



Evaluating evidential pluralism in epidemiology: mechanistic evidence in exposome research

Stefano Canali¹ 

Received: 23 July 2018 / Accepted: 1 February 2019
© The Author(s) 2019

Abstract In current philosophical discussions on evidence in the medical sciences, epidemiology has been used to exemplify a specific version of evidential pluralism. According to this view, known as the Russo–Williamson Thesis, evidence of both difference-making and mechanisms is produced to make causal claims in the health sciences. In this paper, I present an analysis of data and evidence in epidemiological practice, with a special focus on research on the exposome, and I cast doubt on the extent to which evidential pluralism holds in this case. I start by focusing on the claim that molecular data allows for the production of mechanistic evidence. On the basis of a close look at the ways in which molecular data is used in exposome research, I caution against interpretations in terms of mechanistic evidence. Secondly, I expand my critical remarks on the thesis by addressing the conditions under which data is categorised as evidence in exposome research. I argue that these show that the classification of a dataset as a type of evidence is dependent on the ways in which the data is used. This is in contrast with the approach of evidential pluralism, where evidence is classified in different types on the basis of its intrinsic properties. Finally, I come back to what I consider the core of the thesis and suggest that the epidemiological research analysed in the paper indicates different interpretations of evidential pluralism and its applicability in the health sciences.

Keywords Evidence · Data · Difference-making · Mechanism · Evidential pluralism · Exposome

✉ Stefano Canali
stefano.canali@philos.uni-hannover.de

¹ Institute for Philosophy, Leibniz Universität Hannover, Lange Laube 32, 30159 Hannover, Germany

1 Introduction

Evidence is a notion that we associate with multiple activities of our societies, including journalism, law, criminal investigations, etc. Evidence is an important issue when we want to use or produce knowledge. Science is therefore one of the areas where we would intuitively say that evidence plays a crucial role, and one may be even tempted to say that the use of evidence is among the features that characterises practices as scientific in the first place.

The topic of evidence has received philosophical attention primarily in epistemology and philosophy of science (Kelly 2016). In philosophy of science, many discussions have focused on the health sciences. This is partly a consequence of the rise of the methodological approach known as evidence-based medicine, which has brought about new ways of assessing the quality of evidence and provoked many critical reflections in the literature (Worrall 2002; Clarke et al. 2013; Stegenga 2014). As a result of this debate, a number of authors started to discuss how to classify evidence produced in medical research, connecting the issue to more general issues in philosophy of science, such as causality and causal inference (Campaner and Galavotti 2012).

Within this debate on causal inference and evidence classification in medicine, Federica Russo and Jon Williamson introduced a version of evidential pluralism according to which:

To establish causal claims, scientists need the mutual support of mechanisms and dependencies. ... The idea is that probabilistic evidence needs to be accounted for by an underlying mechanism before the causal claim can be established. (Russo and Williamson 2007, p. 159)

The ‘Russo–Williamson Thesis’ (RWT) identifies evidence of two kinds—evidence of causes making a difference to effects and evidence of underlying mechanisms—which has to be provided to make causal claims in the health sciences.¹ A number of examples have been used to illustrate the thesis, including research from epidemiology (Russo and Williamson 2012). Epidemiology can be defined as the area of the health sciences concerned with the study of the distribution and determinants of health states—particularly disease—in human populations (Broadbent 2013). Philosophers of science have recently turned their attention to epidemiology and related problems, including: the notion of cause in epidemiology (Russo 2009; De Vreese 2009); the causal interpretation of epidemiological results (Broadbent 2013); causal explanations (Campaner 2011); and the consequences of methodological novelties on causal inference (Broadbent 2015; Vandenbroucke et al. 2016). In this context, the RWT has played an important role, connecting philosophical work on epidemiology with the broader debate on medical evidence, where epidemiology had

¹ As noted by one reviewer, a few different versions of the thesis have been proposed since the original formulation by Russo and Williamson, and this can generate some confusion. Thus, let me state that this general interpretation of the thesis is the one I shall adopt throughout the article.

largely been overlooked, despite being among the most important sources of medical evidence.

In this paper, I present a critical analysis of the RWT from the perspective of contemporary epidemiological research on the ‘exposome’. The exposome approach combines data on the molecular and external environment to study the internal and external components of exposure and improve the assessment of disease risks associated with environmental elements. It has recently attracted significant attention and has been used as a source of examples in the literature on the RWT (Russo and Williamson 2012; Russo and Vineis 2016). My analysis of this line of research is based on the study of a specific project: EXPOsOMICS.² The project received funding by the European Commission in 2012 and, between 2012 and 2017, it was carried out by as a consortium of research centres in Europe and the US, coordinated by Imperial College London. EXPOsOMICS aimed at improving the assessment of exposure to air and water pollution and their impact on chronic disease, by studying and integrating internal and external components of exposure (Vineis et al. 2017a). The project can be considered a paradigmatic example of exposome research, as it was built on data collected by several other projects studying the exposome, with which it shares approaches and methods.³ At the same time, the aim of integrating data from these diverse sources to produce evidence on environmental exposure makes the project a particularly significant case in the context of discussions on data and evidence (Fleming et al. 2017).

My analysis will proceed as follows. I start by introducing the exposome approach and the way in which the RWT has been presented in this context (Sect. 2). I then turn to my critical remarks on the extent to which the RWT holds up in this case. I present challenges to two main aspects of the interpretation of exposome research based on the RWT, one pertaining to the interpretation of molecular data as mechanistic evidence and another to the way evidence is classified in the RWT framework. Firstly, in Sect. 3, I consider the claim that new sources of molecular data used in exposome research count as the production of mechanistic evidence and that, therefore, exposome research shows the interplay between mechanistic and difference-making evidence suggested by the RWT. On the basis of a detailed look at molecular data in EXPOsOMICS, I show that molecular data is used to study differences and dependencies of exposure and cast doubt on interpretations in terms of mechanistic evidence. Secondly, in Sect. 4, I consider the way in which the RWT distinguishes between different types of evidence. I focus on the ways in which data is used and

² My study is empirically grounded on the study of EXPOsOMICS publications, reports, presentations. In addition, I conducted a series of qualitative interviews, which took place at the MRC-PHE Centre for Environment and Health of Imperial College London and aimed to document practices of data collection, production and integration and notions and classifications of evidence. Quotations from these interviews are taken from transcripts. I will use the empirical material to provide a reconstruction of uses of molecular data and approaches to evidence classification in EXPOsOMICS. A more detailed empirical analysis of the ways in which data is collected, interpreted and used in the project can be found in Canali (under review).

³ See for instance the Helix project, a cohort study looking at the ‘early-life exposome’, which some EXPOsOMICS researchers helped set up (Maitre et al. 2018).

classified as evidence in EXPOsOMICS and I argue that this presents more general challenges to the approach to evidence classification of the RTW framework. On the basis of these criticisms, in Sect. 5, I come back to the ‘core’ of the RWT and ask if it can be reconciled with the challenges highlighted in the paper. I argue that this may be possible and the case of exposome research indicates different ways of interpreting evidential pluralism and its applicability.

2 Exposome research and evidential pluralism

First introduced by Christopher Wild (2005), the exposome refers to the totality of exposures experienced by individuals throughout their lifetime. Totality, here, means that the exposome encompasses exposures experienced at different levels, which include exposures to elements and features of the external environment (such as pollution, diet, stress, etc.) as well as the internal chemical environment (such as inflammation, oxidative stress, gut flora, etc.). These different levels exist because individuals’ exposure to external elements can produce toxicants and/or reactions at the internal level, which can in turn lead to the development of disease. In contrast with the more traditional focus on the external components of exposure, the ‘all-encompassing’ approach to the exposome is often presented as a new paradigm in the context of environmental epidemiology (Rappaport and Smith 2010).

Practically, much of the work on the exposome consists in the search for biomarkers, i.e., elements of the environment or the organism that can be measured and in turn measure the presence of some toxicant. Biomarkers are considered capable of signalling the development of disease, thus guiding researchers in the understanding of the impact of the environment on disease.⁴ In this context, molecular data is among the primary tools used to identify biomarkers of the internal component of exposure. More precisely, exposome researchers use omics technologies to measure the different levels at which exposure from environmental elements leaves traces in the body, including: processes that involve small molecules playing intermediary roles in the functioning of the metabolism (metabolomics); parts of proteins or DNA that are bound to specific chemicals (adductomics); epigenetic changes in the cells’ genetic material (epigenomics); mRNA expressions (transcriptomics) and proteins (proteomics).

Recently, Federica Russo and epidemiologist Paolo Vineis have connected the use of molecular data in exposome research to the RWT. They say that “molecular epidemiology improves on traditional epidemiology also in another important respect: it goes beyond associations and is in principle able to shed light on the *mechanisms* of disease causation” (Russo and Vineis 2016, p. 254; emphasis in original). Similar considerations have been developed elsewhere in the literature on the RWT. In one of the first publications including a philosophical analysis of exposome research, Russo and Williamson (2012) claim that this line of research shows an

⁴ For a discussion of working definitions and the conceptual status of biomarkers, see Strimbu and Tavel (2010).

interplay between difference-making and mechanistic evidence. Together with Phyllis Illari, Russo and Vineis use the exposome approach to cancer research to argue for an information transmission account of causality and talk of the need for evidence of both the difference that exposure makes to the development of cancer and of how exposure leads to the development of cancer (Vineis et al. 2017b).⁵

In this context, the significance of Russo and Vineis' claim is that exposome research can *produce*—and it does not only *use*—mechanistic evidence, and that this production is to be seen in the molecular features of exposome research.⁶ In this sense, exposome research is innovative compared to traditional epidemiology (also) because, thanks to molecular data, it does not only produce new difference-making evidence but also mechanistic evidence. In other words, the claim is not that exposome research *uses* mechanistic evidence that has been gathered in other areas of biomedical research and that this, together with the difference making evidence it produces, shows the interplay maintained by evidential pluralism. Rather, the argument is that the exposome approach improves on traditional epidemiology because it can study the mechanisms of disease causation and, thus, produce mechanistic evidence. It is this claim about the production of a specific kind of evidence that will be the target of my critique in the next section.

Before presenting my remarks, let me briefly summarise the distinction between difference-making and mechanistic evidence developed by the RWT. Evidence of difference-making refers to evidence that is used to show that effects depend on causes and that causes make a difference to effect (Russo and Williamson 2007, p. 158). Randomised controlled trials are among the most cited examples of a methodology producing difference-making evidence. Other examples, from Illari (2011a), include evidence gathered in observational studies, probability distributions in Bayesian networks, as well as counterfactual and invariant relationships. As such, Russo and Williamson say that difference-making often comes in a probabilistic form and is usually employed to make predictions.⁷

As for mechanistic evidence, Russo and Williamson suggest that we should think of it as evidence of the mechanism(s) underlying the relation between causes and effects. Mechanistic evidence is what provides the productive component of causal claims, which gives information on *how* causes bring about effects. They also argue that this is the kind of evidence which plays the crucial role of helping with issues possibly affecting difference-making evidence, such as confounding (Russo and Williamson 2012, pp. 249–250). This notion of mechanistic evidence has often been criticised because of its vagueness and ambiguity, which has in turn led to disambiguating work in the literature on the RWT. For example, Illari argues that we should

⁵ For a presentation of Illari and Russo's information transmission account, see Illari (2011b) and Illari and Russo (2016).

⁶ I will come back this distinction between use and production of evidence and why it is important in Sect. 5.

⁷ Russo and Williamson sometimes use the term *probabilistic* evidence instead of *difference-making* evidence. For matters of consistency, I will follow their original claim on the connection of probabilistic evidence with causes making a difference to effects and, in line with Illari (2011a), I will only use the term *difference-making* evidence.

not think of the distinction between difference-making and mechanistic evidence as a distinction between the methods we use to gather evidence, but rather as a difference between objects of evidence (Illari 2011a). In her view, mechanistic evidence is not necessarily the evidence we get through, say, interventions, as opposed to observational studies. Rather, mechanistic evidence is evidence of the existence of mechanisms, evidence that is about mechanism and has mechanisms as its' object. Following her and Williamson's characterisation, evidence of mechanisms thus refers to evidence of the entities, activities or the way these are organised to produce the phenomenon for which the mechanism is responsible (Illari and Williamson 2012). In this view, potential mechanistic evidence for, say, protein synthesis would be evidence of the organised entities (DNA, RNA, ribosomes, etc.) and activities (e.g. transcription, regulation, etc.) involved in the mechanism that brings about the synthesis of proteins.

Despite these disambiguating efforts, the evidential pluralism proposed by the RWT has received significant criticism in the literature, especially with regard to the strong stance on mechanistic evidence. For example, Howick (2011) has casted doubt on the claim that mechanistic evidence has always to be provided in order to establish causal claims. In the philosophy of epidemiology, Broadbent (2011) has argued against mechanistic interpretations of the causal inference strategies used by epidemiologists. Gillies (2011) has shown that the epidemiological discovery of the causal relation between smoking and lung-cancer was first established on the basis of difference-making evidence only, much earlier than any mechanism was found.

3 Mechanistic evidence in exposome research

The use of molecular data in exposome research has been interpreted as the production of mechanistic evidence, which is considered necessary by the RWT to support a causal claim. I now give a close look at how molecular data is collected and used in exposome research, focusing on examples of the use of exposure profiles from molecular data in EXPOsOMICS. This analysis leads me to argue that, in the context of EXPOsOMICS, evidence from molecular data is not really used as mechanistic evidence, at least if we are to follow the notion of mechanistic evidence developed in the RWT literature.

In EXPOsOMICS, molecular data was collected through the analysis of the microscopic structure of samples from cohort studies. For example, blood plasma samples retrieved from cohort studies were analysed using the adductomics technique, i.e., by collecting data on protein adducts, which are formed by the binding of proteins to chemicals, in order to study the effects of exposure to air pollution on proteins. As one EXPOsOMICS researcher explained:

Adductomics is an omic technique and the specificity of the method is that it measures adducts of human serum albumin and toxicants. The toxicants present in your serum tend to have a very fast turnaround, so they disappear and they become metabolite and they're very difficult to measure. So, for example, if you've been exposed to air pollution it's going to be difficult

to find the related toxicant in your serum. But the same molecules can bind to albumin and then they are conserved in your serum for a longer period of time. So, by measuring these adducts, you can track exposure for longer time. (Researcher A1)

Results consist in spectrometry data, i.e., plots with lines and peaks, where different peaks correspond to different molecules and break down the sample into its molecular elements and compounds. These are usually called ‘exposure profiles’ or ‘omic profiles’. On the basis of exposure profiles, researchers try to identify toxicants whose presence can be considered a consequence of exposures to external pollutants and can in turn be used as an intermediate biomarker to track the effects of exposure and possibly the development of disease. Exposure profiles are usually presented as tables where lines are dedicated to chemicals and the content of the columns provides information on the measured mass of the chemical within the sample.

As we have seen, according to the disambiguated notion provided by Illari, mechanistic evidence is the kind of evidence that has mechanisms as its object (Illari 2011a); and, following Illari’s work with Williamson, a mechanism for a phenomenon consists in the entities and activities organised in order to be responsible for the phenomenon (Illari and Williamson 2012). The first problem I see in interpreting evidence from exposure profiles as mechanistic evidence is precisely that the ‘object’ is not a mechanism, to use Illari’s language. The rationale behind the use of molecular data in exposome research is to look for chemicals whose presence may be due to an exposure to the external environment, not to look for the ways in which these chemicals are organized to form mechanisms. From the perspective of Illari and Williamson’s notion of mechanism, one may say that researchers do look for entities—yet, they do not really look at the activities these carry out, as molecular profiles provide a static picture, and not even at the organisation of these entities as part of a mechanism. In EXPOsOMICS, exposure profiles were rather used to provide evidence for claims about the *dependency* of the molecular presence of toxicants on external exposure and to find out about the *difference* that the environmental exposures experienced by individuals make to the presence of these toxicants. These features place molecular evidence very close to difference-making evidence, of which dependency and difference are core elements, and quite apart from evidence of mechanisms linking exposure and disease.

More generally, the context where molecular evidence was used in EXPOsOMICS is the identification of associations between pollutants and aspects of health and disease, rather than the discovery of mechanisms. Exposure profiles were used in regression models as evidence about toxicants at the molecular level, in order to find associations with the presence of toxicants at the external level. One EXPOsOMICS researcher explained this whilst talking about modelling molecular data from metabolomics to study the effects of environmental exposures on birth weight:

Each model models the relation of birth weight to that metabolic feature, so what’s the *change* of birth weight with that metabolite, once you take into account covariates like the mum’s weight, the mum’s education and all these things. (Researcher B1; my emphasis).

The focus here is on changes and differences, rather than on how these differences are brought about or through which mechanisms.

In addition, one of the characteristic features of mechanistic evidence, as defined in the RWT framework, does not seem to be supported by molecular data. As argued by Illari (2011a), mechanistic evidence is crucial for causal inference because it helps with issues affecting difference-making evidence, including confounding. However, this feature is not supported by exposure profiles, which on the contrary are affected by issues that typically affect difference-making evidence. I take this point from the scientific literature on molecular evidence, where for instance Vineis and Perera say that one of the problems related to the use of omics as intermediate biomarkers is confounding:

One of the main challenges with intermediate biomarkers is to understand whether they belong to the causal pathway between exposure and disease, whether they are simply a side effect of exposure or disease, or whether their measurement is confounded by some other exposure. (Vineis and Perera 2007, p. 1961)

The emphasis on difference making and problems of confounding in exposome research are the reasons why I caution against interpreting molecular data as mechanistic evidence. One may still say that there is a simpler and weaker claim that can be defended, namely that molecular data in exposome research *hints* at where to look for mechanisms and this is enough to say that it counts as mechanistic evidence. In this view, evidence from exposure profiles would count as mechanistic because it hints at the presence of entities involved in underlying mechanisms. The claim, however, is problematic even in this weaker form, as it is not clear how molecular data hinting at a mechanism would count as the kind of mechanistic evidence defined by the RWT. To call mechanistic evidence something that hints at mechanisms would imply to stretch the notion of mechanistic evidence too much, up to a point where the notion itself loses most of its power. It would imply going beyond the notion of mechanistic evidence put forward by Illari (2011a), which is already quite broad. One may say that my contention about mechanistic evidence may be too stringent and suggest that, in constrained areas of research like exposome research, even just providing hints at mechanisms could count as production of mechanistic evidence or mechanistic explanations. Still, as I explain here, I struggle to see how evidence hinting at mechanisms could count as the mechanistic evidence defined by the RWT. If one wants to look at the case from the perspective of mechanistic explanations, I would rather say that the molecular data used in EXPOsOMICS plays the role of a heuristic for finding mechanisms, as no mechanism seems to be discovered directly, or that is part of how-possibly mechanistic explanations (Forber 2010; Reydon 2012). Moreover, mechanisms can often be discovered on the basis of types of evidence such as difference-making evidence and manipulative evidence, rather than mechanistic evidence (Campaner and Galavotti 2012).

4 Classifying Evidence in Exposome Research

Whilst critically analysing interpretations of molecular data as mechanistic evidence, I have put significant emphasis on the context of the *use* of molecular data as evidence in EXPOsOMICS. Building on this approach, in this section I expand my analysis and make a more general point on evidence classification. I intend to show that, in the context of data practices in EXPOsOMICS, what makes a difference for the classification of data as evidence of a specific kind is the way in which the data is used. I argue that this is in contrast with how evidence is differentiated in the RWT framework. I therefore expand my critical remarks on the RWT by pointing to more general issues in the way medical evidence is framed in the thesis.

In the literature on the RWT, a specific type of evidence is usually linked to a specific method for the generation of evidence. This suggests that the notion of evidence employed by the RWT is close to traditional philosophical accounts of evidence and data. In these views, data counts as evidence on the basis of intrinsic properties, that are fixed and inherently local and stand in a representational relation with aspects of reality, independently of the context in which the data is used.⁸ These views have recently been challenged by Sabina Leonelli, who has argued that data should rather be considered a relational notion with a non-representational character (Leonelli 2015, 2016, pp. 69–92). When asked about notions and classifications of evidence, EXPOsOMICS researchers discussed these issues in the context of considerations about the data used in the project, and more specifically the ways in which they use, analyse, aggregate and disseminate their data. Thus, I now introduce the various types and sources of data researchers work with in exposome projects like EXPOsOMICS.

In the previous section, whilst analysing the use of molecular data, I have mentioned that the starting point for those analyses is data retrieved from cohort studies. In EXPOsOMICS (as in most epidemiology), cohort studies were one of the crucial and initial sources of data. Cohort studies follow a population of interest—the cohort of participants—for a long period of time, during which extraction of physical samples, questionnaires and one-to-one interviews are performed to collect data on a number of variables, including clinical states, features of participants' lifestyle, house environment, etc. The types of data used in EXPOsOMICS and retrieved from cohort studies varied significantly, including: quantitative data, such as measures of chemicals and compounds, as well as qualitative data, such as information provided in interviews; data about specific individuals as well as data about the surrounding environment; physical samples stored in low temperature as well as spreadsheets saved in databases.

As we have seen, some of the data retrieved in cohort studies is analysed to study the molecular composition of the internal chemical environment and produce exposure profiles. Together with what is directly acquired through cohort studies, molecular profiles constitute what EXPOsOMICS researchers defined as evidence

⁸ A paper by Bogen and Woodward is often cited as a prime example of this view (Bogen and Woodward 1988). For more recent and nuanced takes on this approach, see e.g. Rheinberger (2011).

on population health, or ‘health data’, which is thus composed of various kinds of data, collected at different levels of abstraction, at different times, and with different kinds of expertise involved. EXPOsOMICS researchers distinguished health data from what they called ‘environment data’, i.e., data on the conditions participants were exposed to during the study period. As with health data, the starting point is cohort studies, where participants are asked about the environment where they live and/or work, their lifestyle, diet, etc. One of the basic questions that participants to cohort studies are asked to provide is the postcode of the area where they live. In EXPOsOMICS, this was used to design geographical information systems to produce new data about the quantity of toxicants in the cohort’s environment (Vineis et al. 2017a, p. 143; Fiorito et al. 2017, pp. 237–238). Starting from postcodes, and on the basis of maps of characteristics of the population (such as density of the population, traffic intensity, whether the area is an industrial area, etc.), a geo-spatial regression model is designed to take into account these variables and differences in populations and assign exposure estimates to every participant in the study (Gulliver et al. 2018).

I argue that the context of data practices in EXPOsOMICS shows that the classification of a dataset as a specific kind of evidence crucially depends on the ways in which the dataset is *used* as evidence. For example, data from cohort studies can be evidence for a variety of phenomena, analyses and claims, depending on the way it is used. In EXPOsOMICS, it was indeed used as evidence for the study of different phenomena, including: as molecular evidence for exposure; as evidence for health states of a population; and, as part of geo-spatial models, to produce further evidence for the environment the cohort could have been exposed to. This is at odds with the idea that evidence should be classified on the basis of the method with which it is produced. Collection methods do play a crucial role for the evidential value of a dataset, but they do not determine it on their own. Another important feature of data practices in EXPOsOMICS is that the evidential content of a specific type of data is determined by the mobilisation and use of other types and sources of data. This is particularly noticeable in the case of environment data, where the evidential value of the initial dataset (postcodes) is determined through the mobilisation of various other sources of data.

I argue that these features point to a general approach to evidence, where considerations about evidence are indeed determined by the data collected and retrieved in specific projects like EXPOsOMICS. Yet, what plays a crucial role is not just the intrinsic properties of a dataset (such as its origin, type, content, format, etc.), but rather the ways in which the dataset is used, fits into research questions and is mobilised and put in relation with other datasets.

This approach to evidence is also reflected in researchers’ reactions when asked about the classification of data as mechanistic evidence. One researcher said that he would find it necessary to perform various further experiments and have other types of data, for molecular data to be used as mechanistic evidence. His way of framing mechanistic evidence closely linked classification to the context of use of data. For example, he pointed to research in the metabolomics literature, published in *Nature Medicine* (Wang et al. 2011). In this work, Wang and colleagues first identified molecules that were correlated with cardiovascular disease through omics analyses of metabolites.

However, the data produced in this way was not used as evidence of mechanisms. In order to test the presence of mechanisms underlying associations and thus whether the metabolites were rightfully predictive of the disease, researchers carried out further tests, including animal studies. It was only as a result of the mobilisation of these other sources of data that researchers claimed they had managed to link molecules both to an external exposure (a dietary source) and disease, thus granting the use of their molecular data as mechanistic evidence. In the words of one EXPOsOMICS researcher:

Through all these different experiments—and you can tell, a *Nature* paper is only two sides, but it brings together lots and lots of work –, using these, they were able to show the dietary source, phosphatidylcholine (which is found in meat, eggs, etc.) and they were able to link it all the way though to these metabolites and also show these metabolites directly caused atherosclerotic, as the blood vessels are directly linked to cardiovascular disease. (Researcher B1)

What I am arguing here is not that a definition of mechanistic evidence closer to Wang and colleagues' work is superior to the RWT. Rather, I am pointing out that views of what counts as a specific kind of evidence can vary significantly and depend on the context of use of data as evidence. Therefore, what counts as a type of evidence may shift considerably and clear-cut distinctions between pieces of evidence, as if these were inherently different entities, seem too strict. The kind of evidence a dataset may provide is not determined *ex ante*, and it rather depends on the heterogeneous ways in which data is used and disseminated across different sites.

The problem of the RWT approach is that it does not seem flexible enough to accommodate this attitude to evidence classification, as the way in which evidence is classified according to the RWT is based on a view of evidence as something that is produced in a form which makes it of a specific and stable kind, more precisely either difference-making or mechanistic evidence. This approach seems to overlook the context of data practices, where considerations about what counts as evidence actually take place in practice; and more generally, questions on the relation between data and evidence do not seem to be considered in the RWT framework. In this context, the work of Illari (2011a) on specifying the notion of mechanistic evidence is particularly significant, as it moves away from a classification based on evidence-gathering methods towards a distinction based on what we use evidence for. This is close to my argument, but Illari also argues that evidence should be defined on the basis of its 'object'. For instance, a piece of evidence is mechanistic evidence insofar as it has the existence of a mechanism as its object. As such, Illari's view seems similarly based on a view of evidence as something with a fixed and representational content, which—I have argued—does not hold in the context of the data practices of exposome research.

5 Exposome research and the core of the Russo–Williamson Thesis

The case of exposome research presents a number of challenges to the conceptual framework of the RWT. On the basis of the critical remarks of the paper, what should we make of the evidential pluralism proposed by the RWT? As we have seen,

discussions on causal inference in the health sciences are the main context of the RWT. The RWT is a claim about what we need to know to make causal claims. The arguably most important and controversial element of the thesis—its core idea—is that production is a necessary aspect for inferring causality and that we need to know about production when making causal inferences. This productive component is framed in terms of mechanistic evidence.

The core of the RWT seems to grasp an idea that is in general terms shared in the health sciences, including exposome research. EXPOsOMICS researchers often commented on their work saying that it is mostly focused on getting information about dependences between variables, but knowing what lies *in between* dependence relations makes a crucial difference and is the ultimate way of testing causal claims:

And the ultimate way of validating your result and making sure that it's something that is general is to understand the biology behind it. We're not at this stage, but the ultimate way would be to do some mechanistic wet lab experiment, to understand the mechanisms. And, once you understand the mechanisms, then there's no reason why there should be differences between one individual and another. So that's your ultimate way of generalising. (Researcher D1)

How can we reconcile this intuition, at the core of the RWT, with the critical remarks made in the paper? The problem is how to interpret the claim on the productive aspect of causality. The RWT frames it in terms of a need for mechanistic evidence. But does that mean that mechanistic evidence should be *produced* when causal claims are made, or is the *use* of mechanistic evidence enough? The RWT is not clear on this issue.

If we take it to mean that mechanistic evidence has to be produced, the thesis seems too strong and hard to maintain in many lines of research in the life sciences. For instance, in exposome research and other areas of epidemiology, it seems often difficult to identify mechanisms and mechanisms are not always the focus of research. At the same time, we will probably want to say that epidemiologists can make causal claims, even if no mechanistic evidence is produced. Many public health policies are based on epidemiological studies where no mechanistic evidence seems to be produced. For example, environmental epidemiologists are often concerned with determining air quality standards, that is measurements of how much air pollution we can and should tolerate.⁹ Public health policies on these standards are developed on the basis of epidemiological studies about, say, the relation between ambient air pollution and the incidence of cardiovascular disease, often with no primary focus on mechanisms. We will probably want to say that these policies are based on causal claims of some sort, as knowledge of causal relations is usually considered to provide among the best justifications for effective policy interventions (see e.g. Cartwright 2011). Hence, requiring mechanistic evidence to be produced, each time causal claims are made in the health sciences, seems problematic. This

⁹ Projects like EXPOsOMICS, which focuses on air pollution and water contamination, are precisely funded with the aim of gathering evidence for updating and revising these standards.

suggests that the core of the RWT should be rather interpreted as the need for use, rather than production, of mechanistic evidence.

At the same time, it is important to stress that, in exposome research and epidemiology more generally, new studies are designed, data is analysed and research is shaped (also) on the basis of mechanistic knowledge. When environmental epidemiologists assess, for instance, associations between air pollution and cardiovascular disease, background knowledge on potential mechanistic pathways linking the two phenomena plays an important role. However, in cases such as exposome research, evidence of mechanisms often comes from research in other areas of the health sciences. This suggests an interdisciplinary interpretation of the thesis. So far, most of the examples used in the literature on the RWT come from single disciplines and/or single projects in the life sciences. Yet the thesis can be interpreted to apply on a higher level: some disciplines may provide evidence for the probabilistic angle of causal inference, whilst others may provide the productive one. In this view, one could say that a project like EXPOsOMICS provides some of the evidence that is necessary to put together to solve the puzzle of disease causation. Other kinds of evidence, possibly including evidence of productive causation, may need to be provided by other disciplines and lines of research. This interpretation would take into account my characterisation of evidence classification in the previous section, and especially the need for mobilising various lines of evidence for data to be used as evidence. As such, the thesis would allow to interpret epidemiology as contributing in part to explanations of disease that, in the long run, may be mechanistic, without running the risk of ‘over-interpreting’ all disciplines or lines of research in productive or mechanistic terms. In other words, in this view one may be able to say that exposome research provides hints at mechanisms without having to say it produces mechanistic evidence.¹⁰

This interpretation of the RWT on the basis of my critical remarks might be a promising project. Still, a few problems remain. One problem concerns the normativity of the thesis. Whilst this interpretation of the thesis would pose weaker requirements on single disciplines or lines of research, the claim on mechanistic evidence may still be too strong. Namely, in some areas of the health sciences, researchers might be tempted to make a causal claim even if they do not have mechanistic evidence and before carrying out research aimed at discovering underlying mechanisms. For example, in the aforementioned case of the epidemiological discovery of the relation between smoking and lung cancer, the mechanism was discovered long after the initial causal claim was made and the initial smoking bans were enforced (Gillies 2011). A second and related problem is the notion of mechanistic evidence, which may not be the best way of framing the need for the productive component of causality. In disciplines like epidemiology, mechanisms are indeed used to shape research. However, more than evidence of mechanisms per se, it is rather knowledge of mechanisms that is used to guide research.

¹⁰ Personally, I find it more charitable to say that epidemiological research on the exposome is more than an instantiation of research intended to describe a mechanism, because it can potentially inform mechanistic research but produces evidence that is also important on its own.

6 Conclusions

The central object of research of this paper has been data practices and ways of classifying and framing evidence in epidemiology, with particular reference to research on the exposome. The principal issue I have dealt with is the extent to which the RWT can be applied to this line of research.

I have cautioned against the claim that the use of molecular evidence can be considered as providing mechanistic evidence as defined by the RWT. I have shown that, in the EXPOsOMICS project, molecular evidence has been used for claims about the difference that exposures make, in contrast with the notion of mechanistic evidence proposed by the RWT. Then, I have focused on another aspect of the RWT, i.e., the way evidence is classified and distinguished between different kinds. On the basis of a reconstruction of data practices in EXPOsOMICS, I have shown that, in contrast with the RWT, there are substantially different views and approaches to what counts as a type of evidence. Finally, I have looked back at the core of the RWT from the perspective of my case study and critical remarks, and I have proposed an interdisciplinary and use-based interpretation of the thesis.

This interpretation takes into account the critical points I have raised in the paper. Additionally, it tries to push the thesis forward, towards a direction where it is capable of accounting for the use of data as evidence in current research and can therefore make fitting suggestions on the consideration and appreciation of a plurality of medical evidence.¹¹ I have pointed to some aspects of the core of the thesis, and the interpretation that I have put forward, that need further review and consideration. It also remains to be seen to which extent these considerations are applicable in other areas of epidemiological and, speaking more broadly, health science research. Another related issue concerns how the integration of evidence from different sources, produced through different methods and mobilised through different pathways, is and should be carried out—a topic that has recently been discussed in the literature, but has yet to be explored in the context of epidemiology.¹²

Acknowledgements I am extremely grateful to EXPOsOMICS researchers for sharing their time and facilities and discussing their fascinating work with me. This research was funded by the *Deutsche Forschungsgemeinschaft* (DFG, German Research Foundation)—Project 254954344/GRK2073, as part of the graduate research group “Integrating Ethics and Epistemology of Scientific Research” (<https://grk2073.org>). I am indebted to many members of this group for insightful discussions and feedback on previous drafts—and particularly to Enno Fischer, David Hopf, Daria Jadreškić, Saana Jukola, Tobias Schönwitz and Martin Wasmer. I thank Egenis, the Centre for the Study of the Life Sciences, at the University of Exeter, and particularly the Data Studies group (<https://datastudies.eu/>) and the Biological Interest Group (currently known as Egenis Research Exchange: <https://exeterbig.wordpress.com>) for hosting me while an early draft of this paper was developed. Warm thanks to Raffaella Campaner, Sabina Leonelli, Simon Lohse and Thomas Reydon for the support and for commenting previous drafts. Thanks

¹¹ In this sense, see also the work on evaluating evidence of mechanisms by the EBM+consortium (Parkkinen et al. 2018).

¹² See the special issues of *Synthese* on “Evidence Amalgamation in the Sciences” (e.g. Bertolaso and Sterpetti 2017) and *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences* on “Integration in Biology: Philosophical Perspectives on the Dynamics of Interdisciplinarity” (Brigandt 2013).

also to the anonymous reviewers and the editor of the journal, Staffan Müller-Wille, for the close reading and insightful and constructive comments. I am grateful for feedback from participants to the SILFS 2017 conference at the University of Bologna (June 2017) and the Big Data in the Social Sciences conference at the University of Kent (June 2017), where I presented earlier versions of this paper.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

- Bertolaso, M., & Sterpetti, F. (2017). Evidence amalgamation, plausibility, and cancer research. *Synthese*. <https://doi.org/10.1007/s11229-017-1591-9>.
- Bogen, J., & Woodward, J. (1988). Saving the phenomena. *The Philosophical Review*, 97(3), 303–352.
- Brigandt, I. (2013). Integration in biology: Philosophical perspectives on the dynamics of interdisciplinarity. *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences*, 44(4), 461–465.
- Broadbent, A. (2011). Inferring causation in epidemiology: Mechanisms, black boxes, and contrasts. In P. Illari, F. Russo, & J. Williamson (Eds.), *Causality in the sciences* (pp. 45–69). Oxford: Oxford University Press.
- Broadbent, A. (2013). *Philosophy of epidemiology*. London: Palgrave Macmillan.
- Broadbent, A. (2015). Causation and prediction in epidemiology: A guide to the “Methodological Revolution”. *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences*, 54, 72–80.
- Campaner, R. (2011). Causality and explanation: Issues from epidemiology. In D. Dieks, W. J. Gonzalez, S. Hartmann, T. Uebel, & M. Weber (Eds.), *Explanation, prediction, and confirmation* (pp. 125–135). Dordrecht: Springer.
- Campaner, R., & Galavotti, M. C. (2012). Evidence and the assessment of causal relations in the health sciences. *International Studies in the Philosophy of Science*, 26(1), 27–45.
- Canali, S. (under review). 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*.
- Cartwright, N. (2011). A philosopher’s view of the long road from RCTs to effectiveness. *The Lancet*, 377(9775), 1400–1401.
- Clarke, B., Gillies, D., Illari, P., Russo, F., & Williamson, J. (2013). The evidence that evidence-based medicine omits. *Preventive Medicine*, 57(6), 745–747.
- De Vreese, L. (2009). Epidemiology and causation. *Medicine, Health Care and Philosophy*, 12(3), 345–353.
- Fiorito, G., Vlaanderen, J., Polidoro, S., Gulliver, J., Galassi, C., Ranzi, A., et al. (2017). Oxidative stress and inflammation mediate the effect of air pollution on cardio- and cerebrovascular disease: A prospective study in nonsmokers. *Environmental and Molecular Mutagenesis*, 59(3), 234–246.
- Fleming, L. E., Tempini, N., Gordon-Brown, H., Nichols, G., Sarra, C., Vineis, P., et al. (2017). *Big Data in environment and human health: Challenges and opportunities: Oxford encyclopaedia for environment and human health*. Oxford: Oxford University Press.
- Forber, P. (2010). Confirmation and explaining how possible. *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences*, 41(1), 32–40.
- Gillies, D. (2011). The Russo–Williamson Thesis and the question of whether smoking causes heart disease. In P. Illari, F. Russo, & J. Williamson (Eds.), *Causality in the sciences* (pp. 110–125). Oxford: Oxford University Press.

- Gulliver, J., Morley, D., Dunster, C., McCrea, A., van Nunen, E., Tsai, M.-Y., et al. (2018). Land use regression models for the oxidative potential of fine particles (PM 2.5) in five European areas. *Environmental Research*, 160, 247–255.
- Howick, J. (2011). Exposing the vanities—and a qualified defense—of mechanistic reasoning in health care decision making. *Philosophy of Science*, 78(5), 926–940.
- Illari, P. M. (2011a). Mechanistic evidence: disambiguating the Russo–Williamson Thesis. *International Studies in the Philosophy of Science*, 25(2), 139–157.
- Illari, P. M. (2011b). Why theories of causality need production: An information transmission account. *Philosophy and Technology*, 24(2), 95–114.
- Illari, P., & Russo, F. (2016). Information channels and biomarkers of disease. *Topoi*, 35(1), 175–190.
- Illari, P. M., & Williamson, J. (2012). What is a mechanism? Thinking about mechanisms *across* the sciences. *European Journal for Philosophy of Science*, 2(1), 119–135.
- Kelly, T. (2016). Evidence. In: E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy*, winter 2016 edn, Metaphysics Research Lab, Stanford University. <https://plato.stanford.edu/entries/evidence/>. Accessed July 2018.
- Leonelli, S. (2015). What counts as scientific data? *A Relational Framework*. *Philosophy of Science*, 82(5), 810–821.
- Leonelli, S. (2016). *Data-centric biology: A philosophical study*. Chicago: University of Chicago Press.
- Maitre, L., de Bont, J., Casas, M., Robinson, O., Aasvang, G. M., Agier, L., et al. (2018). Human Early Life Exposome (HELIX) study: A European population-based exposome cohort. *British Medical Journal Open*, 8(9), e021311.
- Parkkinen, V. P., Wallmann, C., Wilde, M., Clarke, B., Illari, P., Kelly, M.P., et al. (2018). *Evaluating evidence of mechanisms in medicine. Principles and procedures*. Springer Briefs in Philosophy.
- Rappaport, S. M., & Smith, M. T. (2010). Environment and disease risks. *Science*, 330(6003), 460–461.
- Reydon, T. A. C. (2012). How-possibly explanations as genuine explanations and helpful heuristics: A comment on Forber. *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences*, 43, 302–310.
- Rheinberger, H. J. (2011). Infra-experimentality: From traces to data, from data to patterning facts. *History of Science*, 49(3), 337–348.
- Russo, F. (2009). Variational causal claims in epidemiology. *Perspectives in Biology and Medicine*, 52(4), 540–554.
- Russo, F., & Vineis, P. (2016). Opportunities and challenges of molecular epidemiology. In G. Boniolo & M. J. Nathan (Eds.), *Philosophy of molecular medicine: Foundational issues in research and practice* (pp. 252–281). London: Routledge.
- Russo, F., & Williamson, J. (2007). Interpreting causality in the health sciences. *International Studies in the Philosophy of Science*, 21(2), 157–170.
- Russo, F., & Williamson, J. (2012). EnviroGenomarkers: The interplay between mechanisms and difference making in establishing causal claims. *Medicine Studies*, 3(4), 249–262.
- Stegenga, J. (2014). Down with the hierarchies. *Topoi*, 33(2), 313–322.
- Strimbu, K., & Tavel, J. A. (2010). What are biomarkers? *Current Opinion in HIV and AIDS*, 5(6), 463–466.
- Vandenbroucke, J. P., Broadbent, A., & Pearce, N. (2016). Causality and causal inference in epidemiology: The need for a pluralistic approach. *International Journal of Epidemiology*, 45(6), 1776–1786.
- Vineis, P., Chadeau-Hyam, M., Gmuender, H., Gulliver, J., Herceg, Z., Kleinjans, J., et al. (2017a). The exposome in practice: Design of the EXPOmICS project. *International Journal of Hygiene and Environmental Health*, 220(2), 142–151.
- Vineis, P., Illari, P., & Russo, F. (2017b). Causality in cancer research: A journey through models in molecular epidemiology and their philosophical interpretation. *Emerging Themes in Epidemiology*, 14(1), 7.
- Vineis, P., & Perera, F. (2007). Molecular epidemiology and biomarkers in etiologic cancer research: The new in light of the old. *Cancer Epidemiology, Biomarkers and Prevention*, 16(10), 1954–1965.
- Wang, T. J., Larson, M. G., Vasan, R. S., Cheng, S., Rhee, E. P., McCabe, E., et al. (2011). Metabolite profiles and the risk of developing diabetes. *Nature Medicine*, 17(4), 448–453.
- Wild, C. P. (2005). Complementing the genome with an “exposome”: The outstanding challenge of environmental exposure measurement in molecular epidemiology. *Cancer Epidemiology, Biomarkers and Prevention*, 14(8), 1847–1850.
- Worrall, J. (2002). What evidence in evidence-based medicine? *Philosophy of Science*, 69(S3), S316–S330.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.