TECHNICAL REPORT
R-483
August 2018

Causal, Casual and Curious

Judea Pearl*

Does Obesity Shorten Life? Or is it the Soda? On Non-manipulable Causes

https://doi.org/10.1515/jci-2018-2001

Abstract: Non-manipulable factors, such as gender or race have posed conceptual and practical challenges to causal analysts. On the one hand these factors do have consequences, and on the other hand, they do not fit into the experimentalist conception of causation. This paper addresses this challenge in the context of public debates over the health cost of obesity, and offers a new perspective, based on the theory of Structural Causal Models (SCM).

Keywords: Manipulability, consistency, causal effect, counterfactuals, interventions

1 Introduction

Manipulability theories of causation, according to which causes earn their meaning and usefulness by transmitting change from actions to effects have had considerable intuitive appeal among scientists and philosophers [1–4]. The rise of Fisher's RCT to the "gold standard" of experimental science further entrenched manipulability as a prerequisite for causation. In some communities, this entrenchment has turned into a dogma, cast for example in the mantra "no causation without manipulation [5] that has led to cultural prohibition on labeling sex or race as "causes."

Other research camps have been more tolerant to causal labels. In the structural causal model (SCM) framework, for example, manipulations are merely convenient means of interrogating nature, and causal relations enjoy independent existence, oblivious to external interventions [6–9]. In this framework, variables earn causal character through their capacity to sense and respond to changes in other variables. For example the variable "sex" earns the label "cause" by virtue of having responders such as "hormone content" or "height" which are gender dependent.

2 Does obesity kill?

These conceptual differences came to public attention in the context of concerns with obesity and its health care consequences.

A Reuters story published on September 21 2012 [10] cites a report projecting that at least 44 percent of U.S. adults could be obese by 2030, compared to 35.7 percent today, bringing an extra $66 billion a year in obesity-related medical costs. A week earlier, New York City adopted a regulation banning the sale of sugary drinks in containers larger than 16 ounces at restaurants and other outlets regulated by the city health department.

Interestingly, an article published earlier in the *International Journal of Obesity* [11] questions the logic of attributing consequences to obesity. The authors, M. A. Hernán and S. L. Taubman (both of Harvard's School of Public Health) imply that the very notion of "obesity-related medical costs" is undefined, if not misleading and that, instead of speaking of "obesity shortening life" or "obesity raising medical costs" one should be

***Corresponding author: Judea Pearl,** Cognitive Systems Laboratory, Computer Science Department, University of California, Los Angeles, CA 90024, USA, e-mail: judea@cs.ucla.edu

Brought to you by | University of California - Los Angeles - UCLA Library
Authenticated
Download Date | 10/31/18 12:04 AM

speaking of manipulable variables like "life style" or "soda consumption" as causing whatever harm we tend to attribute to obesity.

The technical rationale for these claims is summarized in their abstract: "We argue that observational studies of obesity and mortality violate the condition of consistency of counterfactual (potential) outcomes, a necessary condition for meaningful causal inference, because (1) they do not explicitly specify the interventions on body mass index (BMI) that are being compared and (2) different methods to modify BMI may lead to different counterfactual mortality outcomes, even if they lead to the same BMI value in a given person" [11].

That BMI is merely a coarse proxy of obesity is well taken; obesity should ideally be described by a vector of many factors, some are easy to measure and others are not. But accessibility to measurement has no bearing on whether the effect of that vector of factors on morbidity is "well defined" or whether the condition of *consistency* is violated when we fail to specify the interventions used to regulate those factors.

Consistency asserts that a patient who happened to take treatment x and recovered would also recover "had he taken the treatment" by choice, persuasion or ideal experimental protocol. The idea that consistency is an *assumption* and, hence, that it may be violated in some experimental settings appears only in Neyman-Rubin's potential outcome framework [12] but not in any of the logical accounts of counterfactuals. In both the structural account [13–15] and the closest-world accounts of counterfactuals [16] *consistency* emerges as a theorem in counterfactual logic, not as an assumption that may vary from problem to problem [8, pp. 202–206, 238–240]. According to both accounts, counterfactuals are inherently consistent [17].

A counterfactual may appear to be inconsistent when its antecedent A (as in "had A been true") is conflated with an external intervention I devised to enforce the truth of A. Indeed in Rubin's writings, the antecedent x of the potential outcomes Y_x is invariantly a "treatment" or "treatment assignment," not an event, or a state, as in "has the Temperature X been x." Unwilling to distinguish between the antecedent event $X = x$ and the instruments I deployed to make $X = x$ true, Rubin was compelled to pose consistency as an assumption, (part of SUTVA), in order to prevent side effects of I from contaminating causal effects attributed to $X = x$.

Practical interventions tend indeed to have side effects, and these need to be reckoned with if we insist on equating "the effect of A" with "the effect of I." But counterfactuals and causal effects are defined independently of those interventions and therefore, are not to be denied existence, or rendered "inconsistent" by the latter's imperfections. In [17], I show how consistency is maintained despite such imperfections.

In epidemiology, the word "treatment" is frequently replaced by "exposure" and much research is concerned with observational studies, where "treatments" or "exposures" are not controlled by humans but by Nature. This change in focus has not been sufficient to separate epidemiological research from its potential outcome roots. The clinging to "potential interventions" still dominates epidemiological thinking, vocabulary and education. [18].

3 Obesity, the *Do*-operator and SCM

In this section, we will cast the consideration above in the vocabulary of the *do*-operator, which is used to denote the operation of "holding a variable constant." Translated into this vocabulary, Hernán and Taubman's arguments amount to claiming that the causal effect $P(mortality = y|do(obesity = x))$ is undefined, seemingly because the consequences of obesity depend on how we choose to manipulate it. Since the probability of death will generally depend on whether you manipulate obesity through diet versus, say, exercise. (We assume that we are able to model the joint effect of that vector of factors X that defines "obesity.") Hernán and Taubman conclude that $P(mortality = y|do(obesity = x))$ is not formally a function of x, but a one-to-many mapping.

This contradicts, of course, what the quantity $P(Y = y|do(X = x))$ was meant to represent. In its original conception [19], $P(mortality = y|do(obesity = x))$ does not depend on any choice of intervention; it is defined relative to a hypothetical, minimal intervention needed for establishing $X = x$ and, so, it is defined independently of how the event $obesity = x$ actually came about.

Brought to you by | University of California - Los Angeles - UCLA Library
Authenticated
Download Date | 10/31/18 12:04 AM

While it is true that the probability of death will generally depend on whether we manipulate obesity through diet versus, say, exercise, the quantity $P(mortality = y|do(obesity = x))$ depends only on the levels x of the vector X and the anatomical or social processes that respond to this vector of characteristics X. $Do(obesity = x)$ describes a virtual intervention, by which nature sets obesity to x, independent of diet or exercise, while keeping everything else in tact, especially the processes that respond to X.

Three more points are worth considering:

(1) Causal effects of anatomical conditions draw their meaning from functional dependencies among those conditions and these dependencies were chosen to serve as standards of scientific communication. In ordinary scientific discourse, bodily functions are described in terms of "absence" or "presence" of certain agents not by the means through which those agents are controlled (see Appendix).

(2) If we wish to define a new operator, say $do_a(X = x)$, where a stands for the means used in achieving $X = x$, this can easily be done within the syntax of the *do*-calculus [20]. But that would be a new operator altogether, related to but not equal to $do(X = x)$ which is manipulation-neutral.

(3) An intervention that requires a special treatment involves disjunctive actions, for example, "paint the wall red or blue" or "exercise at least half hour each day." Taken formally, such partly specified actions translate into a counterfactual with disjunctive antecedent, $Y_{A \text{ or } B}$, which cannot be directly computed from the structural definition of counterfactuals. They do not produce a unique value but a range of values [17]. With additional assumptions however a unique value can be defined for disjunctive queries as well [21].

There are several ways of loading the $do(X = x)$ operator with manipulational or observational specificity. In the obesity context, one may wish to consider $P(mortality = y|do(diet = z))$ or $P(mortality = y|do(exercise = w))$ or $P(mortality = y|do(exercise = w), do(diet = z))$ or $P(mortality = y|do(exercise = w), See(diet = z))$ or $P(mortality = y|See(obesity = x), do(diet = z))$. The latter corresponds to the studies criticized by Hernán and Taubman, where one manipulates diet and passively observes obesity. All these variants are legitimate quantities that one may wish to evaluate, if called for, but have little to do with $P(mortality = y|do(obesity = x))$ which is manipulation-neutral.

Under certain conditions we can even infer $P(mortality = y|do(obesity = x))$ from data obtained under dietary controlled experiments (i. e., data governed by $P(mortality = y|See(obesity = x), do(diet = z))$ [22]). But these conditions can only reveal themselves to researchers who acknowledge the legitimacy of $P(mortality = y|do(obesity = x))$, and accept it as a derived quantity of one's model.

All these variants can be defined and evaluated in SCM and, moreover, the modeler need not think about them in the construction of the model, where only one relation matters: who listens to whom.

To summarize:

1. The fact that morbidity varies with the way we choose to manipulate obesity (e. g., diet, exercise) does not diminish our need, or ability to define a manipulation-neutral notion of "the effect of obesity on morbidity." The latter is often the actual target of scientific investigation, and may serve to inform manipulation-specific effects of obesity.
2. In addition to defining and providing identification conditions for the manipulation-neutral notion of "effect of obesity on morbidity" the SCM framework also provides formal definitions and identification conditions for each of the many manipulation-specific effects of obesity, and this can be accomplished through a single SCM model provided that the characteristics of those manipulations are encoded in the model. Concrete examples are presented in Pearl [17].

4 Scientific vs. policy-based causation

The tension between knowledge-based statements (e. g., "obesity kills") and policy-specific statements (e. g., "Soda kills.") goes back at least several decades, and was a subject of a debate I had in 2010 with Nancy Cartwright [23]. In her book, *Hunting Causes and Using Them* [24], Cartwright expresses several objections to

Brought to you by | University of California - Los Angeles - UCLA Library
Authenticated
Download Date | 10/31/18 12:04 AM

the $do(x)$ operator and the "surgery" semantics on which it is based (pp. 72 and 201). One of her objections concerned the fact that the *do*-operator represents an ideal, atomic intervention, different from the one implementable by most policies under evaluation. According to Cartwright, for policy evaluation we generally want to know what would happen were the policy really set in place, and the policy may affect a host of changes in other variables in the system, some envisaged and some not. In my answer to Cartwright [23], I stressed two points. First, the *do*-calculus enables us to evaluate the effect of compound interventions as well, as long as they are described in the model and are not left to guesswork. Second, that in many studies our goal is not to predict the effect of the crude, non-atomic intervention that we are about to implement but, rather, to evaluate an ideal, atomic policy that cannot be implemented given the available tools, but that represents nevertheless scientific knowledge that is pivotal for our understanding of the domain.

The example I used was as follows: Smoking cannot be stopped by any legal or educational means available to us today; cigarette advertising can. That does not stop researchers from aiming to estimate "the effect of smoking on cancer," and doing so from experiments in which they vary the instrument—cigarette advertisement—not smoking. The reason they would be interested in the atomic intervention $P(cancer|do(smoking))$ rather than (or in addition to) $P(cancer|do(advertising))$ is that the former represents a stable biological characteristic of the population, uncontaminated by social factors that affect susceptibility to advertisement, thus rendering it transportable across cultures and environments. With the help of this stable characteristic, one can assess the effects of a wide variety of practical policies, each employing a different smoking-reduction instrument. For example, if careful scientific investigations reveal that smoking has no effect on cancer, we can comfortably conclude that increasing cigarette taxes will not decrease cancer rates and that it is futile for schools to invest resources in anti-smoking educational programs.

There is of course a fundamental difference between smoking and obesity; randomization is physically feasible in the case of smoking (say, in North Korea)—not in the case of obesity. Yet it is not entirely unimaginable. An RCT on obesity requires more creative imagination, invoking not a powerful dictator, but an agent such as Lady Nature herself, who can increase obesity in a well-specified way (say 1 % increase in fat content of a certain muscle) and evaluate its consequences on various body functions.

This is what the *do*-operator does, it simulates an experiment conducted by Lady Nature who, for all that we know is almighty, and can permit all the organisms that are affected by a given change (say in fat content) to respond to that change in the same way that they responded in the past.

Moreover, she is able to do it by an extremely delicate surgery, without touching those variables that we mortals need to change in order to drive BMI up or down.

This is not a new agent by any means, it is the standard agent of science. For example, consider the *ideal gas law*, $PV = nRT$. While Volume (V), Temperature (T) and the amount of gas (n) are independently manipulable, pressure (P) is not. This means that whenever we talk about the pressure changing, it is always accompanied by a change in V, n and/or T which, like diet and exercise, have their own side effects. Does this prevent us from speaking about the causal effect of tire pressure on how bumpy the road is? Must we always mention V, T or n when we speak about the effect of air pressure on the size of the balloon we are blowing? Of course not! Pressure has life of its own (the rate of momentum transfer to a wall that separates two vessels) independent on the means by which we change it.

At this point, the skeptic may argue: "Things are nice in physics, but epidemiology is much more complex, we do not know the equations or the laws, and we will never in our lifetime know the detailed anatomy of the human body." This ignorance-pleading argument always manages to win the hearts of the mystic, especially among researchers who feel uncomfortable encoding partial scientific knowledge in a model. Yet Lady Nature does not wait for us to know things before she makes our heart muscle respond to the fat content in the blood. And we need not know the exact response to postulate that such response exists."

Scientific thinking is not unique to physics. Consider any standard medical test and let's ask ourselves whether the quantities measured have "well-defined causal effects" on the human body. Does "blood pressure" have any effect on anything? Why do we not hear complaints about "blood pressure" being "not well defined?" After all, following the criterion of Hernán and Taubman [11], the "effect of X on Y" is ill-defined whenever Y depends on the means we use to change X. So "blood pressure" has no well defined effect on any organ

Brought to you by | University of California - Los Angeles - UCLA Library
Authenticated
Download Date | 10/31/18 12:04 AM

in the human body. The same goes for "blood count," "kidney function," "Rheumatoid Factor."... If these variables have no effects on anything why do we measure them? Why do physicians communicate with each other through these measurements, instead of through the "interventions" that may change these measurements? The reasons lie, again, in the scientific meaning of these entities and their stability across domains.

5 A recent flair up

The debate concerning non-manipulable factors has not subsided in 2012. A recent version of the debate was re-kindled in *Epidemiology* in 2016.

A series of articles, commentaries, letters, and responses across numerous journals has considered the limits of the potential outcomes framework, among them its dogmatic insistence on tying causal effects to interventions, including the undefined notion of "hypothetical intervention" [25, 26].

While some of the discussants were willing to relax this intervention-centric attitude [27], others defended it. Miguel Hernán [28], for example, insisted on disqualifying non-manipulable states of nature (e. g., being obese, being black) from having causal effects, and goes as far as questioning the scientific legitimacy of SCM for allowing such effects. Quoting: "...different interventions often result in different effect estimates and therefore, it is also unclear whether nonparametric structural equations models are always scientifically meaningful" [28].

In a recent letter to *Epidemiology* [29], Tyler VanderWeele proposes a formal definition of "hypothetical Interventions" which requires essentially that $Y_x(u)$ be uniquely entailed "by the state of the universe and the laws of nature." He states that this proposal is not incompatible with the one implied by the *do*-operator, which allows a unique derivation of every counterfactual $Y_x(u)$ in the system [6, p. 204]. Indeed, since a structural causal model represents (a summary of) the state of the universe and the laws of nature, one would expect the two proposals to be compatible.

However, VanderWeele's proposal still leaves open the question whether constructs such as obesity, or even "Temperature" qualify as "causes." The reason, again, are possible side effect of the needed interventions, this time, Nature's interventions. In his words: "there may have to be numerous things that would have to be different for the temperature to be different, such as the wind speed and air pressure in that town, along with the temperature in spatially contiguous towns, etc."

This is a major departure from the philosophy of the *do*-operator. The latter calls for cutting off the influences of all these factors, freezing them at their current state, modifying Temperature and observing its effect. The net result is well defined and intervention-neutral definition of "the causal effect of Temperature." We do not dismiss the causal effect of X as "undefined" just because X is influenced by things we do not measure. Estimating such effect is precisely what causal inference is set out to do, and Fisherian experiments demonstrate that such influences can be neutralized.

A diametrically opposed position is advanced by Pearl and Mackenzie's recent *Book of Why* [9] where Fisher's RCT obtains its legitimacy for emulating the *do*-operator rather than the other way around.

It is not surprising that this debate is taking place in epidemiology as opposed to other data-intensive fields such as economics or social science. A large segment of epidemiological researchers view their mission as that of "changing the world for the better" and, in that role, they care mostly about treatments (causal variables) that are manipulable. However, to figure out which of those treatments should be applied in any given situation, we need to understand causal relationships between both manipulable and non-manipulable variables. For instance, if someone offers to sell you a new miracle drug that (provenly) reduces obesity, and your scientific understanding is that obesity has no effect whatsoever on anything that is important to you, then, regardless of other means that are available for manipulating obesity you would tell the salesman to go fly a kite. And you would do so regardless of whether those other means produced positive or negative results. The basis for rejecting the new drug is precisely your understanding that "obesity has no effect on outcome," the very quantity that some epidemiologists now wish to purge from science, all in the name of only caring about manipulable treatments.

Brought to you by | University of California - Los Angeles - UCLA Library
Authenticated
Download Date | 10/31/18 12:04 AM

6 Conclusions

Empirical science needs both scientific and clinical knowledge to sustain and communicate that which we have learned and to advance beyond it. While the effects of diet and exercise are important for controlling obesity, the health consequences of obesity are no less important; they constitute legitimate targets of scientific pursuit, regardless of current shortcomings in clinical knowledge.

Funding: This research was supported in parts by grants from International Business Machines Corporation (IBM) [#A1771928], National Science Foundation [#IIS-1527490 and #IIS-1704932], and Office of Naval Research [#N00014-17-S-B001].

Appendix

The purpose of the appendix is to examine the vocabulary used in ordinary scientific discourse, and demonstrate that bodily functions are described in terms of "absence" or "presence" of certain agents, not by the means through which those agents are controlled. The text is taken from Wikipedia article on Cholesterol dated 5/10/15:

> According to the lipid hypothesis, abnormal cholesterol levels (hyperchol esterolemia) or, more properly, higher concentrations of LDL particles and lower concentrations of functional HDL particles are strongly associated with cardiovascular disease because these promote atheroma development in arteries (atherosclerosis). This disease process leads to myocardial infraction (heart attack), stroke, and peripheral vascular disease. Since higher blood LDL, especially higher LDL particle concentrations and smaller LDL particle size, contribute to this process more than the cholesterol content of the HDL particles, LDL particles are often termed "bad cholesterol" because they have been linked to atheroma formation. On the other hand, high concentrations of functional HDL, which can remove cholesterol from cells and atheroma, offer protection and are sometimes referred to as "good cholesterol." These balances are mostly genetically determined, but can be changed by body build, medications, food choices, and other factors. Resistin, a protein secreted by fat tissue, has been shown to increase the production of LDL in human liver cells and also degrades LDL receptors in the liver. As a result, the liver is less able to clear cholesterol from the bloodstream. Resistin accelerates the accumulation of LDL in arteries, increasing the risk of heart disease. Resistin also adversely impacts the effects of statins, the main cholesterol-reducing drug used in the treatment and prevention of cardiovascular disease.

My point in quoting this paragraph is to show that, even in "clinical significance" sections, most of the relationships are predicated upon states of variables, as opposed to manipulations of variables. They talk about being "present" or "absent", being at high concentration or low concentration, smaller particles or larger particles; they talk about variables "enabling," "disabling," "promoting," "leading to," "contributing to," etc. Only two of the sentences refer directly to exogenous manipulations, as in "can be changed by body build, medications, food choices…"

References

1. Collingwood R. An essay on metaphysics. Oxford: Clarendon Press; 1940.
2. von Wright GH. Explanation and understanding. Ithaca, NY: Cornell University Press; 1971.
3. Cook TD, Campbell DT. Quasi-experimentation: design and analysis issues for field settings. Boston: Houghton Miflin Company; 1979.
4. Woodward J. Making things happen: a theory of causal explanation. Oxford: Oxford University Press; 2003.
5. Holland P. Statistics and causal inference. J Am Stat Assoc. 1986;81:945–60.
6. Pearl J. Causality: models, reasoning, and inference. 2nd ed. New York: Cambridge University Press; 2000. 2009.
7. Pearl J. Causal inference in statistics: an overview. Stat Surv. 2009;3:96–146.
8. Pearl J. Causality: models, reasoning, and inference. 2nd ed. New York: Cambridge University Press; 2009.
9. Pearl J, Mackenzie D. The book of why: the new science of cause and effect. New York: Basic Books; 2018.

Brought to you by | University of California - Los Angeles - UCLA Library
Authenticated
Download Date | 10/31/18 12:04 AM

10. Begley S. "Can it! Soda studies cite stronger link to obesityulation," Reuters: Health News. 2012. http://www.reuters.com/article/2012/09/21/us-obesity-soda-idUSBRE88K17820120921. September 21.
11. Hernán MA, Taubman SL. Does obesity shorten life? The importance of well-defined interventions to answer causal questions. Int J Obes. 2008;32:S8. https://doi.org/10.1038/ijo.2008.82. EP.
12. Rubin D. Estimating causal effects of treatments in randomized and nonrandomized studies. J Educ Psychol. 1974;66:688–701.
13. Balke A, Pearl J. Counterfactual probabilities: computational methods, bounds, and applications. In: de Mantaras RL, Poole D, editors. Uncertainty in artificial intelligence 10. San Mateo, CA: Morgan Kaufmann; 1994. p. 46–54.
14. Balke A, Pearl J. Probabilistic evaluation of counterfactual queries. In: Proceedings of the twelfth national conference on artificial intelligence, volume I. Menlo Park, CA: MIT Press; 1994. p. 230–7.
15. Halpern J. Axiomatizing causal reasoning. In: Cooper G, Moral S, editors. Uncertainty in artificial intelligence. San Francisco, CA: Morgan Kaufmann; 1998. p. 202–10, also, J Artif Intell Res. 2000;12(3):17–37.
16. Lewis D. Counterfactuals. Cambridge, MA: Harvard University Press; 1973.
17. Pearl J. On the consistency rule in causal inference: an axiom, definition, assumption, or a theorem? Epidemiology. 2011;21:872–5.
18. VanderWeele TJ. Explanation in causal inference: methods for mediation and interaction. New York: Oxford University Press; 2015.
19. Pearl J. Comment: graphical models, causality, and intervention. Stat Sci. 1993;8:266–9.
20. Petersen ML. Compound treatments, transportability, and the structural causal model: the power and simplicity of causal graphs. Epidemiology. 2011;22:378–81.
21. Pearl J. Physical and metaphysical counterfactuals: evaluating disjunctive actions. J Causal Inference. 2017;5. Causal, Casual, and Curious Section.
22. Bareinboim E, Pearl J. Causal inference by surrogate experiments: *z*-identifiability. In: de Freitas N, Murphy K, editors. Proceedings of the twenty-eighth conference on uncertainty in artificial intelligence. Corvallis, OR: AUAI Press; 2012. p. 113–20.
23. Pearl J. Review of N. Cartwright 'Hunting causes and using them'. Econ Philos. 2010;26:69–77.
24. Cartwright N. Hunting causes and using them: approaches in philosophy and economics. New York, NY: Cambridge University Press; 2007.
25. Broadbent A, Vandenbroucke J, Pearce N. Formalism or pluralism? a reply to commentaries on 'Causality and causal inference in epidemiology'. Int J Epidemiol. 2016;45:1841–51.
26. Krieger N, Davey Smith G. The tale wagged by the DAG: broadening the scope of causal inference and explanation for epidemiology. Int J Epidemiol. 2016;45:1787–808.
27. Daniel R, De Stavola B, Vansteelandt S. The formal approach to quantitative causal inference in epidemiology: misguided or misrepresented? Int J Epidemiol. 2016;45:1817–29.
28. Hernán M. Does water kill? A call for less casual causal inferences. Ann Epidemiol. 2016;26:674–80.
29. VanderWeele TJ. On well-defined hypothetical interventions in the potential outcomes framework. Epidemiology. 2018;29:e24–5.

Brought to you by | University of California - Los Angeles - UCLA Library
Authenticated
Download Date | 10/31/18 12:04 AM