

DEAN RICKLES

PUBLIC HEALTH

ABSTRACT:

Public health involves the application of a wide variety of scientific and non-scientific disciplines to the very practical problems of improving population health and preventing disease. Public health has received surprisingly little attention from philosophers of science. In this chapter we consider some neglected but important philosophical aspects of the *science* of public health.

To appear in F. Gifford (ed.), *Handbook of the Philosophy of Science, Vol. 16: Philosophy of Medicine*. Elsevier.

1

WHITHER THE PHILOSOPHY OF PUBLIC HEALTH?

Public health concerns the health of *populations* of people, rather than *individual* people, as is the case in clinical medicine (where these individuals are dealt with as ‘patients’) or in biomedicine more generally (where they are treated as ‘subjects’).¹ It deals with *aggregates* of measurements of properties of individuals and is therefore a statistical science, facing the many (technical, epistemological, and metaphysical) problems that this inevitably involves. It prizes the *prevention* (of disease, disability, and premature death) over cure, and is therefore rather more difficult to assess than clinical medicine or biomedical science in terms of its success or failure, since one has ‘counterfactual’ successes: e.g. ‘if this policy (or intervention) *hadn’t* been in place Jones would have perished’.² It is a massively interdisciplinary field (perhaps the *most* interdisciplinary subject there is—*cf.* [Afifi and Breslow, 1994], pp. 225–6), incorporating epidemiology, statistics, biology, informatics, sociology, economics, psychology, environmental science, civic planning, architecture, engineering, and more, making public health a rather unwieldy and complex discipline—indeed, there isn’t much that public health doesn’t (or couldn’t) utilize in some way to achieve its aims. What is considered part of the domain of applicability of public health is flexible to the point of near universal inclusivity: almost anything can be viewed as a public health issue. Furthermore, what is included in public health (and therefore the understanding of what public health *is*) has changed over time, adapting to the changing conditions in society

¹The population can be defined in any number of ways, and need not refer to geographical boundaries. For example, the population might be the ‘scattered object’ that consists of all smokers under the age of eighteen years. We might, then, take individual-based clinical medicine to constitute a degenerate case of population-based public health (i.e. for cases where the number of individuals in the population is 1).

²It is in this regard that C. -E. A. Winslow refers to the “silent victories of public health” ([Winslow, 1923], p. 65). Likewise, Bernard Turnock writes that “when public health efforts are successful, nothing happens. Events that don’t occur don’t attract attention. ... Indeed, the vast majority of those who will ultimately benefit from the efforts of past and present public health workers are yet to be born” ([Turnock, 2006], p. 1). This is particularly so when (on the basis of risk-benefit assessments) interventions are imposed to stop technologies that *might* have otherwise been developed.

that often bring new diseases, and adapting to the time-varying concepts of disease and its determinants. These features make public health an especially challenging field for philosophers of science various with novel issues not to be found in the study of biomedical science.³

Despite the many conceptually interesting features alluded to above, and despite its age and importance, the field of public health has received virtually no attention from philosophers of science, especially those belonging to the ‘analytical’ school—one exception is Douglas Weed (a professional epidemiologist and ‘amateur’ philosopher!): he too bemoans the absence of philosophical work on public health *qua* scientific discipline [Weed, 1999; 2004]. Much of the work available tends to follow a ‘Continental path’, which has a tendency to ignore the scientific aspects of public health (in favour of considerations of *power relations*, à la Foucault, for example). The work that falls outside of the Continental tradition is primarily located within the fields of Science Studies and Bioethics, which tend to have different agendas to the philosophy of science.⁴

My interest is with highlighting issues that are of relevance and importance to (analytic) philosophers of science. I make no apologies for this lopsided approach; I think it is necessary to provide balance to the debate as a whole. Here, then, I attempt to redress the balance by offering a simple field guide to some philosophical aspects of public health, covering in (in a fairly preliminary way) a variety of topics that really ought to have been better studied (or just studied, *period*) by philosophers of science. Because public health includes in its domain many other fields from medicine it treads on the toes of a vast array of issues that are dealt with by other contributors to this volume. For this reason, when overlap is an issue I avoid the nitty gritty details and paint the broad picture, showing how it connects

³This is not to say that all of the issues are novel: for example, since we are dealing with the health and disease of the public we still have to say what we mean by ‘health’ and ‘disease’, and many of the same issues are thrown up in this context as are thrown up in biomedicine: there are broadly ‘normativist’ (or subjectivist) and broadly ‘naturalist’ (or objectivist) approaches. This issue has ramifications for how we measure health and disease which, in turn, has ramifications on how resources are distributed in health care systems—hence, these are not idle issues of no practical consequence. There are also issues to do with causal inference, evidence, the nature of theories, and so on, that are more or less on a par with those from biomedicine. However, even these issues do take on a very different flavour on account of the fact that the systems of interest are populations and the quantities of interest belong to these populations.

⁴‘Public health ethics’ has recently emerged as a specialized sub-discipline of bioethics devoted to “those ethical issues and perspectives that may be said to be distinctive to public health ... apart from the perspective of clinical medicine” ([Bayer *et al.*, 2006], p. 4)—see [Dawson and Verweij, 2007] for a nice collection of essays on the subject (many of which are by philosophers). The ethical implications flow fairly ineluctably and quite obviously: the systems of investigation in public health are often very large, making it impossible to gain consent for some intervention. Decisions, over such interventions, are made by external agencies (government agencies, local boards, etc.). How far should one take this often involuntary enforcement of public health interventions? To the point of involuntary inoculation *for the greater good*? Involuntary fluoridation for the common good? This is but one kind of issue; there are many more. Ethics flows into politics too when we consider that the policies thus imposed often constrain the liberties of people in some way or other (e.g. the enforcing of seat-belt wearing; the banning of smoking in public places, etc.). Hence, though related to issues found in clinical medicine, public health throws up issues that appear to be *sui generis*.

to public health. Further, in order to keep the scale of this chapter manageable I restrict my attention to those issues that are peculiar to public health.

I begin with a brief historical review of some key elements of public health with the aim of building up a picture of what public health is and what kinds of phenomena it deals with (a surprisingly difficult issue). I then consider definitional issues, focusing in particular on the possible meanings of ‘public’ and ‘health’—I briefly discuss the ‘demarcation question’ too; that is: Is public health a science at all? This leads in to a discussion of the concepts of ‘health’ and ‘disease’ in the public health (population) context, which in turn leads into a discussion of the measurement of health states and disease states, and the construction of population (or summary) measures. I then consider several epistemological issues, having to do with causality and causal inference (using public health interventions), and the concept of evidence in public health contexts.

2

THE HISTORICAL DEVELOPMENT OF PUBLIC HEALTH

It is difficult to say exactly when and where public health as a distinctive discipline began.⁵ In order to answer these questions we need to know *what* public health is. This in itself is very difficult since the nature of public health has evolved considerably over time. Some prefer to focus on the quantification of properties (relevant to health) for large groups of people, and their ‘surveillance’, providing an evidence-base for informed interventions. Others focus on the particular ‘cartographic’ methodology of finding the determinants of diseases by ‘mapping’ their spread and charting their evolution. Both of these understandings overlap significantly with epidemiology: in both cases the aim is to identify the causes of health and disease (given their variation in populations). However, still others view the connection of health and disease with social and societal issues as the defining characteristic, so that public health involves social (or community) *action*. Really, we ought to view all of these as essential components of modern public health, for it can be viewed as placing epidemiology and statistics in the service of the wider community (extending as far as the global community).

Let us briefly review some of these ideas in historical context to gain a better feel for the kinds of conceptual issues that can arise in public health. In the following historical remarks I do not aim for any kind of completeness, nor do I present the episodes in chronological sequence. Rather, the remarks are grouped thematically—note, however, as intimated above, that there are multiple interactions between the groups.

⁵Good general histories of public health (portraying very different aspects) are: [Hamlin, 1998; 2005; Leavitt and Numbers, 1997; Porter, 1994; 1999; Rosen, 1993; Rosenberg, 1992; Ward and Warren, 2006]. As the manifold differences in these books reveal, the history of public health is an exceedingly complex thing to unpack. A superb recent reader, tracing the development of public health as a discipline, is [Schneider and Lilienfeld, 2008].

2.1 From Vital Statistics to Biostatistics

Public health concerns large numbers of interacting people and systems. As with any theory involving systems composed of very many subunits, public health is a statistical science (or a discipline based on statistical science), dealing with coarse-grained properties of wholes, rather than specific details of the parts. In this sense public health resembles statistical physics, only now the ‘particles’ are patients, GPs, institutions, and so on. The population-level nature of public health, and the necessity to utilize statistics,⁶ was well understood by William Augustus Guy in 1870:

As hygiene deals with mankind not one by one, but in masses, its scientific method can be no other than that numerical method so often confounded with its leading application—statistics. (Guy, 1870)

However, the first numerical approach to epidemiology was John Grant’s development of ‘vital statistics’ in the seventeenth century—as laid out in his *Natural and Political Observations on the Bills of Mortality*. The idea of making records of deaths (“books of the dead”) began with the plague spanning the 14th to the 16th century. These records were used to identify and track epidemics. Incidentally, the plague also led to another common public health measure: quarantine. This is a fine example of a social intervention: by preventing certain interactions from occurring one modifies the social network and thereby prevents the spread of disease in a population. Moreover, the death rates were one of the earliest methods of measuring the health of the public (of populations). That is, vital statistics give one a (rough) numerical reading of the health of populations. However, the census affords perhaps the best overview of the public’s health, enabling stratification by race, social groupings, education, gender, and so on.⁷ This provides a good starting basis for considering interventions to determine whether any apparent links between properties and categories are causally implicated in health states.

Graunt’s ideas are an integral part of modern epidemiology and public health. For example, Graunt used statistical data to monitor the health of populations: using it to identify potential public health problems, to alert the state to such problems, or else to show that a problem was subsiding.⁸ He also pioneered the use

⁶An excellent summary of the history of statistical methods in public health is [Stroup and Berkelman, 1998]. The Wellcome library has a good brief historical overview of the use of statistics in public health—see: http://library.wellcome.ac.uk/doc_WTL038911.html.

⁷A census provides us with a maximal roster of the individuals’ properties of interest. They give us an idea of the state of the population \mathcal{P} and its dynamics $\partial \text{dist}(\mathcal{P})/\partial t$. The census can provide snapshots that provide the population distribution of properties $\text{dist}(\mathcal{P})$. Some quantities will be constant, $d\theta/dt = 0$, and these allow us to parameterize the other varying quantities. The mean μ and variance σ^2 are constant, and they determine the shape of the distribution. So: one draws up a census, and from this one extracts the parameters θ , and this gives us the information we need to assess such things as the ‘health state’ of a population. Of course, in reality it will be difficult if not impossible to perform a census for all members of a population. The statisticians’ trick is to draw a subset that will, to varying degrees, represent the population as a whole: this is, of course, the sample.

⁸This monitoring was later made a central part of Johann Peter Frank’s ‘medical police’. Both are used as tools of the state in some sense. The idea is to have a monitoring system in place at the heart

of the data to monitor variations in health between different populations, a notion central to research on health inequalities (perhaps the most pressing issue facing contemporary population health researchers—see §5). William Petty, a friend of Graunt's, further grounded the basis of modern public health by enjoining public health with political and economic issues—Petty referred to the study of mortality in populations as “political arithmetic”.⁹

In the nineteenth century, both William Farr and Edwin Chadwick (see next subsection) were interested in the social interactions of health and society. Farr's interests were grounded more in statistics: like Graunt he viewed statistics as the basis for social action. But, like Chadwick, he viewed the societal conditions as in large part responsible for the health state of a population, and for differences between the health of populations. Farr was statistician to the General Registrar Office, a post which saw him formulate and investigate many basic principles of epidemiology and public health. His basic aim was fundamentally a public health related one: to prevent and control disease (*cf.* [Adelstein and Susser, 1976], p. iii).

Francis Galton and Karl Pearson were responsible for the development of the field of biostatistics which replaced vital statistics. Biostatistics deals with data derived from all manner of studies pertaining to medicine and biology. Its lessons were propagated into the epidemiological community, and so applied to public health-related studies, by Major Greenwood, Wade Hampton Frost, Bradford Hill and others. Biostatistics was primarily a clinical affair. However, the introduction of computers and the availability of vast databases and the potential to simulate complex processes, and make forecasts of complex processes, has led to the field of ‘public health informatics’. This is an amalgam of public health, engineering and technology. This technology is, and will continue to increase so as to include more variables of interest. For example, there has recently been an integration of GIS (geographical information systems) technology to allow for more expansive surveillance.¹⁰ The internet too allows for more detailed comparisons of statistics. A substantial component of biostatistics is public health surveillance. See [O'Carroll *et al.*, 2002] for more details on these recent developments.

Another recent development in the use of statistics in public health is in the area of meta-analysis; namely, the statistical synthesis of evidence gathered from multi-

of the health system, and on whose data one can act: in Frank's era this involved the quarantining and disinfecting of the ‘unclean’ elements in order to have a growing, fit population. In the former case it is in place to alert the health care agencies to public health threats. For a recent collection of technical articles dealing with aspects of public health surveillance, see [Teutsch and Churchill, 2000]—see also [Thacker *et al.*, 1989] for a good brief review.

⁹Another interesting development was Christian Huygen's development of ‘life tables’ to determine the life expectancy of individuals (at any age), for use in computing life annuities. See Chapter 8 of [Hald, 2003] for an excellent account of this history.

¹⁰See [Cromley and McLafferty, 2002] and [Elliot *et al.*, 2001] for accessible introductions to the applications of GIS to public health. For a historical survey of the ‘mapping’ of disease, up to and including GIS, see [Koch, 2005].

ple independent studies.¹¹ There are many issues of philosophical interest lurking in the way the evidence is synthesized in these analyses—see §3 of [Worrall, 2007] for a good discussion of the conceptual issues that arise.

2.2 *Social and Environmental Dimensions of Health and Disease*

In an early discussion of the role of statistics and statisticians in public health, Edgar Sydenstricker argues that although “medicine” can be viewed as “synonymous with public health”, the latter has in addition “a social objective” ([Sydenstricker, 1928], p. 116). The understanding of public health has in very large part been guided by societal issues. Abdelmonem and Breslow refer, in this regard, to the “dynamic nature of public health” ([1994], p. 224). The emergence of certain novel types of behaviour (due to evolving norms) and the development of new technologies can bring with them new threats to health that simply did not previously exist. For example, industrialization brings with it a greater likelihood of certain kinds of epidemic. The invention of new forms of transport will modify epidemiology and public health due to the new types of injury that can occur. Moreover, greater success in healthcare can, ironically, bring with it its own problems, such as an increased aging population (demanding new specialisms such as gerontology) and a population explosion.

Of course, it isn’t only changes in society that can cause the emergence of new public health threats. The physical environment is implicated too and can radically alter the distribution of health and disease (famine is an obvious example of this). Moreover, the social and the physical are often bound together, so that changes in one will modify the other. As Geoffrey Rose writes, the “scale and pattern of disease reflect the way that people live and their social, economic, and environmental circumstances, and all of these can change quickly” ([Rose, 1992], p. 1). These elements have been investigated and conceptualized in a variety of ways. For example, the emerging diseases of nineteenth century England were hypothesized to be a result of the insanitary conditions that resulted from overcrowding. Likewise, overwork, malnutrition and other (what we would now think of as) ‘social dimensions of disease’ were isolated as part of public health in the Chadwick report [Chadwick, 1842]—similar conclusions were made slightly later in the US in the Shattuck report [Shattuck, 1850]. The isolation and analysis of such problems (using primarily epidemiological studies) were intimately linked with a plan of action to intervene for the betterment of society.

Edwin Chadwick is often taken to have been an advocate of social-wide environmental interventions via sanitary reform. In this goal he was aided by the statistical work of William Farr (see §2.1). However, Sylvia Tesh [1995] has argued that that Chadwick was not so disposed, and that his concerns were firmly

¹¹Note that this is not the same as a systematic review, which does not necessarily involve statistical manipulation: a meta-analysis constitutes a particular kind of systematic review. See [Egger *et al.*, 2001] for the canonical text on systematic reviews (including meta-analyses).

grounded in ‘miasma theory’. This often pointed to interventions that had their location in conditions of poverty, to do with water supply and sewers. The target was miasma producing things, not the social setup *per se*. Tesh’s claim is, then, that the miasma theory led to the particular public health prevention measures (many that were, indeed, loaded with societal significance). Hence, though sanitarianism led to what looks like social reform, it was more a technical fix rather than a matter of social justice.¹²

Hamlin [1995] agrees with this general idea, but takes issue with this claim that “aetiological theory drives preventative strategy”, arguing that there are plenty of cases of under-determination (in which one theory generates multiple preventative strategies) and over-determination (in which multiple theories correspond to the same strategy). His examples focus on cases whereby miasmatic theory can be dealt with either at the level of social conditions (poverty reduction and so on) or at more direct level of contagion (sanitation upgrades and so on).

Mendelsohn [1995], on the other hand, argues that the whole distinction between ‘social’ and ‘physical’ (environmental) is without a real difference, or is, at least, not so easy to support as most discussions suppose. That is, it is easy to assign ‘poverty’ to either the social side or the physical side; it can be viewed as a social condition, a physiological, or a physical condition. Given this, the distinction cannot do any real work in this context. However, the idea that individuals’ health status is connected to the social structures they find themselves in has recently become very fashionable. Social epidemiology, for example, highlights just this ‘social embeddedness’ of individuals. The health states of individuals are not *intrinsic* properties, but are determined by the social networks in which the individuals find themselves: transporting an individual to another context would change the health profile of that individual (*cf.* [Galea and Putnam, 2007], p. 7).¹³ The key role of public health on this account is, then, to modify the environment (or ‘context’) in such a way as to benefit the individuals occupying it. The environment itself then becomes the subject of the adjectives ‘sick’ and ‘healthy’—see [Rose, 2001].

Not surprisingly, this modification of the social and physical environment is what those interested in the ethical consequences of public health find objectionable: the modification is enforced, and not to the benefit of all (most often, in fact,

¹²Virginia Berridge [2007] discusses the Health of Towns Association—an environmental public health intervention advocacy group—that was formed after the publication of Chadwick’s report. This group was instrumental in the promotion of sanitary reform and is often held up as a particularly ‘moral’ movement. However, as with Chadwick, Berridge argues (following the work of Chris Hamlin), the group was convinced that “the problem was sewers and not deprivation” (p. 22): it was firmly grounded in the (erroneous) miasmatic theory of disease.

¹³There is a whiff of the debate between (semantic) ‘externalists’ and ‘internalists’ here. That debate concerns what a speaker means by some word: is it determined by social and physical factors external to the speaker, or is it determined by factors about the speaker? Or, to put it another way: do physical duplicates always mean the same thing by their words regardless of the external social and physical environment? The social epidemiologist is, in this sense, a kind of externalist about health: an individual’s health does not supervene on its intrinsic properties. See [Putnam, 1975] for the original source of this debate.

for economic reasons). For example, if I stop a shuttle service taking students between campuses, and build a nice connecting footpath, then that will force them to look for alternative methods, with the hope that many will walk or use a bicycle, and therefore reduce the burden of disease for the population as a whole. This amounts to an involuntary modification in the behaviour patterns of individuals. The benefit to the population as a whole might be worth the cost (i.e. in health-economic terms), but it might well also be the case that numerous individuals are adversely affected—perhaps dropping out of college, and so on.¹⁴

In many ways, the recent introduction of this ‘social epidemiology’ harks back to the very earliest public health work of Hippocrates, who also considered the impact of the social and physical environments of health (in his work *Airs, Waters, Places* for example).¹⁵ His evidential basis was observational: certain places, with certain social systems and physical conditions appear to be correlated with quite specific health conditions. Weather, for example, was found to be correlated with specific disease patterns, as was the status of the water supply. Of course, the precise nature of the effects of the natural environment on health (that is, the *mechanisms* responsible) were not known at the time—for that, one requires the integration of this ‘social medicine’ with biomedicine. Before the idea of local contagion theory (according to which an ‘agent’ is passed between individual or from a source to an individual), leading to germ theory, the common view was that ‘miasma’ was responsible for the spread of disease—simply by inhaling foul air, one would be exposed. Though this was the ‘wrong theory’, it can nonetheless prove effective in reducing the incidence of disease, since one will tend to cover the airways, and so reduce the risk of infection, and remove the sources of the stench which can simultaneously serve to remove the bacteria (the ‘true cause’). Given this, one could be forgiven for thinking that interventions based on the miasmatic theory, being successful, constituted good evidence for the theory (given the state of knowledge at the time)—that is to say, there were good *rational* reasons for believing the theory.

Before leaving this subsection, I briefly note that the so-called ‘new public health’ bears many similarities to social epidemiology: it is a sociological approach. The basic idea is that threats to health go beyond infectious diseases and lifestyle risks, and can originate in social organization and structures (*cf.* [Baum, 1990]). This belief motivates advocacy of implementing structural social changes in order to improve health. Again, it suggests a kind of externalist conception of health states. The new public health is also characterized by a greater linkage to the state than previously. Legislation in the form of public health acts serve to control risky activities. Risk is also at the root of legislative action in the context of infec-

¹⁴This is known as ‘the prevention paradox’, though it isn’t really a paradox as such: ‘prevention dilemma’ would perhaps be a more appropriate label. See §3.5 for more on this matter.

¹⁵Note, however, that Hippocrates advocated the ‘humoral theory of disease’, according to which disease comes about when there is an imbalance amongst the four humors, blood, phlegm, yellow bile, and black bile. This theory suggests physiological treatments that restore the balance (*cf.* [Thagard, 2005], p. 48).

tious individuals; the state can detain such individuals if they are deemed to pose a sufficient threat to the public's health. In this sense public health is an object in its own right; something to be moulded and altered. However, as we saw, this political connection can be found in the work of Chadwick, who was in fact employed as assistant to the royal commission set up to investigate the Poor Laws. This work unearthed the terrible conditions that England's poor were subject to. However, as mentioned previously, Chadwick's concerns were more economic than than social justice ones. But the results were nonetheless consequential in social justice terms: by eliminating the filth that encouraged disease cholera, for one, was brought under control.¹⁶ Squalor, rather than poverty, was thought to be the root of the problem. Disease, rather, was thought to be a cause of poverty. Hence, the various reforms, associated with sanitary engineering, were intended to reduce disease with the aim of stimulating the economy.

2.3 *The Germ Theory of Disease*

The bacteriological theory of disease, of Pasteur, Koch and others¹⁷, identified the precise biological organisms responsible for the transmission of infectious diseases. Germ theory reduced the spread of disease to the transmission of these bacteria.¹⁸ Hence, the causes of diseases were conceptualized as local biological impingements. A key move was Koch's isolation and culturing of the tuberculosis virus, and his demonstration that tuberculosis could be artificially induced in animals. This engineered *production* of disease (or rather, the fact that it served to establish the germ theory of disease) appears to constitute an instance of Hacking's 'entity realist' stance: causality, manipulability, and reality were bound together—see [Marcum, 2008] (pp. 33–48) for more along these lines.

Public health, *qua* non-local theory of the determinants of human health, suffered somewhat at the hands of this new local and individualistic theory of disease and illness.¹⁹ However, public health has been concerned throughout its existence (however blurry the origins might be) with disease (however that might be understood) and advances in microbiology were quickly integrated into public health

¹⁶Though not, of course, *caused* disease, as per the theories of disease aetiology at the time—see the next subsection.

¹⁷I should perhaps point out that Koch was not much impressed with Pasteur's methodology. As Latour points out, Koch thought that Pasteur's generalization from his vaccinated sheep to a "*general* method, applicable to all infectious diseases" was somewhat "hasty" ([Latour, 1988], p. 29; *cf.* [Guala, 2003], p. 1195). Koch, of course, developed precise postulates that gave the necessary and sufficient conditions for inferring bacterial causation [Koch, 1882]. [Evans, 1993] gives a good historical survey of the interconnections between theories of disease and theories of causation. For a more 'revisionist' history, see [Worboy, 2000].

¹⁸As Thagard ([Thagard, 2005], p. 48) notes, the germ theory is a class of theories that applies to multiple specific disease types, each with their own specific infectious agents, from protozoa, to fungi, to prions.

¹⁹As Mervyn Susser [1985] points out, this can be explained in Kuhnian terms: the non-local, population-based ideas (the miasma paradigm) was replaced by a paradigm based on the Henle-Koch postulates involving the determination of disease-causation.

(again highlighting the latter's adaptability) and used as both a way of identifying disease and as a means of intervening so as to eliminate it. One of the most significant developments along these lines was the establishing of local and government public health departments whose initial role was to keep an eye on the status of populations *vis-à-vis* communicable diseases. The infrastructure cemented for this purpose was quickly expanded for other means, such as screening programmes. The idea of an agency responsible for the control and monitoring of infectious diseases was later, in 1948, implemented at the international level by the United Nations in the form of The World Health Organization [WHO]—of course, the remit of the WHO is now way beyond infectious diseases, extending to various demographic issues that are believed to have health impacts.²⁰

Epidemiology and public health are concerned with the spread of disease. This has often been done at a more coarse-grained level. Molecular epidemiology links molecular biology to epidemiology and public health by providing transmission mechanisms for the spread of disease and potential responses to stop the transmission. Genetic screening and engineering also offers the promise of greater control over the spread of disease, and even the promise to entirely wipe out certain diseases. Less grandiosely it offers the potential for more targeted intervention strategies. Hence, the local, biomedical model is utilized by public health and can be merged with the more global issues more characteristic of the latter's way of operating.

2.4 *Epidemiology and Public Health*

In their article on the connections between epidemiology and public health, Abraham and David Lilienfeld conclude with the statement that “without public health, there is no epidemiology” ([Lilienfeld and Lilienfeld, 1982], p. 148). They trace this intimate connection largely through the public health movement, which includes figures we have already discussed, notably W. A. Guy. The Lilienfeld's argued that there was an increasing disconnect between the two fields (*ibid.*, p. 147). Milton Terris [1987] ultimately concurs with the Lilienfeld's about the tight relations between epidemiology and public health, but argues that there is an emerging tightening once again in what he calls a “second epidemiologic revolution” (p. 327), in which the domain of epidemiology is expanded to include all manner of non-infectious diseases.

Combining the numerical methods (quantifying cases) with information about the populations from which the cases are drawn from is characteristic of modern epidemiology and really began with John Snow's isolation of the Bow Street Pump as the source of a cholera epidemic propagated through human waste. The basic methodological details of this approach were made more rigorous by William Farr, who formulated many more concepts of modern epidemiology—specifically,

²⁰The WHO has an excellent online historical collection, accessible at: <http://www.who.int/library/collections/historical/en/>. This resource has material going back to 1507.

Snow utilized Farr's tabulations relating cholera mortality to water supply and based his hypothesis—about the nature of cholera and its propagation via an organism passed through human waste—on this data. The hypothesis was tested using the weekly mortality tables supplied by Farr. Of course, as mentioned above, at this stage it was not known that the cholera bacillus was the agent responsible for the disease, but the success of the intervention did not depend on this knowledge.

Though controlled experiments in public health contexts are extraordinarily difficult, for a variety of reasons (e.g. complexity, ethics, etc.), when they are possible they constitute the most reliable way of generating evidence and testing hypotheses—if conducted correctly, that is, with the right number of trial arms and the right experimental subjects (or units). The modern idea of performing trials (or sampling) to test determinants of health and disease was instigated in the scurvy trial of James Lind. Lind gathered a group of sailors stricken with scurvy and split them up into various pairs, each pair given a different supplement to their usual diet. The pair that was given citrus fruits recovered. This result was used to infer the causal efficacy of citrus fruit in curing scurvy—later biomedical work determined that it was specifically the vitamin C component of citrus fruits that did the work. This evidence could then be used to *prevent* cases of scurvy from occurring, thus demonstrating how a clinical trial can be used in the service of public health.

A similar trial was conducted by Ignaz Semmelweiss (in the 1840s) to test a hypothesis about the causal factor responsible for the difference in the incidence of puerperal fever within maternity wards operated by midwives (in training) and those run by physicians (who also conducted autopsies). Semmelweiss conjectured that there was some infection as a result of the autopsy work (transmitted by 'cadaverous particles'). To test this he simply instigated measures to cleanse hands prior to deliveries. The rates between the wards balanced out. It is worth spelling out the details of this case.

In fact, Hempel considers this case in his *Philosophy of Natural Science* [Hempel, 1966]. Donald Gillies [2005] examines Hempel's analysis, and argues that it needs to be supplemented by elements of Kuhnian philosophy (in order to make sense of Semmelweiss' failure to convince the wider medical community that his ideas were sound). Gillies notes that Semmelweiss' methodology was largely in line with Popper's model of conjectures and refutations: hypotheses were suggested and then tested experimentally by appropriate interventions.²¹ The first hypothesis was that "atmospheric-cosmic-terrestrial" factors were responsible—as Gillies notes, this is another way of referring to the miasmatic theory (*ibid.*, pp. 161-162). This was quickly rejected by noting that these would be constant across the wards (given their close proximity), and so could not be called upon to account for the observed differences in death rates. A government commission established to investigate the curious differences came up with the theory that differences in the way the patients were handled were responsible, in one case involving rough

²¹ Peter Lipton also considered Hempel's treatment of Semmelweiss and argued that it constituted an instance of inference to the best explanation ([Lipton, 1991], pp. 75-98).

student medics (often foreign) and in the other case involving more delicate (non-foreign!) student midwives. The rationale is that this was another difference between the groups, and one needs a difference in the cause to get a difference in the effect. However, Semmelweiss performed a test involving the canceling out of the suggested causal factors (by making the groups comparable), so that the wards were balanced with respect to them, but found no significant reduction.²²

It was new background evidence—concerning the death of a colleague by an infection from a knife wound during an autopsy resulting in something similar to puerperal fever—that brought about the breakthrough. He surmised that since some of the doctors go from autopsy to clinic they might be transmitting the same material during their examinations. The midwives, on the other hand, do not conduct autopsies. By implementing a trial of thorough disinfection of the hands following autopsies Semmelweiss was finally able to balance out the death rates. The disinfection was generalized from post-autopsy situations to any medical examination involving “ichor” (that is, discharge emanating from a wound). Further generalization to airborne transmission of “ichorous particles” followed other incidents, resulting in the isolation of those presenting with such wounds.

Gillies, as mentioned, views the hypothetico-deductive account espoused by Hempel to be incomplete: it doesn’t offer an account of why Semmelweis’ theory was not adopted given the strength of evidence. Gillies argues that Kuhn’s model can provide an easy answer: “the theory ... was rejected because it contradicted the then dominant paradigm concerning the causation of disease” ([Gillies, 2005], p. 171)—it contradicted both the miasmatic theory and the contagion theory, making Semmelweis a revolutionary.

However, both of these trials are rather small-scale, though the Semmelweis trial was carried out at the level of wards (a population of sorts), rather than individuals. Modern day public health trials can involve entire neighbourhoods! In such cases, observational studies and natural experiments (in which one utilizes coincidentally matching systems) are the alternative. In many respects Snow’s determination of the Broad Street pump as the source of a cholera epidemic was a perfect natural experiment: that is, a confluence of circumstances that have conspired, naturally (without intervention), to bring about what looks like a controlled experiment, with multiple groups one of which has ‘the intervention’ and other which doesn’t. In this case the groups (the population at risk) are those who subscribed to water from the Southwark and Vauxhall Company (with contaminated water from the Thames) and those subscribing to water from the Lambeth Company (with water fed from further upstream, and without sewage contamination). What’s more, the groups appear to be well-matched in other covariates, so that no confounding would seem to be at work—i.e. if the Lambeth Company customers

²²Gillies also notes an even more surprising hypothesis involving a Priest ringing a bell on the way to give the last sacrament to a dying patient (ibid, p. 163). There were differences in the Priest’s trajectory to patients brought about by the arrangement of the rooms, that were correlated with the difference in death rates: the experience of the ringing bell was thought to have some deathly psychological effect. Again, Semmelweiss controlled for this and again found no difference.

had to pay more for that service then they would likely be richer and in better health anyway. The choice of service provider was fairly random.

Indeed, one could not hope for a better experimental setup if one tried, and Snow was well aware of this, writing:

Now it must be evident that, if the diminution of cholera, in the districts partly supplied with improved water, depended on this supply, the houses receiving it would be the houses enjoying the whole benefit of the diminution of the malady, whilst the houses supplied with the [impure] water from Battersea Fields would suffer the same mortality as they would if the improved supply did not exist at all. As there is no difference whatever in the houses or the people receiving the supply of the two Water Companies, or in any of the physical conditions with which they are surrounded, it is obvious that no experiment could have been devised which would more thoroughly test the effect of water supply on the progress of cholera than this, which circumstances placed ready made before the observer. ([Snow, 1855], p. 74)

By finding out how many houses each company supplied water to, and then finding out how many deaths occurred in each (Lambeth supplied versus Southwark and Vauxhall supplied) Snow was able to prove that if one took water from Southwark and Vauxhall one was 14 times more likely to suffer a fatal infection. Although no experimental control is exerted here, given the matching (in ‘relevant respects’), inferences are well supported. Of course, though an intervention suggests itself, there is no mechanism specified, and no ‘low level’ underlying theory. Hence, this constitutes a paradigm instance of ‘black box’ epidemiology (also known as ‘risk factor epidemiology’). It is a perfectly legitimate and often very effective way of stopping the transmission of diseases—see [Greenland *et al.*, 2004] for a recent defense. One might not know the exact mechanism underlying the transmission, or what is being transmitted (‘the agent’), but if one knows that certain behaviours or events lead to transmission then then is sufficient to be able to put an intervention into operation. Naturally, a ‘deeper’ account would most likely result in more effective (and more efficient) preventative measures, and perhaps eradication of the disease.²³ However, if necessary, black box studies can often point the way to such studies.

2.5 *Modern Public Health*

The previous discussion highlights the fact that there are two broad, apparently *competing* strands of public health: a biomedical strand and a socio-economic strand. These are woven together by epidemiology, which utilizes results concerning one to impact on the other. Modern public health places more emphasis on the ‘macrosocial determinants’ of health, than was common during the earlier parts of the twentieth century, with its focus tending to be more on the biomedical aspects of health. We saw that the view that the social and physical environment plays a role in determining health has played a crucial role in the development of

²³Note that this is a moot point. Social epidemiologists would argue that the best preventative measures come from imposing measures ‘upstream’ (i.e. focusing on more distal social causes). We shall return to this controversy several times in subsequent sections.

public health and epidemiology. Writing specifically about the treatment of tuberculosis, Winslow argued that intervening in social aspects, through the education of individuals for example, “has proved almost as far reaching in its results as the discovery of the germ theory of disease thirty years before” [Winslow, 1923]. Moreover, new kinds of theories about non-infectious diseases have emerged in the twentieth century, such as lupus, genetic diseases, and cancer. Theories concerning the social causes of diseases have begun to take shape recently too. This suggests that all components, statistics, sociology, bacteriology, and epidemiology, are necessary for the proper functioning of public health. The current trend appears to suggest greater integration of these elements in the future.

3

WHAT IS PUBLIC HEALTH?

In the previous section we got to grips with several key themes in the historical development of public health as a discipline. In this section we consider the question of how to define public health, and what assumptions and implications the proposed definitions involve.

3.1 *A Catalogue of Definitions*

In the latest edition of *The Oxford Textbook of Public Health*, public health is defined as “*the process of mobilizing and engaging local, state, national, and international resources to assure the conditions in which people can be healthy*” (IDetels and Breslow, 2005], p. 3). This emphasis on social engineering, and community action, points to a notion of ‘infrastructure’: public health provides some underlying structure necessary to support the health of the overlying public. It is a substructure that enables to prevention of disease, the support of the sick, and responses to emergencies. The substructure is often more intangible than bricks and mortar: social norms are required to make many public health ventures go.

This definition is more or less identical to the US Institute of Medicine’s, which states that the “mission” of public health involves “fulfilling society’s interest in assuring conditions in which people can be healthy” ([U.S. Institute of Medicine, 1988], p. 40). This is clearly intended to be normative as well as descriptive. The mission is implemented through “organized community efforts aimed at the prevention of disease and the promotion of health” (*ibid.*, p. 41). The organizational framework within which this is carried out includes “activities undertaken within the formal structure of government and the associated efforts of private and voluntary organizations and individuals” (*ibid.*, p. 42). Again, the emphasis is on community and action—see [Beauchamp, 1985] for an engaging study of these aspects.

The model for these definitions, and almost all recent definitions, is C. -E. A. Winslow’s canonical definition as:

The science and the art of preventing disease, prolonging life, and promoting physical health and efficiency through organized community efforts for the sanitation of the environment, the control of community infections, the education of the individual in principles of personal hygiene, the organization of medical and nursing services for the early diagnosis and preventative treatment of disease, and the development of the social machinery which will ensure to every individual in the community a standard of living adequate for the maintenance of health. ([Winslow, 1923], p.1)

The implementation of this vision is aided by informatics and epidemiology. Often there is public health action without identification of the actual causes of some phenomenon. One does not need a mechanism to use epidemiological results in decisions. Public health often proceeds without theory. This much paints an ‘aim oriented’ picture of public health.

The early WHO definition of health as “a state of complete physical, mental, and social well-being” ([1947], p. 1) has been the subject of a lot of criticism. The main objection is that it is too strong to ever be satisfied by actual individuals. However I think this misses the point somewhat. The definition should be viewed as an ideal towards which our public health activities should strive—in this sense it too is aim-oriented. Naturally, any individual’s health state will only ever be an approximation of ‘complete well-being’, but, the idea would seem to be, the state can nonetheless be shifted closer to the ideal. However, understood this way it nonetheless has its problems. For example, it does not give any information on how to achieve (or approximate) this golden state—in short it makes absolutely no reference to the determinants of health and to the methods by which one might intervene. Even as an ideal then, it is useless. Moreover, we might rightly inquire as to what “well-being” is, if not health by a different name. In the next subsections we examine the notion of public health in more detail.

3.2 *Narrow Versus Broad Conceptions*

Verweij and Dawson [2003] distinguish “broad” versus “narrow” conceptions of public health. Their focus is a demarcation issue: what counts as a public health problem? On the one hand, according to a narrow (or traditional) conception, public health is concerned with such things as the environment’s impact on health, screening programmes, infectious diseases, education campaigns, and so on. A broader conception would focus on less direct aspects, such as socioeconomic and cultural factors. I prefer to call these ‘local’ and ‘nonlocal’ since they concern factors that act directly on individuals in the former case and more indirectly in the latter case.

As Verweij and Dawson point out, the problem the nonlocalists have with the local account is that the latter fails to capture all possible (reasonable) determinants of health: “[if] public health is primarily about prevention in its widest sense [then] true prevention will have to focus on all of the causes of public health problems” (*ibid.*, p. 16). In other words, according to the nonlocal account, *everything* (no matter how non-proximal) will be included in public health if it does indeed have an impact on the health of individuals. This all-inclusive approach clearly has

its own problems. Verweij and Dawson object that “such a conception of public health could be limitless, as almost all human activities (and many inactivities) may affect health” (*ibid.*, p. 17). This is true, but one might respond by pointing to the fact that not all contributions to public health will be equally weighted. That is, the fact that public health problems are essentially limitless does not mean that they cannot be prioritized. Some factors will play a much lesser role than others, and so one should naturally devote more attention to those. Likewise, some factors may be easier to intervene in than others and so one should focus on those. There are no doubt many other prioritization criteria that one could employ to decide which factors to focus on, and these will most likely be suggested by the context.

3.3 *The ‘Public’ Aspect of Public Health*

Verweij and Dawson go on to distinguish two senses of ‘public’ in ‘public health’. Firstly, they follow Geoffrey Rose [2001] who takes ‘public’ to refer to a population of individuals so that public health refers to the collective health state of the population. In other words, the basic system of public health is the population rather than the individual. This system, like an individual, can have a health state that can be measured, tracked, compared to other systems, and modified. Secondly, they point to the methodology of implementing *interventions* at the population level, or through collective (public) action: “Taken as a whole, we propose that the practice of public health (roughly) consists of *collective interventions that aim to promote the health of the public*” ([Verweij and Dawson, 2003], p. 21). This seems to follow Winslow’s lead once again; however, in this case we have implicitly spelt out what is meant by ‘public’. The sense so given contrasts with Winslow’s idea of ensuring good health for the *individuals* in a community, for here we have what sounds like a more utilitarian notion of collective health in which the individual members are secondary.

This conception of public health seems to map onto Beauchamp’s understanding. He argues that public health ought to be an instrument of social justice [Beauchamp, 1976]. His target is the notion of ‘market justice’ (this should really be *free* market justice), the idea that a system should be left to its own devices to self-organize, and this way individual freedoms are left alone. However, as in the economic sphere, what we find in market justice worlds are (power law) inequalities: most people have very poor health and living conditions while very few (in terms of percentages) have exceptional health. A social justice understanding would seek to lessen this inequality to the point at which inequalities are no longer inequities (minimal conditions necessary for good health). The right to some specified minimal level is fundamental according to social justice approaches.

3.4 *The ‘Health’ Aspect of Public Health*

Christopher Boorse treats the problem of defining ‘health’ in great detail in his chapter; for now we just briefly discuss a few salient points as they relate to the

definition of public health—the shift to public health does introduce novel aspects to this much discussed problem. The key issue to consider here, I think, is whether there is a plurality of concepts (of health and disease), depending on whether one is studying health-related phenomena at the individual or population level, or whether these concepts are of the ‘one size fits all’ variety. We return to this in subsequent sections, for now it will be instructive to see if the standard answers to the question of what health is are equally applicable in public health.

Kitcher distinguishes between two broad conceptions in the understanding of disease (and so, by extension, health):

Objectivism: “there are facts about the human body on which the notion of disease is founded, and that with a clear grasp of those facts we would have no trouble drawing lines, even in challenging cases” ([Kitcher, 1997], p. 208).

Constructivism: objectivism “is an illusion ... the disputed cases reveal how the values of different social groups conflict, rather than exposing any ignorance of fact, and that disagreement is sometimes even produced because of universal acceptance of a system of values” (*ibid.*, p. 209).

One can find many (not necessarily equivalent) variations on this distinction in the literature. For example, Lennox [1995] speaks in terms of “reductionists” and “relativists”. One can also find a distinction between “naturalists” and “normativists” (see [Amundson, 2000] for example). However, at the root of all of these is the ‘fact/value’ distinction: does our understanding of disease and health involve value-judgements, and if so do these judgements cloud the objectivity of our talk of health and disease or can fact and value peacefully coexist? Naturalists, objectivists, and reductionists will generally wish to say that the concepts of health and disease are value-free theoretical concepts that occur in the health sciences: they will likely wish to base their philosophical understanding of the concepts on the scientific understanding of them. Normativists, relativists, and constructivists will deny this, taking the concepts of health and disease to be value-laden (with attributions of ‘disease’ reflecting our disapproval and ‘health’ reflecting what we find desirable), subject to change, and in no way ‘carving nature at its joints’. Does this debate transfer over to the public health context?

Clearly, many of the definitions of health that have been offered in the biomedical realm will not be applicable in the public health context on the grounds that many of them make specific reference to humans and the human body or to the biological function of organisms. One might extend some of the definitions by taking ‘body’ and ‘organism’ to mean any organized system involving biological components. However, if taken to be a defining condition of the general concept of disease it threatens to be over-applicable (on the grounds that there are many systems with biological components, not all of which deserve the attributions of health and disease states). Having said this, depending on our philosophical proclivities, we might wish to generalize the concepts even wider, so that they are applicable to any system whatsoever. I mentioned earlier, for example, that we

speak of the health of financial markets. However, this is usually understood to be a metaphor rather than any indication of the nature of the reality of markets.

The idea that health and disease at the population level are value-laden might be considered more appropriate since populations do not seem to be natural kinds in the way that organisms are often taken to be. However, we need not necessarily make reference to natural kinds. According to Boorse's naturalistic theory, for example, disease, is biological disfunction (and health the absence of it). Let us present the four components of his 'Biostatistical Theory' ([Boorse, 1997], p. 7, as presented in [Schwartz, 2007], p. 52):

1. The *reference class* is a natural of organisms of uniform functional design; specifically, an age group or a sex of a species.
2. A *normal function* of a part or process within members of the reference class is a statistically typical contribution by it to their individual survival and reproduction.
3. A *disease* is a type of internal state, which is either an impairment of normal functional ability, i.e. a reduction of one or more functional abilities below typical efficiency, or a limitation on functional ability caused by environmental agents.
4. *Health* is the absence of disease.

Hence, we get the view that health is simply the absence of disease, where disease is given by statistical subnormality of biological function (defined by reference to survival and reproduction) in a (stratified) reference class of organisms.

The idea that disease is abnormal function (if we consider 'abnormal' to be deviation from a normal curve) might look initially appealing since public health is based on statistics. We might take populations with health properties that are normally distributed to be 'healthy' ones. If we have a skewed distribution or one with fat tails then this points to inequalities and, therefore (perhaps) functional inefficiency. One can consider the distributions of certain properties to be indicators of healthy and unhealthy populations. As we see in §4, something like this forms the basis of so-called 'summary measures of population health', namely measures of health that roll up individual-level health data into a single number taken to be representative of health and disease in the global system.

However, so understood, this proposal faces a simple problem in that the determination of health status is a relative matter: one could have a distribution reflecting very little by way of inequality, and yet in which the individual events making up the distribution are all very low (in the sense of low life expectancy, for example)—or, in other words, the distribution is distributed 'healthily' according to this proposal, but the individual health levels are very low. Conversely, in a society where the majority of people suffer extreme depression (because of some tragic accident perhaps), then this registers itself in the statistics: it is 'normal' to have extreme depression. Those who *don't* suffer are diseased. There is an easy response to such problems: health and disease are independent of our labels. Just because we do and do not choose to call certain patterns and states 'diseased' and 'healthy' does not mean that there is a genuine correspondence. However, Boorse's definition involves the notion of normal *function*, not normality *simpliciter*. Mere

statistical normality is not sufficient to determine health; one must look at how some state is linked up with matters of function and efficiency. Lennox points to the fact that one might need to look to the population level in order to determine these latter aspects ([1995], p. 508), by looking at correlations between the variable of interest and variables that are directly linked to mortality (for it is the maintenance of life that guides Lennox's approach). However, we approach the problem, it is clear that values will enter at some point, if only in the weighting of health states in terms of severity.

A more pragmatic approach, one that I will flesh out later on, would argue that it is nonsense to try to pin down a unique definition of health and disease independently of one's interests and the uses to which the concepts are being put to. This filters through into the (operational) construction of population measures of health and disease—that is, the approach maps onto actual scientific practice. Here there are very many such measures, and one can pick and choose according to task: if one is interested in resource allocation then one can focus on an approach to health that is insensitive to many aspects that one could not ignore if one was interested in equity issues, or in whether it was right to intervene in some property. We return to this issue again in the context of health measures in §4.

3.5 *Prevention versus Cure*

A corollary to the individual/population distinction (discussed in §3.3), though not a strictly necessary one, is the distinct aims that are associated with individual and population level approaches to health. There is an epistemic difference embedded in this difference in aims: the preventative measure concentrates as much focus on 'unknowns' as it does on 'knowns'. In the clinical encounter, the focus is on some presentation of disease in an individual with the aim of diagnosing the disease and finding a cure. Public health, on the other hand, will tend to focus on a disease-free population, with the aim of keeping it that way. The notion of prevention in public health overlaps with the issue of aetiology. A public health programme will usually isolate causes in a different way than to clinical health practice in the sense that a more 'distal' cause will be deemed responsible for some health problem so that interventions (if used) will be applied at different sites.

Furthermore, as Rose points out ([Rose, 1992], pp. 12-13), according to the preventative strategy the benefit to individuals can be very small (and even negative: say a loss of earnings). As Rose puts it himself: "*a preventative measure that brings large benefits to the community offers little to each participating individual*" (p. 12). The gain is at the population level. Rose accepts the tough side of prevention, including the seeming necessity of alteration of norms of behaviour and the social fabric to achieve the desired effects.²⁴ Naturally this level of control invites ethical commentary—but that is not my concern here. Rose's point is that a

²⁴For example, by altering the social status quo for smoking one does not need the distant incentive of better health, one can rely on the immediate social disapproval that smoking generates.

truly preventative science must involve a thorough knowledge of the determinants of what it is one seeks to prevent: disease. And, likewise, of what one seeks to promote: health. Once these have been isolated then one can intervene so as to get the optimal situation. In the ideal case one will no longer need such strategies as screening programmes and such like: the root cause will longer be in operation.

This seems rather over-optimistic. Even if the causes could be isolated, it isn't clear that one could persuade a whole population (or even most of a population) to engage in practices that are often of minimal individual benefit. On the latter problem Rose suggests that one adopt a "high-risk strategy" involving the specific targeting "of those individuals who are judged most likely to develop disease" (*ibid.*, p. 13). The strategy depends then, as one might expect, on the distribution of risk in a population. If the risk is spread fairly uniformly over a population then a "population strategy" (a mass strategy) is appropriate. If one can find clusters of risk, then one can adopt a targeted approach.

Rose adopts (though doesn't argue directly) for a holistic position well opposed to methodological individualism. This, he says, provides the "sociological basis" for his idea of a population-level prevention strategy ([Rose, 1992], p. 95). As he notes, the view has noble ancestry that he traces to Durkheim. In the health context it involves the thesis that "healthiness is a characteristic of the population as a whole and not simply of its individual members" (*ibid.*, p. 62). The problem is, how to turn this from mere talk into a practical framework. How, for example, does one measure this state if not by measuring the states of the individuals and aggregating the results? Even in the case of Durkheim's classic analysis of suicide the measurement and construction of population-level properties is done 'upwardly' via the individual members.

The division common in the social science between methodological individualism and holism arises in the field of public health, then, and can have a bearing on practical matters. For example, if one adopts a holistic approach, then one will focus on the population as the object of investigation, with its own properties to be measured, intervened in, and evaluated (*cf.* [Weed, 2004], p. 532). If, on the other hand, one adopts an individualistic approach, then one will focus on the individual people in a population, and measure and intervene in their properties. This, of course, colours the explanation one gives of the causes of health and disease. A methodological individualist will seek to explain by drawing attention to factors having to do with the behaviour of people; holism will look to collective factors, or the wider social and physical environment, that drive that individual behaviour.

3.6 *Scientific Foundations of Public Health*

Epidemiology and Biostatistics form the scientific core of public health, the foundations on which decisions and actions are made. As Detels and Breslow state, epidemiology "is the scientific method used to describe the distribution, dynamics, and determinants of disease and health in human populations" ([Detels and

Breslow, 2005], p. 10).²⁵ Epidemiology is the obvious choice for the scientific basis of public health: both take populations as their primary objects.²⁶ Public health is grounded in statistical methods. Given aggregates of measurements of health-related properties they enable the identification of health problems (bad trends, inequalities, etc.), which in turn allows for evidence-based promotion and intervention, for informing efficient health economic policy, and assessing the impact of interventions. Epidemiology is in the business of inferring conclusions about the distribution of health (and disease) in a population. Part of this involves uncovering the determinants of health and disease. To do this epidemiologists attend to relative frequencies between various patterns (smoking and lung cancer, for example). We can roughly view the epidemiologist as giving us a measure of public health (or at least some components) in the sense that we are able to see how these components are distributed over the individuals in some population (whose members are not necessarily geographically or temporally coincident). As mentioned, one can look at relationships between components, and their trajectories over time and place, and form hypotheses about the connections. It is these hypotheses that can often lead to public health action in the form of interventions (or in the form of analytical studies).

Public health and epidemiology are, then, very tightly woven together. John Last, in his dictionary of epidemiology, broadens the fairly standard definition given above to include *action*, defining it as “the study of the distribution and determinants of health related states or events in specified populations and the application of this study to the control of health problems” ([Last, 1995], p. 42). That closes the gap between public health and epidemiology making them almost identical. It is more common, I think, to view epidemiology as a providing an evidential basis on which public health acts and bases its decisions (concerning the management of resources, and so on). Milton Terris makes the ‘evidential-basis’ role particularly explicit:

1. To discover the agent, host and environmental factors which affect health, in order to provide the scientific basis for the prevention of disease and injury and the promotion of health.
2. To determine the relative importance of causes of illness, disability, and death, in order to establish priorities for research and action.
3. To identify those sections of the population which have the greatest risk from specific causes of ill health, in order that the indicated action may be directed appropriately.
4. To evaluate the effectiveness of health programs and services in improving the health of the population. ([Terris, 1993], p. 142)

Put in this way, public health is nothing but applied epidemiology. However, Sander Greenland strongly distinguishes epidemiology from public health, latching on to exactly the action-based component raised by Last:

²⁵Note that “disease” here is an umbrella term covering all manner of health-related events or phenomena: HIV, smoking, teenage pregnancy, bullying, and so on. In fact, it needn’t be a negative event: one might be interested in what is causing some positive trend, such as an increase in exercise amongst young people.

²⁶A lovely little book (just 69 pages) that introduces the essentials of epidemiology for the ‘uninitiated’ is [Coggon *et al.*, 1997].

[P]ublic health is not a science, but a form of social activism, one whose benefits appear profound enough to society that it is institutionalized and heavily subsidized by governments. A public health activist promoting or searching for an action will be concerned with communicating, his or her own opinions, evaluating the opinions of colleagues, and influencing the opinions of governmental figures and the public. ([Greenland, 1988], p. 96)

Epidemiology, by contrast, is seen as unbiased, objective, and unfettered by any kind of social incursion—thus following Weber’s dictum that “it is the duty of the man of science to remain silent ... on value questions” ([Weber, 1920], p. 188). However, this is a rather naive view of how science works, as countless philosophers, historians, and sociologists of science have demonstrated, whether through feministic challenges, Kuhnian challenges, or many others. Epidemiology, more so than many other sciences (given its statistical basis), is very much invested with values. Indeed, if we are persuaded by Donald MacKenzie [1981] arguments, even the mathematical foundations of epidemiology are infiltrated with ‘interests’ due to their incursion into the foundations of modern statistical theory.

4

HEALTH MEASUREMENT AND HEALTH MEASURES

Health scientists and professionals want to be able to measure health for a variety of reasons: to track changes in health, to identify problems, to identify causes and risk factors, to check how an intervention has performed, and to perform cross-comparisons between groups. To do this we need to have a clear idea of what we are measuring.

4.1 *Measurement and Standardization*

Measurement is, as Grigory Barenblatt succinctly puts it, “the direct or indirect comparison of a certain quantity with an appropriate standard, or, to put it another way, with an appropriate *unit of measurement*” ([Barenblatt, 2003], p. 12). If we are talking about measuring the health of the public this must involve comparison with a standard too: a ‘unit of health’. It is clear that this is not going to be a *fundamental* unit; rather, it will be a complex derived (by aggregation) from other, more fundamental units. However, it is still a problem to say what this thing to be measured is. The method of definition in the context of public health is to give an operational definition. In order to ensure objectivity (or as near to objectivity as possible) the focus is on the individual, and the wider social context is ignored: health is viewed as what something that happens under the skin. Even the health of populations is to be reduced to the functioning of individual bodies (a form of methodological individualism in the context of the health sciences). That is, ill health (at least according to most measures) is taken to be the reduction in individual human function caused by disability or, alternatively, a reduction in the well-being of individuals.

This a natural position to adopt: one demands standardization in measurement. The environment (be it physical or social) varies considerably, so it is desirable to have a measure that does not take account of that, or that is ‘insensitive’ to it. This is clearly especially vital if one wishes to conduct comparative work on different populations, to measure health inequalities for example. However, this standardization misses out on the crucial role played by context (see [Allotey *et al.*, 2003]): a broken leg in Canada is a fairly trivial matter, and one expects the health discounting to be minimal. However, a broken leg in a developing country is a far more significant matter: there may be no sickness benefit to draw from, no easily accessible health care services, and so on. In this case the impact ought to involve a weightier discounting. Yet we are dealing with the same event, from the point of view of the measure: a broken leg. The individualistic health measure will treat the cases as the same in terms of their value.

One can come up with more examples that do not involve inter-population comparisons. Consider two individuals from the same population, both with a sprained finger, one of whom is a concert pianist, the other is, say, a teacher. Clearly these individuals will weight the severity in different ways. Severity is not an objective fact of the matter, it depends very much on the individual. Clearly then, one’s health measures have to take account of more than the local considerations of individual bodies. However, the determination to restrict health to individuals resulted in the WHO introducing an idealized “uniform environment”, and then considering (in the population health case) the “capacities” of individuals within this environment. Health is then defined as having the capacity to perform certain tasks within an idealized (though not necessarily non-actual) environment. In other words, the environment is introduced into the definition of health, but it itself is standardized. The idea is to switch off the role played by the environment. We return to the problems with such measures in §4.4, before we get to that let us first say some more about the health measures.

4.2 *Summary Measures of Health*

Health measures aim to give a numerical representation of health, be it in an individual or in a population. A summary measure of health will summarize the health states of the individuals in a population. Health status indicators (those features that might go into a health measure) come in various kinds, and can refer to individual or population properties, and such things as waiting times, resources-to-demand, and so on. Aggregate measures might take a number of these and average over them to produce an index akin to the Dow Jones Industrial Average [DJIA].²⁷ Amongst other things, these measures are used to assess the impact of

²⁷For those not acquainted with the basics of financial markets, the DJIA is an index composed of thirty blue chip stocks. Initially, the value was computed by simply adding together all of the company stock prices and dividing by 30. The number is is then taken to provide a (fairly rough) measure of the economy’s state—it is sometimes said to provide a measure of the economy’s health! There are many such indices one can use, largely depending upon one’s interests.

interventions and policy. Though a little dated, [Bergner, 1985] gives an excellent overview of the measurement of health status.

In public health, then, (or population health, more generally) one does not (thus far) measure a property of some system ‘the public’. Rather, one measures properties of individuals and then aggregates the data that results. One is then left with a single number, a statistic, that is intended to provide the requisite (‘summary’) measure of the larger system. This can be understood in terms of ‘social indicators’, i.e. statistics that are intended to be calibrated to the quality of life of the individuals whose relevant properties are aggregated in some way (*cf.* [Michalos, 2006], p. 344). Vital statistics would be an example of social indicators, as would the various financial indices. They would be examples of *objective* indicators since they rely on facts that are independent of ‘internal’ states of individuals²⁸—whether they are ‘value-free’ is another matter, one that we return to below. Quality of life indicators might refer to the subjective reports of individuals, such as degree of happiness, and in this case they are *subjective* indicators. As we saw, the job of representing (or indicating) the overall health state of a population was, from very early on, filled by mortality rates. However, this misses a major component: morbidity. Summary measures of population health were devised in the 1960s to take account of both mortality and morbidity.

One unit for measuring health (and so the effectiveness of an intervention) via the notion of quality of life, is the ‘quality-adjusted life year’ [QALY]. This gives a measure of the quality and quantity of life.²⁹ These depart from the more objective measure in terms of life expectancy (alive = 1 or dead = 0) by introducing a continuum of states between 0 and 1, where the value 1 is taken to represent ‘perfect health’ and 0 still represents death. More precisely, the value 1 is assigned to one year of perfect health-life expectancy. If a year contains less than perfect health-life expectancy, then a value of less than 1 is assigned. Computations are then very straightforward. For example, if we intervene to extend a patient’s life by 5 years, but the quality of life for those 5 years is half of the perfect quality, then we simply compute the QALYs gained in the intervention as 5×0.5 . Hence, 2.5 QALYs will have been generated by the intervention. Such medical mathematics informs medical decisions, since one can use the values so computed to work out which interventions will have the greatest ‘yield’ in terms of QALYs.

In [Murray *et al.*, 2000] a distinction is made between ‘ideal’ and ‘actual’ health. “Ideal” has two senses here: it signifies the health state we *want* individuals to attain (namely a full life at full health: with full defined appropriately); and it also serves as kind of limiting case of real health. It is ideal in that it doesn’t exist in reality, but only as a concept, or perhaps in some other possible world. An aggregate (or summary) population measure can be constructed from measurements of individuals’ by simply adding all of the differences between the ideal and actual health of individuals. This gives the years of life lost in a population [YLL]

²⁸See [Stouman and Falk, 1937] for a thorough, early review of objective indices of health.

²⁹The on-line resource *Bandolier* refers to the QALY as “a slightly mythical creature of dubious parentage” (<http://www.jr2.ox.ac.uk/bandolier/band24/b24-7.html>).

and can be used to measure the level of disease. The way of calculating a figure is fairly straightforward in practice, if not philosophically. Given a specified ‘ideal’ age (80 for males; 80.2 for females), the YLL is the difference between the actual life and the ideal age. Each subsequent year is weighted slightly less than the previous year. This way of speaking (i.e. of ‘ideal’ states) fits in with the WHO’s definition of health as “a state of complete physical, mental, and social well-being and not merely the absence of disease and infirmity”. Naturally this is a limiting case: no individual, I presume, could ever hope to attain such a state.

One can also consider a measure involving morbidity the years lost due to disability (with mortality as a limiting case). In this case one reduces the value assigned to each year for which the individual was disabled (with an ordering of severities of disability corresponding to an ordering of the amount of the reduction). Population health is then the sum of the years lost to premature mortality and to morbidity. For a given population of individuals (including those who are recently deceased), when we sum this figure for all events during a single year then we get the measure known as the DALY: the Disability Adjusted Life Year. One DALY represents the loss of one year’s worth of healthy life. The DALY again involves an ideal reference population with life expectancy (at birth) as above. The burden of disease is then computed as the difference between this ideal state and the state of a population’s health (recorded in DALYs).³⁰

There are numerous problems with these measures, largely stemming from the lack of an objectively-agreed upon weighting of events. For example, there might well be states of life in which the suffering involved is worse than death. The measures that are constructed are clearly not carving social reality at the joints (although they may serve to define those joints arbitrarily, or rather, by convention). The DALY is most decidedly not a natural kind.³¹

4.3 *Classifying and Measuring Health States*

Summary measures of population health perform multiple useful roles: they can enable the cross-time or cross-place comparison of health state; they can identify problems, and they can tell us if an intervention worked. Dennis Fryback ([1998], p. 43) notes that there are three steps involved in the construction of health measures:

1. decide on the aspects of health that will be included in the (discrete) classification scheme
2. construct a mapping between the health of humans and the states in the

³⁰Note that QALYs are a subclass of the more general measures known as HALYs (‘health-adjusted life years’). A nice review of these issues can be found in [Reidpath, 2007]. A compendious volume dealing with a host of issues relating to population health measures is [Murray *et al.*, 2002b].

³¹Of course, ethical issues loom large; however, I want to steer clear from these in this chapter. For an excellent review of the moral implications of summary measures, see [Brock, 1998].

classification scheme³²

3. assign weights to each health state included in the classification to be used to compute population health

The classification scheme is clearly going to involve massive abstraction from ‘real’ human health states. How much abstraction will be determined by the use to which the measure is put. The earliest measures simply classified health according to two values, ‘alive’ and ‘dead’. This measure will be adequate for any task that requires only mortality rate data. A finer measure will need to differentiate various sub-states within ‘alive’—clearly ‘dead’ has no relevant fine structure. Again, the amount of differentiation will be a choice determined by the level of detail one needs for the use.³³ If one wants to know the best way to deliver some mental health intervention, then one will demand a measure that takes account of this. Health, in other words, is multi-dimensional. There are different perspectives that one can adopt towards the health of a system. Given this pragmatic way of conceptualizing health, I don’t see that it makes much sense to try to adopt a single definition, as is the trend in philosophical discussions. Biological disfunction, for example, is but one aspect of health. In some cases it might be an appropriate definition or measure of health, in others it will not be. A genuinely naturalistic approach to the question of what health is ought to follow what our best science has to say on the matter, and it appears that this pragmatic multifaceted approach is the answer given by that science.³⁴

One can usefully view this situation through the lens of Giere’s scientific perspectivism. The idea is to view the various measures as so many scientific instruments restricted to ‘viewing’ only certain aspects of the systems they are directed at. Here too, I think, “one’s theoretical perspective ... depends on the kind of problem one faces”. Different problems demand different perspectives.

The old problems, however, do set in when it comes to weighting the various health states separated by the classification system. There is no objective way of doing this, and an infinite variety is clearly possible. But I think it is best to run the same argument just given: how one assigns health state weights depends on what problem one has in mind. For example, if one is concerned with the

³²Given that the classification system results in a measurement system, we then get an operationalization of our health concept.

³³For example, the (summary) health measure known as ‘HUI-Mark III’ (where ‘HUI’ = ‘Health Utility Index’) has the ability to distinguish between 972,000 distinct health states pertaining to the physical, mental and social dimensions of health. For many purposes this amount of complexity would be simply unnecessary.

³⁴Sander Greenland [2002] argues that health’s multidimensional nature ought to be reflected in the scrapping of scalar measures in favour of multidimensional measures (vectors whose components represent the different aspects of health). Hausman [2002] argues that while there might be some theoretical attractiveness to this proposal, it is not practical and would most likely be ignored by health policy makers. That is probably correct, however, such a measure would have many uses not covered by unidimensional measures. Unless we wish to fall under the philosopher’s spell of the one unique measure that *best* represents the *true* health state, then we ought to accept these measures as providing a perfectly acceptable additional perspective on the health state of a population.

‘productivity’ of the population, then physical disabilities will presumably be more heavily weighted than self-esteem, say. If one wants a measure that will appear the most democratic, then one might wish to base the weightings on average values assigned to various health states as taken from a survey. The QALY discussed above is based on cost-effectiveness issues and so naturally it bases health state weights on utility. If one is happy to say that there is a plurality of health systems (i.e. that there is no classification system that is the most objectively true) then the values that inevitably go into the weightings do not cause the kinds of problems they cause for other so-called naturalistic accounts: one is not privileging one system as ‘fact’ in the first place.

4.4 *The Inadequacy of Aggregative Measures*

As we have seen, aggregate measures are based on the view that the health of a system is determined solely by the health of its individual parts; if one knows the latter then one knows the former. Daniel Reidpath [2005] argues that aggregate measures of public health are inadequate on the grounds that the aggregate does not provide any information on how health is spread out over the population. To work this out one needs to look at the shape of the probability distribution of health states over the individuals. Presumably we would consider a population in which 5 % of the population have enormously high health values (relative to some measure, life expectancy, say) and the remaining 95 % have relatively poor health (giving a very skewed, fat-tailed distribution) to warrant a lower health value than one in which there is a relatively high and even (or normal) distribution of health. However, one can construct all manner of distributions of health of a population of individuals many of which will be grossly iniquitous, yet that correspond to one and the same aggregate value, on account of possessing the same average. That is to say, the value assigned to a summary measure of health is multiply realizable by (infinitely) many ‘spreadings’ of health and disease over the population, (infinitely) many of which are grossly iniquitous. An individual asked to choose which of the populations they would like to belong to would not be indifferent, therefore, argues Reidpath, the distribution is relevant to the way we go about measuring the health of a population.

Take a simple toy example. Suppose we have two populations with the same number of people in each. Suppose that we aggregate the health of the individuals and come up with figures of 100 QALYs as the aggregate measure in both cases. Now, if we were to use this to determine which population were healthier would we have to say that they were equal. The measure is not sensitive enough to detect finer details. However, the finer details are all important. In one population 80% of the QALYs might be generated by 20% of the people (so that most people are very unhealthy), whereas in the other population the situation is more balanced, with, say, 80% of the QALYs generated by 80% of the people. Rather more visually, we might consider two populations, one of which were composed of two types of human, giants and dwarfs, and the other with a broader range of

heights. We could set this up so that the average heights of the populations were identical, and yet the average figure is not giving us the kind of information we want from a measure of, say, what the sizes of the people are like in the respective populations.

The problem, then, is to do with the aggregative methodology which involves simply taking the individual level data and summarizing it. The methodology is individualistic: population-level phenomena are seen as nothing but the synthesis of individual-level data. The business of the distribution of health and disease is then seen as a quite separate issue. Reasons given, for this separability of level of health and distribution of health, by those who construct the measures range from ‘tradition’ (i.e. the health statistics tradition stemming from mortality rates—see [Murray *et al.*, 2002a], p. 752) to ‘communicability’ (in the sense that it is easily assimilated by the general public—*ibid.*).

Reidpath argues that this is the wrong way to conceptualize health at the population level. He does not suggest a holistic approach as such, but instead one based on development economics, that blends population level data about the distribution of health (or, rather, well-being) with the individual level data. That is, the distribution of socially relevant properties over the individuals in populations should not be separated from the measurement of that property at the population level. The implication of this seems to be that aggregate measures miss out on ‘emergent’ features of population health. As Reidpath puts it, “there is information relevant to the health of a population that can only be derived from the gestalt that cannot be ascertained from the sum of its parts” ([Reidpath, 2005], p. 879). More is different in population health, we might say (here following [Anderson, 1972]).

Reidpath is certainly right that the summary measures are not sufficient for many purposes. However, clearly sufficiency depends on the task to which the measure is being put. Sometimes a more coarse-grained measure of health might be all that is needed. Other times, this will be inadequate and the fine structure of a population’s health will need to be incorporated. That is, one cannot, as Reidpath does, speak of insufficiency *simpliciter*; insufficiency is tied to a specific goal.

4.5 *Cross-Comparison of Health Categories and the Ranking of Health States*

Daniel Hausman draws attention to a problem of the cross-comparison of what appear to be incommensurable categories of health state. He presents the example of comparing an individual with a mild learning disability to an individual with quadriplegia. How is this comparison to be made? As Hausmann puts it: “How can one compare units of mobility with units of cognitive functioning? How can one measure the ‘distance’ of health states from H [complete health]?” ([2006], p. 251). As he points out, the usual way of comparing is via *evaluation*. One makes a value judgement about *which would be better or worse*. Hausman sees no way past this state of affairs: “Measurements of population health are measures of how

good or *bad* population health is, and the goodness or badness of health depends on the physical, technological, and social environment and on the characteristics of people's activities and objectives as much as they depend on facts about stomachs or brains" (*ibid.*, p. 252). Indeed, the DALY, and many other health measures, involve a ranking of disabilities according to the impact on functional capacity (the capacity to do certain things, like walking a certain distance). That this weighting is explicit is touted as an advantage over QALYs since the value judgements are open to view and modification—see [Murray, 1994].

John Broome argues that the ranking of health states is done according to how the state contributes "to well-being" ([2002], p. 94). Well-being is a bad basis for the evaluation of health states since it is too vague. Hausman argues, instead, that the ranking is, in practice, done according to *preference* (*ibid.*, p. 253). This is a "faulty" method on account of problems (false beliefs and cognitive deficiencies) with preference (*ibid.*, p. 264). Preferences are the outcome of some other reasoning processes: these, Hausman argues, ought to be investigated by those concerned with the evaluation of health states. This all points to the fact that measuring and comparing health states is a difficult enterprise, not just technically and conceptually, but also morally.

5

HEALTH INEQUALITIES

Health inequalities refer to differences in health state between units, or the variation in a population of subunits. A large part of public health is devoted to reducing such inequalities. Many of the same problems faced with measuring health *per se* can be found in the context of measuring health inequalities. The WHO's measures use a variation on the Gini coefficient (an index used to measure wealth inequalities) in which the health of every individual in a population is compared to every other from the same population:

$$(1) \quad \mathcal{I}(\alpha, \beta) = \frac{\sum_{i=1}^n \sum_{j=1}^n \|y_i - y_j\|}{2n^2 \mu^\beta}$$

Here, the parameters α and β control the contribution of the absolute difference pairs of individuals and the weight of the mean respectively. The individuals i and j can be people, social groups, or entire populations, as appropriate, and y_k represents a health measurement outcome performed on individual k —the health measure can take a variety of forms, as discussed previously. The term μ represents the average health (or expected health, relative to some measure) of the entire population (or, given very large n , some well-chosen sample). The value is proportional to the difference between the (area under the) perfect equality curve \mathcal{E} and the (area under the) Lorenz curve \mathcal{L} : $G = 1 - \frac{A_{\mathcal{L}}}{A_{\mathcal{E}}}$. A value $G = 0$ represents a situation with perfect equality and the value $G = 1$ represents perfect inequality.

Inequality, formally, concerns the distribution of a property over a population of individual units or between populations. What measure one chooses will to a large extent depend on the units in question. Individual people, cities, social groups, gender, countries, etc... So the health measure will be guided by context: if one wishes to compare countries, and search for inequalities at this level, then one might use an index built by averaging over a bunch of health related properties. In this case one is treating the countries as individuals with their own properties. One is then concerned with the distribution of the values of this property over a domain of countries. Hausman *et al.* point out that focusing on “contrasts between social groups ... hides inequalities within groups” ([Hausman *et al.*, 2002], p. 184). The point here is that one could have a pair of countries with the same value relative to our chosen measure, but have radically different distributions of individual health within the population.³⁵ (We have been here before, of course: this is essentially the same objection raised by Reidpath in §4.4.) Writing on the subject of economic inequalities, Charles Wheelan writes: “If the pie is growing, how much should we care about the size of the pieces?” ([Wheelan, 2003], p. 115). In other words, if the economy is steadily growing, so that everyone is becoming ‘better off’ in an absolute sense, then does it matter that the *spread* between rich and poor is simultaneously increasing? There are many things one can say about this: one might argue that the reason the economy is growing steadily is precisely because of those at the top earning more: they are the companies and individuals who are investing more, in research, technology, and enterprise. On the other hand, you could argue that the more people there are earning more pushes prices of items up that lie out of the range of those poorer people, thus making the inequalities even more extreme. Either way, the challenge of Reidpath and Hausman *et al.* needs to be answered.

Beyond these issues of distribution, the key philosophical problems with the research on health inequalities are to do with the level of support the data give to the possible explanations of the inequalities. There seems to be some genuine underdetermination going on. This seems to be what underlies virtually all of the objections raised by Forbes and Wainwright [2001] in their philosophical investigation of health inequality explanations: the data does not uniquely determine one explanation. Nor, they imply, can one give, on methodological grounds, an inference to the best explanation. Forbes and Wainwright further contend that a latent positivism underlies much of the health inequality literature. They argue that the extant explanations of health inequalities are too data-dependent. They claim to argue for a ‘realist’ position in [Wainwright and Forbes, 2000]; however, the fact that they deny the links between the mathematical representations used by health scientists and the reality it is supposed to be representing does not appear to match any of the standard brands of scientific realism. They seem to confuse realism

³⁵Here they are following the analysis of Murray *et al.* [1999] according to which “health inequality should be defined in terms of inequality across individuals” (p. 541). I say we should take ‘individual’ as applying to people, groups, and populations in general according to ones interests and goals—this seems to be implied in the way the WHO measure is constructed.

with a belief in unobservables that have no observable effect whatsoever, direct or indirect. Of course, realist positions are committed to the reality of unobservables, however, they will only be committed to such things when they generate some kinds of effect or are necessary to explain some observable phenomenon. Completely detaching from data is a dangerous position to espouse in any science, let alone in health research. For this reason, philosophers would do well to scrutinize these arguments.

6

HEALTH MEASURES AND NATURAL KINDS

In his classic discussion of the problem of defining health and disease Lester King wrote that “Science, in studying relations within the total environment, cares not a whit about ‘health’” ([1954], p. 193). What he meant by this is that the data alone do not represent a state of disease until we have given it that interpretation:

Disease is the aggregate of those conditions which, judged by the prevailing culture, are deemed painful, or disabling, and which, at the same time, deviate from either the statistical norm or from some idealized status. Health, the opposite, is the state of well-being conforming to the ideals of the prevailing culture, or to the statistical norm. The ideal itself is derived in part from the statistical norm, and in part from the ab-normal which seems particularly desirable. ([King, 1954], p. 197)

The problem of how to measure health (or disease) is intimately connected with what we take health (or disease) to be. Even if we can agree on the definition of individual health, it is a further difficulty to figure out how to aggregate health to get a handle on the health of a population. Furthermore, the debate between normativists and naturalists reasserts itself at this level. Naturalists will claim that the aggregate measures satisfy value-free construction methods, whereas normativists will disagree. An important and interesting question for philosophers to tackle would be to investigate the links between one’s position with respect to individual health and disease and health and disease and the aggregate level. Intuitively one would expect that if one adopts the position that the definition of health and disease at the individual level is value-laden then this would be transmitted to the aggregate definition. However, given the fact that the measures involve weightings of specific diseases and disabilities, one might expect that the transmission must fail. The weightings and measures, as I argued above, are not themselves stalking out natural kinds, they are bound to the use to which they will be put.

Daniel Sulmasy [2005] argues, in any case, that disease itself is not a natural kind, but that it involves *reference* to natural kinds. Humans, for example, are natural kinds and disease “is a classification of a certain state of affairs that can occur in members” of this kind (p. 496). Although Sulmasy argues that one must refer to multiple individuals in order to infer that some phenomenon (some illness) is a disease, the account is nonetheless an individuals-based account. It is not easy to see how this could be extended to public health. The only way I can see how it would be possible would be to argue that certain populations are natural kinds, and

then one might identify certain patterns than arise repeatedly in such populations as diseases.

Beyond this, it has been argued that population-level features can have a direct bearing on the way we conceive of these notions:

Characteristics of populations also influence our very definitions of what is health and what is disease. Rose notes that what we consider normal is influenced by what is prevalent. 'What is common is all right, we presume.' One implication of this is that social facts may also influence disease incidence in the broadest sense, by determining what we consider to be a disease. Social facts influence our expectations of how many aches and pains are normal, how long we expect to live and what we expect our bodies to look like and our minds to accomplish. Bodily aberrations and biological variants can come to be defined as diseases or redefined as normal. Obesity, intersexed conditions, senility, acne, post-traumatic stress and gender identity disorder are just a few examples. ([Schwartz and Diez-Roux, 2001], p. 439).

In other words, since values are linked to the population, we cannot escape considerations of population in the debate over the nature and definition of health and disease.

I have been hinting at a pragmatic response to the debate in much of the preceding discussion. This would involve a rejection of the distinction between facts and values on which the debate rests. An earlier attempt at a pragmatic definition was made by Fanshel [1972]. However, his approach puts the pressure on the notion of functional states (and dysfunctional states) and then weights these, using a notion of ideal function (thus bringing in the debate once again). The philosophical debate over the ontological status of disease and health involves the extremes of realism and relativism. I think the work on population measures of health and disease suggests a pluralistic approach that avoids the excesses of these extremes. The approach is well-stated by Giere (here in the context of modeling water):

consider the simple case of water. If one is studying diffusion or Brownian motion, one adopts a molecular perspective in which water is regarded as a collection of particles. But the situation is far too complex to adopt a Newtonian perspective for individual particles. Instead, one adopts a statistical perspective in which the primary variables are things like mean free path (the average distance a particle travels between collisions). However, if ones concern is the behavior of water flowing through pipes, the best fitting models are generated within a perspective that models water as a continuous fluid. Thus, ones theoretical perspective on the nature of water depends on the kind of problem one faces. Here employing a plurality of perspectives has a solid *pragmatic* justification. There are different problems to be solved and neither perspective by itself provides adequate resources for solving all the problems. ([Giere, 2006], pp. 33–34)

There are multifarious uses to which health measures are put: cross comparison of health, determination of healthcare expenditure, targeting of interventions, etc. I advocate a transferal, of this way of thinking about health measures, to the philosophical problem of understanding what health and disease are. There no reason to think that the measure used in each case need be the same in all cases. To think this would be to suppose that there is some 'One True Health State' in the world that these measures are trying to latch on to.

CAUSALITY IN PUBLIC HEALTH

Public health, even more so than medicine, is fundamentally concerned with finding the causes of those phenomena that cause disease and impact on health.³⁶ Such phenomena can be singular or general. That is, one might wish to ascertain the cause of a particular outbreak of SARs or one might want to understand the general mechanism by which SARs outbreaks occur and propagate themselves. The population-level focus alters the causation debate's compass somewhat. In the public health context, as Maxwell Parkin and Bray put it:

"Cause" is a relative concept, that only has meaning [in] terms of its removal being associated with diminished risk of the disease, and, in this context, it is just as relevant to improve educational levels in a population as a means of reducing infection by HIV as it is to identify the mechanisms by which the virus enters the host cell. ([Maxwell Parkin and Bray, 2005], pp. 158–9)

Nancy Cartwright makes a similar point, stating that "although causes may not be universally conjoined with their effects, at least they should increase their frequency" ([Cartwright, 1989], p. 55). This is, of course, the hallmark of a *probabilistic* conception of causality. The approach to causation in the context of health intervention research is certainly probabilistic (or *statistical*): one often works 'backwards' from data, containing patterns of association between variables over a sample (e.g. joint distributions), to causes. As Holland explains, the emphasis in statistical models of causation is on "*measuring the effects of causes*" rather than "the causes of effects" ([Holland, 1986], p. 945 - emphasis in original).

However, as Judea Pearl points out, the data isn't sufficient, by itself, to permit causal inferences:

There is nothing in the joint distribution of symptoms and diseases to tell us that curing the former would or would not cure the latter. [Moreover,] there is nothing in a distribution function to tell us how that distribution would differ if external conditions were to change ... because the laws of probability theory do not dictate how one property of a distribution ought to change when another property is modified. ([Pearl, 2001], p.191)

It is clearly crucial that we know the direction of an association, and how the association would change given different background conditions: the success of intervention research *depends* on such information. Hence, some extra piece needs to be added to the puzzle in order to allow for the extraction of valid causal inferences from mere statistical data. In *The Oxford Dictionary of Statistical Terms* 'causality' is defined as follows:

Philosophically difficult notion of relation between an explanatory variable and a response. Older discussions were non-statistical and involved some notion of necessary and sufficient condition for the response. There are a number of variants of a statistical definition of causality (Holland, 1986 ["Statistics and Causal Inference", *Journal of the American Statistical Association* **81**: pp. 945-960]). In one the cause must be in some sense prior to the response and alternative allowable explanations of the statistical independence involved must be excluded. In another there is a notion that the possible cause can conceptually be manipulated with a consequent systematic effect on the response. ([Dodge, 2003], p. 59)

³⁶I say 'even more so than medicine' because medicine is generally more concerned with treatment than prevention.

The latter aspect, pertaining to manipulability, has formed the basis of many contemporary accounts of causation, philosophical and otherwise and gives us the “extra piece” alluded to above. For example, Woodward and Hausmann write:

[M]anipulation is crucial to our conception of causation and to the contrast between causation and mere correlation. When X and Y are correlated and X does not cause Y , one expects that when one manipulates X , the correlation will break down. By contrast, if X causes Y , one expects that for some range of values of X , if one is able to manipulate those values, one can thereby control the value of Y . (Hausman and Woodward, 2004, p. 847)

Following Glymour and his team, Woodward and Hausmann understand interventions as processes that (directly) manipulate some variable (the response variable) so as to ‘detach’ the manipulated variable from its other causes (i.e. its ‘parents’)—i.e. the response variable is rendered *probabilistically* independent of any other causes. This condition they call ‘modularity’, and, along with its close relative the ‘causal Markov condition’, it has been the subject of much recent controversy (primarily having to do with whether real systems of interest are themselves modular).

This manipulationist account is not new, however. In their popular epidemiology textbook, MacMahon and Pugh define a causal association as one in which “an alteration in the frequency or quality of one category is followed by a change in the other” ([MacMahon and Pugh, 1970], pp. 17-8; cited in [Schaffner, 1991], p. 206). Likewise, Rubin ([1986], p. 962) holds up the motto “no causation without manipulation” as a “critical guideline for clear thinking in empirical studies for causal effects”. The manipulationist account has also been central to the study of experimental design. For example, Cook and Campbell ([1979], p. 36) write that a “paradigmatic assertion” regarding causal relationships is that by *manipulating* a cause we will manipulate the effect: “Causation implies that by varying one factor I can make another vary”.³⁷

In the context of public health research, at least, when causes get more indirect or ‘distal’, they are labeled ‘risk factors’—*cf.* Schaffner ([1991], p. 206). In other words, many health researchers are reluctant to use the term ‘cause’ when the association is probabilistic. For example, Kleinbaum *et al.* write that:

Because of the lack of certainty in our results, epidemiologists generally use the term risk factor instead of cause to indicate a variable that is believed to be related to the probability of an individual’s developing the disease prior to the point of irreversibility. ([Kleinbaum *et al.*, 1982], p. 29; quoted in Schaffner [1991], p. 206)

This attitude continues to be seen in many areas of health research. In the context of public health, however, it has transformed into the concept of a *determinant*. In this wider context the determinants are often social, which greatly increases

³⁷In his lecture “The Environment and Disease: Association or Causation”, Sir Austin Bradford Hill considered some rules of thumb that might enable causal inference in difficult situations. Hill did the wise thing and tried to avoid a philosophical discussion of causality. Of course, one can’t really engage in a discussion of causality without slipping into philosophical issues. In laying out his view of causality he too clearly defends a probabilistic, manipulationist account. This can be clearly discerned in his claim that the decisive question in causality research is “whether the frequency of the undesirable event B would be influenced by a change in the environmental feature A”.

the complexity of issues to do with causality, and *prima facie* decreases the applicability of the manipulationist account. Public health focuses primarily on these social determinants of health, and, inasmuch as manipulability is involved at all, the approach to interventions that it underwrites is one that seeks to manipulate the very fabric of society; that is, to shift the *patterns* of disease and their distributions in the population as a whole. Often, however, direct manipulation (i.e. control) is not possible, and so one has to resort to observational studies. Furthermore, as Weed points out, manipulability misses out on certain important factors that cannot be controlled: gender and ethnicity. These are called ‘categorical properties’ in the literature of statistical theories of causality. In fact I think when merged with the counterfactual account we can make good sense of manipulating categorical properties too. That is, we can imagine (or give state descriptions of) worlds in which gender is swapped, and so on. Even in the case where one *can* manipulate, the evidence that it was one’s manipulations that caused some outcome, and not some other factor, is not easy to assess because of the complexity of the systems involved. Understanding this latter aspect is the task of ‘evaluation’. A very serious problem with evaluation in this context is just such ‘fat hand’ intervention features—see [Scheines, 2005] for more on fat hand interventions.

There is an old debate, as we have seen, over whether a social or biomedical (natural) cause is responsible for some disease. Causation in public health looks at the determinants of disease and health at the level of the population. The incidence of disease in a population is given by the averaging out of individual cases over the population. However, in doing this we can see patterns that can point to causes for the incidence of disease than cannot be gleaned from measuring the individual cases themselves. Social epidemiologists often use this to argue that the biomedical model of disease causation ought to be replaced by a socio-medical one involving what they call “upstream” or “distal” causes. That is, if we want to have causal explanations of disease, then the place to look for the fundamental causes is not at the level of local biological phenomena but at the level society and the social networks in which individuals find themselves because it is only there that manipulation will lead to elimination of the mechanisms that lead to disease—see, for example, Link and Phelan [1995].

Michael Root moves this debate into an interesting direction, linking it up with the issue of natural and artificial kinds and classifications that we discussed earlier. Root notes that there are clear disparities in the health states of black and white people: “blacks are seven times more likely to die of tuberculosis than whites, three times more likely to die of H.I.V.-A.I.D.S and twice as likely to die of diabetes” ([Root, 2000], p. S629). The diseases themselves, Root argues, are biological while the racial differences are social; and yet a social factor here appears to be determining biological factors: racial factors appear to be resulting in differences in the rates of disease. A possible sociological explanation, then, for the health inequalities between blacks and white suggests itself: it is known that high stress levels can suppress the immune system, and being black is stressful (at least in the U.S.). This is an explanation of disease distribution, and health inequality, that

is broadly in line with those given by social epidemiologists. The idea is that it is social factors, rather than biological factors or mechanisms, that are ultimately responsible for the distribution of disease (despite the fact that some biological explanation can be given for the disease occurrence in some individual). However, Root assumes that the explanation is the correct one without argument or evidence, relying on plausibility alone.

Albert Mosley attempts to deflate the ‘biomedical versus social’ debate by drawing attention to the distinction between the distribution of disease in a population and the occurrence of disease in an individual. He argues—in the context of the debate over whether HIV or poverty causes AIDS—that the answer one gives “is relative to whether the inquirer is interested in the disease or the epidemic, with a focus on individuals or populations” ([2004], p. 412). In other words, in a sense *both* cause AIDS, but the notion of ‘cause’ and the notion of the disease are different in each case: “HIV and poverty are different kinds of causes that operate on different levels of inquiry” (p. 413). This approach is more or less equivalent to that of Stallones who argues that we should understand causation of health and disease in “two modes”, both in terms of the production of illness in individuals and in terms of the generation of patterns illness in populations ([1980], p. 73).³⁸ I think this kind of pluralism offers a good way to cut through some of the dense and seemingly interminable debate over social or biological causation. However, though I think the distinction is needed, it leaves us no better off in terms of our understanding of causation at either the individual or the population level.

8

STUDY DESIGN AND EVIDENCE-BASED MEDICINE

Study designs are intended to get the best evidence for a given context, with the ultimate hope of enabling good causal inferences to be made (or at least to suggest causal hypotheses).³⁹ That is, one designs a study to investigate correlations between variables (‘exposure variables’ on the one hand, and variables associated with disease on the other). There are two broad categories of study: ‘experimental’ and ‘observational’.

Experimental Studies. Experimental studies or ‘intervention trials’ involve the active intervention into the system of interest; this is often compared with a

³⁸Note, however, that Russo et al. [2006] argue that the fact that one speaks about “two levels of causation” does not thereby commit one to saying that causation acts differently at these two levels. Russo and Williamson argue that while causal monism is false, causal pluralism is not right either, since it involves a conflation “the *evidence* from which causal relations are drawn with the very notion of cause” ([2007], p. 169). In this case, however, it is clear that evidence is not at issue: it is the *level* at which causality is being considered.

³⁹Weed identifies as his “keystone” issue for a philosophy of public health the question: “What justifies the decision to implement a preventative intervention, to move, in other words, from scientific evidence to public health action?” ([Weed, 2004], p. 531). This is essentially a policy issue: how does evidence get translated into a concrete result?

control which does not receive an intervention (but may receive a placebo). The reasoning is that if the incidence of some disease is reduced following the intervention (or if there is a difference between the intervention and control groups) then, adopting probabilistic causal reasoning, there is a causal relationship between the intervention and the outcome of interest.

The most superior form of intervention is the RCT in which the treatment allocation is randomized. In public health these will most often take the form of ‘preventative trials’ (the most prevalent of which is the population or group-level community intervention trial). Of course, given the nature of many of the hypotheses relevant in public health (involving the gender or race of individuals, or involving children, for example), experimental studies are rare.⁴⁰

Observational Studies. An observational study is broadly descriptive, aiming to find out how health and disease are distributed over a population. The data can often reveal correlations between variables which can in turn point the way to testing of causal hypotheses (either statistically or experimentally). Observational studies split into two broad types: descriptive and analytic.

1. *Descriptive Studies:*

- **Cross-Sectional Studies:** ‘Health Statics’. Cross-sectional studies (otherwise known as cross-sectional field surveys) are descriptive studies intended to give an instantaneous picture (or a ‘thin-sandwich’ picture) of some system (specifically of the *prevalence* of disease) by investigating survey data for the members of the population of interest. Though the study is considered too weak (for making causal inferences) by modern standards, it can lead to the development of such hypotheses. Note that it also led to both the development of both case-control and cohort designs (*qua* repeated cross-sectional surveys)—*cf.* [Susser, 1985], pp. 28–31.
- **Ecological Study.** An ecological study is a descriptive study taking populations as its units of analysis. It looks for correlations at the population level that might pave the way to more detailed causal investigations.

2. *Analytic Studies:*

⁴⁰The units of intervention in public health contexts are often, as mentioned, higher-level entities, such as schools, hospitals, and other large groups of individuals (including entire communities in the case of community intervention trials). It is clear that if it is needed at the level of individual people randomization is also needed at this level too, for confounding will be just as possible here. For example, an intervention to reduce the incidence of skin cancer in a community by the application of a new sun lotion might be confounded by a number of factors: behaviour modification resulting in less frequent exposure to the sun (perhaps as a result of the idea of risk of skin cancer suggested by the trial itself) and a mild summer are two possibilities. To avoid confounding at this level one randomizes ‘clusters’ of individuals—hence, this study design is known as “cluster randomization”.

- **Case-Control Studies.** Case-control studies (also known as ‘retrospective’ or ‘case-referent’ studies) focus on individuals who develop a disease (the “cases” in question), after which one examines their past histories in order to seek out relevant differences between their histories and the histories of individuals without the disease—that is, one looks for a greater frequency in the presence of some risk factor. This method is a fairly effective way of discovering causes of rare diseases. But it is clearly restricted to small scale. Moreover, it is highly fallible due to the fact that the histories most often involve the individual patient’s memory recall.
- **Longitudinal Studies: ‘Health Dynamics’.** In a longitudinal study the aim is to build up a picture of the evolution of some system over time.⁴¹ Hence, one gathers information about the system (usually via the sample members) at multiple times. In other words, a longitudinal study works by piecing together the snapshots from cross-sectional studies. One can use such studies to identify seasonal effects in health, and to get a firmer grip on potential correlations between variables. They are the obvious tool to assess the health impact of interventions. Longitudinal studies also enable one to chart the history of phenomena of interest, say the spread of a disease.

A *cohort study* (also called ‘follow-up’ or ‘prospective’ studies) is the most well-known longitudinal study design, focusing on a group (or ‘cohort’) of individuals who do not have the disease, but in which there are both exposed and non-exposed members.⁴² If there is an increase in the risk in the exposed subset then this is taken to be indicative of a causal relationship. The major problem with this method is that of confounding factors. In order to draw solid causal conclusions about the exposure, the exposure must be the sole difference between the members. Clearly this is never the case in real-world situations. One needs to supplement the account with some other factors.

Quasi-Experimental Studies. Quasi-experiments or ‘natural experiments’ rely on natural variation with respect to exposure (and matching in all other relevant respects) in and across populations. However, this means that there is no experimental control in who is exposed.

⁴¹Strictly speaking, of course, longitudinal studies are not restricted to observational studies: one could perform repeated experimental studies to determine the dynamics of a population too. However, in practice, given the expense required, they are most often conducted observationally.

⁴²Snow’s investigation of the cholera epidemic (discussed in §2.4) has elements of a cohort study: Snow divided his units (households) into two groups according to their exposure to water from either one or the other water companies. Grouped in this way the data revealed a clear connection between the water supply and cholera mortality.

As with clinical medicine, there is believed to be a ‘hierarchy’ of evidence, with RCTs (randomized controlled trials) at the top—in fact, the systematic review is seen as being at the top of the evidential hierarchy, since this synthesizes the results from multiple RCTs. Many novel problems emerge when one considers large-scale public health interventions. For example, there is a serious difficulty in external validity (or generalizability from one context to another) on account of the complexity of the environment and the problem of shielding the study from interference effects (between groups) and attrition.

In his analysis of the (in-) efficiency of the NHS (the UK’s national health service)—in his lecture series in 1971 entitled *Effectiveness and Efficiency: Random Reflections on Health Services*—the epidemiologist A. L. Cochrane identified the use of ineffective treatments as one of the primary sources of inefficiency in the health system. In response he argued that the treatments ought to be evaluated scientifically, using the best available evidence. In particular, they ought to be run as randomized controlled trials [RCTs] as a matter of course since these trials eliminate bias and are the most systematic method available. We can trace the evidence-based medicine movement back to this point.⁴³ In both cases there is an underlying ethical imperative to minimize harm, in this case by subjecting treatments to better evaluation. In the context of evidence-based medicine there is a notion of a ‘hierarchy of evidence’ which (qualitatively) ranks various forms of evidential support for hypotheses by their level of susceptibility to bias and confounding more generally.⁴⁴ As John Worrall points out, RCTs are not always deemed necessary; sometimes the efficacy of a treatment will be obvious, such as when it prevents otherwise fatal conditions ([Worrall, 2000], p. S319).⁴⁵

Cochrane, however, was concerned as much with the efficiency (i.e. the real benefit of an intervention outside the confines of the RCT) aspect as he was with efficacy (i.e. the maximum possible effect) of treatments. In other words, health economics was given equal weight. Efficiency and efficacy go hand in hand: in getting rid of ineffective ‘treatments’, the burden on the system is reduced. Thus, health and economics go hand in hand. Cochrane’s early advocacy of RCTs, an axiom of evidence-based medicine, was encapsulated in the development of an international database of RCT evidence: *The Cochrane Collaboration*. Multiple RCTs on the same hypothesis are statistically analysed via meta-analysis. This can strengthen or weaken the level of evidence. The evidence is graded according to ‘quality’, and often, if there are multiple studies of ‘higher quality’ any studies of ‘lower quality’ will be ignored (*cf.* [Doyle *et*

⁴³The canonical definition of evidence-based medicine is: “the conscientious, explicit and judicious use of current best evidence in making decisions about the care of individual patients” ([Sackett *et al.*, 1996], p. 71). Though well known in the context of clinical medicine, it has only recently been expanded into public health, where “the patient” transforms into “the public”.

⁴⁴See [Ashcroft, 2004] for a good review of several epistemological issues concerning evidence-based medicine (and RCTs).

⁴⁵Worrall (*ibid.*, p. S328) also notes a curious inconsistency in the reasoning for the high status of RCTs, namely their *reliability*. In meta-analyses of RCTs there is significant divergence over the effectiveness of treatments.

al., 2008], p. 214). This privileging of RCTs has been widely questioned, and there is some lively debate in the philosophical literature: e.g. [Worrall, 2000; 2007] and [Grossman and Mackenzie, 2005]. Grossman and Mackenzie argue that insufficient caution is used in assessing RCTs, as compared to the excessive caution used when assessing observational studies. The case hasn't been made, they say, for the *general* superiority of RCTs over other study designs, and that observational studies are sometimes better, as in the context of public health interventions, for example. Rather, the study design should be matched to the research question, and this will sometimes mean that RCTs are most appropriate, but not always. Worrall argues that the idea that RCTs control for all factors is a "will-o'-the-wisp": without supplementing an inference with background knowledge, about plausible mechanisms and so on, there will be the potential for (plausible) alternative causal factors underlying any evidence.

The notion of a controlled population-level experiment to intervene in all but the simplest health-states is fraught with difficulties. For example, extremely large samples are needed to detect even very small effects. Even when one can implement such an experiment drawing causal inferences from them is incredibly difficult. Suppose you want to alter the distribution of weight, so that there are fewer anorexic and obese individuals, then one can see how to go about designing an intervention to do this, and then measure the effect. Weight is a simple additive factor so one can weigh a sample to see if there is a reduction following the intervention (as compared to data gathered before the intervention). However, one simply will not have the ability to control elements of the social and physical environment to test whether the intervention worked—one major problem along these lines is that one cannot ensure perfect compliance with the randomization.⁴⁶ That is, even if there are significant differences in the weight distribution after the intervention has been implemented one cannot be sure that it was the intervention that was responsible. To assume otherwise is to commit the *post hoc ergo propter hoc* fallacy: there is a strong possibility that some other factor was responsible.⁴⁷

This is, of course, the problem of underdetermination of theory by data. Weed [1997] discusses this problem in the context of a suggested relationship between induced abortion and breast cancer. Another, classic, example is William Farr's study of the possible influence of marriage on mortality [1858]. The underdetermination in this case concerned the issue of whether the observation that married people tend to live longer than unmarried people was due to selection (the fact that healthier people have a tendency to marry) or causation (married life leads to a healthier life: "marriage protection"). Farr argued that it was a selection effect, and recent statistical studies (involving longitudinal rather than cross-sectional tech-

⁴⁶See [Kaufman *et al.*, 2003] for a survey of the problems and prospects for conducting RCTs of social interventions in the context of social epidemiology.

⁴⁷A further factor that causes problems when conducting complex social interventions is that there is often no standardization of intervention categories. [Doyle *et al.*, 2008] give the example of differing definitions of 'smoker' 'ex-smoker', 'quitter' and so on, that can lead to complications in comparing results.

niques) seem to confirm this: see, e.g., [Goldman, 1993]. Of course, this is not the kind of thing one can experimentally determine!

Underdetermination can be broken by appealing to other factors external to the data. Such factors may lead us to prefer one theory over another. However, Weed notes that the ‘criteria’ used by epidemiologists to break the underdetermination are rather weak. The standard method is to invoke Bradford Hill’s so-called causal-criteria: specificity, strength of association, consistency, coherence, temporality, dose-response, biological plausibility, experimentation, and analogy. These are not intended to be necessary and sufficient conditions but rules of thumb. With the exception of temporality (assuming the absence of retro-causality), all of the criteria could be violated without ruling out causality—*cf.* [Weed, 1997], p. 113.⁴⁸

Daniel Little suggests that the problem posed by confounding variables might be resolved by invoking mechanisms and argues that the notion of a ‘plausible mechanism’ could rule out some hypotheses:

We can best exclude the possibility of a spurious correlation between variables by forming a hypothesis about the mechanisms at work in the circumstances. If we conclude that there is no plausible mechanism linking nicotine stains to lung cancer, then we can also conclude that the observed correlation is spurious. (Little, 1991, pp. 24–25)

This approach appears to be more or less in line with Bradford Hill’s methodology. I think Steel [2004] offers a definitive dismissal of the ‘plausible mechanism’ method of breaking this inferential deadlock, at least concerning ‘negative’ explanations (i.e. those showing that we can infer that X is *not* a cause of Y when there is no plausible mechanism connecting them). He argues that it is in reality extremely difficult to think up *any* case of a social phenomenon that could not be explained via some plausible mechanism (*ibid.*, p. 65). Moreover, it seems rather odd to think that the inability to imagine a mechanism generating some phenomenon in some ways aids causal inference (*ibid.*, p. 66). We might also point to the looseness in the notion of ‘plausibility’ here.

9

INTERNAL AND EXTERNAL VALIDITY

The internal validity of a trial concerns the extent to which differences between the trial arms can be attributed to the intervention. As Guala puts it, internal validity

⁴⁸Note that Bradford Hill wished explicitly to avoid philosophical issues, and furthermore made no claim that he was presenting criteria for causation. He states that he is presenting aspects of associations that would lead us to conclude that causation “is the most likely interpretation” ([Bradford Hill, 1965], p. 295). This does not imply that meeting all of the aspects is definite evidence of causality at play. Bradford Hill was concerned with grounds for public health action over perfect knowledge: “All scientific work is incomplete—whether it be observational or experimental. All scientific work is liable to be upset or modified by advancing knowledge. That does not confer upon us a freedom to ignore the knowledge we already have, or to postpone the action that it appears to demand at a given time” (*ibid.*, p. 300). It is interesting to speculate on what the health-research landscape might look like had researchers followed this message rather than Cochrane’s.

“is achieved when the structure and behaviour of a laboratory system (its main causal factors, the ways they interact, and the phenomena they bring about) have been properly understood by the experimenter” ([2003], p. 1198). It is a causal principle: low internal validity means that we can’t tell whether some other factors infected the trial and caused the differences. Such methodology makes these trials practically difficult, for one needs a large number of entire groups. Naturally, the group, being composed of individuals, depends on these individuals so that blinding is done at the lower level, and expected to transfer to the group level. However, the blinding procedure is especially problematic in cluster trials.

The fundamental idea of trials is to generate evidence on which to base or withhold some intervention. In some cases the intervention will be applied at one site only. However, more often the intervention will be applied to multiple subjects, be they individual patients or hospital wards or entire cities. That is, the results of the trial are generalized away from the original test site, despite a host of differences between them. External validity refers to the generality of the results: trials with high external validity are most likely to have their results replicated in diverse contexts.⁴⁹

Guala labels ‘radical localism’ the view that “Experimental results do not apply to the world out of the laboratory” ([Guala, 2003], p. 1196; see also [Guala, 1999])—there is a close resemblance between this view and Nancy Cartwright’s idea of a ‘dappled world’ according to which we do not have grounds for believing that laws of nature transfer from “the highly contrived environments of a laboratory [to] less regulated settings” ([Cartwright, 1999], p. 25). This notion seems particularly appropriate in the context of group-level or population-level health intervention research, of the kind appearing in public health. In such cases the context (the social and physical environment) can interfere with the experiment—so-called “neighbourhood effects” (i.e. the effects of social context, such as a residential community, on health outcomes) constitute one instance of this phenomena.⁵⁰ One way of attempting to correct for neighbourhood effects is to employ multi-level modeling and multilevel analysis, involving the treatment of neighbourhoods as contexts with individuals nested within, in order to separate out the effects (see [Diez-Roux, 2000], especially pp. 180–183). However, there are many problems of validity and causal inference that remain to be worked out in this context in order to have applicability in public health: see [Oakes, 2004] for a review of these issues.

⁴⁹Meta-analysis is intended to provide a quantitative estimate of the degree of replication or consistency across several trials that aim to test the same hypothesis. More importantly, perhaps, it allows for the investigation of trials that yield *inconsistent* results. Weed [2000] argues that meta-analysis can offer a better estimate of an effect than the usual ‘criteria based’ rule of thumb employed.

⁵⁰This idea can, perhaps, trace its ancestry back to Durkheim and Weber’s studies of the ways in which social forces and factors can influence various behaviours.

CONCLUSION

I hope to have shown in this brief guide that the philosophy of public health has many untilled fields ripe for cultivating. The mixture of concepts and techniques (from statistics, epidemiology, demography, and so on) used in public health result in novel philosophical issues not to be found in the study of clinical medicine (or, at least, not in the same form). For this reason it ought to be studied by philosophers alongside clinical medicine, with equal vigour.

BIBLIOGRAPHY

- [Adelstein and Susser, 1976] A. Adelstein and M. Susser. "An Introduction." In N. A. Humphreys (ed.), *Vital Statistics: A Memorial Volume of Selections from the Reports and Writings of William Farr* (pp. iii-xiv). Scarecrow Press, 1976.
- [Afifi and Breslow, 1994] A. A. Afifi and L. Breslow. "The Maturing Paradigm of Public Health." *Annual Reviews of Public Health* **15**, 1994: 223-235.
- [Allotey et al., 2003] P. Allotey, D. Reidpath, A. Kouame, and R. Cummins. "The DALY, Context and the Determinants of the Severity of Disease: An Exploratory Comparison of Paraplegia in Australia and Cameroon." *Social Science and Medicine* **57**(5), 2003: 949-958.
- [Amundson, 2000] R. Amundson. "Against Normal Function." *Studies in History and Philosophy of Biological and Biomedical Sciences* **31**(1), 2000: 33-53.
- [Anderson, 1972] P. W. Anderson. "More Is Different." *Science* **177**(4047), 1972: 393-396.
- [Ashcroft, 2004] R. E. Ashcroft. "Current epistemological problems in evidence based medicine." *Journal of Medical Ethics* **30**, 2004: 131-135.
- [Barenblatt, 2003] G. I. Barenblatt. *Scaling* Cambridge University Press, 2003.
- [Bauer, 2006] S. Bauer. "The population as a laboratory - Epistemic and visual cultures of epidemiology, 1955-2005." *Årsskrift for Medicinsk Museion* **3** 2006: 24-34.
- [Baum, 1990] F. Baum. "The New Public Health: Force for Change or Reaction?" *Health Promotion International* **5**(2), 1990: 145-150.
- [Bayer et al., 2006] R. Bayer, L. O. Gostin, B. Jennings and B. Steinbock. *Public Health Ethics: Theory, Policy, and Practice*. Oxford University Press, 2006.
- [Beauchamp, 1976] D. E. Beauchamp. "Public Health and Social Justice." *Inquiry* **13**, 1976: 1-14.
- [Beauchamp, 1985] D. E. Beauchamp. "Community: The Neglected Tradition of Public Health." *The Hastings Center Report* **15**(6), 1985: 28-36.
- [Bergner, 1985] M. Bergner. "Measurement of Health Status." *Medical* **23**(5), 1985: 696-704.
- [Berridge, 2007] V. Berridge. "Public Health Activism." *British Medical Journal* **335**, 2007: 22-29.
- [Boorse, 1997] C. Boorse. "A Rebuttal on Health." In J. M. Humber and R. F. Almeder (eds.), *What is Disease?* (pp. 1-134). Humana Press.
- [Bradford Hill, 1965] A. Bradford Hill. "The Environment and Disease: Association or Causation?" *Proc R Soc Med.* (1965) **58**(5): 295-300
- [Brock, 1998] D. W. Brock. "Ethical Issues in the Development of Summary Measures of Population Health Status." In MJ. J. Field and M. R. Gold (eds.), *Summarizing Population Health: Directions for the Development and Application of Population Metrics* (pp. 73-81). The National Academic Press, 1998.
- [Broome, 2002] J. Broome. Measuring the burden of disease by aggregating wellbeing. In C. J. L. Murray, J. A. Salomon, C. Mathers and A. D. Lopez (eds), *Summary Measures of Population Health: Concepts, Ethics, Measurement and Applications* (pp. 91-113). Geneva: World Health Organization, 2002.
- [Cartwright, 1999] N. Cartwright. *The Dappled World*. Cambridge University Press, 1989.
- [Cartwright, 1989] N. Cartwright. *Nature's Capacities and their Measurements*. Cambridge University Press, 1989.
- [Chadwick, 1842] E. Chadwick. *Report on the Sanitary Conditions of the Labouring Population of Great Britain by Edwin Chadwick*. Edited by M. W. Flinn. Edinburgh University Press, 1965.

- [Coggon *et al.*, 1997] D. Coggon and G. Rose and D. J. P. Barker. *Epidemiology for the Uninitiated* (4th Edition). BMJ Publishing Group, 1997.
- [Cook and Campbell, 1979] T. D. Cook and D. T. Campbell. *Quasi Experimentation: Design and Analysis Issues for Field Settings*. Chicago: Rand McNally, 1979.
- [Cromley and McLafferty, 2002] E. K. Cromley and S. L. McLafferty. *GIS and Public Health*. The Guilford Press, 2002.
- [Dawson and Verweij, 2007] A. Dawson and M. Verweij (eds.). *Ethics, Prevention, and Public Health*. Oxford University Press, 2007.
- [Detels and Breslow, 2005] R. Detels and L. Breslow. "Current Scope and Concerns in Public Health." In R. Detels, J. McEwen, R. Beaglehole, and H. Tanaka (eds.), *Oxford Textbook of Public Health 4th Ed. Vol. 1. The Scope of Public Health* (pp. 3-20). Oxford University Press, 2005.
- [Dodge, 2003] Y. Dodge (ed.) *The Oxford Dictionary of Statistical Terms*. Oxford University Press, 2003.
- [Doyle *et al.*, 2008] J. Doyle, R. Armstrong, and E. Waters. "Issues Raised in Systematic Reviews of Complex Multisectoral and Community Based Interventions." *Journal of Public Health* **30**(2), 2008: 213–215.
- [Egger *et al.*, 2001] M. Egger, G. D. Smith, and D. Altman (eds.) *Systematic Reviews in Health Care: Meta-Analysis in Context*. BMJ Books, 2001.
- [Elliot *et al.*, 2001] P. Elliott, J. Wakefield, N. Best, D. Briggs (eds.). *Spatial Epidemiology: Methods and Applications*. Oxford University Press, 2001.
- [Evans, 1993] A. S. Evans. *Causation and Disease: A Chronological Journey*. Plenum Press, 1993.
- [Fanshel, 1972] S. Fanshel. "A Meaningful Measure of Health for Epidemiology." *International Journal of Epidemiology* **1**(4), 1972: 319–337.
- [Farr, 1858] W. Farr. "The Influence of Marriage on the Mortality of the French People." In G. W. Hastings (ed.), *Transactions of the National Association for the Promotion of Social Science* (pp. 504–513). London: John W. Parker.
- [Forbes and Wainwright, 2001] A. Forbes and S. P. Wainwright. "On the Methodological, Theoretical and Philosophical Context of Health Inequality Research: A Critique." *Social Science and Medicine* **53**, 2001: 801–816.
- [Fryback, 1998] D. G. Fryback. "Methodological Issues in Measuring Health Status and Health-Related Quality of Life for Population Health Measures." In MJ. J. Field and M. R. Gold (eds.), *Summarizing Population Health: Directions for the Development and Application of Population Metrics* (pp. 39–57). The National Academic Press, 1998.
- [Galea and Putnam, 2007] S. Galea and S. Putnam. "The Role of Macrosocial Determinants in Shaping the Health of Populations." In S. Galea (ed.), *Macrosocial Determinants of Population Health* (pp. 3-12). Springer, 2007.
- [Giare, 2006] R. N. Giare. "Perspectival Pluralism." In S. H. Kellert, H. E. Longino, and C. K. Waters (eds.), *Scientific Pluralism* (pp. 26-41). University of Minnesota Press, 2006.
- [Gillies, 2005] D. Gillies. "Hempel and Kuhnian Approaches in the Philosophy of Science: The Semmelweis Case." *Studies in the History and Philosophy of Biology and the Biomedical Sciences* **36**, 2005: 159–181.
- [Goldman, 1993] N. Goldman. "Marriage Selection and Mortality Patterns: Inferences and Fallacies." *Demography* **30**(2), 1993: 189–208.
- [Greenland, 1988] S. Greenland. "Probability Versus Popper: An Elaboration of the Insufficiency of Current Popperian Approaches for Epidemiologic Analysis." In K. Rothman (ed.), *Causal Inference* (pp. 95-104). Epidemiology Resources Inc., 1988.
- [Greenland, 2002] S. Greenland. "Causality Theory for Policy Uses of Epidemiological Measures." In C. J. L. Murray, J. A. Salomon, C. Mathers and A. D. Lopez (eds), *Summary Measures of Population Health: Concepts, Ethics, Measurement and Applications* (pp. 291–301). Geneva: World Health Organization, 2002.
- [Greenland *et al.*, 2004] S. Greenland, M. Gago-Dominguez, and J. E. Castela. "The Value of Risk-Factor ("Black-Box") Epidemiology." *Epidemiology* **15**(5), 2004: 529–535.
- [Grossman and Mackenzie, 2005] J. Grossman and F. Mackenzie. "The Randomized Controlled Trial: Gold Standard, or merely Standard?." *Perspectives in Biology and Medicine* **48**(4), 2005: 516–34.
- [Guala, 1999] F. Guala. "The Problem of External Validity (or "Parallelism") in Experimental Economics." *Social Science Information* **38**, 1999: 555–573.
- [Guala, 2003] F. Guala. "Experimental Localism and External Validity." *Philosophy of Science* **70**, 2003: 1195–1205.

- [Guy, 1870] W. A. Guy. *Public Health* Renshaw Press, 1870.
- [Hald, 2003] A. Hald. *A History of Probability and Statistics and Their Applications before 1750*. Wiley-Interscience, 2003.
- [Hamlin, 1995] C. Hamlin. "Finding a Function for Public Health: Disease Theory or Political Philosophy?" *Journal of Health Politics, Policy and Law* **20**(4): 1025–1031.
- [Hamlin, 1998] C. Hamlin. *Public Health and Social Justice in the Age of Chadwick: Britain, 1800–1854*. Cambridge University Press, 1998.
- [Hamlin, 2005] C. Hamlin. "The History and Development of Public Health in Developed Countries." In R. Detels, J. McEwen, R. Beaglehole, and H. Tanaka (eds.), *Oxford Textbook of Public Health 4th Ed. Vol. 1. The Scope of Public Health* (pp. 21–37). Oxford University Press, 2005.
- [Hausman and Woodward, 2004] D. M. Hausman and J. Woodward. "Manipulation and the Causal Markov Condition." *Philosophy of Science* **71** 2004: 846–856.
- [Hausman, 2006] D. M. Hausman. "Valuing Health." *Philosophy & Public Affairs* **34**(3), 2006: 246–274.
- [Hausman et al., 2002] D. M. Hausman, Y. Asada and T. Hedemann. "Health Inequalities and Why They Matter." *Health Care Analysis* **10**, 2002: 177–191.
- [Hausman, 2002] D. Hausman. "Causality and Counterfactual Analysis." In C. J. L. Murray, J. A. Salomon, C. Mathers and A. D. Lopez (eds), *Summary Measures of Population Health: Concepts, Ethics, Measurement and Applications* (pp. 309–313). Geneva: World Health Organization, 2002.
- [Hempel, 1966] C. G. Hempel. *Philosophy of Natural Science*. Englewood Cliffs, NJ: Prentice-Hall, 1966.
- [Holland, 1986] P. W. Holland. "Statistics and Causal Inference." *Journal of the American Statistical Association* **81**, 1986: 945–960.
- [Kaufman et al., 2003] J. S. Kaufman, S. Kaufman, and C. Poole. "Causal Inference from Randomized Trials in Social Epidemiology." *Social Science and Medicine* **57**, 2003: 2397–2409.
- [King, 1954] L. S. King. "What is Disease?" *Philosophy of Science* **21**(3), 1954: 193–203.
- [Kitcher, 1997] P. Kitcher. *The Lives to Come: The Genetic Revolution and Human Possibilities*. Simon and Schuster, 1997.
- [Kleinbaum et al., 1982] D. G. Kleinbaum, L. L. Kupper, and H. Morgerstern. *Epidemiologic Research*. Belmont, CA: Lifetime Learning, 1982.
- [Koch, 2005] T. Koch. *Cartographies of Disease: Maps, Mapping, and Medicine*. ESRI Press, 2005.
- [Koch, 1882] R. Koch. "Die Aetiologie der Tuberculose." *Berl. klin. Wchnschr.* **19**, 1882: 221–230.
- [Last, 1995] J.M. Last (ed.) *A Dictionary of Epidemiology*. Oxford University Press, 1995.
- [Latour, 1988] B. Latour (ed.) *The Pasteurisation of France*. Harvard University Press, 1988.
- [Leavitt and Numbers, 1997] J. W. Leavitt and R. L. Numbers (eds.) *Sickness and Health in America: Readings in the History of Medicine and Public Health*. University of Wisconsin Press, 1997.
- [Lennox, 1995] J. Lennox. "Health as an Objective Value." *Journal of Medicine and Philosophy* **20**(5), 1995: 499–511.
- [Lilienfeld and Lilienfeld, 1982] A. M. Lilienfeld and D. E. Lilienfeld. "Epidemiology and the Public Health Movement: A Historical Perspective." *Journal of Public Health Policy* **3**(2), 1982: 140–149.
- [Link and Phelan, 1995] B. G. Link and J. Phelan. "Social Conditions as Fundamental Causes of Diseases." *Journal of Health and Social Behaviour* **35**, 1995: 80–94.
- [Lipton, 1991] P. Lipton. *Inference to the Best Explanation*. Routledge, 1991.
- [Little, 1991] D. Little. *Varieties of Social Explanation*. Westview Press, 1991.
- [MacMahon and Pugh, 1970] T. F. MacMahon and B. Pugh (eds.) *Epidemiology: Principles and Methods*. Boston: Little Brown & Co., 1970.
- [Mackenzie, 1981] D. Mackenzie *Statistics in Britain, 1865–1930: The Social Construction of Scientific Knowledge*. Edinburgh University Press, 1981.
- [Marcum, 2008] J. A. Marcum. *An Introductory Philosophy of Medicine. Humanizing Modern Medicine*. Springer, 2008.
- [Maxwell Parkin and Bray, 2005] D. Maxwell Parkin and F. Bray. "Descriptive Studies." In W. Ahrens and I. Pigeot (eds.), *Handbook of Epidemiology* (pp. 157–230). Springer, 2005.
- [Mendelsohn, 1995] J. Andrew Mendelsohn. "'Typhoid Mary' Strikes Again: The Social and the Scientific in the Making of Modern Public Health." *Isis* **86**(2), 1995: 268–277.
- [Michalos, 2006] A. C. Michalos. "The Quality-Of-Life (QOL) Research Movement: Past, Present, and Future. Conceptual and Philosophical Foundations." *Social Indicators Research* **76**, 2006: 343–363.

- [Mosley, 2004] A. Mosley. "Does HIV or Poverty Cause AIDS? Biomedical and Epidemiological Perspectives." *Theoretical Medicine* **25** 2004: 399-421.
- [Murray, 1994] C. J. L. Murray. "Quantifying the Burden of Disease: The Technical Basis for Disability-Adjusted Life Years." *Bulletin of the World Health Organization* **72**, 1994: 429-424.
- [Murray *et al.*, 1999] C. J. L. Murray, E. E. Gakidou and J. Frenk. "Health Inequalities and Social Group Differences: What Should we Measure?" *Bulletin of the World Health Organization* **77**(7), 1999: 537-543.
- [Murray *et al.*, 2000] C. J. L. Murray, J. A. Salomon and C. Mathers. "A critical examination of summary measures of population health." *Bulletin of the World Health Organization* **78**(8), 2000: 981-994.
- [Murray *et al.*, 2002a] C. J. L. Murray, J. A. Salomon, C. Mathers and A. D. Lopez. "Summary Measures of Population Health: Conclusions and Recommendations." In C. J. L. Murray, J. A. Salomon, C. Mathers and A. D. Lopez (eds), *Summary Measures of Population Health: Concepts, Ethics, Measurement and Applications* (pp. 731-756). Geneva: World Health Organization, 2002.
- [Murray *et al.*, 2002b] C. J. L. Murray, J. A. Salomon, C. Mathers and A. D. Lopez (eds). *Public Health Informatics and Information Systems*. Geneva: World Health Organization, 2002.
- [Oakes, 2004] J. M. Oakes. "The (Mis)Estimation of Neighborhood Effects: Causal Inference for a Practicable Social Epidemiology." *Social Science and Medicine* **58** 2004: 1929-1952.
- [O'Carroll *et al.*, 2002] P. W. O'Carroll, W. A. Yasnoff, M. E. Ward, L. H. Ripp, and E. L. Martin (eds). *Summary Measures of Population Health: Concepts, Ethics, Measurement and Applications*. Springer, 2002.
- [Pearl, 2001] J. Pearl. "Causal Inference in the Health Sciences: A Conceptual Introduction." *Health Services and Outcomes Research Methodology* **2**(3/4) 2001: 189-220.
- [Porter, 1994] D. Porter (ed.). *The History of Public Health and the Modern State*. Editions Rodopi, 1994.
- [Porter, 1999] D. Porter. *Health, Civilization and the State: A History of Public Health from Ancient to Modern Times*. Taylor & Francis, 1999.
- [Powers and Faden, 2006] M. Powers and R. Faden (ed.). *Social Justice: The Moral Foundations of Public Health and Health Policy*. Oxford University Press, 2006.
- [Putnam, 1975] H. Putnam. "The Meaning of 'Meaning'." In H. Putnam (ed.), *Mind, Language and Reality: Philosophical Papers, Volume 2* (pp. 215-71). Cambridge University Press, 1975.
- [Reidpath, 2005] D. D. Reidpath. "Population Health: More than the Sum of the Parts?" *Journal of Epidemiology and Community Health* **59**: 877-880.
- [Reidpath, 2007] D. D. Reidpath. "Summary Measures of Population Health: Controversies and New Directions." In I. Kawachi and S. Wamala (eds.), *Globalization and Health* (pp. 187-200). Oxford University Press, 2007.
- [Root, 2000] M. Root "How we Divide the World." *Philosophy of Science, Vol. 67, Supplement. Proceedings of the 1998 Biennial Meetings of the Philosophy of Science Association. Part II: Symposia Papers*. 2000: S628-S639.
- [Rose, 2001] G. Rose. "Sick individuals and sick populations." *International Journal of Epidemiology* **30** 2001: 427-432.
- [Rose, 1992] G. Rose. *The Strategy of Preventive Medicine*. Oxford University Press, 1992.
- [Rosen, 1993] G. Rosen. *A History of Public Health*. The Johns Hopkins University Press, 1993.
- [Rosenberg, 1992] C. E. Rosenberg. *Explaining Epidemics and Other Studies in the History of Medicine*. Cambridge University Press, 1992.
- [Diez-Roux, 2000] A. V. Diez-Roux. "Multilevel Analysis in Public Health Research." *Annual Review of Public Health* **21**: 171-92.
- [Rusnock, 2002] A. A. Rusnock. *Vital Accounts: Quantifying Health and Population in Eighteenth-Century England and France*. Cambridge University Press, 2002.
- [Rubin, 1986] D. B. Rubin "Statistics and Causal Inference: Comment: Which Ifs Have Causal Answers?" *Journal of the American Statistical Association* **81**(396) 1986: 961-962.
- [Russo *et al.*, 2006] F. Russo, M. Mouchart, M. Ghins, and G. Wunsch. "Statistical Modelling and Causality in the Social Sciences." *Institute de Statistique Discussion Paper* 0601, 2006.
- [Russo and Williamson, 2007] F. Russo and J. Williamson. "Interpreting Causality in the Health Sciences." *International Studies in the Philosophy of Science* **21**(2), 2007: 157-170.
- [Rychetnik *et al.*, 2004] L. Rychetnik, P. Hawe, E. Waters, A. Barratt, and M. Frommer. "A Glossary for Evidence Based Public Health" *Journal of Epidemiology and Community Health* **58**: 538-545.

- [Sackett *et al.*, 1996] D. L. Sackett, W. M. C. Rosenberg, J. A. Muir Gray, R. Brian Haynes, W. Scott Richardson. "Evidence based medicine: what it is and what it isn't" *British Medical Journal* **312**(7023): 71–72.
- [Schaffner, 1991] K. F. Schaffner. "Causing Harm: Epidemiological and Physiological Concepts of Causation." In D. G. Mayo and R. D. Hollander (eds.), *Acceptable Evidence: Science and Values in Risk Management* (pp. 204–217). New York: Oxford University Press, 1991.
- [Scheines, 2005] R. Scheines. "The Similarity of Causal Inference in Experimental and Non-experimental Studies." *Philosophy of Science* **72**, 2005: 927–40.
- [Schneider and Lilienfeld, 2008] Schneider, D. and D. E. Lilienfeld (eds). *Public Health: The Development of a Discipline*. Rutgers University Press, 2008.
- [Schwartz, 2007] P. H. Schwartz. "Decision and Discovery in Defining 'Disease'." In H. Kincaid and J. McKittrick (eds.), *Establishing Medical Reality* (pp. 47–63). Springer, 2007.
- [Schwartz and Diez-Roux, 2001] S. Schwartz and R. Diez-Roux. "Commentary: Causes of Incidence and Causes of Cases—A Durkheimian Perspective on Rose." *International Journal of Epidemiology* **30**, 2001: 435–439.
- [Shattuck, 1850] L. Shattuck. *Report on a General Plan for the Promotion of Public and Personal Health*. Boston: Dutton & Wentworth, 1850.
- [Snow, 1855] J. Snow. *On the Mode of Communication of Cholera*. New York: Hafner, 1965.
- [Stallones, 1980] R. A. Stallones. "To Advance Epidemiology." *Annual Review of Public Health* **1**, 1980: 69–82.
- [Steel, 2004] D. Steel. "Social Mechanisms and Causal Inference." *Philosophy of the Social Sciences* **34**, 2004: 55–78.
- [Stigler, 1994] S. M. Stigler. "Some Correspondence on Methodology Between Milton Friedman and Edwin B. Wilson; November-December 1946." *Journal of Economic Literature* **32**(3) 1994: 1197–1203.
- [Stouman and Falk, 1937] K. Stouman and I. S. Falk. (1937). "Health Indices: A Study of Objective Indices of Health in Relation to Environment and Sanitation." *The Milbank Memorial Fund Quarterly* **15**(2): 173–195.
- [Stroup and Berkelman, 1998] D. F. Stroup and R. L. Berkelman. "History of Statistical Methods in Public Health." In D. F. Stroup and S. M. Teutsch (eds.), *Statistics in Public Health: Qualitative Approaches to Public Health Problems* (pp. 1–18). Oxford University Press, 1998.
- [Sulmasy, 2005] D. P. Sulmasy. "Diseases and Natural Kinds." *Theoretical Medicine and Bioethics* (2005) **26**: 487–513.
- [Susser, 1985] M. Susser. "Epidemiology in the United States after World War II: The Evolution of Technique." *Epidemiologic Review* (1985) **7**: 147–77. Reprinted in his *Epidemiology, Health, and Society: Selected Papers* (pp. 22–49). Oxford University Press, 1987.
- [Sydenstricker, 1928] E. Sydenstricker. "The Statistician's Place in Public Health Work." *Journal of the American Statistical Association* **23**(162), 1928: 115–1203.
- [Tesh, 1995] S. N. Tesh. "Miasma and 'Social Factors' in Disease Causality: Lessons from the Nineteenth Century." *Journal of Health Politics, Policy and Law* **20**(4): 1001–1024.
- [Terris, 1985] M. Terris. "The Changing Relationships of Epidemiology and Society." *Journal of Public Health Policy* **6**, 1985: 15–34.
- [Terris, 1987] M. Terris. "Epidemiology and the Public Health Movement." *Journal of Public Health Policy* **8**(3), 1987: 315–329.
- [Terris, 1993] M. Terris. "The Society for Epidemiological Research and the Future of Epidemiology." *Journal of Public Health Policy* **14**(2): 137–148.
- [Teutsch and Churchill, 2000] S. M. Teutsch and R. E. Churchill. *Principles and Practice of Public Health Surveillance, 2nd Edition*. Oxford University Press, 2000.
- [Thacker *et al.*, 1989] S. B. Thacker, R. L. Berkelman, and D. F. Stroup. "The Science of Public Health Surveillance." *Journal of Public Health Policy* **10**(2): 187–203.
- [Thagard, 2005] Thagard, P. "What is a Medical Theory?" In L. McNamara (ed.), *Multidisciplinary Approaches to Theory in Medicine, Vol. 3* (pp. 47–62). Elsevier, 2005.
- [Turnock, 2006] B. J. Turnock. *Public Health*. Jones and Bartlett Publishers, 2006.
- [U.S. Institute of Medicine, 1988] U.S. Institute of Medicine, Committee for the Study of the Future of Public Health. *The Future of Public Health*. National Academy Press, 1988.
- [Verweij and Dawson, 2003] M. Verweij and A. Dawson. "The Meaning of 'Public' in 'Public Health'." In A. Dawson and M. Verweij (eds.), *Ethics, Prevention, and Public Health* (pp. 13–29). Oxford University Press, 2007.

- [Wainwright and Forbes, 2000] S. P. Wainwright and A. Forbes. "Philosophical Problems with Social Research on Health Inequalities." *Health Care Analysis* **8**, 2000: 259–277.
- [Ward and Warren, 2006] J. W. Ward and C. Warren. *Silent Victories: The History and Practice of Public Health in Twentieth-Century America*. Oxford University Press, 2006.
- [Weber, 1920] M. Weber. "Die Protestantische Ethik und der Geist des Kapitalismus." In *Gesammelte Aufsätze zur Religionssoziologie*. Tübingen, 1920.
- [Weed, 1997] D. L. Weed. "Underdetermination and Incommensurability in Contemporary Epidemiology." *Kennedy Institute of Ethics Journal* **7**(2), 1997: 107–124.
- [Weed, 1999] D. L. Weed. "Towards a Philosophy of Public Health." *Journal of Epidemiology and Community Health* **53**, 1999: 99–104.
- [Weed, 2000] D. L. Weed. "Interpreting Epidemiological Evidence: How Meta-analysis and Causal Inference Methods are Related." *International Journal of Epidemiology* **29**(3), 2000: 387–390.
- [Weed, 2004] D. L. Weed. "Ethics and Philosophy of Public Health." In G. Khushf (eds.), *Handbook of Bioethics* (pp. 525–547). Kluwer Academic Publishers, 2004.
- [Wheelan, 2003] C. Wheelan. *Naked Economics: Undressing the Dismal Science*. W. W. Norton & Company, 2003.
- [WHO, 1947] World Health Organisation. *The constitution of the World Health Organisation*. WHO Chronicle, 1947.
- [WHO, 2000] World Health Organisation. *The World Health Report, 2000*. Geneva: WHO, 2000.
- [Worrall, 2000] J. Worrall. "What Evidence in Evidence-Based Medicine?" *Philosophy of Science* **69**, 2002: S316–S330.
- [Worrall, 2007] J. Worrall. "Evidence in Medicine and Evidence-Based Medicine?" *Philosophy Compass* **2**(6), 2007: 981–1022.
- [Winslow, 1923] C. -E. A. Winslow. *The Evolution and Significance of the Modern Public Health Campaign*. Yale University Press, 1923.
- [Winslow, 1929] C. -E. A. Winslow. "The Contribution of Hermann Biggs to Public Health: The 1928 Biggs Memorial Lecture" *The American Review of Tuberculosis* **20**, 1929: 1–28.
- [Worboy, 2000] M. Worboy. *Spreading Germs: Disease Theories and Medical Practice in Britain, 1865–1900*. Cambridge University Press, 2000.