Preprint of the chapter in the volume *Integrative Approaches in Environmental Health and Exposome Research: Epistemological and Practical Issues* (E. Giroux, F. Merlin, Y. Fayet Eds.)
https://doi.org/10.1007/978-3-031-28432-8_5

Which Integration for Health? Comparing Integrative Approaches for Epidemiology

Stefano Canali

Department of Electronics, Information and Bioengineering & META—Social Sciences and Humanities for Science and Technology, Politecnico Di Milano, Milan, Italy

1. Introduction

Biomedical research and clinical care are constantly filled with new promises, trends, and keywords that frame health and disease in new and different terms. These keywords intersect at various levels, appearing in biomedical research and funding schemes, in policy discussions and political speeches, as well as in the media and public discourses. Some of these concepts often end up having a short lifespan, but others stay and sometimes get entrenched within specific research communities or even ways in which health policy is conducted. The relevance and importance of these concepts and their epistemological consequences merit our attention – in particular if we want to make sense, conceptually, of what these mean and promise and when they have shaped lines of research and can further do so in the future. In philosophy of medicine, significant work has focused on the merits and pitfalls of approaches and notions including evidence-based medicine (Worrall 2002; 2007; Stegenga et al. 2017), personalised and precision medicine (Vogt, Hofmann, and Getz 2016; Prainsack 2020; Plutynski 2022), genomics and postgenomics (Richardson and Stevens 2015; Hilgartner 2017; Gibbon et al. 2020). Less attention has instead been given to conceptualisations of health and disease in the context of epidemiology and public health research, but the area is ripe with new and different ways of discussing and investigating health and disease, from new types of collection and use of health data to the new methodological and statistical treatment of health data, from new conceptualisations of health to the employment of new types of causal inference.

In this chapter, I analyse a specific type of concepts: concepts aimed at framing the relations between the health of populations and the environment. As we will see, the focus on these relations is a recent shift in epidemiology, but one that has gained increasing momentum in recent years. I will specify my discussion on the following three concepts: the exposome, a notion aimed at capturing all the exposures that individuals and populations experience; planetary health, a conceptualisation of the impact of local and multispecies environments on human health; and global health, a concept aimed at framing the health of populations as the result of different social factors of different and diverse populations. In analysing these new concepts, I will apply a specific methodological focus by looking at data integration, as the ensemble of data practices aimed at using different types and bodies of

data for the study of specific phenomena. This focus is grounded on recent philosophical literature on the epistemology of integration in the life and health sciences and the epistemology of scientific data. As I will show, this focus enables me to identify significant assumptions and commitments, merits, and limitations of the three concepts.

The chapter is structured as follows. I start by motivating my focus on a specific type of integration – data integration – in relation to ongoing research in philosophy of science and medicine (Sect.2). I then use this focus as a conceptual lens to analyse and discuss three crucial keywords that frame health and biomedical research in specific ways and, as I argue, are distinct approaches to the integration of different types of data. I then focus on the exposome (Sect. 3), global health (Sect. 4), and planetary health (Sect. 5), critically analysing their approaches and identifying limits their limitation.

2. The Many Faces of Integration: Why Data Integration Matters

Integration is a significant focus of recent philosophy of science, particularly in the context of discussions on the metaphysical and epistemological foundations of the life and physical sciences and their distinctions. The latter have traditionally been discussed as particularly fragmented, with highly diverse communities, approaches, methodologies, theories, styles of explanation, commitments, and goals. In this sense, the analysis of these aspects of the life sciences has been intertwined with questions of reduction, unification, and indeed integration, with a focus on the extent to which elements of biological theory or ontology can be integrated and reduced to physics, the plurality and diversity of the life sciences is something that can and should be integrated and reduced, or should lead to unifications at the conceptual or methodological level (Dupré 1996). Several philosophers of biology have argued that that different theories in biology cannot be unified nor reduced to more fundamental ones (Mitchell and Dietrich 2006) and the use of diverse methods, models and representations is crucial in a number of areas of the life sciences (Mitchell and Gronenborn 2017), as many questions and problems require explanations developed in different biological disciplines and with different scientific aims (Brigandt 2010).

The question of integration thus brings substantial results from the philosophical literature and can be analysed from several different points of view in the context of contemporary life and health sciences, and epidemiology in particular (Giroux, Fayet, and Serviant-Fine 2021). In the context of this chapter, I will focus on a specific type of integration: data integration. Data integration can be defined as the set of data practices involved in making different types of data usable as single bodies of evidence. For example, data integration can entail working on samples collected at different points of data collection, for instance blood samples collected in a longitudinal study, to make sure that all the samples can be compared and used as evidence for a broader study, for instance a study of cardiovascular disease in the general population. At the same time, as a result of the aforementioned

fragmentation of the life and health sciences, in this context the notion of data can refer to many different objects, usually comes in significantly different types and is collected by different communities, for different purposes and with diverse commitments. In addition, the life and health sciences present an interesting case in which the standardised production of large volumes of data has often been mentioned as a potential game changer (Golub 2010; Weinberg 2010), but practices of data handling, storing and analysis have strong continuity with longstanding approaches (Müller-Wille and Charmantier 2012; Leonelli 2016; Strasser 2019). This entails that substantial work needs to be conducted on preparing diverse datasets for data analysis, in order to compare different findings, building a more robust evidential basis for a specific claim, contrasting quality from various research settings. Data also sit at the crossroads of many current trends of the field, especially in the biomedical sciences. For example, personalised and precision medicine, as the attempts to take into account individual variables into the study and prevention of disease, are largely built on the assumption that the use of large datasets "can account for an increasing number of factors that influence health and disease, and that these data can be used to stratify the population and health problems according to various characteristics" (Green and Vogt 2016). Similarly, the molecularisation of the health sciences (Boniolo and Nathan 2017) and postgenomics (Richardson and Stevens 2015) are largely based on the use and integration of new datasets, at different levels of abstraction and at increasing volumes.

These features suggest that data integration involves technical considerations about the intrinsic properties of the data, such as their interoperability, quality, resolution, etc. However, results from the philosophy of scientific data have shown that data integration also involves significant inferences, assumptions, strategies that provide a significant window into the epistemology of biomedical research and data (Leonelli 2013). In this sense, focusing on data integration is an important analytical choice for philosophy of science, but is particularly interesting in the context of epidemiology, health, and especially the approaches I discuss in this chapter.

Epidemiological research is traditionally concerned with diverse and large types of data (Morabia 2004), but the recent raise in the volume, diversity, heterogeneity of data in the sciences (Leonelli 2016; Leonelli and Tempini 2020) has often been discussed as a substantial novelty for the field, potentially opening new and uncharted territories for both research and policy-making (Holmberg, Bischof, and Bauer 2013; Hogle 2016; Fleming et al. 2017). For example, the increasing availability of new data on the environment and climate is opening new opportunities for the study of the impact of different types of environments on population health and shaping the ways in which the environment is conceptualised and operationalised across approaches in epidemiology (Canali and Leonelli 2022). At the same time, the availability of new sources of data on population behaviour, for instance through social media, personal health applications, and monitoring devices, is supposed to create new disciplines and areas of research in the life sciences, such as digital epidemiology,

relying on the increasing datafication of health and medicalisation of population activities through data coming from outside epidemiology and scientific research more generally (Salathé 2018; Mittelstadt et al. 2018; Klingwort and Schnell 2020). Data integration thus gives us a window into the approaches, assumptions, and conceptualisations that are crucial aspects of the epistemology of contemporary epidemiology (Giroux, this volume).

It is on this basis that in the chapter I focus on the integration of data as an analytical lens to discuss the different approaches to integration, identify assumptions and commitments, and discuss conceptual gaps and methodological issues. As we will see, the approaches that I discuss in the chapter are all aimed at a certain type of integration at the level of health and environment, which elicits questions on which type of integration they aim for, the extent to which such integration is actually achievable or achieved, and what this implies more generally. In particular, the concept of the exposome is a concept that aims to integrate different dimensions and categories of exposure; the concept of planetary health attempts to integrate different dimensions of the health of populations and environments; and the concept of global health is a way of integrating different types and sources of health at a global level. In the context of this chapter I will approach them with the methodological choice of focusing on data integration and discussing the different types of data integration that are part of these approaches to environmental research. There are clearly various other aspects of integration that merit discussion in the context of these three approaches and various other chapters of this book go in that direction (see Meloni, Russo, Giroux, this volume). My choice of focusing on data integration builds on this work and aims to identify epistemic assumptions, features, and issues that are involved in these three approaches and emerge at the level of data practices.

Framing these concepts as data integration approaches is a particular way of studying them and specifying philosophical questions of integration with a focus on an increasingly important area of scientific practice. This is a specific methodological choice that frames and specifies the work I do in the chapter. Other concepts, frameworks, approaches could be chosen to discuss integration and focus on data integration, including for instance social determinants of health and disease, precision medicine, toxicology. Yet in this chapter I focus on attempts at integration between health and environment in the study of population health and look at their relations with current research projects and data practices that have been largely successful at securing funding and establishing research centres in the last few years. As I will discuss at several points of the remainder of the chapter, while I do not aim for a broad generalisation of my findings, the approaches I chose to focus on also share significant features with the aforementioned and other approaches in contemporary epidemiology.

Preprint of the chapter in the volume *Integrative Approaches in Environmental Health and Exposome Research: Epistemological and Practical Issues* (E. Giroux, F. Merlin, Y. Fayet Eds.) https://doi.org/10.1007/978-3-031-28432-8_5

3. The Exposome: Integrating Molecular and Environmental Data

The exposome has raised as a specific approach, concept, and area of research in the last decade in epidemiology and beyond. The exposome was first conceived as specific notion, which conceptually would entail a new expansion of the traditional notion of exposure by adding more specific dimensions at the external level and a new dimension at the internal level (Wild 2005). The rationale here was that epidemiologists needed to expand and specify the ways in which exposure and the environment were considered and different types of data were collected; a crucial element of inspiration was genomics and the Human Genomic Project, both in the sense of providing molecular tools for data collection and analysis and in the sense of adding the new dimensions to internal exposure. As such, conceptually the exposome is presented as the totality of all exposures that individuals experience in a specific point in time and cumulatively though all their life and methodologically is made of three types and categories of exposure: general external exposure, specific external exposure, and internal exposure (Wild 2012). The idea here is that individuals and populations can be exposed to environmental agents at different levels, both internally and externally: for instance at the general external level with processes connected to socio-economic status, at a specific external level with processes such as specific chemical pollutants and infectious agents, or at the internal level with processes including endogenous responses to external exposure (e.g. inflammation, oxidative stress, etc.). In this sense, conceptually the exposome has been framed as a counterpart to the genome – i.e., taking inspiration from the methodological and conceptual status of the genome whilst intending to critically complement genomic research with the study of health and disease with epidemiological tools and approaches, in particular individual and cohort studies.

The exposome as an area of research has indeed relied extensively on molecular data collection and analytic tools, by integrating omics techniques into epidemiology (Boniolo and Nathan 2017). Omics are molecular techniques that in the genomic context are used to study molecules and processes within a cell, including for example metabolism and protein development and quantify and collect data on these processes. In the exposome context, omics data have been used to analyse the internal dimension of exposure and develop exposure profiles, which measure processes and entities at the internal level and can be analysed in comparison and contrast with external exposure data (Russo & Vineis, 2016). For instance, omics techniques can be used to analyse blood samples collected in longitudinal studies and measure adducts at the molecular and internal level whose presence can be connected to external toxicants from e.g. pollution. The resulting exposure profiles can be contrasted and compared with quantitative data on the presence of these pollution toxicants and thus data on external exposure at different levels, as a way to study the influence on pollution on the internal molecular environment and more generally the effects of pollution on the development of health and disease.

Preprint of the chapter in the volume *Integrative Approaches in Environmental Health and Exposome Research: Epistemological and Practical Issues* (E. Giroux, F. Merlin, Y. Fayet Eds.) https://doi.org/10.1007/978-3-031-28432-8_5

The extension of these genomic tools and methods into epidemiology has contributed to the raising influence of the exposome in contemporary epidemiology (Canali 2020a). As such, research on the exposome has received substantial and specific funding, for instance with dedicated programmes in the EU, and several research centres and units have been established, for example in the US. These processes have contributed to establishing the exposome as a way of doing epidemiology, based on the integration of different types of exposures, that are measured and quantified through molecular and data-intensive tools.[1]

As a result, the exposome approach is an important point of focus for philosophical, methodological, historical analyses of epidemiology (e.g. Illari and Russo 2016; Giroux, Fayet, and Serviant-Fine 2021; Russo, Giroux, Vineis in this volume). In this sense, in the context of this chapter, I frame the exposome as a significant and specific approach to the integration of data in epidemiology. Following Giroux (this volume), there are several different dimensions of integration that need to be discussed in the case of the exposome – data integration is a specific and particularly interesting type in this context.

As I have mentioned, the data integration that takes place in exposome research is largely based on the extension of genomic solutions and techniques to epidemiology. Practically, this mostly means integration between molecular data on the internal dimension of exposure and environmental data on various types of external exposure, which frames the use of genomic data as a platform to link biological and environmental data to study exposure. What count as biological and environmental data can vary rather substantially depending on the specific focus of research groups and projects. In the context of exposome research, biological data have mostly been identified with molecular and omics data, whilst environmental data have varied from climatic data and environmental sampling to the collection of questionnaires data in cohort studies and social scientific data more generally.[2] This has led to the development of new epistemic strategies for the integration of biological and environmental data in exposome research and their alignment with existing methodological approaches in epidemiology (Canali 2020b). In particular, the use of omics and molecular data has pushed for a microscopic focus on internal exposure as a central aspect of epidemiological research – to the point that this focus has led to changes also in

[1] These features have led to significant differences between traditional approaches to exposure in epidemiology and the exposome, for instance at the level of the conceptualisation of internal exposure, the use of omics and other molecular tools, and related types of funding required and sought after. Yet the exposome research has been only partially successful at securing long term funding and, as helpfully suggested by one anonymous reviewer, the exposome approach has not taken over epidemiology in general (Canali 2020a).

[2] This variety has consequences for the conceptual dimensions and boundaries of crucial notions for epidemiology, such as exposure and environment. As one anonymous reviewer helpfully suggested, the concept of environment is often quite vague in the context of exposome research: see Canali & Leonelli (2022) for more on the concept of the environment in exposome research and data-intensive epidemiology more generally and its relation with the availability of new environmental data.

the way external exposure is studied and environmental data are collected. In general terms, the focus has led to a push for more standardisation and quantification of the presence and effects of environmental exposure. For instance, exposome researchers have started to collaborate with disciplines in computer and information science, such as information systems, to quantify and estimate the presence of specific chemicals and pullulations in the external environment. On this basis, for example, the individual concentration of chemicals connected to air pollution, such as particulate air matter, is available to be compared and contrasted with molecular data on the presence of chemicals in the body and reactions that can be connected to health and disease (see e.g. Gulliver et al. 2018).[3]

The focus on molecular and internal components of exposure and the push for standardisation and quantification at the external level of exposure is not surprising. Part of the rationale and one of the aims of the original introduction of the exposome was the need for more standardisation in environmental sampling and the idea of transferring genomic solutions in epidemiology was indeed to match the level of quantification and precision of genomics for epidemiological and environmental research (Wild 2005; Rappaport and Smith 2010). The need for a movement in this direction was elicited by a general lack of focus on the impact of the environment of changes in health and disease in populations, but also the scattered and irregular state of environmental sampling and difficulties with data integration and collaboration across the health and environmental sciences.[4]

The approach applied by the exposome is thus based on the use of molecular data as a platform and basis to use environmental and particularly climatic data in the study of population health. In this direction, exposure profiles developed through omics techniques are the basis for the collection of data on external exposure of a similar, comparable level of abstraction and resolution. For instance, this has led to the collection of geographic data on the presence of pollutants and toxicants in the environment surrounding participants to a cohort studies, such as data on particulate air matter and other air pollutants (Canali 2020b). In these ways, one of the epistemic goals of the exposome approach has been to analyse correlations between molecular and environmental data with the aim of identifying causal relations between health, social and environmental features and study their effects on health. In this direction, in the exposome context, several statistical methods and approaches have been developed and used, such as the "meet-in-the-middle approach”

[3] As a result, in exposome research environmental data are often equated with climatic and geographical data. This is not just a feature of exposome research, as similar considerations can be traced to other data-intensive approaches in epidemiology as well as other areas of the life and health sciences, such as toxicology, exposure science, biomarkers research (Canali and Leonelli 2022).

[4] This is connected to more general aspects of the history and development of the health sciences, epidemiology in particular, and the environmental sciences. For more on this in the context of history of notion of environment and research in the environmental sciences, see Warde et al. (2018).

(Chadeau-Hyam et al. 2011; 2013). These have been interpreted philosophically as ways to move beyond the sole focus on associations – a typical feature of epidemiological research – and collect elements of mechanistic evidence, which together with difference-making evidence are necessary to develop robust causal claims in the health sciences (Russo and Williamson 2007; Russo 2009; Illari and Russo 2016; Russo and Vineis 2016).

These have clearly been steps in a new and promising direction, but this approach to data integration also suffers from significant limitations – as exposome researchers are aware. The reliance on genomic solutions and the use molecular and omics data as proxies to study environmental features has been considered an extension of the reductionism of genomics to new areas of the health sciences (Landecker and Panofsky 2013; Shostak and Moinester 2015). Using, for instance, inflammation and oxidative stress at the molecular data as proxies for the impact of socio-economic and other environmental conditions can be seen as a problematic way of quantifying qualitative processes. As a response to this criticism, it is important to note that the data integrated through the exposome approach can be used to develop mechanistic explanations of the impact of the environment on health and disease (Illari and Russo 2016; Russo and Vineis 2016). Yet using molecular data to build mechanistic evidence faces challenges as well. Omics techniques are used in exposome research to look for specific entities whose presence may be due to and depend on external exposure, rather than the activities and dynamics that can make up mechanistic explanations – to the point that exposome researcher are usually extremely cautious when it comes to mechanistic explanations and, more generally, causal claims on the basis of molecular data (Canali 2019). In addition, more recent exposome projects have relied on genomic data as a potential proxy for the study of some social features and determinants of health and disease (Ghiara and Russo 2019). The collection and development of mechanistic evidence about the pathways through which social and environmental features can have an impact on health and disease is crucial for the ability to act on those pathways and improve health for a population. Yet, often, mechanistic evidence is not sufficient here: many environmental and social dynamics that have a significant impact on a population operate through various different pathways and mechanisms. For instance, racism and poverty can operate in various different ways and may be sources of exposure at the level of education, pollution, diet, mentorship, etc. (Valles 2019). Knowing about one or more of these is crucial, but in many cases it can be more important to know the differences and dependencies that these causes have with disease and act on them, instead of getting precise pictures of how these have an impact (Giroux, this volume; Valles 2021).

4. Planetary Health: Integrating Diverse Environmental Data

The need to expand the integration of data on the environment in epidemiology is a central goal and defining feature of the exposome approach. Yet the exposome as an approach and area of research is not alone in the landscape of contemporary epidemiology – various

approaches and new conceptualisations aim to capture the needs and types of integration of the environment for health. *Planetary health* is one of these novelties and is an increasingly important concept for epidemiology and the life sciences more generally.[5]

Planetary health has been introduced in the last decade as a conceptual and methodological expansion for the field of public health. As an approach, planetary health has encouraged an inclusion of more and different human, physical, animal, climate environments as the types of environments that should be analysed in the study of population health. As discussed by Richard Horton and colleagues in the introduction of the concept in 2014, planetary health is supposed to direct more light on the "fragility of our planet and our obligation to safeguard the physical and human environments" (Horton et al. 2014, 847). Planetary health is thus based on a twofold realisation: health and environment are intrinsically connected phenomena, as changes and features of different environments have an impact on health at various levels; conversely, the environment is clearly shaped by changes and attitudes of populations related to health. As a result, in population health the state of health and disease of environments and populations is equated, with the goals of looking at health as a complex, multi-level, but essentially unitary phenomenon. Health is seen as a multi-level property precisely because of the interactions and interrelations between environments, species, populations, and individuals. But it is unitary in the sense that health is a property, according to the planetary point of view, of both individuals and populations on the one hand and environments and ecosystems on the other. Hence the need to consider these complex phenomena and study them together, for the improvement of both.

Since the introduction of the notion, planetary health has gained momentum with the expansion of new declarations and campaigns, such as the São Paulo Declaration on Planetary Health from the fall of 2021 (Myers, Pivor, and Saraiva 2021), and has now dedicated journals, including for instance a speciality journal published by *The Lancet*.[6] This shows a first, significant difference between the planetary health and conceptualisations of health and disease of the exposome. Planetary health has been first and foremost a contribution for political discussions at the level of policy on the environment, biodiversity, social determinants of health and, as such, consequently on the management, organisation, and funding on these issues. This means that planetary health is currently a conceptual framework and political project, rather than a very concrete methodological approach such as the exposome. This does not mean, however, that the concept should be dismissed as something that is significant 'only' at the policy level and is thus not particularly interesting for discussions on the epistemology and methodology of the health sciences. On the

[5] In turn, planetary health shares significant features with and is close to other conceptualisations, such as 'one health' and 'spatial epidemiology'.

[6] See *The Lancet Planetary Health*: https://www.thelancet.com/journals/lanplh/home (accessed October 2022). As a write this, Richard Horton, the author of the paper introducing the concept of planetary health (Horton et al. 2014) is editor in chief of *The Lancet*.

contrary, conceptualisations, values, and decisions at the political and economic level of the sciences clearly have a significant epistemic impact on the ways in which research is funded, managed, directed but also conducted at a concrete and practical level. This is why the focus on data integration can be helpful to analyse and discuss the assumptions, goals, and limits of an approach and conceptualisation such as planetary health.

The approach to data integration for health applied by planetary health is based on an expansion of the types of data that are considered and used as environmental data, with a push for more focus on different environments and the interactions between different species that inhabit these environments. This might seem like a clearly important focus and almost trivial realisation, but is already quite different from the approach of exposome research, where the expansion is mostly focused on introducing genomic techniques and data in epidemiology and thus integrating population health and the environment on the basis of an expansion of internal exposure and health data. As we have seen with the discussion of the exposome, studies of population health based on environmental data are far from abundant in both the environmental and health sciences. On the one hand, research on environmental changes and more generally the idea that the environment is something that changes and should be studied for health is relatively recent. Before the in the mid-1800s, the notion of 'environment' referred to the set of background circumstances, conditions and stimuli that shape the character of an individual, situation or phenomenon (Warde, Robin, and Sörlin 2018). According to this meaning of the notion, the environment was always the environment *of* something; and the focus was on this something, i.e. the individual, situation or phenomenon affected by the environment. While the notion also referred to conditions in the environment that needed to be acted upon, these conditions were mostly considered unchangeable and immutable per se – the focus of possible changes was again on the subject, rather than on the environment itself. In epidemiology, this focus changed with the hygiene and sanitary movements of the second half of the nineteenth century, as agents were identified as causes of infectious disease and loci of intervention beyond the subject. Historically, the development and success of epidemiology is deeply connected to these policy interventions, but the study of population health has long focused its priority on individual behaviours and lifestyle choices, and thus not necessarily on the environment (Morabia 2004). At the same time, in the natural and physical sciences, the shift towards viewing the environment as a subject of specific interest and agency only happened in the second half of the twentieth century, in connection with various other trends in the sciences, politics, and society more generally. As a result, the study of the environment and population health have interacted relatively rarely.

This background on the interactions between the study of the environment and population health has concrete consequences on the limitations of the planetary health approach to data integration. The studies and methods used to analyse the relations between population health and its environment in epidemiology are configured in ways that allow for

the integration of exposure data from different sources and their assemblage in specific configurations to study health and disease conditions of interest (Bauer 2013). These approaches to the collection and analysis of environmental data enable epidemiologists to study populations over extended periods of time and compare them over different configurations through "probabilistic thinking" (Morabia 2004). As a result, these methods are usually coordinated and centralised efforts that generate observational data about the environment, but in ways that tend to be quite different and largely independent from the experiments, models, and simulations employed in environmental research. Therefore, in epidemiology the study of the environment in relation to population health is normally based on data that are highly related to a population and environmental exposure and only partially to the direct study of the environment. In most cases, no direct sampling of environmental pollutants can be conducted and data about the internal biochemical environment are used as a proxy for variables tracking specific features and changes in the external environment.

A more direct focus on the environment for population health is thus made difficult by differences between the types of data collected for environmental and health research, their varying time scales, and frequencies (Fleming et al. 2017). In addition, epidemiological methods usually employ regression approaches that focus on few exposure factors for small groups and single health outcomes, which is problematic when trying to address the overall impact of the environment on health. While epidemiologists are clearly interested in monitoring and surveilling population health, climate and environmental scientists have developed tools and methods for the estimate and prediction of climate events (Parker 2018). This is an issue for planetary health research, particularly at the level of gaps in the collection of data that are considered crucial for the study of health and disease in their connection with diverse types of environments. Similar gaps and issues are connected to disciplinary boundaries between epidemiology and environmental research, which make collaborations and interdisciplinarity difficult in this context. The epistemic impact of these issues, moreover, is particularly severe at the level of data interpretation and use as evidence, particularly when it comes to causal analysis and inference. The aforementioned gaps in the causal attribution of environmental determinants of health and disease are more severe in the case of missing and unclear data collection on the environment.

5. Global Health: Integrating Diverse Health Data

With both the introduction of the exposome and planetary health concepts we see new and different approaches to the integration of health and environmental data to study the relation between health and the environment. In this context, an additional approach that is important to discuss from a data integration perspective is *global health*. Global health is the consideration of health in global terms: a view of health and disease as the result of the

differently and unequally distributed needs and issues of the various population that live in different parts of the world and in related environments.[7]

The concept emerged in the late 1990s as a way of expanding the notion of international health, that was extensively used until then with reference to the issue of epidemics and their control beyond the borders of individual nation states. Global health was introduced in this context as a way of expressing new interest and concern for health needs and inequalities of the whole global population, beyond the need to control specific issues such as epidemics and pandemics (Brown, Cueto, and Fee 2006). Various health institutions and national and translational political bodies now use global health as a framing to discuss health policy and related issues. Similarly to planetary health, global health is based on a relatively recent realisation in the health sciences and epidemiology in particular – the idea that socio-economic factors have a crucial impact on the development of health and disease at the individual and population level, to the point that significant differences and inequalities in these socio-economic factors can have a significant impact on health and disease. This might seem like a trivial realisation by epidemiologists and health policy-makers, considering that epidemiology is traditionally the area of biomedical research that directly focuses on the distribution and determinants of disease and health in populations (Broadbent 2013) and the historical development of the discipline has as such been tied to various public health measures and interventions on the environment, including socio-economic environments (Morabia 2004). Indeed, the study of socioeconomic status is a traditional and historical feature of public health and arguably one of the defining characteristics at the origin of the field, but the notions of social determinants of health is relatively recent and connected to research in the 1970s (such as the famous Whitehall studies on grade of employment and cardiovascular disease).

However, similarly to planetary health, contemporary epidemiology and the focus on risk factors have rarely led to collaborations and interactions with the social sciences. The reason is evident when looking at data integration as the main focus of our analysis. Evidential standards tend to be significantly different between the health and social sciences, as contemporary epidemiological research is usually based on individuals and not populations as the main units of data collection and analysis (Kelly and Russo 2021), curation and annotation of metadata are differently advanced in the two disciplines (Boumans and Leonelli 2020), and the social sciences often lack the data and storing infrastructures that are typical of the biomedical sciences (Ankeny and Leonelli 2016). This has led to little similarity and compatibility at the level of data and evidence across the life, health, and social sciences. Epistemological and knowledge standards can vary greatly between the health and social sciences, with the latter focusing on qualitative and often descriptive goals whilst the former aim for the quantitative study of the determinants of health and disease

[7] See Giroux (2021) on conceptual approaches and limits to the idea of populations being healthy.

and knowledge is normally presented in quantitative terms with the crucial goal of generalisation. As a result, the interactions between environments, societies, populations, and individuals are significantly understudied and rarely with the goal of causal inference in mind (Ghiara and Russo 2019).[8] A general yet substantial disregard for socio-economic determinants of health and disease is thus crucial background for the introduction of the concept of global health. But this is even more significant when considering a more general and widespread disregard for socio-economic features and causes of health and disease in other areas of the world than the West.

The approach of global health is based on the use of health data from all the different populations of the globe for the study of population health. Global health is in this sense primarily a framework coming from health and economic policy, gaining its roots and background in planetary health, based on the economic and political role of the World Bank and the United Nations, and now emerging at the intersection of these and other public institutions such as the World Health Organisation and nongovernmental and translational players such as corporations and non-profit foundations (Reubi 2018). This is also the context where the collection and integration of global health data has taken place, with the development of large databases infrastructures and related data analysis, integration, and visualisation tools such as the Global Burden of Disease, a research centre and database located at the University of Washington and funded by the Gates Foundation. Yet this is also the context where we see the limitations of the approach to data integration of global health.

The emergence and role of databases such as the Global Burden of Disease has been analysed by several historians and sociologists interested in the political and economic shifts around global health (Brown, Cueto, and Fee 2006; Birn 2009; Reubi 2018).[9] As discussed by Jean-Paul Gaudilliere and Camille Gasnier, the collection and use of global health data in the context of the Global Burden of Disease initiative was originally aimed at policy, particularly economic interventions for economic growth (Gaudilliere and Gasnier 2020). The initial development of the Global Burden of Disease was framed in particular around the comparison between different health interventions, with the goal of using health data as a basis to identify the most economically effective and efficient interventions. However, the shift towards the global health framework has also implied a shift in the goals of data collection and integration in the Global Burden of Disease, whereby global health data are increasingly used as single indicators of health and disease in different geographical distributions, rather than relating data to economic considerations about growth and

[8] The COVID-19 pandemic is a source of examples in this direction, where these differences are partly responsible for the fact that public health policy-making has rarely been grounded in social scientific knowledge and evidence (Lohse and Canali 2021).

[9] See https://www.thelancet.com/gbd (accessed October 2022).

efficiency (Gaudilliere and Gasnier 2020, 362–66). In other words, this has increasingly rendered global health data as primarily *health* data, to be used for the analysis and study of health and disease in the global population. A similar shift in the use of large datasets at the global level in epidemiology has taken place with the emergence of visualisation tools connected to these data, for instance with dedicated dashboards and maps that can track the spreading and impact of specific diseases and display these phenomena in constantly updated maps at a global scale.[10]

With global health we thus see the repurposing of data collected at the global level for health purposes and their integration with other available data for the tracking and analysis of health and disease at the global scale. This shift has had significant consequences at the research and policy level: for instance, the use of data from the Global Burden of Disease has successfully shown the extent to which mental health and mental disorders are increasingly present in the global South. Yet this repurposing for data integration has significant limitations too. The use of Global Burden of Disease data for the study of mental health has mostly not been based on local studies in the global South, but rather on the "complex set of correlations between the burden of mental health disorders and various epidemiological, social and economical variables worked out in countries benefiting from more reliable statistics" (Gaudilliere and Gasnier 2020, 365). The type of data integration that has been elicited by global health initiatives is thus often based on data that are not actually global, which is a crucial problem for the intended aim of global health to study health and disease of the whole global population, including usually neglected populations. The problem here is that in many cases data on health and disease in areas of the world such as the Global South are just not available and initiatives to increase local data collection are scarce.[11] More recently, a number of project in global health have tried to implement an approach presented as "precision global health" (Flahault et al. 2020; Sheath et al. 2020). The goal of these project can be seen as precisely tapping into the issues discussed in this section – the lack of local and indigenous data on population health – through the collection of health data from digital devices such as smartphones, wearables, trackers. Setting aside ethical and social considerations on the use of these types of data as evidence for the study of population health, there is a broader issue that affects data integration for global health. In most cases, data from neglected areas of the world and low-resourced research environments tend to be perceived as low quality, which are not up to being integrated with other data (Leonelli 2017). This is among the reasons why more advanced technologies, such as digital devices, are proposed as ways of dealing with these issues. The use of digital technology as a way of including areas of the world that are underrepresented and excluded areas from health data collection is promising (Celi 2022),

[10] These dashboards and maps have gained increasing political and epistemic importance in the context of the COVID-19 pandemic, see a critical analysis of this by Susanne Bauer (2021).

[11] See similar considerations by Rachel Ankeny on the need to bring data "out of the shadows" (Ankeny 2017).

but the quality of data collected through these devices is often unclear and not transparent because of commercial interests. Data collection for global health has thus a crucial clash at its centre – the contrast between the need for more consideration of the global determinants of health and disease and evidential standards of what count as high-quality data across the globe.

Approaches	**Features**	**Limitations**
Exposome	- Expansion of health data (e.g. molecular, omics, climatic) - Focus on genomics and omics techniques	- Reductionism of genomic approaches - Genomic data as a proxy for environmental and social features
Planetary health	- Expansion of environmental data (geographical, climatic, animal) - Focus on different environments and interactions between different species	- Gaps in data collection - Disciplinary boundaries
Global health	- Repurposing of health data at the global level (local data, data from the Global South) - Focus on health needs and inequalities of the global population	- Missing data - Varying approaches towards data quality

Table 1. Approaches to data integration analysed with respect to their specific features and limitations

6. Conclusions

In this chapter, I have analysed three approaches to the study of the relations between environment and population health, which have recently emerged in the context of various social and political discussions on health and biomedical research. Using the conceptual lens and methodological choice of focusing on data integration, I have looked at the exposome,

planetary health, and global health as specific and distinct approaches to integrating different types of environmental and health data (see Table 1). With the introduction of these three concepts, we see new and different approaches to the integration of health and environmental data to study the relation between health and the environment, but also a conceptual and methodological expansion of what count as environment and health in relation to what count as environmental and health data and the ways these should be used and integrated for research in the health sciences.

The exposome approach shows an expansion at the internal level of the environment and health, with the transfer of the genomic and molecular scaffolding for the study of internal exposure, their extension to epidemiology with an increasing focus on the individual level of environmental exposure, and thus a *vertical* expansion of the diversity of health data. As we have seen, this has significant consequences at the level of methodological approaches as well as disciplinary collaborations for the contemporary landscape of molecular and environmental epidemiology. The introduction of the concept of planetary health is a further expansion of these boundaries, but at the external level of the environment: a *horizontal* expansion of the types of environments that need to be analysed for health research, with a focus on the physical features of different environments and the interactions between different species that inhabit these environments and, therefore, also on different ways of collecting and integrating data on these new aspects of types of environment and health. As a consequence, the raise of this concept has significant consequences for the disciplinary boundaries of epidemiology, with the inclusion of further political and policy movements and the development of causal inference reflections on the influence of different environments on the health and disease of human and multispecies populations. The raise of the concept of global health has instead pushed for an expansion of the collection of health data – a *horizontal* expansion of the types of data that are considered necessary and the populations that need to be monitored and included in studies of population health and as such need to be integrated with environmental data. As a result, the extension of global health as a conceptual framework and approach is tied with the development of data infrastructures and partnerships across health and political institutions in different parts of the world.

The focus on data integration sheds light on the concrete implications of following these conceptual and methodological frameworks, thus enabling more understanding of the epistemic implications of otherwise abstract and sometimes vague conceptualisations related to health, disease, and the environment. As we have seen throughout the chapter, this also includes shedding light on the limitations of these approaches. A common theme on limitations emerges as a result of the analysis in this chapter, which are related to the use of data as an asset to build collaborations between the health and environmental sciences. One of the consequences of shifts in the volume and diversity of scientific data in the last two decades has been the increasing value of scientific data as epistemic, social, and

political assets (Leonelli 2019). In this sense it is unsurprising that changes in conceptual and methodological frameworks of the health sciences are crucially tied to choices of which data to integrate and how. Yet this renewed value of scientific data lies in the extraction of evidence and knowledge from data as an asset, which is not an automatic nor neutral act. As we have seen, the use of specific types of health and environmental data (e.g. planetary health and epidemiological data) is tied with methodological choices and epistemic assumptions of the collection and interpretation of the same data – data are no ready-made solutions for the study of relations between environment and health and the collaboration between fields therein. The assumptions, methods, judgements that form the contextual features of e.g. molecular data need to be taken into account when data are integrated and aligned with other and existing assumptions, methods, judgements of the different disciplinary approaches in e.g. exposome research. This is even more significant when, as we have seen for instance with global data, data are not available and other data need to be repurposed for new uses. Data is hence a crucial asset for the study of the relations between environment and health in contemporary epidemiology – but attention to the contextual and epistemic implications of using different data is in turn crucial in order to fulfil these aims.

Preprint of the chapter in the volume *Integrative Approaches in Environmental Health and Exposome Research: Epistemological and Practical Issues* (E. Giroux, F. Merlin, Y. Fayet Eds.) https://doi.org/10.1007/978-3-031-28432-8_5

References

Ankeny, Rachel A. 2017. 'Bringing Data Out of the Shadows'. *Science, Technology, & Human Values* 42 (2): 306–10. https://doi.org/10.1177/0162243916689138.

Ankeny, Rachel A., and Sabina Leonelli. 2016. 'Repertoires: A Post-Kuhnian Perspective on Scientific Change and Collaborative Research'. *Studies in History and Philosophy of Science Part A* 60 (December): 18–28. https://doi.org/10.1016/j.shpsa.2016.08.003.

Bauer, Susanne. 2013. 'Modeling Population Health: Reflections on the Performativity of Epidemiological Techniques in the Age of Genomics'. *Medical Anthropology Quarterly* 27 (4): 510–30. https://doi.org/10.1111/maq.12054.

———. 2021. 'Pandemic Infrastructure: Epidemiology as Compartmentalization'. *Mefisto. Rivista Di Medicina, Filosofia, Storia* 5 (1): 79–104.

Birn, Anne-Emanuelle. 2009. 'The Stages of International (Global) Health: Histories of Success or Successes of History?' *Global Public Health* 4 (1): 50–68. https://doi.org/10.1080/17441690802017797.

Boniolo, Giovanni, and Marco J. Nathan, eds. 2017. *Philosophy of Molecular Medicine: Foundational Issues in Research and Practice*. New York: Routledge, Taylor & Francis Group.

Boumans, Marcel, and Sabina Leonelli. 2020. 'From Dirty Data to Tidy Facts: Clustering Practices in Plant Phenomics and Business Cycle Analysis'. In *Data Journeys in the Sciences*, edited by Sabina Leonelli and Niccolò Tempini, 79–101. Cham: Springer International Publishing. https://doi.org/10.1007/978-3-030-37177-7_5.

Brigandt, Ingo. 2010. 'Beyond Reduction and Pluralism: Toward an Epistemology of Explanatory Integration in Biology'. *Erkenntnis* 73 (3): 295–311. https://doi.org/10.1007/s10670-010-9233-3.

Broadbent, Alex. 2013. *Philosophy of Epidemiology*. London: Palgrave Macmillan UK. https://doi.org/10.1057/9781137315601.

Brown, Theodore M., Marcos Cueto, and Elizabeth Fee. 2006. 'The World Health Organization and the Transition From "International" to "Global" Public Health'. *American Journal of Public Health* 96 (1): 62–72. https://doi.org/10.2105/AJPH.2004.050831.

Canali, Stefano. 2019. 'Evaluating Evidential Pluralism in Epidemiology: Mechanistic Evidence in Exposome Research'. *History and Philosophy of the Life Sciences* 41 (1): 4. https://doi.org/10.1007/s40656-019-0241-6.

———. 2020a. 'What Is New about the Exposome? Exploring Scientific Change in Contemporary Epidemiology'. *International Journal of Environmental Research and Public Health* 17 (8): 2879. https://doi.org/10.3390/ijerph17082879.

———. 2020b. 'Making Evidential Claims in Epidemiology: Three Strategies for the Study of the Exposome'. *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences* 82 (August): 101248. https://doi.org/10.1016/j.shpsc.2019.101248.

Canali, Stefano, and Sabina Leonelli. 2022. 'Reframing the Environment in Data-Intensive Health Sciences'. *Studies in History and Philosophy of Science* 93 (June): 203–14. https://doi.org/10.1016/j.shpsa.2022.04.006.

Preprint of the chapter in the volume *Integrative Approaches in Environmental Health and Exposome Research: Epistemological and Practical Issues* (E. Giroux, F. Merlin, Y. Fayet Eds.) https://doi.org/10.1007/978-3-031-28432-8_5

Celi, Leo Anthony. 2022. 'PLOS Digital Health, a New Journal Driving Transformation in the Delivery of Equitable and Unbiased Healthcare'. *PLOS Digital Health* 1 (1): e0000009. https://doi.org/10.1371/journal.pdig.0000009.

Chadeau-Hyam, Marc, Toby J. Athersuch, Hector C. Keun, Maria De Iorio, Timothy M.D. Ebbels, Mazda Jenab, Carlotta Sacerdote, Stephen J Bruce, Elaine Holmes, and Paolo Vineis. 2011. 'Meeting-in-the-Middle Using Metabolic Profiling – a Strategy for the Identification of Intermediate Biomarkers in Cohort Studies'. *Biomarkers* 16 (1): 83–88. https://doi.org/10.3109/1354750X.2010.533285.

Chadeau-Hyam, Marc, Gianluca Campanella, Thibaut Jombart, Leonardo Bottolo, Lutzen Portengen, Paolo Vineis, Benoit Liquet, and Roel C.H. Vermeulen. 2013. 'Deciphering the Complex: Methodological Overview of Statistical Models to Derive OMICS-Based Biomarkers: Statistical Approaches for OMICS-Based Biomarkers'. *Environmental and Molecular Mutagenesis* 54 (7): 542–57. https://doi.org/10.1002/em.21797.

Dupré, John. 1996. *The Disorder of Things: Metaphysical Foundations of the Disunity of Science*. 1. paperback ed., Nachdr. Cambridge, Mass.: Harvard Univ. Press.

Flahault, Antoine, Jürg Utzinger, Isabella Eckerle, Danny J Sheath, Rafael Ruiz de Castañeda, Isabelle Bolon, Nefti-Eboni Bempong, and Fred Andayi. 2020. 'Precision Global Health for Real-Time Action'. *The Lancet Digital Health* 2 (2): e58–59. https://doi.org/10.1016/S2589-7500(19)30240-7.

Fleming, Lora, Niccolò Tempini, Harriet Gordon-Brown, Gordon L. Nichols, Christophe Sarran, Paolo Vineis, Giovanni Leonardi, et al. 2017. 'Big Data in Environment and Human Health'. In *Oxford Research Encyclopedia of Environmental Science*, by Lora Fleming, Niccolò Tempini, Harriet Gordon-Brown, Gordon L. Nichols, Christophe Sarran, Paolo Vineis, Giovanni Leonardi, et al. Oxford University Press. https://doi.org/10.1093/acrefore/9780199389414.013.541.

Gaudilliere, Jean-Paul, and Camille Gasnier. 2020. 'From Washington DC to Washington State: The Global Burden of Diseases Data Basis and the Political Economy of Global Health'. In *Data Journeys in the Sciences*, edited by Sabina Leonelli and Niccolò Tempini, 351–69. Cham: Springer International Publishing. https://doi.org/10.1007/978-3-030-37177-7_18.

Ghiara, Virginia, and Federica Russo. 2019. 'Reconstructing the Mixed Mechanisms of Health: The Role of Bio- and Sociomarkers'. *Longitudinal and Life Course Studies* 10 (1): 7–25. https://doi.org/10.1332/175795919X15468755933353.

Gibbon, Sahra, Barbara Prainsack, Stephen Hilgartner, and Janelle Lamoreaux. 2020. *Routledge Handbook of Genomics, Health and Society*.

Giroux, Élodie. 2021. 'Can Populations Be Healthy? Perspectives from Georges Canguilhem and Geoffrey Rose'. *History and Philosophy of the Life Sciences* 43 (4): 111. https://doi.org/10.1007/s40656-021-00463-x.

Giroux, Élodie, Yohan Fayet, and Thibaut Serviant-Fine. 2021. 'L'Exposome: Tensions entre holisme et réductionnisme'. *médecine/sciences* 37 (8–9): 774–78. https://doi.org/10.1051/medsci/2021092.

Golub, Todd. 2010. 'Counterpoint: Data First'. *Nature* 464 (7289): 679–679. https://doi.org/10.1038/464679a.

Green, Sara, and Henrik Vogt. 2016. 'Personalizing Medicine: Disease Prevention in Silico and in Socio'. *HUMANA.MENTE Journal of Philosophical Studies* 9 (30): 42.

Preprint of the chapter in the volume *Integrative Approaches in Environmental Health and Exposome Research: Epistemological and Practical Issues* (E. Giroux, F. Merlin, Y. Fayet Eds.) https://doi.org/10.1007/978-3-031-28432-8_5

Gulliver, John, David Morley, Chrissi Dunster, Adrienne McCrea, Erik van Nunen, Ming-Yi Tsai, Nicoltae Probst-Hensch, et al. 2018. 'Land Use Regression Models for the Oxidative Potential of Fine Particles (PM 2.5) in Five European Areas'. *Environmental Research* 160 (January): 247–55. https://doi.org/10.1016/j.envres.2017.10.002.

Hilgartner, Stephen. 2017. *Reordering Life: Knowledge and Control in the Genomics Revolution*. Inside Technology. Cambridge, Massachusetts: The MIT Press.

Hogle, Linda F. 2016. 'Data-Intensive Resourcing in Healthcare'. *BioSocieties* 11 (3): 372–93. https://doi.org/10.1057/s41292-016-0004-5.

Holmberg, Christine, Christine Bischof, and Susanne Bauer. 2013. 'Making Predictions: Computing Populations'. *Science, Technology, & Human Values* 38 (3): 398–420. https://doi.org/10.1177/0162243912439610.

Horton, Richard, Robert Beaglehole, Ruth Bonita, John Raeburn, Martin McKee, and Stig Wall. 2014. 'From Public to Planetary Health: A Manifesto'. *The Lancet* 383 (9920): 847. https://doi.org/10.1016/S0140-6736(14)60409-8.

Illari, Phyllis, and Federica Russo. 2016. 'Information Channels and Biomarkers of Disease'. *Topoi* 35 (1): 175–90. https://doi.org/10.1007/s11245-013-9228-1.

Kelly, Michael P, and Federica Russo. 2021. 'The Epistemic Values at the Basis of Epidemiology and Public Health'. *Mefisto. Rivista Di Medicina, Filosofia, Storia* 5 (1): 105–20.

Klingwort, Jonas, and Rainer Schnell. 2020. 'Critical Limitations of Digital Epidemiology'. *Survey Research Methods*, June, 95-101 Pages. https://doi.org/10.18148/SRM/2020.V14I2.7726.

Landecker, Hannah, and Aaron Panofsky. 2013. 'From Social Structure to Gene Regulation, and Back: A Critical Introduction to Environmental Epigenetics for Sociology'. *Annual Review of Sociology* 39 (1): 333–57. https://doi.org/10.1146/annurev-soc-071312-145707.

Leonelli, Sabina. 2013. 'Integrating Data to Acquire New Knowledge: Three Modes of Integration in Plant Science'. *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences* 44 (4): 503–14. https://doi.org/10.1016/j.shpsc.2013.03.020.

———. 2016. *Data-Centric Biology: A Philosophical Study*. Chicago London: The University of Chicago Press.

———. 2017. 'Global Data Quality Assessment and the Situated Nature of "Best" Research Practices in Biology'. *Data Science Journal* 16 (June): 32. https://doi.org/10.5334/dsj-2017-032.

———. 2019. 'Data — from Objects to Assets'. *Nature* 574 (7778): 317–20. https://doi.org/10.1038/d41586-019-03062-w.

Leonelli, Sabina, and Niccolò Tempini, eds. 2020. *Data Journeys in the Sciences*. Cham: Springer International Publishing. https://doi.org/10.1007/978-3-030-37177-7.

Lohse, Simon, and Stefano Canali. 2021. 'Follow *the* Science? On the Marginal Role of the Social Sciences in the COVID-19 Pandemic'. *European Journal for Philosophy of Science* 11 (4): 99. https://doi.org/10.1007/s13194-021-00416-y.

Mitchell, Sandra D., and Angela M. Gronenborn. 2017. 'After Fifty Years, Why Are Protein X-Ray Crystallographers Still in Business?' *The British Journal for the Philosophy of Science* 68 (3): 703–23. https://doi.org/10.1093/bjps/axv051.

Mitchell, Sandra D., and Michael R. Dietrich. 2006. 'Integration without Unification: An Argument for Pluralism in the Biological Sciences'. *The American Naturalist* 168 (S6): S73–79. https://doi.org/10.1086/509050.

Mittelstadt, Brent, Justus Benzler, Lukas Engelmann, Barbara Prainsack, and Effy Vayena. 2018. 'Is There a Duty to Participate in Digital Epidemiology?' *Life Sciences, Society and Policy* 14 (1): 9. https://doi.org/10.1186/s40504-018-0074-1.

Morabia, Alfredo, ed. 2004. *A History of Epidemiologic Methods and Concepts*. Basel: Birkhäuser Basel. https://doi.org/10.1007/978-3-0348-7603-2.

Müller-Wille, Staffan, and Isabelle Charmantier. 2012. 'Natural History and Information Overload: The Case of Linnaeus'. *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences* 43 (1): 4–15. https://doi.org/10.1016/j.shpsc.2011.10.021.

Myers, Samuel S, Jeremy I Pivor, and Antonio M Saraiva. 2021. 'The São Paulo Declaration on Planetary Health'. *The Lancet*, October, S0140673621021814. https://doi.org/10.1016/S0140-6736(21)02181-4.

Parker, Wendy. 2018. 'Climate Science'. In *Stanford Encyclopedia of Philosophy*. Vol. Summer 2018 Edition. Stanford (CA): Metaphysics Research Lab, Stanford University. https://plato.stanford.edu/archives/sum2018/entries/climate-science/.

Plutynski, Anya. 2022. 'Why Precision Oncology Is Not Very Precise (and Why This Should Not Surprise Us)'. In *Personalized Medicine in the Making. Philosophical Perspectives from Biology to Healthcare*, edited by Marta Bertolaso and Chiara Beneduce, 30. Springer.

Prainsack, Barbara. 2020. 'The Meaning and Enactment of Openness in Personalised and Precision Medicine'. *Science and Public Policy* 47 (5): 647–54. https://doi.org/10.1093/scipol/scaa013.

Rappaport, Stephen M., and Martyn T. Smith. 2010. 'Environment and Disease Risks'. *Science* 330 (6003): 460–61. https://doi.org/10.1126/science.1192603.

Reubi, David. 2018. 'Epidemiological Accountability: Philanthropists, Global Health and the Audit of Saving Lives'. *Economy and Society* 47 (1): 83–110. https://doi.org/10.1080/03085147.2018.1433359.

Richardson, Sarah S., and Hallam Stevens, eds. 2015. *Postgenomics: Perspectives on Biology after the Genome*. Durham: Duke University Press.

Russo, Federica. 2009. 'Variational Causal Claims in Epidemiology'. *Perspectives in Biology and Medicine* 52 (4): 540–54. https://doi.org/10.1353/pbm.0.0118.

Russo, Federica, and Paolo Vineis. 2016. 'Opportunities and Challenges of Molecular Epidemiology'. In *Philosophy of Molecular Medicine*, edited by Giovanni Boniolo and Marco J. Nathan. Taylor & Francis.

Russo, Federica, and Jon Williamson. 2007. 'Interpreting Causality in the Health Sciences'. *International Studies in the Philosophy of Science* 21 (2): 157–70. https://doi.org/10.1080/02698590701498084.

Salathé, Marcel. 2018. 'Digital Epidemiology: What Is It, and Where Is It Going?' *Life Sciences, Society and Policy* 14 (1): 1. https://doi.org/10.1186/s40504-017-0065-7.

Sheath, Danny J., Rafael Ruiz de Castañeda, Nefti-Eboni Bempong, Mario Raviglione, Catherine Machalaba, Michael S. Pepper, Effy Vayena, et al. 2020. 'Precision Global Health: A Roadmap for Augmented Action'. *Journal of Public Health and Emergency* 4 (March): 5–5. https://doi.org/10.21037/jphe.2020.01.01.

Preprint of the chapter in the volume *Integrative Approaches in Environmental Health and Exposome Research: Epistemological and Practical Issues* (E. Giroux, F. Merlin, Y. Fayet Eds.) https://doi.org/10.1007/978-3-031-28432-8_5

Shostak, Sara, and Margot Moinester. 2015. 'The Missing Piece of the Puzzle? Measuring the Environment in the Postgenomic Moment'. In *Postgenomics: Perspectives on Biology after the Genome*, edited by Sarah S. Richardson and Hallam Stevens. London: Duke University Press.

Stegenga, Jacob, Ashley Graham, Şerife Tekin, Saana Jukola, and Robin Bluhm. 2017. 'New Directions in Philosophy of Medicine'. In *The Bloomsbury Companion to Contemporary Philosophy of Medicine*, edited by James Marcum, 26. Bloomsbury Academic.

Strasser, Bruno J. 2019. *Collecting Experiments: Making Big Data Biology*. Chicago: The University of Chicago Press.

Valles, Sean A. 2019. *Philosophy of Population Health Science: Philosophy for a New Public Health Era*. Routledge.

Valles, Sean A. 2021. 'A Pluralistic and Socially Responsible Philosophy of Epidemiology Field Should Actively Engage with Social Determinants of Health and Health Disparities'. *Synthese* 198: 2589–2611. https://doi.org/10.1007/s11229-019-02161-5.

Vogt, Henrik, Bjørn Hofmann, and Linn Getz. 2016. 'The New Holism: P4 Systems Medicine and the Medicalization of Health and Life Itself'. *Medicine, Health Care and Philosophy* 19 (2): 307–23. https://doi.org/10.1007/s11019-016-9683-8.

Warde, Paul, Libby Robin, and Sverker Sörlin. 2018. *The Environment: A History of the Idea*. Baltimore, Maryland.

Weinberg, Robert. 2010. 'Point: Hypotheses First'. *Nature* 464 (7289): 678–678. https://doi.org/10.1038/464678a.

Wild, Christopher Paul. 2012. 'The Exposome: From Concept to Utility'. *International Journal of Epidemiology* 41 (1): 24–32. https://doi.org/10.1093/ije/dyr236.

Wild, Cristopher Paul. 2005. 'Complementing the Genome with an "Exposome": The Outstanding Challenge of Environmental Exposure Measurement in Molecular Epidemiology'. *Cancer Epidemiology Biomarkers & Prevention* 14 (8): 1847–50. https://doi.org/10.1158/1055-9965.EPI-05-0456.

Worrall, John. 2002. 'What Evidence in Evidence-Based Medicine?' *Philosophy of Science* 69 (S3): S316–30. https://doi.org/10.1086/341855.

———. 2007. 'Evidence in Medicine and Evidence-Based Medicine'. *Philosophy Compass*, 42.