Global Philosophy (2024) 34:25
https://doi.org/10.1007/s10516-024-09723-4

ORIGINAL PAPER

Evidence and Computer Simulations in Public Health

Valeriano Iranzo[1] · **Saúl Pérez-González**[1]

Received: 18 January 2024 / Accepted: 12 September 2024

Abstract
In the last decades, the evidence-based approach has become dominant in the biomedical sciences. Its influence has extended from clinical practice to other areas such as drug regulation and public health. Nonetheless, given the limitations of the evidence-based framework, a more pluralist approach towards evidence has been demanded. In this sense, evidential pluralism holds that evidence of mechanisms should complement statistical evidence. This paper aims to explore and discuss the prospects of evidential pluralism in public health. First, we will analyse the evidential standards of the International Agency for Research on Cancer. Those standards are argued to be in line with the pluralist approach. Second, we will consider the main challenges for gathering evidence of mechanisms in public health. For that purpose, diverse types of evidence of mechanisms will be identified. And third, we will assess the role of agent-based models as mechanistic studies. Based on a case study, we will discuss whether agent-based models can provide evidence of mechanisms and mitigate the abovementioned challenges.

Keywords Public health · Evidence · Evidence-based medicine · IARC · Mechanism · Agent-based model

1 Introduction

The aim of this paper is to explore and discuss the prospects for a pluralist approach concerning evidence in public health. Evidential pluralism has emerged as a reaction against the dominant approach for contemporary biomedical practice, that is, evidence-based medicine (EBM). EBM was initially applied for management and decisions in the clinical encounter between the patient and the physician, but soon it

✉ Valeriano Iranzo
iranzov@uv.es

[1] Department of Philosophy, University of Valencia, Valencia, Spain

Published online: 21 November 2024

was extended to related fields as nursing and physiotherapy, and later to public health (Jenicek 1997; Brownson et al. 2009, 2017).

EBM holds that biomedical decisions must be based on the best available evidence. In order to avoid ineffective or harmful interventions, favouring an evidence-based guidance seems, in principle, a sensible advice. However, this general claim is vacuous without further criteria to discern what should be labelled as the best available evidence and how to combine different sorts of evidence to inform decisions. Regarding this, EBM definitely favours statistical evidence obtained through controlled experiments over any other sort of evidence. But in the last decades alternative approaches to evidence have emerged. In this sense, evidential pluralism criticises the standard view and claims that evidence of mechanisms, should be judged as similarly reliable and useful. In this paper, we will explore the potential application of evidential pluralism in public health. Moreover, we will draw attention to the fact that computer simulations should be counted among the resources that advocates of evidential pluralism have at their disposal.

The paper is structured as follows. In Sect. 2, we present the main features of evidence-based medicine and introduce evidential pluralism as an alternative to the standard evidence-based approach. According to pluralists, statistical evidence should be complemented by evidence of mechanisms. In the following sections we discuss the prospects of evidential pluralism in public health. In Sect. 3, the evidential standards of the International Agency for Research on Cancer (IARC) are scrutinized. IARC's methodology is considered to be in line with the core principles of evidential pluralism. Subsequently, in Sect. 4, we identify and characterise diverse sorts of evidence of mechanisms. Based on that characterisation, we identify the main challenges for obtaining evidence of mechanisms in public health. In Sect. 5, we address computer simulations, particularly agent-based modelling. We discuss whether agent-based models can provide evidence of mechanisms and mitigate the challenges faced by evidential pluralism in public health. Finally, Sect. 6 is devoted to conclusions.

2 Public Health and Evidential Pluralism

In the 1990s, as a result of initiatives such as the founding of the Cochrane Collaboration, so called 'evidence-based medicine' emerged. The general motto for it is that medical practice should be guided by the best available evidence. Since this is a comparative claim, EBM is committed to ranking those procedures used in biomedical sciences for obtaining evidence. Particularly, and according to its original definition, "Evidence-based medicine de-emphasizes intuition, unsystematic clinical experience, and pathophysiological rationale as sufficient grounds for clinical decision making and stresses the examination of evidence from clinical research." (Guyatt et al. 1992).[1]

[1] For a discussion of this definition, see (Howick 2011; Sect. 2.3).

The EBM approach is, then, committed to an evidential hierarchy.[2] It is usually depicted with a pyramid shape (see Fig. 1). Different methods and strategies are ranked, supposedly, in terms of their ability to avoid biases and provide reliable inferences from the evidence obtained through them. At the top of the hierarchy are meta-analyses and randomized controlled trials (RCTs). Just below them are observational studies at the population level, such as cohort studies, and then, case-control studies (i.e., prospective and retrospective studies, respectively). Finally, basic research (e.g., in-vitro studies, animal trials, and mechanism-based reasoning) and medical expertise are confined in the bottom levels.

The EBM hierarchy offers a guiding framework for discerning causal relations in a systematic way. This sort of research "does not rely directly on *how* the intervention might produce an outcome, and instead involves directly observing the putative outcome relative to the putative outcome produced by a control treatment." (Howick 2011, p. 16). The general procedure is, then, discerning causal relations from spurious correlations in line with standard experimental designs. EBM favours statistical evidence above any other sorts of evidence and, particularly, that statistical evidence obtained by combining randomization and comparison between the experimental and the control group.

However, insofar as medical decisions are based, partly at least, on our causal knowledge, EBM is concerned not only about inference but also about decision. It can be used, and indeed it is routinely used, to inform policy-making. For instance, the EBM-hierarchy provides rules to set specific requirements for commercialization of drugs. The alleged effectiveness of the chemical compound should be supported by evidence obtained at the top levels of the hierarchy.[3]

Public health is, in fact, a further field where demands for grounding decisions on evidence provided by experimental research have been levered. As it is claimed by Mercury and Upshur (2022, p. 147), even though "the concern that drove the demand for EBM also drove the demand for EBPH [Evidence-based Public Health]"

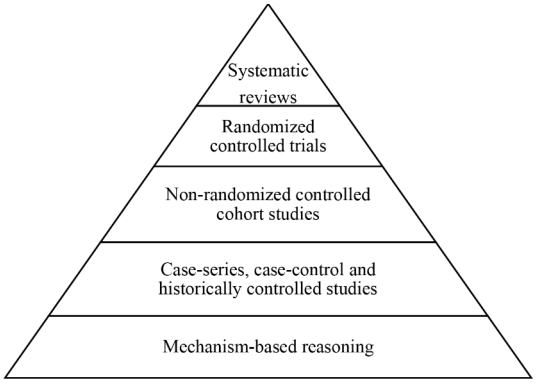

Fig. 1 Levels of evidence according to "The Oxford 2011 Levels of Evidence" (OCEBM Levels of Evidence Working Group 2011)

[2] Some well-known characterisations are the GRADE rating for the quality of evidence (Guyatt et al. 2011) and the classification of levels of evidence from the Oxford Centre for Evidence-Based Medicine (OCEBM Levels of Evidence Working Group 2011).

[3] Both the FDA (Food and Drug Administration) and the EMA (European Medicine Agency) appeal to clinical trials and take for granted the evidential hierarchy defended by EBM.

they are distinct in both scope and practice. Public health is not primarily interested in care of individuals. Rather it endorses a social perspective since typical interventions (disease prevention, hygiene…) in public health are evaluated by taking into account their global impact in the community. The crucial question here is anticipating their effects on a target population. Consequently, decisions in public health usually involve governmental agencies. Besides, they are constrained by a wide array of factors, not only medical ones, but also economic, ethical, political… Think, for instance, of the mandatory use of masks to restrain the spread of COVID-19.[4]

Nonetheless, the evidence-based approach and its application in public health have also raised concerns. The idea that RCTs and studies based on them are the most reliable sources of evidence has been contested by claiming that, after all, those strategies could not be so reliable concerning detection of causal links (Worrall 2007; Cartwright and Munro 2010; Stegenga 2011). When designing the experiment an exhaustive list of potential confounding factors may be difficult to get in advance. This can make it difficult to determine the size of the effect attributed to the alleged cause (Kincaid 2012). Furthermore, meta-analyses do not always give an unequivocal verdict because different (extensive and careful) RCTs may afford different conclusions. In fact, although a causal relationship is confidently established in the experimental population, extrapolation to the target population, that is, to the real population outside the lab, may be risky (Victora et al. 2004). The expected effects found in the experimental population may be severely distorted by uncontrolled operating factors in the target population. In other words, even though the external validity of a clinical trial is what really matters in medical practice, sometimes it can hardly be anticipated from its internal validity.

Moreover, in the context of evidence-based policy, RCTs are usually unavailable or inadequate for establishing causal claims (Pérez-González 2024).[5] First, many questions relevant for policy-makers cannot be addressed by means of RCTs. There are interventions that cannot be tested experimentally, e.g., national-level food withdraw (Khosrowi and Reiss 2019). Furthermore, RCTs are usually focused on the average effect and provide little information about the distribution of the effect (Gamoran 2018). Second, when units of interventions are communities (e.g., schools), it is very challenging to design and conduct RCTs. Randomisation is unlikely to result in balanced intervention and control groups. Finally, it may be unethical to test certain interventions by means of RCTs (Barrett and Carter 2010). For instance, conducting RCTs about the negative effects of alcohol abuse would be considered unethical.

In response to those limitations of the evidence-based approach, several authors have called for a broader approach towards evidence (Clarke et al. 2014; Gillies 2018; Anjum et al. 2020). Regarding public health, it has been argued that evidential pluralism should be adopted (Parkkinen et al. 2018; Russo 2022). *Evidential pluralism* claims that both statistical evidence and evidence of mechanisms are crucial for addressing causality in the health sciences (Russo and Williamson 2007; Clarke et

[4] On the general difficulties for making decisions in public health contexts see (Norris et al. 2019).

[5] In the decision-making process, there are also non-empirical uncertainties (Ongaro and Andreoletti 2022). For instance, decision-makers are often unsure about the desirability of alternative interventions. Those uncertainties are not entirely reducible to the absence of information or evidence.

al. 2014; Williamson 2019a). The theoretical basis of evidential pluralism is summarised in the Russo-Williamson Thesis (Russo and Williamson 2007). According to this epistemological thesis, both evidence of correlations and evidence of mechanisms are normally needed to establish a causal claim in medicine. Current evidential pluralism, nonetheless, maintains that both kinds of evidence are also required for extrapolating causal claims to target populations or individuals (see, for instance, Williamson 2019a).

The main reason in favour of evidential pluralism is that evidence of correlations and evidence of mechanisms complement each other (Clarke et al. 2014; Williamson 2019a). Evidential pluralists argue, then, that evidence of mechanisms can address and mitigate the main limitations of evidence of correlations and vice versa.

Evidential pluralism can be illustrated by the well-known debate about the causal connection between smoking and lung cancer. Since the 1950s, strong correlational evidence linking smoking and lung cancer was available (Proctor 2012). This evidence was provided by diverse methods such as population studies and animal experimentation. Nonetheless, the causal link between smoking and lung cancer was still contested. It was argued that the established correlation might be non-causal and that alternative explanations should be sought. For example, Fisher (1957) famously argued that there may be a gene that both caused some people to smoke cigarettes and made them more likely to develop lung cancer. In the 1990s, the mechanisms through which tobacco causes lung cancer was discovered (Proctor 2001). This complementary evidence of mechanisms was crucial to dismiss alternative interpretations of the correlation and establish a strong causal relationship between smoking and lung cancer.

In sum, given the limitations of RCTs and other methods privileged by hierarchies of evidence, evidential pluralism has been proposed as an alternative to the standard evidence-based approach. Evidential pluralism holds, rather, that statistical evidence and evidence of mechanisms somehow complement each other. Therefore both are crucial for addressing causality in the health sciences. In the following sections we will discuss the application of evidential pluralism in public health.

3 Evidential Pluralism Put into Practice: The IARC Case

Advocates of evidential pluralism have proposed it as a useful framework to assess evidential bases and establish causal claims in public health (Parkkinen et al. 2018; Russo 2022). Nonetheless, they acknowledge that "further work should tailor evidential pluralism to the specific needs of PH [Public Health]" (Russo 2022, p. 106) since there are relevant divergences between public health and medicine. When considering the application of evidential pluralism to public health, evidential pluralists often refer to the International Agency for Research on Cancer (Leuridan and Weber 2011; Parkkinen et al. 2018; Williamson 2019b). They argue that IARC methodology embraces the general principles of evidential pluralism. In this sense, it would provide a valuable guideline for implementing those principles in public health.

IARC has been engaged for decades in the evaluation of agents suspected of causing cancer. Panels of experts appointed by IARC make a systematic review of extant

studies, that is, an assessment of the available evidence, and are expected to reach a verdict on the carcinogenic nature of the substance in question. Results are published as open access monographs since 1972.

According to IARC, evaluation of suspicious agents should follow a standardised procedure for assessing the available evidence. Initially, the reports surveyed epidemiological studies about humans and experimental research with laboratory animals. Since results of randomized trials could rarely be obtained, evidence on humans was basically gathered from cohort studies, case-control studies, correlation studies, and case-studies. And they were ranked according to the EBM hierarchy (see Fig. 1):

> The uncertainties surrounding interpretation of case reports and correlation studies make them inadequate, except in rare instances, to form the sole basis for inferring a causal relationship. When taken together with case-control and cohort studies, however, relevant case reports or correlation studies may add materially to the judgement that a causal relationship is present. (IARC 1990, pp. 21–22)

Concerning evidence about mechanisms, IARC monographies maintained a rather strict positioning until 1991, as can be read on the "Preamble" included in volume 50:

> In the present state of knowledge, the aim of the Monographs is to evaluate evidence of carcinogenicity at any stage in the carcinogenic process *independently of the underlying mechanism involved*. There is as yet insufficient information to implement classification according to mechanisms of action. Definitive evidence of carcinogenicity in humans can be provided only by epidemiological studies. Evidence relevant to human carcinogenicity may also be provided by experimental studies of carcinogenicity in animaIs and by other biological data, particularly those relating to humans. (IARC 1990, p. 6; our emphasis)

It is widely accepted, however, that there are important difficulties for gathering reliable statistical evidence in this context. RCTs are usually ruled out for moral reasons. Moreover, in many cases, there is a long period (even decades) between exposure to the supposedly harmful agent and the development of the disease. Cohort studies, unfortunately, do not usually extend for such a long time. Supporting a negative verdict, consequently, is complicated since "latency periods substantially shorter than about 30 years cannot provide evidence of lack of carcinogenicity" (IARC 2024, p. 25). Although these limitations cannot be considered intrinsic to the standard statistical methodology, since they depend rather on the particularities of the effect studied (cancer in humans caused by exposure to certain agents), they encourage the search for complementary sources of evidence.

The fact is that since 1992 the IARC procedure gradually tried to integrate evidence of mechanisms. Changes implemented in the "Preamble" included in the monographs reflect this turn. After volume 54, published in 1992, a short subsection on "Inferences about mechanism of action" was included in the section devoted to "Evidence for Carcinogenicity in Humans". This subsection changed into a separate section, entitled "Other Data Relevant to an Evaluation of Carcinogenicity and its

Mechanisms", appended for the first time in volume 86, published in 2006. Subsequently, this section evolved to its present form. It is worth noticing that IARC's concern now is not so much the inferential status of mechanisms, which could explain those misgivings in the eighties and nineties, as the proper epistemic value of evidence about mechanisms. Thus, we find a section entitled "Mechanistic and other relevant data" in the Preamble of volume 116 (2018), which was renamed as "Mechanistic evidence" since monograph 124 (IARC 2020). This section still stands in the latest monograph published in 2024 (volume 134) and that would be our main reference in what follows.

But, how does IARC incorporate evidence of mechanisms in verdicts of carcinogenity at present? How do experts weight the relative value of different sources of evidence?

It is worth emphasizing that, in contrast to what happens with statistical evidence, there is no principled corpus of standardized procedures for gathering, analysing, and evaluating evidence of mechanisms. In addition to this, carcinogenic agents may cause cancer through multiple mechanisms. Searching for evidence of mechanisms brings its own problems, but the way in which IARC deals with them is instructive.

To begin with, the IARC procedure points at different sorts of evidence of mechanisms. First of all, information about how the agent is incorporated in the organism (i.e., absorption, distribution, metabolism, and excretion of the agent) is scrutinized. Identifying the metabolites involved and their chemical reactivity is useful to track the metabolic pathways and products. Given that confirming carcinogenicity requires further evidence about potential or actual damages, additional evidence of mechanisms is evaluated:

(a) Evidence that the agent belongs, based on mechanistic considerations, to a class of agents for which one or more members have been classified as carcinogenic or probably carcinogenic to humans.
(b) Evidence that the agent exhibits "key characteristics" of carcinogens in different targets (exposed humans, human primary cells/tissues, or experimental systems).
(c) Evidence that the mechanism of carcinogenicity operating in experimental animals does not operate in humans.

Notice that (a) and (c) could properly be deemed to be *indirect* evidence of mechanisms since both are exclusively based on analogies. Thus, the former is grounded on analogies between agents (e.g., chemical compounds) and the latter on the analogy animal/human. But IARC's strategy also takes into account direct evidence of mechanisms for carcinogenicity through (b), that is, by detecting key characteristics. The set of key characteristics includes ten general properties regularly associated with those agents previously classified as carcinogenic by IARC—although it is not required that carcinogenic agents exhibit them all. Examples of key characteristics are "induces epigenetic alterations", "is immunosuppresive", and "alters cell proliferation, cell death or nutrient supply". The evidential relevance of detecting key characteristics varies depending on the targets where they are found. Detecting them in exposed humans should be assigned more importance than detecting them just in experimental systems.

But why this detour through key characteristics? Actually, they are distinct from the hypothesized mechanistic pathways which describe a particular sequence of biological events postulated to occur during carcinogenesis. Focusing on the latter would be a very narrow perspective. Key characteristics, in contrast, implement a more general description for the agent's carcinogenic properties. Furthermore, key characteristics are not a set of necessary/sufficient conditions for positive carcinogenicity verdicts. It is not explicitly stated that there is a minimum number of them to give a verdict of carcinogenicity. And, of course, the list could change as new human carcinogens are identified.

Once it is acknowledged that evidence about mechanisms is relevant and that there are several routes for obtaining it, two questions are still pending. Firstly, a quality standard for evidence of mechanisms is demanded. The answer with respect to statistical evidence is clear. Remember that EBM's hierarchy provides a quality graded standard for statistical evidence. Correspondingly, criteria for discerning better/worse bits of evidence of mechanisms are required to reach reliable verdicts about carcinogenicity. The second question is how to combine both sorts of evidence (i.e., statistical and mechanistic). What to do, for instance, if one of them tentatively points at a positive carcinogenicity verdict while the other is silent or non-conclusive at all? What is the relative weight that should be assigned to each one?

In IARC evaluations, the quality of the various types of evidence is separately and qualitatively evaluated. Statistical evidence, which may be on humans and also on animals, can be categorized as sufficient, limited, inadequate or contrary (that is, evidence suggesting lack of carcinogenicity). Evidence of mechanisms, on the other hand, is identified as strong, limited or inadequate. Strong mechanistic evidence does not require evidence from the three sources—(a), (b) and (c)—aforementioned. In principle, it could be obtained through any of these routes. After each sort of evidence is assessed, there is a further evaluation stage that consists in a decision table with takes all possible combinations as input, and which issues as output the verdict on the carcinogenic nature of the agent:

Now, could we affirm that the IARC procedure puts into practice evidential pluralism? A closer look at Table 1 is compulsory before answering to this question. According to the table:

- When statistical evidence in humans is considered sufficient, that alone is enough to conclude that the agent is carcinogenic, without any further qualification (line 1). In contrast, strong evidence of mechanisms is not enough by itself to infer that the agent is carcinogenic, since additional sufficient statistical evidence is required, if not obtained with humans, at least that obtained with animals (line 2).
- Evidence of mechanisms *sometimes* may be crucial to conclusively establish that an agent is carcinogenic (particularly in combination with evidence in animals; line 2).
- For non-conclusive verdicts (e.g., groups 2A, 2B and 3), strong evidence of mechanisms is less determinative than statistical evidence labelled as inadequate or less than sufficient. Even though there is strong mechanistic evidence in favour of the carcinogenic nature of the agent, the agent is not classified as such. It could even be classified as just possibly carcinogenic (see lines 8, 9 and 10).

Table 1 Integration of streams of evidence in reaching overall classifications (the evidence in *bold italics* represents the basis of the overall evaluation) (IARC 2024, p. 36)

Stream of evidence			Classification based on strength of evidence
Evidence of cancer in humans	Evidence of cancer in experimental animals	Mechanistic evidence	
Sufficient	Not necessary	Not necessary	**Carcinogenic to humans (Group 1)**
Limited or Inadequate	***Sufficient***	***Strong (b)(1) (exposed humans)***	
Limited	***Sufficient***	Strong (b)(2–3), Limited, or Inadequate	**Probably carcinogenic to humans (Group 2 A)**
Inadequate	***Sufficient***	***Strong (b)(2) (human cells or tissues)***	
Limited	Less than Sufficient	***Strong (b)(1–3)***	
Limited or Inadequate	Not necessary	***Strong (a) (mechanistic class)***	
Limited	Less than Sufficient	Limited, or Inadequate	**Possibly carcinogenic to humans (Group 2B)**
Inadequate	***Sufficient***	Strong (b)(3), Limited, or Inadequate	
Inadequate	Less than Sufficient	***Strong b(1–3)***	
Limited	***Sufficient***	***Strong (c) (does not operate in humans)***	
Inadequate	***Sufficient***	***Strong (c) (does not operate in humans)***	**Not classifiable as to its carcinogenicity to humans (Group 3)**
All other situations not listed above			

Summing up, IARC evidential policy accepts that evidence of mechanisms should be taken into account along with statistical evidence and that it is not always worse than statistical evidence. It is also acknowledged that *sometimes* evidence of mechanisms can be crucial to modulate the verdict. A literal application of the EBM hierarchy is, then, ruled out.

Still, IARC evidential policy undeniably gives some priority to statistical evidence. For instance, concerning a crucial category as Group 1—i.e., "carcinogenic"—, evidence of mechanisms is not necessary to get a conclusive result, while statistical evidence in humans is indispensable. Consequently, IARC guidelines hardly fulfil the Russo-Williamson thesis. Besides, when both sorts of evidence collide, statistical evidence is the decisive factor to favour one or other verdict. See line 9, for instance, where after finding key characteristics in exposed humans, cell tissues and experimental animals, the verdict is just "possibly carcinogenic". The fact is that the relative weight given to statistical evidence is greater than that given to evidence of mechanisms. The former sometimes is sufficient; the latter never is. And the strongest positive verdict can be obtained without any support at all in terms of evidence of mechanisms.[6]

[6] For a different view on this issue—i.e., to what extent the latest version of the IARC monographs' preamble accommodates evidence of mechanisms—, see (Williamson 2019b).

Therefore, if we take that evidential pluralism entails the Russo-Williamson thesis, the IARC procedure does not instantiate, strictly speaking, evidential pluralism. Nonetheless, it is a decision procedure where the summative effect of different sorts of evidence is carefully assessed, particularly evidence of cancer in humans, evidence of cancer in animal models, and evidence of mechanisms. To that extent it can properly be regarded a *pluralist* alternative to the EBM hierarchy, albeit a weaker version than that entailed by the Russo-Williamson thesis. Furthermore, it is a living example effectively applied in practice and closely related to public health concerns.

However, even though the IARC procedure goes in the right direction according to a pluralist viewpoint,[7] it is specifically devised for a particular context. Linking evidence of mechanisms to key characteristics may be appropriate regarding carcinogenity, but it cannot be extrapolated to other contexts of research. Pluralists who defend the import of evidence of mechanisms in public health have to deal with quite different mechanisms than those scrutinized in the IARC procedure, as we will see in the following sections. Hence, addressing evidence of mechanisms from a more general viewpoint seems desirable. That is the subject of the next section.

4 Types of Evidence of Mechanisms

In public health, as it has been noted, research has traditionally focused on obtaining evidence of correlations according to EBM guidelines. To assess the application of evidential pluralism in public health, it is important to consider in detail the particularities of the involved mechanisms and the evidence about them. As advocates of evidential pluralism have acknowledged, there are important divergences between public health and other biomedical fields (Brownson et al. 2017; Russo 2022). For that purpose, it is crucial to start from a solid account of (the diverse types of) evidence of mechanism.

Evidence of mechanisms is characterised (and distinguished from evidence of correlations) based on the object of evidence, i.e., the thing we get evidence of. Evidence of mechanisms is commonly understood as evidence of the existence or the properties of mechanisms in the domain of inquiry (Illari 2011). In this context, a broad characterisation of mechanisms is usually adopted (Parkkinen et al. 2018; Williamson 2019a). A mechanism can be a complex system, a causal process, or a combination of both. In this sense, evidential pluralism is closely related to minimal characterisations of mechanisms. Those accounts try to capture the common features of mechanisms across the sciences. For example, Glennan argues that a "mechanism for a phenomenon consists of entities (or parts) whose activities and interactions are organized so as to be responsible for the phenomenon" (Glennan 2017, p. 17; see also Illari and Williamson 2012). Some examples of mechanisms are the mechanism

[7] Leuridan and Weber suggest three ways to improve the IARC procedure by increasing the relative weight of evidence of mechanisms. We could sympathise with any of them, for instance with performing randomized in vitro experiments about mechanisms (Leuridan and Weber 2011, p. 97). But our main goal here are the specific difficulties for a fruitful use of evidence of mechanisms in public health, and not so much a detailed recipe to combine different sorts of evidence.

of protein synthesis, the mechanism of neural communication, or the mechanism of competitive exclusion.

Evidence of mechanisms, consequently, is not associated to a specific evidence-gathering method. A mechanistic study is any study that provides information about mechanisms in the domain of interest. Mechanistic studies can take the form, for instance, of process tracing, in-vitro studies, or even RCTs. Gathering evidence of mechanisms, then, does not necessarily rule out those strategies usually associated to statistical evidence (see, for instance, footnote 7).

Given that evidence of mechanisms is characterised by the object of evidence, i.e., mechanisms, depending on which aspects of mechanisms are scrutinized, we can distinguish diverse sorts of evidence of mechanisms:

- *Evidence of the existence of a mechanism*. This evidence supports the existence of some mechanism linking two phenomena or variables. Nevertheless, it provides no information about the properties or components of the linking mechanism. For instance, think about a RCT that identifies a strong and reliable correlation between C and E. This study provides evidence of the existence of a mechanism linking C and E. Given the identified correlation, it is very likely that there is a mechanism that links C and E. However, that mechanism remains a black box. The study provides no additional information about its internal workings.

- *Evidence of a property of a mechanism*. This evidence illuminates a particular property of a mechanism linking two phenomena or variables. It does not refer to a specific part of the mechanism, but to the mechanism as a whole. For example, consider a series of in-vitro experiments that study the output of a mechanism in diverse environmental conditions. If the detected output remains similar across environmental conditions, the study supports the stability of the mechanism at issue. Evidence of a property of a mechanism may also indirectly provide evidence of certain component. In those cases in which a property requires the presence of certain entities or activities, evidence of the property also supports the presence of those entities or activities.

- *Evidence of a component of a mechanism*. This evidence supports that certain entity or activity takes part in a mechanism linking two phenomena or variables. Evidence about components includes not only the role played by components, but also their relevant properties for the mechanism outcome. To illustrate this kind of evidence, think of an autopsy of a patient with the condition C and certain injury E. This study can provide evidence of the entities whose actions and interactions have resulted in E. In some cases, evidence of components may indirectly provide evidence of a property of a mechanism. When the presence of certain components gives rise to certain property (e.g., absence of stability), evidence of those components also supports the presence of that property.

- *Evidence of an interaction between mechanisms*. This evidence supports that some mechanisms in the domain of interest interact in a specific way. The interaction may involve two or more mechanisms. Consider, for instance, a process

tracing of a mechanism M_1 linking C and E. This study can make salient that, at certain step, a further disturbing mechanism M_2 partially masks mechanism M_1 and undermines its influence of C. Evidence of an interaction between mechanisms may indirectly provide evidence of properties or components of the operating mechanisms. In some cases, the interaction between two mechanisms requires that the mechanisms have certain properties or components. In those cases, evidence of the interaction also provides evidence of properties or components of the operating mechanisms.

4.1 Evidence of Mechanisms in Public Health

The mechanisms that underlie public health interventions are remarkably complex. Public health mechanisms have a high number of (very diverse) component entities (e.g., civil servants, health workers, children, institutions, etc.) and activities (e.g., supply of goods, information exchange, advertising campaigns, etc.). They are not just biological mechanisms, but often also involve (among others) social, political, and legal aspects. For example, think of the mechanisms underlying lockdown interventions during the COVID-19 pandemic. Certainly, policy makers' took into account biological aspects related to disease: ways of transmission, infectivity rate, recovery rate… But social and political aspects such as the impact on the labour market, on the traffic and supply of goods, and even on citizens' own willingness to adhere to the new regulations, were also considered.

The point is that, given the complexity of mechanisms and the unavoidable practical limitations, appealing to evidence of mechanisms in public health faces important challenges.

First, indirect evidence of the existence of mechanisms is rarely available. High-quality evidence of correlations can provide evidence of the existence of mechanisms (Williamson 2019a). Thus, when reliable evidence-gathering methods (e.g., RCTs with low risk of bias) establish a correlation between C and E, they indirectly provide evidence of the existence of a mechanism linking C and E. However, in public health, high-quality evidence of correlations about the target population is rarely available in advance. Because of practical and ethical limitations, interventions can rarely be assessed by means of RCTs. Randomised trials may not be feasible or advisable in the target population. Consequently, most evidence about the target population is obtained by means of suboptimal evidence-gathering methods.

Second, in principle experimental studies allow us to intervene in mechanisms and study their behaviour (as a whole) in diverse conditions. In this sense, in-vitro experiments and in-vivo studies are very useful to obtain evidence of properties of mechanisms in the biomedical sciences. Nonetheless, such experimental studies are rarely available in public health. It is highly unfeasible to isolate the relevant mechanisms, intervene on them, and observe them in a controlled environment. They are often integrated in larger and complex mechanisms and we can only make a limited set of interventions.

Third, extensive and reliable evidence about the main components of the relevant mechanisms can hardly be obtained. As it has been noted, public health mechanisms

include many and very diverse entities and activities. Those components may be biological, social, political, etc. Moreover, given practical limitations involved in public health, we may have important constrains to study them. For example, given the urgency and potential danger of a threat, time constraints may exist. Consequently, our evidence of components is always limited and fragmentary. In most cases, we have evidence just of some of the component entities and activities. Furthermore, given the complexity of public health mechanisms, it is difficult to assess the extent of our evidence. We can hardly know how exhaustive is our knowledge about the components (and their relevant properties).

Fourth, evidence of detailed interactions between mechanisms is usually limited. In public health, causal relations and interventions are influenced by many diverse mechanisms. When some mechanisms are identified and studied, unknown disturbing mechanisms may be present. Those disturbing mechanisms may influence the outcome and even mask the identified mechanisms (Clarke et al. 2014). Besides, given the complex network of causal relationships, even when the most relevant mechanisms are known, it may be difficult to anticipate the effects of a particular intervention or change in the environment.

To conclude, there are remarkable difficulties to collect evidence of mechanisms and they are not restricted to a particular type of evidence of mechanisms. Evidential pluralism is challenged in public health due to the complexity of the mechanisms involved and the practical constraints that shape the research. In the next section, we will explore whether computer simulations, and particularly agent-based modelling, could provide complementary evidence of mechanisms.

5 Agent-Based Models

Agent-based models (ABMs) have been vindicated as an important tool for public health (Auchincloss and Diez Roux 2008; Maglio and Mabry 2011; Tracy et al. 2018; Silverman et al. 2021). They are heavily associated with research in infectious disease epidemiology, namely, disease spread, vaccination strategies, or prevention measures. Nonetheless, ABMs have also been applied to other fields such as non-communicable disease epidemiology, health behaviour, and social epidemiology (Nianogo and Arah 2015; Tracy et al. 2018).

ABMs are computer simulations of autonomous agents that interact according to a set of predefined rules and within a specific environment (Epstein 2006). A set of individual agents is the starting point for ABMs. Agents are specific software objects and ABMs allow for a great flexibility regarding them. They normally represent people, but they can also stand for institutions, companies, or even towns. Each individual agent is characterised by a vector of attributes (de Marchi and Page 2014). Those attributes may be not only biological properties (e.g., age), but also other kinds of properties (e.g., job, studies, savings). Moreover, attributes may remain invariable during the entire simulation or be susceptible to change.

Individual agents are guided by behavioural rules in their interactions among them and with the environment. Those rules guide individual agents' behaviour on the basis of the input they receive from the environment. This input may comprise

past states, current states, or potential states of other agents, the environment, or even agents themselves. States of the environment usually refer to structural aspects (e.g., topologies). Regarding behavioural rules, ABMs also allow for great diversity. Behavioural rules are independently assigned to each individual agent. They may be the same for all agents or change across subpopulations. As with attributes, rules are also susceptible to change as the simulation proceeds.

Although ABMs are commonly understood as "fine-grained" models, they can vary a lot regarding their degree of detail (de Marchi and Page 2014; Bruch and Atwell 2015; Iranzo and Pérez-González 2021). They can range from very simple to highly complex models. Great flexibility is allowed about the number of agents, the (kinds of) attributes, the interactions, etc. ABMs are also diverse concerning their resemblance to the target population. Some models make many unrealistic idealisations and abstractions, while others are empirically calibrated according to large databases. The adequate level of complexity and realism of a particular ABM will depend on the purpose of the model.

ABMs, unlike other kinds of models, pay attention to how individual behaviour results in system-level outcomes. They aim to identify the link between certain individuals' traits and behaviours and the system-level outcomes. Given the dynamic nature of ABMs, starting from the set of initial conditions, individual agents interact guided by the behavioural rules and the input they receive from the environment. The autonomous (but interdependent) behaviour of individual agents aggregates in emergent or bottom-up system-level outcomes (Epstein 2006; Macy and Flache 2009). Precisely the most attractive aspect of this type of simulation is that it sometimes generates multi-scale patterns that, in principle, are not predictable from the data and behavioural rules implemented at the individual level. In these cases, the interest of the results obtained through simulation lies not only in their potential predictive value, but also in the fact that they make it possible to anticipate novel scenarios that might otherwise have been completely overlooked.

ABMs are particularly useful to simulate interventions and policies (Nianogo and Arah 2015; Tracy et al. 2018; Iranzo and Pérez-González 2021). They can explore modifications in any attribute of (any group of) individual agents or parameters of the environment. Multiple simulations, which diverge just in details, can be conducted and compared in terms of their results, that is, the hypothetical scenarios, arrived at. Furthermore, given their dynamic nature, ABMs can study changes in agents' behaviour resulting from the introduction of the intervention (Bruch and Atwell 2015). As a result, ABMs can be an effective tool to address complex issues in research on public health such as emergence, non-linearity, and adaptive behaviour.

In order to illustrate the evidential role of ABMs in public health, let us briefly analyse an example about the relationship between adolescent smoking and friendship networks.

The model developed by Schaefer et al. (2013) is a particular implementation of a Stochastic Actor-Based Model (SABM).[8] This model is composed of 509 agents

[8] SABM, also referred to as SIENA, is one option among a vast array of ABMs now available. SABMs are especially designed to cope with the reciprocal interaction between social networks and behaviour. For technical details see (Snijders 2017).

representing individual students. Each student is characterised by a vector of attributes including age, delinquency, alcohol consumption, or parent smoking. Some attributes are assigned to an individual (e.g., age), while others are dyadic (e.g., age similarity). Initially, agents' attributes are calibrated using data from the National Longitudinal Study of Adolescent Health. It should be noted that some attributes are invariant during the whole simulation (e.g., parent smoking), while others could vary (e.g., smoking).

This ABM simultaneously uses two different behavioural rules or functions to study the interactions between smoking and social influence. The smoking function reflects changes in smoking over time (e.g., prevalence, initiation, and cessation). The network function governs modifications in network structure of friendship relations over time (e.g., creation of new ties). Some agents' attributes are included in the smoking function (e.g., parent smoking), while others are included in the network function (e.g., age similarity). At any step, each individual agent may change a tie, change smoking behaviour, or make no changes.

There is a substantial amount of empirical research about the effects of social influence on smoking behaviour. More specifically, Schaefer and his collaborators were interested in the micro-level processes underlying smoking behaviour in adolescence. Peer-networks may be crucial here. Supposedly, the fact that your close friends are, or are not, smokers influence your smoking behaviour; and the more popular are smokers at school, the higher the probability that other students imitate their behaviour. According to this, Schaefer analysed social influence at school in relation to two different parameters: peer influence and smoker popularity. Then, the questions to be investigated through this simulation were, firstly, how both factors, measured in discrete values, affect population-level smoking outcomes and, secondly, which interventions could be recommended to decrease the proportion of smokers.

Consequently, peer influence and smoker popularity were manipulated. The authors developed scenarios with alternative values for peer influence (while smoker popularity remained constant at the observed value), scenarios with alternative values for smoker popularity (while peer influence remained constant at the observed value), and scenarios with alternative values for both peer influence and smoker popularity. A thousand simulation runs were performed for each combination of peer influence and smoking-based popularity, that is, sixty-three thousand in total.

The results suggest that "changes in PI [peer influence] and smoker popularity can affect smoking behaviour, but their effects are contingent to one another" (Schaefer et al. 2013, p. 295). Both factors do not contribute to smoking behaviour in the same extent. Higher values for peer influence, while keeping constant smokers popularity, did not have substantial consequences on smoking prevalence. However, holding peer influence constant, as smoking-based popularity was increased, smoking prevalence and initiations increased while cessations decreased. Manipulating peer influence and smokers popularity shows that when the former is at its minimum value and students do not try to imitate their peers behaviour, the popularity of smokers has no effect on smoking prevalence. The authors conclude that "[…] both peer influence *and* a nonzero level of smoking-based popularity may be necessary for changes in the other parameter to affect smoking" (Schaefer et al. 2013, p. 28 S). For instance, higher levels of peer influence, when smoker popularity is positive, increase smok-

ing. But when smoker popularity is negative, higher levels in peer influence decrease smoking. It should be added that the joint-effect of both parameters is non-additive.

In line with public health issues, if the goal is smoking reduction, these results throw some light on the micro-level processes linking friendship and smoking in adolescence. Furthermore, the alternative settings generated by successive changes in the parameter values can be understood as virtual post-intervention scenarios. They can offer, then, a guide for modulating the direction and intensity of real interventions.

5.1 Agent-Based Models as Mechanistic Studies

Returning now to the question that concerns us here, namely, evidence of mechanisms, we have already noted that it is not necessarily associated to a specific evidence-gathering method. Could we consider that a computer simulation like that carried out by Schaefer and his collaborators provides us with such sort of evidence? Evidential pluralists have argued, indeed, that ABMs may be understood as mechanistic studies that provide such evidence (Clarke et al. 2014; Parkkinen et al. 2018). Researchers in public health have pronounced in a similar vein: "[t]hese models provide insight into the underlying mechanisms that give rise health behaviours and outcomes" (Tracy et al. 2018, p. 84). In fact, Schaefer and his collaborators contemplate their work as an investigation "that emphasizes micro-level mechanisms that may underlie population-level outcomes" (Schaefer et al. 2013, p. 25 S).

Let us consider again their model from a different perspective, that is, as a potential source of evidence of mechanisms. Particularly, two mechanisms are scrutinized in this simulation. First, the model includes the mechanism responsible for adolescent smoking. By means of individual agents, it represents the component entities of the mechanism (i.e., students) and those properties considered relevant such as age, alcohol, and parent smoking. The model also accounts for actions of students such as starting or quitting smoking. These actions are represented through the behavioural rule of smoking function. But the model also includes the mechanism responsible for friendship networks. Its component entities are the same (i.e., students) but the relevant properties here are popularity, age similarity, and smoking ego. Interactions between students (e.g., creation of a new tie) are also taken into account. They are represented through the behavioural rule of network function.

The model, then, assumes that there are two relevant mechanisms involved in adolescent smoking behaviour. The existence of those mechanisms is taken for granted as a working hypothesis at least. This is not, of course, an ungrounded point of departure, since there is previous empirical evidence, mentioned by Schaefer and their collaborators, that favours the idea that friendship networks are crucial in this context. To observe and study the interactions between the mechanisms of adolescent smoking and friendship networks, the simulation is conducted. The empirically-calibrated model is run to observe the evolution of the adolescent smoking and friendship networks. Both phenomena do not evolve independently. In this sense, the authors say that "the SABM represents the coevolution of peer networks and smoking" (Schaefer et al. 2013, p. 255).

Recall now our classification about sorts of evidence of mechanisms at Sect. 4. No evidence of the existence of any mechanism is provided by Schaefer's model.

It is built and empirically calibrated to represent two mechanisms which are taken for granted as a working hypotheses. On the other side, evidence of components of the mechanisms—e.g, a database about adolescent health—is *used* as input to feed the model. The profile of the components of the mechanism, that is, the set of those properties considered relevant, is fixed from the outset. Subsequently, simulations are conducted to observe how both mechanisms evolve and interact with each other in terms of the smoking behaviour and incidence. As it has been noticed, the study's main target is the *interaction* between both mechanisms in diverse scenarios. The values of peer influence and smoker popularity are modified in subsequent runs and the effects of those interventions on smoking behaviour are estimated taken separately but also jointly. The simulation outcome revealed that there is no pure additive effect. This result illuminates the sort of interaction between the operating mechanisms. Lastly, regarding evidence of a property of a mechanism, the study remains silent, since it is primarily focused on the interactive aspects of the process.

This example shows, on our view, that ABMs can provide evidence of mechanisms that could be helpful to overcome the abovementioned limitations to obtain evidence in public health. Nonetheless, as mechanistic studies, ABMs also face important challenges. First, as the previous example illustrates, to provide evidence of mechanisms, ABMs often need to be empirically calibrated and this empirical calibration usually relies on previous evidence of mechanisms. For instance, to study the properties of a mechanism, an ABM requires previous evidence about the component entities and activities of the mechanism. It must be calibrated to adequately represent the mechanism. Nonetheless, in public health, as it has been noted, evidence of mechanisms is often scarce and empirical calibration is not always possible.

Secondly, it is not clear that ABMs may provide some types of evidence of mechanisms. As Schaefer's model shows, ABMs can provide evidence of interactions between mechanisms. They also seem able to supply evidence of components and properties of mechanisms. However, given their bottom-up nature, ABMs can hardly provide evidence of the existence of a mechanism. As it has been explained, evidence of the existence of a mechanism just supports the existence of some mechanism linking two phenomena or variables. It remains silent about the properties or components of the mechanism. Nonetheless, ABMs always make assumptions about the components of the mechanism since individual agents and behavioural rules are their basic building blocks.

We wish to emphasize here, nonetheless, that ABM-modelling is a viable strategy for obtaining evidence of mechanisms, at least in some of its dimensions, in public health. It can be considered, then, as a source of evidence of mechanisms to complement available statistical evidence.

6 Conclusions

In the last few decades, the evidence-based approach has become dominant in the biomedical sciences. According to it, statistical evidence provided by experimental studies is strongly favoured. Nonetheless, despite its popularity, remarkable limitations of standard evidence-based practice have been denounced. In this context, evi-

dential pluralism has recently been developed as an alternative standpoint. Evidential pluralism holds that both statistical evidence and evidence of mechanisms are crucial to establish causal claims in the biomedical sciences. Advocates of evidential pluralism argue that statistical evidence and evidence of mechanisms complement each other.

In this paper, we have analysed the application and prospects of evidential pluralism in public health. Regarding this matter, some initial conclusions have been reached. IARC's framework incorporates a pluralistic perspective on evidence, but its evidentiary procedure is not straightforwardly generalizable to the field of public health. Given its focus on biological mechanisms related to carcinogenesis, the standards proposed by IARC seem rather parochial. Therefore, coping with evidence of mechanisms in public health requires a broader perspective which extends to a wider array of mechanisms.

Accordingly, we have evaluated the prospects for obtaining evidence of mechanisms in the more general framework of public health. Hence, we discern several dimensions of evidence of mechanisms which are, in principle, independent of the research context. In addition, we contend that computational simulation should be acknowledged as a further source for obtaining evidence of mechanisms. In the field of public health in particular, ABM modelling is an appealing strategy given its bottom-up approach, from individual agents to collective outcomes, and its flexibility to integrate very diverse bits of information.

On our view, decision-making should be approached from a pluralist approach to evidence provided that reliable and complementary evidence on the efficacy of interventions can be obtained. We have argued that evidential pluralism may be useful to alleviate those empirical uncertainties related to public health. It must be acknowledged, nonetheless, that decision-making is also affected by non-empirical uncertainties, e.g., ethical uncertainties. They do not concern the actual state of the world and, unfortunately, cannot be dispelled by providing additional evidence. Contexts where public health decision-making is involved are complex. However, even though evidential pluralism does not necessarily provide an ultimate answer for them, we believe it is a significant advance.

Acknowledgements Previous drafts of this paper were presented at Prague-Valencia Workshop on Analytic Philosophy, 2022 (Czech Academy of Sciences); International Workshop on Logic, Philosophy and History of Medicine, 2022 (University of Seville); V Jornadas Novatores, 2023 (University of Salamanca). We are grateful to the audiences. This work was supported by the Spanish Ministry of Science and Innovation [grant number PID2021-125936NB-I00].

Funding Open Access funding provided thanks to the CRUE-CSIC agreement with Springer Nature.

Declarations

Competing Interests The authors have no competing interests to declare that are relevant to the content of this article.

References

Anjum RL, Copeland S, Rocca E (2020) Medical scientists and philosophers worldwide appeal to EBM to expand the notion of 'evidence'. BMJ Evid-Based Med 25:6–8. https://doi.org/10.1136/bmjebm-2018-111092

Auchincloss AH, Diez Roux AV (2008) A New Tool for Epidemiology: the usefulness of dynamic-Agent models in understanding Place effects on Health. Am J Epidemiol 168:1–8. https://doi.org/10.1093/aje/kwn118

Barrett CB, Carter MR (2010) The power and pitfalls of experiments in Development Economics: some non-random reflections. Appl Econ Perspect Policy 32:515–548. https://doi.org/10.1093/aepp/ppq023

Brownson RC, Fielding JE, Maylahn CM (2009) Evidence-based Public Health: a Fundamental Concept for Public Health Practice. Annu Rev Public Health 30:175–201. https://doi.org/10.1146/annurev.publhealth.031308.100134

Brownson RC, Baker EA, Deshpande AD, Gillespie KN (2017) Evidence-based public health, 3 rd. Oxford University Press

Bruch E, Atwell J (2015) Agent-based models in empirical Social Research. Sociol Methods Res 44:186–221. https://doi.org/10.1177/0049124113506405

Cartwright N, Munro E (2010) The limitations of randomized controlled trials in predicting effectiveness. J Eval Clin Pract 16:260–266. https://doi.org/10.1111/j.1365-2753.2010.01382.x

Clarke B, Gillies D, Illari P et al (2014) Mechanisms and the evidence hierarchy. Topoi 33:339–360. https://doi.org/10.1007/s11245-013-9220-9

de Marchi S, Page SE (2014) Agent-based models. Annu Rev Polit Sci 17:1–20. https://doi.org/10.1146/annurev-polisci-080812-191558

Epstein JM (2006) Generative Social Science: studies in Agent-based computational modeling. Princeton University Press

Fisher RA (1957) Dangers of cigarette-smoking. Br Med J 2:297–298. https://doi.org/10.1136/bmj.2.5039.297-b

Gamoran A (2018) Evidence-based policy in the Real World: a cautionary view. Ann Am Acad Pol Soc Sci 678:180–191. https://doi.org/10.1177/0002716218770138

Gillies D (2018) Causality, probability, and Medicine. Routledge, Abingdon, UK

Glennan S (2017) The new mechanical philosophy. Oxford University Press

Guyatt G, Cairns J, Churchill D et al (1992) Evidence-based medicine: a New Approach to Teaching the practice of Medicine. JAMA 268:2420–2425. https://doi.org/10.1001/jama.1992.03490170092032

Guyatt GH, Oxman AD, Schünemann HJ et al (2011) GRADE guidelines: a new series of articles in the Journal of Clinical Epidemiology. J Clin Epidemiol 64:380–382. https://doi.org/10.1016/j.jclinepi.2010.09.011

Howick J (2011) The philosophy of evidence-based medicine. Wiley-Blackwell, Oxford, UK

IARC (1990) IARC monographs on the identification of carcinogenic hazards to humans, vol 50. Pharmaceutical Drugs. International Agency for Research on Cancer, Lyon, France

IARC (2020) IARC monographs on the identification of carcinogenic hazards to humans: Vol. 124. Night Shift Work. International Agency for Research on Cancer, Lyon, France

IARC (2024) IARC monographs on the identification of carcinogenic hazards to humans: Vol. 134. Aspartame, Methyleugenol, and Isoeugenol. International Agency for Research on Cancer, Lyon, France

Illari PM (2011) Mechanistic evidence: disambiguating the russo–Williamson thesis. Int Stud Philos Sci 25:139–157. https://doi.org/10.1080/02698595.2011.574856

Illari PM, Williamson J (2012) What is a mechanism? Thinking about mechanisms across the sciences. Eur J Philos Sci 2:119–135. https://doi.org/10.1007/s13194-011-0038-2

Iranzo V, Pérez-González S (2021) Epidemiological models and COVID-19: a comparative view. Hist Philos Life Sci 43:104. https://doi.org/10.1007/s40656-021-00457-9
Jenicek M (1997) Epidemiology, evidenced-based Medicine, and evidence-based Public Health. J Epidemiol 7:187–197. https://doi.org/10.2188/jea.7.187
Khosrowi D, Reiss J (2019) Evidence-based policy: the Tension between the epistemic and the normative. Crit Rev 31:179–197. https://doi.org/10.1080/08913811.2019.1688520
Kincaid H (2012) Mechanisms, causal modeling, and the limitations of traditional multiple regression. In: Kincaid H (ed) The Oxford Handbook of Philosophy of Social Science. Oxford University Press, pp 46–74
Leuridan B, Weber E (2011) The IARC and mechanistic evidence. In: Illari PM, Russo F, Williamson J (eds) Causality in the sciences. Oxford University Press, Oxford, UK, pp 91–109
Macy M, Flache A (2009) Social dynamics from the bottom up: agent-based models of social interaction. In: Hedström P, Bearman PS (eds) The Oxford Handbook of Analytical Sociology. Oxford University Press, pp 245–268
Maglio PP, Mabry PL (2011) Agent-based models and Systems Science Approaches to Public Health. Am J Prev Med 40:392–394. https://doi.org/10.1016/j.amepre.2010.11.010
Mercury M, Upshur REG (2022) Evidence-Based Medicine and Public Health. In: Venkatapuram S, Broadbent A (eds) The Routledge Handbook of Philosophy of Public Health, 1st edition. Routledge, Abingdon, Oxon; New York, NY, pp 143–160
Nianogo RA, Arah OA (2015) Agent-based modeling of noncommunicable diseases: a systematic review. Am J Public Health 105:e20–e31. https://doi.org/10.2105/AJPH.2014.302426
Norris SL, Rehfuess EA, Smith H et al (2019) Complex health interventions in complex systems: improving the process and methods for evidence-informed health decisions. BMJ Glob Health 4:e000963. https://doi.org/10.1136/bmjgh-2018-000963
OCEBM Levels of Evidence Working Group (2011) The Oxford 2011 levels of evidence
Ongaro M, Andreoletti M (2022) Non-empirical uncertainties in evidence-based decision making. Perspect Sci 30:305–320. https://doi.org/10.1162/posc_a_00416
Parkkinen V-P, Wallmann C, Wilde M et al (2018) Evaluating evidence of mechanisms in medicine: principles and procedures. Springer, Cham, Switzerland
Pérez-González S (2024) Evidence of mechanisms in evidence-based policy. Stud Hist Philos Sci 103:95–104. https://doi.org/10.1016/j.shpsa.2023.11.006
Proctor RN (2001) Tobacco and the global lung cancer epidemic. Nat Rev Cancer 1:82–86. https://doi.org/10.1038/35094091
Proctor RN (2012) The history of the discovery of the cigarette–lung cancer link: evidentiary traditions, corporate denial, global toll. Tob Control 21:87–91. https://doi.org/10.1136/tobaccocontrol-2011-050338
Russo F (2022) Causal Pluralism and Public Health. In: Venkatapuram S, Broadbent A (eds) The Routledge Handbook of Philosophy of Public Health, 1st edition. Routledge, Abingdon, Oxon; New York, NY, pp 98–113
Russo F, Williamson J (2007) Interpreting causality in the health sciences. Int Stud Philos Sci 21:157–170. https://doi.org/10.1080/02698590701498084
Schaefer DR, Adams J, Haas SA (2013) Social Networks and Smoking: exploring the effects of peer influence and smoker popularity through simulations. Health Educ Behav 40:24S–32S. https://doi.org/10.1177/1090198113493091
Silverman E, Gostoli U, Picascia S et al (2021) Situating agent-based modelling in population health research. Emerg Themes Epidemiol 18:10. https://doi.org/10.1186/s12982-021-00102-7
Snijders Table (2017) Siena: Statistical Modeling of Longitudinal Network Data. In: Alhajj R, Rokne J (eds) Encyclopedia of Social Network Analysis and Mining. Springer, New York, NY, pp 1–9
Stegenga J (2011) Is meta-analysis the platinum standard of evidence? Stud Hist Philos Sci Part C Stud Hist. Philos Biol Biomed Sci 42:497–507. https://doi.org/10.1016/j.shpsc.2011.07.003
Tracy M, Cerdá M, Keyes KM (2018) Agent-based modeling in Public Health: current applications and future directions. Annu Rev Public Health 39:77–94. https://doi.org/10.1146/annurev-publhealth-040617-014317
Victora CG, Habicht J-P, Bryce J (2004) Evidence-based public health: moving beyond randomized trials. Am J Public Health 94:400–405
Williamson J (2019a) Establishing Causal claims in Medicine. Int Stud Philos Sci 32:33–61. https://doi.org/10.1080/02698595.2019.1630927

Williamson J (2019b) Evidential Proximity, Independence, and the evaluation of carcinogenicity. J Eval Clin Pract 25:955–961. https://doi.org/10.1111/jep.13226
Worrall J (2007) Why there's no cause to Randomize. Br J Philos Sci 58:451–488. https://doi.org/10.1093/bjps/axm024