguish two kinds of aetiological question. The first seeks the causes of cases, and the second seeks the causes of incidence” (2001, 428). Put differently, Rose claims that the causes of differences in cases within populations are typically distinct from the causes of differences in incidence between populations. For instance, he suggests that variation in genetic causes tends to be much greater among individuals within a population, while variation in environmental causes tends to be much greater between two populations. In his Intersalt study of 32 countries (Intersalt Cooperative Research Group 1988), Rose found that the curve representing the distribution of blood pressure values in a population had a similar shape across countries but varied in its

¹ I will use this definition of the incidence throughout the paper. A related measure is the ‘incidence proportion’ or ‘attack rate’, which expresses the incidence as a proportion of the total population at the start of the time interval (CDC 2012).

placement along the horizontal blood pressure axis, and thus the mean blood pressure varied too. Rose (2008) suggested that these horizontal shifts in population distribution “manifestly reflect characteristics of populations and not characteristics of individuals” (2008, 92).

Rose’s second principle is that a ‘population strategy’ in which one intervenes at a population level is often more effective in prevention than a ‘high-risk strategy’ in which one intervenes to prevent disease in highly susceptible (high-risk) individuals. The two approaches map onto Rose’s distinction between causes of cases versus causes of incidence: “These two approaches to aetiology—the individual and the population-based—have their counterparts in prevention. In the first, preventive strategy seeks to identify high-risk susceptible individuals and to offer them some individual protection. In contrast, the ‘population strategy’ seeks to control the determinants of incidence in the population as a whole” (2001, 429). Thus, the high-risk strategy controls causes in individuals (e.g. cholesterol or blood pressure) and truncates the risk distribution, changing its shape. Meanwhile, the population strategy controls determinants of population health such as environmental or societal factors and shifts the entire population distribution, changing its incidence or mean without necessarily changing the shape of the curve. In summary, there are two conceptual components of Rose’s population perspective: a population perspective on etiology and a population perspective on intervention. Rose fleshes out each component by way of a comparison with a more individualistic perspective on etiology and a more individualistic perspective on intervention, respectively. I will argue that Rose’s population perspectives on etiology and intervention go hand-in-hand.

A major philosophical difficulty with each of Rose’s claims is that each sharply distinguishes populations (and their causes and prevention) from individuals (and *their* causes and prevention). Yet populations are composed of individuals; they are even plausibly

ontologically reducible to a collection of interacting individuals. At least that seems to be the case with epidemiologic populations as epidemiologists and population health scientists usually conceive of them. For instance, Keyes and Galea write, “Typically, we think of populations as collections of people or other organisms that share common characteristics, most often a specific location they inhabit” (2016, 4). Thus, without clarification or revision, Rose’s population perspective risks being incoherent or saddling epidemiology with excess metaphysical baggage.

As with most of the preventive medicine literature, the main diseases Rose had in mind were noncommunicable. However, the importance of the population perspective on epidemics has not been completely overlooked within epidemiology. Khaw and Marmot list herd immunity, an important idea in the science of epidemics, among three examples that “requires us to go beyond the level of the individual” (2008, 18). Pearce claims that “analysis at the individual level cannot explain epidemic spread at the group level and cannot even fully explain the spread of infections between individuals” (1996, 682). And Galea and Keyes (2020) examine the COVID-19 pandemic through a population health science perspective.

I will analyze the population perspective in sections 4 and 5 through the example of infectious disease epidemics. I will do so indirectly by examining Rose’s distinctions between the causes of cases and causes of incidence (section 4), and the high-risk and population strategies for prevention (section 5). First, I will introduce some of the contemporary science of epidemics.

3. Contemporary epidemiology of epidemics

An epidemic is an increase within a large geographical area in the number of cases of a disease above the normal baseline. When an epidemic spreads to several countries and affects a great number of people, it becomes a pandemic (CDC 2012). Although ‘epidemic’ and ‘pandemic’ are sometimes used to describe a noncommunicable disease or health state like obesity, here I will use these terms to refer only to infectious diseases. In these cases, epidemics are due to infectious spread of a pathogen, usually a virus, bacterium or parasite. Epidemics may be common-source or propagated (CDC 2012). In a common source epidemic, infections result from exposure to the same source, such as a contaminated food or water supply. In a propagated epidemic, infections spread from person to person. Once a human population has been seeded with the initial infections, a propagated epidemic starts to grow as the pathogen spreads to a greater number of people with each successive cycle or ‘generation’ of infections.

Many health sciences (and social sciences) are relevant to the scientific understanding of epidemics, including microbiology, immunology and infectious disease epidemiology. Epidemiologists describe, explain and predict the spread of an epidemic and predict and evaluate the effect of interventions to counteract the epidemic. These tasks are greatly aided by the use of epidemic models, a family of mathematical models that model the dynamics of an epidemic. Two kinds of epidemic models are often used: microsimulation models and compartment models (Vynnycky & White 2010). Microsimulation or individual-based models are computer simulation models of an epidemic in a large population (with hundreds, thousands or even millions of individuals); they model the contacts and health states of each member of the population individually.

In contrast, much of the focus of discussion in the rest of this paper will involve compartment models: older, equation-based models that track changes in the size of different

subgroups or ‘compartments’ of the population as an epidemic evolves. Compartment models can be fully deterministic, or they can have stochastic elements. In an SIR model, the population is partitioned into three compartments: one for individuals who are susceptible to infection at time t , one for individuals who are infectious to others at t , and one for individuals who are recovered from and/or are immune from infection (or who have died) at t . There are many variations on this model (Vynnycky & White 2010).

We will consider a variation in which newly infected individuals transition through a pre-infectious or ‘exposed’ compartment E before they become infectious: a SEIR model. Figure 1 provides a schematic of the flow of individuals through the compartments of a SEIR model.

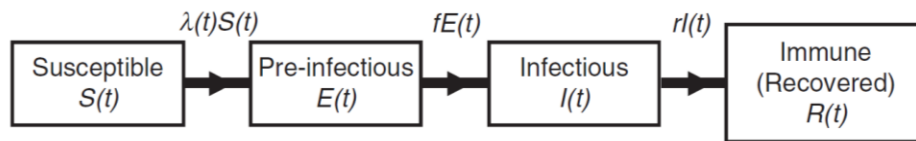


Figure 1. Structure of a simple SEIR model. Assuming random or homogeneous mixing of all members of the population, $\lambda(t) = \beta I(t)$. From Vynnycky & White (2010), *An Introduction to Infectious Disease Modelling*. © Oxford University Press. Reproduced with permission of the Licensor through PLSclear.

Compartment models usually take the form of a system of ordinary differential equations (though difference equations are sometimes used instead). A simple deterministic SEIR model of a closed population with a negligible number of births and deaths is composed of four equations, as follows.

$$\frac{dS(t)}{dt} = -\beta I(t)S(t) \quad 1$$

$$\frac{dE(t)}{dt} = \beta I(t)S(t) - fE(t) \quad 2$$

$$\frac{dI(t)}{dt} = fE(t) - rI(t) \quad 3$$

$$\frac{dR(t)}{dt} = rI(t) \quad 4$$

More complicated models are produced by building in stochasticity or stratifying the population into segments with unique parameters. $S(t)$, $E(t)$, $I(t)$ and $R(t)$ are the number of individuals in the susceptible, pre-infectious, infectious and recovered compartments, respectively. β is the effective contact rate, the rate at which an infectious individual contacts a susceptible individual and transmits infection to them. f is the rate at which a pre-infectious individual becomes infectious. Finally, r is the rate at which an infectious individual recovers. Neither β , f , nor r is a constant; they can vary depending on the pathogen and the population. For example, β will tend to be greater in crowded areas and for diseases in which a greater number of individual pathogens are shed by each infectious person; and r will tend to be smaller in populations that lack curative treatment and for diseases that have a long natural duration of infectiousness.

An important parameter in describing epidemics is the reproduction number. The basic reproduction number R_0 is the average number of secondary infectious individuals resulting from an initial infectious individual in an otherwise completely susceptible population. It can be defined by the equation $R_0 = (\beta N)/r$, where N is the total population size. After the initial

infection has spread and the susceptible compartment has begun to be depleted, the effective or net reproduction number R_n describes the average number of secondary infections resulting from an infectious individual as the epidemic evolves. It is calculated using the equation $R_n = R_0 s$, where s is the proportion of the population that is susceptible. When R_n is greater than 1, the number of new infectious individuals is growing as each infectious individual transmits infection to more than one susceptible individual on average. When R_n is equal to 1, the number of new infectious individuals is unchanging. When R_n is less than 1, the number of new infectious individuals is shrinking. Simple deterministic SIR or SEIR models generate a set of epidemic curves with a classic shape (Figure 2).

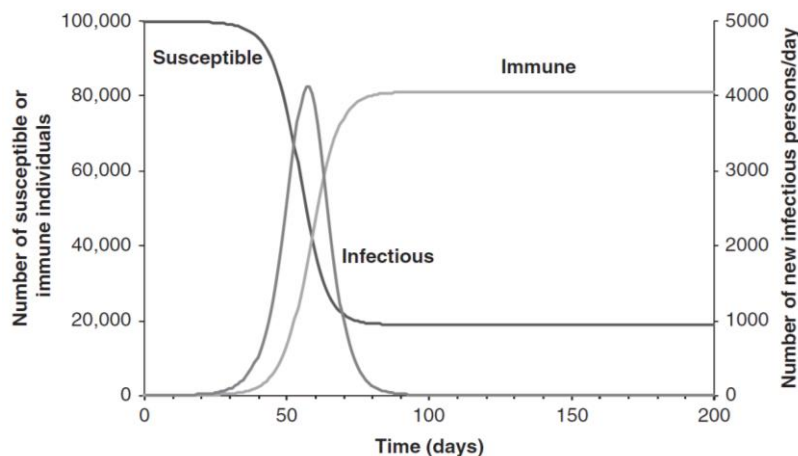


Figure 2. Epidemic curves predicted by a deterministic SEIR model with the following parameters. Total population N (all susceptible at start) = 100,000, $\beta = 0.00001$, $f = 0.5$, $r = 0.5$, $R_0 = 2$. From Vynnycky & White (2010), *An Introduction to Infectious Disease Modelling*. © Oxford University Press. Reproduced with permission of the Licensor through PLSclear.

In examining the population perspective on epidemics, I will use as my running example the COVID-19 pandemic. The pandemic started in December 2019 in Wuhan, China. The

responsible pathogen is severe acute respiratory disease coronavirus 2 (SARS-CoV-2), a novel human coronavirus (Andersen et al. 2020). SARS-CoV-2 is transmitted primarily through respiratory droplets or aerosols that are produced by coughing, sneezing, singing or talking. At the time of writing, the CDC's best estimate for SARS-CoV-2's R_0 is 2.5 (CDC 2020). Many SARS-CoV-2 infections in humans are asymptomatic, but others result in the disease COVID-19, which has a wide spectrum of severity from mild respiratory symptoms to severe pneumonia requiring respiratory support, including mechanical ventilation. From Wuhan in December 2019, the epidemic spread to the rest of China and worldwide. On 11 March 2020, the WHO declared a pandemic.

Many epidemic models have been used to project the growth of the pandemic and the potential for non-pharmaceutical strategies to mitigate it. In March 2020, modelers at the MRC Centre for Global Infectious Disease Analysis at Imperial College London used a microsimulation model to project the evolution of the epidemics in the U.S. and the U.K. and the influence of strategies to reduce person-to-person contact (Ferguson et al. 2020). That same month, they used a SEIR model to extend their analysis to 200 other countries (Walker et al. 2020). The latter model projected 7.0 billion infections and 40 million deaths worldwide in an unmitigated pandemic, but that an early transmission suppression strategy could save 38.7 million lives. It was partly on the basis of epidemic model projections that countries around the world introduced 'lockdown' measures (including government stay-at-home orders) and social distancing policies.

Having introduced some of the science of epidemics, I will now examine the population perspective on epidemics, beginning with the etiology of an epidemic. How should we understand Rose's principles, both in general and as applied to an epidemic?

4. Epidemics and infections have distinct explanations

The first central principle underlying Rose's (1985, 1992) population perspective is his dictum that the causes of cases are different than the causes of incidence – in an epidemic, the causes of cases of infectious disease are different from the causes of incidence of infectious disease. This slogan requires revision but leads us to an important point about explaining epidemics.

On a literal ontic construal, Rose's distinction between the causes of cases and the causes of incidence is not as sharp as Rose intends. In an epidemic, we can see this quite clearly by focusing on the cause that all cases of the epidemic disease share: a specific pathogen. For instance, all cases of COVID-19 are caused by SARS-CoV-2. The incidence of a disease is an aggregate measure of the number of cases. So, an incidence of, say, one million in the pandemic represents one million cases of COVID-19. Each of these cases is caused by SARS-CoV-2. Therefore, an incidence of one million cases is caused by SARS-CoV-2 in the sense that:

$$C_1 \rightarrow E_1, C_2 \rightarrow E_2, \dots, C_{1,000,000} \rightarrow E_{1,000,000} .$$

Each E is a token case of COVID-19, each C is a token SARS-CoV-2 infection, and the arrows represent the relation 'caused'. The causes responsible for the incidence of COVID-19 are the same as the causes responsible for the cases of COVID-19 because the incidence aggregates over the cases.

This result is not an artifact of the COVID-19 example generated by the feature that – like other infectious diseases – COVID-19 has a universal infectious cause, one that is common to each case and necessary for the occurrence of a case of that disease type (Broadbent 2013; Fuller 2018). For any communicable or noncommunicable disease E , we can instead describe a

conjunction of causes for each case, where every conjunction may or not share the same salient universal cause. Then, the ‘complete causes’ of one million cases can be represented as:

$$C_{1,1} \cdot C_{2,1} \cdot \dots \rightarrow E_1, C_{1,2} \cdot C_{2,2} \cdot \dots \rightarrow E_2, \dots, C_{1,1,000,000} \cdot C_{2,1,000,000} \cdot \dots \rightarrow E_{1,000,000} .$$

Here again, the causes responsible for the incidence are the same as the causes responsible for the cases. The causes of the incidence are contained within this list of the causes of the cases.

Imagine that a new token complete cause was introduced that increased the incidence to 1,000,001. Based on the definition of the incidence, this change in incidence must be accompanied by a change in the number of cases: one new case must have occurred. This new case must have had some causes because a case of disease is not the kind of thing that occurs uncaused. However, according to my construction of this scenario, the only new causes are the ones responsible for the change in incidence. Thus, any causes responsible for a change in incidence must also be responsible for new cases. It could be that the causes of the cases are re-described when talking about the causes of the incidence, but this re-description refers to the same actual worldly causes.

One promising way of interpreting ‘the causes of incidence’ is as selecting certain salient causes from among the causes of the cases. What criterion of causal selection might Rose have in mind? One candidate would be a cause that appears in all or most of the complete causes of the cases: in our example, SARS-CoV-2. However, other examples may lack such a universal cause, especially the noncommunicable diseases in which Rose was especially interested.

Another way that Rose (1985, 1992) framed his distinction is that the causes of differences in outcomes within a population are typically different compared to the causes of differences in incidence between two populations. For instance, he claims that genetic factors are often responsible for variation within populations, while environmental factors are usually

responsible for variation between populations. These claims, taken as empirical generalizations, fail to hold in the example of an epidemic because a specific environmental pathogen is a cause of all differences in occurrence of disease both within and between populations. However, if we understand Rose's distinction differently, then we can salvage a deeper generalization, one that is primarily epistemic rather than an empirical statement about the kinds of worldly factors that cause the cases versus the incidence.

Instead of a difference between the causes of cases and the causes of incidence, Rose's distinction is better understood as a difference between the causal *explanation* for cases and the causal *explanation* for the incidence. In a few places, Rose approaches the latter distinction: "I find it increasingly helpful to distinguish two kinds of aetiological question. The first seeks the causes of cases, and the second seeks the causes of incidence. 'Why do some individuals have hypertension?' is a quite different question from 'Why do some populations have much hypertension, whilst in others it is rare?'" (2001, 428). As Valles argues (2018), these two different why-questions demand different types of causal explanation in answer: the former asking for an explanatory difference between hypertensive and non-hypertensive individuals, and the latter asking for an explanatory difference between high incidence and low incidence populations. However, in what ways does the latter explanation differ from the former? Before offering an answer, let us first consider what these two kinds of explanation have in common.

Both questions seek a cause that explains the contrast in question. They thus call for a contrastive causal explanation. For instance, drawing on Peter Lipton's (2004) model of contrastive causal explanation, an answer to Rose's first question will cite a cause of hypertension in some individuals that is absent from (actual) contrasting non-hypertensive individuals. An answer to Rose's second question will list a cause of a high incidence of

hypertension in some populations that differs in (actual) contrasting populations that have a lower incidence.²

Alex Broadbent (2013) similarly makes contrastive causal explanation a central pillar in his account of causation in epidemiology and argues that not enough attention has been given to explanation by epidemiologists. A contrastive causal model of explanation suits epidemiology well because epidemiology is consumed with contrastive causal analysis, whether comparing cases to controls, different population cohorts with one another, or clinical trial groups. We don't have to assume that a contrastive model exhausts all the kinds of explanations that are sought after in epidemiology, but the model does capture many of them. It is also compatible with one of Rose's (1985, 1992) main concerns in drawing attention to the causes of incidence, which was to account for differences when comparing populations. In fact, it provides a useful causal selection criterion for the causes of incidence: 'the cause' of the incidence is that cause (or causes) which explain(s) the relevant population contrast.

To illustrate the contrastive model, take the question 'Why did these individuals develop COVID-19?'. If the questioner is an epidemiologist trying to identify the cause of a new syndrome in Wuhan at the beginning of the COVID-19 pandemic, the relevant contrast class contains individuals without COVID-19 symptoms ('controls'), and a good answer would cite an exposure common to all the cases of COVID-19 and absent from the controls (i.e. SARS-CoV-2). If instead a doctor wanted to understand why some individuals infected with SARS-CoV-2 develop symptomatic COVID-19 while others do not, the relevant contrast is people infected

² On contrastive accounts of causation (e.g. Schaffer 2005), causal claims (e.g. 'a high salt diet caused their hypertension') are themselves (often implicitly) contrastive. A contrastive causal explanation is not necessarily committed to a contrastive interpretation of causal claims. However, even if one held a contrastive view of causal claims, that would not dissolve my distinction between causes of cases (incidence) and contrastive explanations for cases (incidence) or render the distinction trivial or uninteresting by somehow making the causes of cases also contrastive because by 'causes of cases' I mean those features of the world that are represented in causal claims rather than the causal claims we make about them.

with SARS-CoV-2 who did not develop symptomatic COVID-19, and a good answer would identify viral or physiologic determinants of disease pathology (such as the body's inflammatory response).

Now take a question suggested by Galea and Keyes (2020): 'Why was the incidence of COVID-19 increasing in Florida in June 2020?'. Their explicit contrast is Massachusetts, in which the incidence of COVID-19 was decreasing in June 2020. Unlike the previous question, this why-question seeks an explanation for the incidence, not an explanation for the cases. A why-question about the incidence (as opposed to a why-question about the cases) treats the incidence as the fact to be explained rather than any of the specific cases into which the incidence can be decomposed. The relevant contrast is not between a set of specific cases of COVID-19 and a set of specific controls, but rather between two populations that vary in terms of a population characteristic: namely, the incidence. The relevant explanation will cite other population characteristic(s) that differed between the two populations.

By 'population characteristic'³, I mean a feature that is described at the level of the population as a whole – whether that feature is internal or external to the population's individuals – and does not necessarily apply to a specific individual member of that population. For instance, the incidence is a characteristic of a population rather than an individual. So are the parameters of an epidemic model such as β , the effective contact rate. A population characteristic often aggregates, averages or otherwise abstracts over individual-level characteristics: the incidence aggregates over cases of disease, while β abstracts over individual disease transmission events. A relevant explanation for a population characteristic will cite another population characteristic that

³ I use the term 'characteristic' rather than 'property' to remain noncommittal about whether these characteristics are real properties of populations and whether populations can bear properties at all.

was causally responsible for the difference between the two populations. A relevant explanatory population characteristic might be something that we simply cannot attribute to a specific individual or their context; for instance, it would be a category mistake to attribute β , as a rate, to an individual contact event. However, sometimes a relevant population characteristic could equally describe a specific individual, as when it refers to a universal or ubiquitous individual exposure that we could reasonably say is also an exposure for the population as a whole. One of Rose's (1985) examples was the hardness of a population's water supply, which we can say is an exposure for the population by virtue of the fact that it is an exposure for all or most individuals in that population. However, in the previous example of an increasing incidence of COVID-19, SARS-CoV-2 is not a population characteristic because the entire population is not (as a matter of fact) exposed to SARS-CoV-2, and thus SARS-CoV-2 does not provide an explanation for the difference in incidence.

When a population characteristic is a universal or ubiquitous exposure for individuals in a population (and absent from a contrasting population), it cannot explain case-control contrasts within the exposed population. A hard water supply might explain why the first population has a higher incidence of disease but it cannot explain why certain individuals in that population have the disease compared to others who lack the disease because – as a universal exposure – water hardness is not a difference between the cases and the controls. (This phenomenon greatly worried Rose: a case-control study done in the universally exposed population would fail to identify the exposure as a cause – even if the exposure caused many cases – because there are no

unexposed controls⁴.) However, when an explanatory population variable abstracts away from individuals as does β , it could be that some cause – in this case, infection transmission – is cited by individual explanations as well. Infection transmission represented by β will explain a higher incidence of COVID-19, while infection transmission represented at the level of particular SARS-CoV-2 transmission events will explain individual cases of COVID-19. So, as a general principle it is not always true that the causes represented in explanations for the incidence are different from the causes represented in explanations for the cases.

Instead, Rose's claim should be rephrased to read: *contrastive explanations for the incidence are distinct from contrastive explanations for the cases*. By 'explanation', I am referring here to a fully articulated proposition (the explanans). My claim should be read as a statement about the explanans as a proposition rather than the worldly causes cited in the explanans, which are not always distinct (as I just pointed out). In specifying that the explanation is fully articulated, I also wish to preclude shorthand expressions like 'Because the tap water is hard'. It might seem as though this expression could explain a difference in incidence *or* a difference in cases, but it really stands in for some more fully articulated explanation such as 'Because the tap water in Akron is hard while the tap water in Seattle is not' or 'Because Ramesh drinks hard tap water while Vlada does not'. Not only do I want to claim that *particular* contrastive explanations for the incidence are distinct from *particular* contrastive explanations for its constitutive cases; but also, contrastive explanations for the incidence are different in *kind* in that they typically describe population characteristics in both explanandum and explanans, and

⁴ To further illustrate this idea, a study comparing lit and unlit firewood seeking to identify the causes of firewood lighting that was conducted entirely outdoors would fail to uncover the causal role of oxygen because exposure to oxygen would be common to every lit and unlit heap of firewood.

thus explain phenomena at a different level of decomposition compared to explanations for cases.

One might wonder just how distinct explanations for the incidence and explanations for the cases really are given that the incidence aggregates over cases. In fact, Broadbent (2013) proposes that to explain a difference n in aggregate outcome between two epidemiological groups is to cite a difference in the level of an exposure that caused at least n cases of the outcome in the group with greater exposure. For example, to explain why one population had a 10,000 greater incidence of COVID-19 compared to a second population, we would cite a cause of at least 10,000 cases of COVID-19 in the first population that was absent from the second population. Broadbent's interpretation of aggregate epidemiologic outcomes is plausible. Using his interpretation, is this explanation for the COVID-19 incidence not also an explanation for the cases? In my sense it is not, in that it does not explain specific cases of COVID-19. It explains why *some* individuals contracted COVID-19 while others did not, but not why *particular* individuals contracted COVID-19. It could be that the incidence in the population is 1,000,000, and therefore the cause that explained some 10,000 of those cases (and thus explained the difference in incidence) does not explain the other 990,000.

This epistemic difficulty of identifying the specific individuals affected is also one reason why an explanation for the incidence must typically cite a population characteristic rather than a collection of specific individual-level causes. Assume that relaxed social distancing policies caused exactly 10,000 COVID-19 cases in the population with an incidence of 1,000,000. In principle, if we knew which 10,000 cases were caused by the new policies, we could list these specific causings, which would indeed account for the difference in incidence (bar re-openings caused Abe to contract COVID-19, relaxed restrictions on crowd sizes caused Abby to contract

COVID-19, and so on). In practice, we can typically do no better than list the net difference in incidence explained by a population-wide difference (in this case, in social distancing policies).

Moreover, there are reasons to think that an incidence explanation describing a population characteristic is more explanatory for an epidemiologist or public health authority than an explanation citing specific individual-level causings, at least sometimes. For one thing, an epidemiologist is often interested in general relationships (Broadbent 2013), those that apply to other contexts (e.g. relaxing social distancing too quickly causes a spike in cases), rather than unique happenings (e.g. Abe going to the local pub caused him to contract infection), and population characteristics that abstract away from individuals are more likely to participate in general relationships because they are multiply realized (it makes no difference to the incidence if Abe was infected rather than Abby). Moreover, population characteristics often give public health authorities a handle for the purposes of intervening (e.g. by changing social distancing policies), a pragmatic virtue of these explanations. As we will see, Rose's 'population strategy' for intervention dovetails with his 'causes of incidence'.

Applied to epidemics, we can translate my principle to read that *epidemics and infections have distinct contrastive explanations*, with all my previous clarifications in place. Epidemic explanations explain population characteristics like the incidence, while infection explanations explain individual characteristics like infection.

For example, an answer to the question 'Why was the incidence of COVID-19 increasing in Florida in June 2020 (while in Massachusetts it was decreasing)?' could use a compartment model to help satisfy the inquiry. The incidence of COVID-19 is driven by the rate of new infections, which is given by $\beta I(t)S(t)$ (see equation 2 from section 3). With an immunizing infection in a closed population, generally the number of susceptibles $S(t)$ is decreasing. Thus,

increases in the rate of infection are driven by increases in the number of infectious individuals $I(t)$ and/or in β . Galea and Keyes (2020) suggest one plausible reason why the incidence was increasing in Florida but not Massachusetts: due to quickly loosened social distancing policies and relaxed adherence to social distancing guidelines in Florida (which would account for a higher β in Florida). For the sake of illustration, a different (currently unsupported) explanation for the difference in β would point to a mutant strain of the virus present in Florida but absent in Massachusetts with more efficient viral attachment to the body's ACE2 receptor. In comparison, an explanation for the large number of cases of COVID-19 in assisted-living facilities in either state (compared to controls living elsewhere) might refer to differences in the presence of serious comorbid conditions and in social contact patterns (Galea and Keyes 2020).

A compartment model can often provide a relevant epidemic explanation or contribute to a relevant explanation because the parameters in a compartment model represent population characteristics. Compartment model equations relate the rate of change in a population compartment to the size of that compartment and to other population characteristics like β . A compartment model thus shows how differences in this rate depend on differences in other population parameters.⁵ On the other hand, a microsimulation model might fail to deliver relevant epidemic explanations despite the fact that a compartment model and a microsimulation model can be used to model the exact same epidemic (such was the case with the two models

⁵ The idea that compartment models capture difference-making relationships justifies my claim that they provide or contribute to causal explanations. Here I will follow Meyer (2020), who argues that dynamical models provide causal explanations when their variables satisfy Woodward's (2003) manipulability criterion for causal relevance. Compartment models are often treated by modelers as if they do, as when modelers predict the causal effectiveness of interventions by manipulating a model parameter and quantifying the effect on cumulative infections (Walker et al. 2020). One might worry that this example shows only that compartment models are *predictive* causal models rather than explanatory. After all, compartment models are typically used to make 'projections' (Adams 2020). However, as Schroeder (2021) argues, epidemic model projections can be understood as conditionals ('If $\beta = x \dots$, then $\frac{dE(t)}{dt} = y$ '). These conditionals can be used to causally explain the past and not only to causally predict the future because they are agnostic as to whether $\frac{dE(t)}{dt} = y$ in the past or in the future.

used by the Imperial College team to predict the spread of COVID-19 in the U.K. and the U.S. (Ferguson et al. 2020; Walker et al. 2020)). The reason that a microsimulation model might fail is not because it leaves out the causes of the incidence. A microsimulation model *does* model the causes of the incidence, just with less abstraction: for instance, rather than representing infection transmission on average with β , it represents infection transmission events individually. The reason it might fail to deliver relevant epidemic explanations is precisely because it models interactions among individuals rather than modeling population characteristics, thus missing the epidemic for the trees.

Other distinctions –between reductive and non-reductive explanations, or between biological and social explanations – fail to reliably distinguish epidemic explanations from infection explanations. For example, a relevant explanation for the incidence citing policies around social distancing could reasonably be described as ‘non-reductive’ because it refers to societal-level policies as well as ‘social’ because it refers to social behavior. Meanwhile, a different but no less relevant explanation for the incidence citing a new dominant viral strain could be described as ‘reductive’ because it refers to an entity at a sub-population level as well as ‘biological’ because viruses are objects of biological research.

The distinction between epidemic explanations and infection explanations is in the first place about different levels of contrastive causal explanation. As its name suggests, the population perspective on the causes of an epidemic differs from the individual perspective in perspective. ‘Epidemic causes’ and ‘infection causes’ are often two perspectives on the same perspective-independent thing.⁶ An infection explanation decomposes an epidemic into its cases and explains individual cases (perhaps using biological models of disease), while an epidemic

⁶ On recent work on perspectivism in philosophy of science, see Massimi (2018).

explanation regards the population afflicted with the epidemic as the unit of analysis and explains its population characteristics (perhaps using a compartment model). These perspectives offer complementary rather than competing explanations, differing in how they characterize an epidemic and its causes. We do not need to posit additional causes for the incidence on top of the causes of the cases. Moreover, the particular causes that an epidemic explanation selects might be the same causes represented in an infection explanation.

The first core presupposition of the population perspective on epidemics is that the causes of cases in an epidemic are different than the causes of incidence. On a literal construal, the claim is false because the incidence aggregates over the cases. An improved restatement of the idea is that contrastive explanations for the incidence are distinct from contrastive explanations for the cases. Applied to an epidemic, epidemics and infections have distinct contrastive explanations, the former explanations describing characteristics of epidemic populations and the latter describing characteristics of particular infected individuals. I will now show how epidemic explanations play an important role in the population strategy of intervention.

5. The population strategy for epidemics: targeting dynamical population parameters

The second central principle underlying Rose's (1985, 1992) population perspective is his distinction between 'high-risk' and 'population' strategies for prevention. In an epidemic, the former would protect those at high risk for infection or complications of infection while the latter would protect the population. I will argue that the distinction is not about targeting different individuals, but rather about different ways of conceptualizing the target of intervention.

In the high-risk strategy for prevention, preventive interventions are targeted at individuals who are at high risk of some bad outcome. Rose identifies this strategy as the approach of clinical medicine. For an individual with high blood pressure at high risk of a heart attack or stroke, an antihypertensive is provided that may prevent that individual from having a heart attack or stroke. These interventions are often viewed as ‘lowering the patient’s risk’. I will set aside questions about the meaning of the patient’s individual risk and whether medical interventions are best understood in this way (but see Fuller 2020). The important feature of the high-risk approach for our purposes is that interventions target individuals.

The main consideration Rose puts forward in favor of the population strategy is “one of the most fundamental axioms in preventive medicine: a large number of people exposed to a small risk may generate many more cases than a small number exposed to a high risk” (2008, 59), and so there may be more cases to prevent for a population strategy compared with a high-risk strategy. Rose’s ‘fundamental axiom’ follows from the mathematics of the ‘absolute risk’ in epidemiology: the number of cases is equal to the product of the absolute risk (the proportion of the population that is diseased) and the size of the population; thus, a smaller risk may generate more cases if the population is larger.

Rose describes a corollary that he calls the ‘prevention paradox’: “A preventive measure which brings much benefit to the population offers little to each participating individual” (2001, 432). In the population strategy, the incidence (total number of cases) may be greatly reduced, while the population’s individuals generally experience little benefit because their risk is small. Although not a logical paradox, this idea seems strange when cast as pitting individual benefit against ‘population benefit’, considering that the population is composed of individuals.

However, one way of interpreting the population strategy is as a strategy of intervening in the health states of all *individuals* in the population, or of those individuals who are not in the high-risk subgroup. Stephen John seems to understand the population strategy in this way, as “targeting the (many) members of moderate- or low-risk groups” (2011, 250). John thus interprets Rose’s prevention paradox as trading between two different senses of ‘benefit’: benefit as a reduction in individual risk of harm, and benefit as prevention of realized individual harm. The population strategy may benefit a greater number of individuals than the high-risk strategy in the sense of preventing a greater number of cases of realized harm, while bringing less benefit to each individual in the sense of reducing their individual risk by a smaller amount. Understood in this way, the population strategy for intervention is a variation on the high-risk strategy that targets individuals at lower risk, rather than a genuine population perspective on prevention in any meaningful sense.

Yet Rose (1985) portrays his fundamental axiom as merely one motivation for the population strategy rather than as describing the target of intervention in the population strategy. Rather than lower-risk individuals, the target of intervention in Rose’s population strategy is the population incidence: “These two approaches to aetiology—the individual and the population-based—have their counterparts in prevention. In the first, preventive strategy seeks to identify high-risk susceptible individuals and to offer them some individual protection. In contrast, the ‘population strategy’ seeks to control the determinants of incidence in the population as a whole” (Rose 2001, 429). Rose’s population strategy involves targeting the ‘causes of incidence’. In the previous section, I argued that we should understand the ‘causes of incidence’ as referring to contrastive (causal) explanations for the incidence. Therefore, we can understand *the population*

strategy for intervention as targeting population characteristics that contrastively explain the incidence.

This interpretation of the population strategy piggybacks on my interpretation of ‘the causes of incidence’. In explaining the incidence in a population, we cite a cause of the incidence in that population – a population characteristic – that differed in a contrast population with a different incidence. The contrast population is usually an actual population, whether a different population or the same population at a different time. Similarly, in the population strategy for intervention, we target a population characteristic, manipulating its value, in order to produce a difference in incidence between two populations. This time, the relevant contrast is a counterfactual comparison of the same population under different values for the population characteristic. In explaining, we locate a cause responsible for a difference. In intervening, we manipulate a cause to produce a difference. This difference can then be explained by citing the manipulated cause.

‘Targeting a population characteristic’ is an intellectual exercise: we conceptualize our intervention as manipulating causes that are described as population characteristics, and we predict resulting differences in other population characteristics. It is not that interventions targeting population characteristics affect only the population without affecting its individuals. After all, some of these population characteristics (like the incidence) are mathematically defined in terms of individual characteristics (like the presence of disease). Others (like the basic reproduction number) are defined in terms of other population characteristics, yet they nonetheless supervene on individual characteristics (like individual behavior). In either case, a change in the population characteristic will be reflected in changes in individual characteristics. Unless we think these changes in individual characteristics are uncaused or that they are

mysteriously brought about by some other cause, our population intervention must be responsible for them.

On my interpretation, Rose's prevention paradox is best seen as a conflict between individual benefit and population benefit, rather than as one between the two distinct senses of individual benefit that John (2011) proposes. The population strategy often brings little benefit to most individuals (whether individual benefit is understood as a small risk reduction or as prevention of a realized harm). At the same time, the population strategy often brings much benefit to the population in the sense of greatly affecting population characteristics such as the incidence. It is beyond the scope of the present paper to probe these concepts of 'individual benefit' versus 'population benefit' (but see John (2014) on the ethics of high-risk versus population approaches, and Broadbent (2020) on the prevention paradox and ethical trade-offs in COVID-19). Again, the upshot for the population perspective is that the population strategy targets population characteristics rather than individual characteristics.

One might argue that even if the conceptual target of intervention is different in the population strategy, practically speaking it involves extending the use of an intervention for *many individuals* beyond the high-risk group. In some examples like blood pressure lowering or water fluoridation that may sometimes be true: the population strategy might recommend antihypertensives for the many individuals without severe hypertension and fluoride for the many individuals who are not at high risk of dental decay. This scenario is readily understood in terms of my proposal: an antihypertensive serves as a universal or ubiquitous exposure (an exposure for the population as a whole) that targets the mean blood pressure in the population (a population characteristic); and this lowered mean blood pressure then explains the lower incidence of heart attacks and strokes than would have otherwise occurred. Or, to use Rose's

more visual interpretation, the intervention targets and shifts the entire blood pressure distribution.

Moreover, other times the conceptual strategy of targeting population characteristics that drive differences in the incidence may recommend different interventions compared to the high-risk strategy and thus cannot be seen as extending an intervention to many individuals. For instance, in controlling an epidemic, the high-risk strategy might recommend reducing contacts for those at highest risk of death to prevent them from being infected, while the population strategy might recommend banning large gatherings at which infection ‘super-spreading’ events may occur in order to reduce the general incidence of infection and prevent deaths that way. The population strategy might sometimes even recommend an intervention that is not easily seen as aimed at specific individuals at all, such as influencing social norms (Rose 1992). For instance, changing norms around diet and exercise might shift the blood pressure distribution in the population. Similarly, influencing norms around facemask wearing might alter average adherence to facemask recommendations and could have effects on infection transmission at the community level.

Applied to epidemics, the population strategy targets population parameters that explain differences in the incidence of infection or other population characteristics. These are the kinds of parameters that appear in a compartment model of the dynamics of an epidemic. In short, *the population strategy for intervening in an epidemic involves targeting its dynamical population parameters*. It involves predicting the effect on an epidemic’s dynamics of intervening on a population parameter that explains those dynamics. One can predict the effect of such an intervention by tweaking only the value of that parameter in a compartment model and observing the change in some other population parameter. This manipulation allows one to observe the

benefit brought by certain changes in the targeted parameter. One can then consider a real-world intervention corresponding to the hypothetical intervention they implemented in the compartment model.

Population-wide vaccination programs illustrate the population strategy for epidemics. The theory behind vaccination is to move individuals from the susceptible to the recovered compartment, bypassing the pre-infectious and infectious compartments. There are different vaccination strategies available to snuff out an epidemic, including vaccinating those exposed to the pathogen, those at-risk of serious complications from infection, or the general population (Grüne-Yanoff 2018). Often, vaccinating the general population will provide the most population benefit in terms of reducing the cumulative incidence of infections by providing the population with herd immunity. When a population has sufficient herd immunity, the proportion of the population that is susceptible s is too small for the epidemic to grow. Recall the equation for the effective reproduction number: $R_n = R_0 s$. When s is less than the inverse of the basic reproduction number ($s < 1/R_0$), then R_n is less than 1. When R_n is less than 1, the number of new infections decreases. As the number of new infections decreases, the cumulative number of infections is lowered and those at risk of complications are statistically protected from infection. The point at which $s = 1/R_0$ is called the ‘herd immunity threshold’ and is a conceptual target that many mass vaccination campaigns seek to surpass.

Targeting at-risk or exposed individuals to provide them with individual immunity is the high-risk strategy, while targeting s (a dynamical population parameter) to provide the population with herd immunity is the population strategy. Mass vaccination illustrates Rose’s prevention paradox because it greatly affects population parameters while bringing little individual benefit to most individuals in the community. It adds an interesting layer to the

paradox because herd immunity benefits individuals in the high-risk group whether or not they receive the intervention. It thus illustrates that the distinction between the high-risk and population strategies is not necessarily about which individuals (higher risk vs. lower risk) are benefited by the intervention.

An example of the population strategy at work during the COVID-19 pandemic is the use of compartment models to predict the influence of social distancing policies on epidemic parameters. Imperial College modelers used this approach with their SEIR model to predict the effect on incidence and deaths from infection (Walker et al. 2020). We can illustrate the thinking with our simpler SEIR model from section 3. Remember from equation 2 that the incidence of new infections is driven partly by the effective contact rate β . We can break β down into two components via the equation $\beta = mq$, where m is the average number of contacts between an infectious and a susceptible individual and q is the probability of transmission given contact (Vynnycky & White 2010). Social distancing decreases m ⁷ (while mask-wearing decreases q). Combing the equations for R_0 and R_n introduced in section 3 with the equation for β above, $R_n = (mqNs)/r$. Therefore, decreasing m decreases R_n . In a vivid demonstration of the population strategy, based on a retrospective study of social distancing and lockdown measures implemented across Europe, the Imperial College researchers concluded: “current interventions have been sufficient to drive the reproduction number [R_n] below 1 (probability [R_n] < 1.0 is 99.9% across all countries we consider) and achieve epidemic control” (Flaxman et al. 2020, 2).

The value of the population strategy for epidemics is that it can potentially achieve large benefits for the population by greatly altering epidemic dynamics. These dynamics determine the

⁷ That is, if we define a contact as an interaction in which two individuals come within (say) six feet of one another, and social distancing means avoiding contact within (say) six feet.

number of cases, and thus a large number of cases may be averted through changes in dynamical population parameters. The strategy can make use of contrastive explanations for the incidence, which speaks to the value of those explanations. By explaining a difference in incidence between two actual populations, one can often identify fruitful causes on which to intervene to lower the incidence in one of those populations.

Once again, Rose's distinction – this time between the population and high-risk strategies – invokes a change in perspective rather than different kinds of causes that operate exclusively at different levels of decomposition. An intervention cannot effect a change in a population without effecting a change in that population's individuals, assuming supervenience of population characteristics on individual characteristics. 'Targeting a population parameter' refers to a way of conceptualizing or modeling an intervention as intervening on a population parameter; it does not invoke a distinction between interventions that affect only the population and interventions that affect only the individuals. Nor does the population strategy amount to targeting one set of individuals rather than another – the sets of individuals impacted might overlap, sometimes greatly. However, the strategies for intervening suggested by these different perspectives may be vastly different and have vastly different impacts (which is why these two strategies are non-redundant).

The second central idea behind Rose's population perspective is the distinction between population and high-risk strategies. Rather than as intervening for different individuals, these strategies conceptualize the target of intervention differently: the high-risk strategy targets high-risk individuals, while the population strategy targets population characteristics that explain differences in the incidence. In the population strategy for epidemics, the relevant population characteristics are dynamical parameters of the kind found in a compartment model.

6. Conclusion

Two central principles for the population perspective in epidemiology are Rose's distinction between the causes of cases and the causes of incidence, as well as his distinction between the population strategy and the high-risk strategy of intervention. I argued that the former principle should be interpreted as the claim that contrastive causal explanations for the incidence are distinct from contrastive causal explanations for the cases. Only sometimes will these explanations refer to distinct causes. The latter principle should be understood as a distinction between a strategy of targeting high-risk individuals versus one of targeting population characteristics that explain differences in the incidence. Rose's prevention paradox can then be seen as trading small benefit with respect to individuals for large benefit with respect to population characteristics. The population perspective prefers contrastive explanations for the incidence as well as population strategies that target population characteristics. It differs from the individual perspective primarily in terms of the level of contrastive causal analysis adopted.

Two corresponding principles underlying the population perspective on epidemics are that epidemics and infections have distinct contrastive explanations, and that the population strategy for intervening in an epidemic involves targeting its dynamical population parameters (versus the high-risk strategy of targeting parameters at the level of infections). Under the high-powered lens of a microsimulation model, an epidemic is a collection of individuals interacting to transmit infection from one to another with each infection evolving individually, and our interventions can be visualized as disrupting transmission among specific individuals. Under the lower magnification of a compartment model, an epidemic is an infected population evolving

collectively, growing or shrinking at shifting rates, and our interventions can be viewed as altering its global characteristics – population pathology and prevention.

References

Adams, J. 2020. “What are COVID-19 models modeling?” *The Society Pages*. 8 April 2020. <https://thesocietypages.org/specials/what-are-covid-19-models-modeling/>. Accessed 29 December 2020.

Andersen, K. G., A. Rambaut, W. I. Lipkin, E. C. Holmes, and R. F. Garry. 2020. "The proximal origin of SARS-CoV-2." *Nat Med* 26 (4): 450-452.

Broadbent, Alex. 2013. *Philosophy of Epidemiology*. Basingstoke: Palgrave Macmillan.

Broadbent, Alex. 2020. “Thinking rationally about coronavirus COVID-19.” *Daily Nous*. <https://dailynous.com/2020/03/09/thinking-rationally-coronavirus-covid-19-guest-post-alex-broadbent/>. Accessed 6 February 2020.

CDC. 2020. COVID-19 Pandemic Planning Scenarios. Centers for Disease Control and Prevention. <https://www.cdc.gov/coronavirus/2019-ncov/hcp/planning-scenarios.html>. Updated July 10, 2020.

CDC. 2012. *Principles of Epidemiology in Public Health Practice: Centers for Disease Control and Prevention*. <https://www.cdc.gov/csels/dsepd/ss1978/index.html>.

Ferguson, Neil et al. 2020. "Impact of non-pharmaceutical interventions (NPIs) to reduce COVID-19 mortality and healthcare demand." Imperial College London.

Flaxman, Seth, et al. 2020. "Estimating the effects of non-pharmaceutical interventions on COVID-19 in Europe." *Nature*. <https://doi.org/10.1038/s41586-020-2405-7>.

Fuller, Jonathan. 2018. "Universal etiology, multifactorial diseases and the constitutive model of disease classification." *Stud Hist Philos Biol Biomed Sci* 67: 8-15.

Fuller, Jonathan. 2020. "Epidemiological evidence: use at your 'own risk'?" *Philosophy of Science* 87: 1119–29.

Galea, Sandro and Katherine M. Keyes. 2020. "Understanding the Covid-19 pandemic through the lens of population health science." *American Journal of Epidemiology*.

<https://doi.org/10.1093/aje/kwaa142>.

Grüne-Yanoff, Till. 2018. "Choosing the right model for policy decision-making: the case of smallpox epidemiology." *Synthese*. <https://doi.org/10.1007/s11229-018-1827-3>.

Intersalt Cooperative Research Group 1988. "Intersalt: an international study of electrolyte excretion and blood pressure. Results for 24 hour urinary sodium and potassium excretion." *BMJ* 297: 319–28.

John, Stephen. 2011. "Why the prevention paradox is a paradox, and why we should solve it: a philosophical view." *Preventive Medicine* 53 (4-5): 250-2.

John, Stephen. 2014. "Risk, contractualism and Rose's prevention paradox." *Social Theory and Practice* 40 (1): 28-50.

Keyes, Katherine M. and Sandro Galea. 2016. *Population Health Science*. New York: Oxford University Press.

Khaw, Kay-Tee, and Michael Marmot. 2008. "Commentary." In *Rose's Strategy of Preventive Medicine*, edited by Geoffrey Rose, Kay-Tee Khaw and Michael Marmot. Oxford Scholarship Online.

Krieger, Nancy. 2011. *Epidemiology and the People's Health*. New York: Oxford University Press.

Lipton, Peter. 2004. *Inference to the Best Explanation*, 2nd ed. London: Routledge.

Massimi, Michela. 2018. "Perspectivism." In J. Saatsi (ed) *The Routledge Handbook of Scientific Realism* (pp.164-175). Oxford: Routledge.

Meyer, Russell. 2020. "The Non-mechanistic Option: Defending Dynamical Explanations." *British Journal for the Philosophy of Science* 71 (3): 959-85.

Pearce, N. 1996. "Traditional epidemiology, modern epidemiology, and public health." *Am J Public Health* 86 (5): 678-83.

Rose, Geoffrey. 2001 (1985). "Sick individuals and sick populations." *Int J Epidemiol* 30 (3): 427-32.

Rose, Geoffrey. 2008 (1992). *The Strategy of Preventive Medicine*. Oxford Scholarship Online.

Rosenberg, Charles E. 1992. *Explaining Epidemics and Other Studies in the History of Medicine*. Cambridge: Cambridge University Press.

Schaffer, Jonathan. 2005. "Contrastive Causation." *The Philosophical Review* 114 (3): 327-358.

Schroeder, Andrew. 2021. "How to interpret COVID-19 predictions: reassessing the IHME's model." *Philosophy of Medicine* 1 (1): 1-4.

Valles, Sean. 2018. *Philosophy of Population Health: Philosophy for a New Public Health Era*. New York: Routledge.

Vynnycky, Emilia and Richard White. 2010. *An Introduction to Infectious Disease Modeling*. Oxford: Oxford University Press.

Walker, Patrick G.T., et al. 2020. “The global impact of COVID-19 and strategies for mitigation and suppression.” Imperial College London.

Woodward, James. 2003. *Making Things Happen: A Theory of Causal Explanation*. New York: Oxford University Press.