

# Causal Inference from Big Data? A Reply to Pietsch

SERENA GALLI

doi:10.48106/dial.v77.i3.04

Serena Galli. 2023. "Causal Inference from Big Data? A Reply to Pietsch." *Dialectica* 77(3): 1–1.  
doi:10.48106/dial.v77.i3.04.



# Causal Inference from Big Data?

## A Reply to Pietsch

SERENA GALLI

In his book *Big Data* (2021), Wolfgang Pietsch defends the view that variational induction, which stands in the tradition of Mill's methods, allows us to derive conclusions about causal relationships from observational data and that the algorithms that are most successfully applied for big data analysis implement variational induction. In his opinion, the fact that the analysis of big data by machine learning algorithms enables reliable predictions and effective interventions in the world supports the assumption that these algorithms correctly identify causal relationships. In the first part of the paper, I argue that attempts to infer causal relationships from observational data by variational induction face fundamental difficulties. Furthermore, I contend that these difficulties are not due to the specific way in which the method is spelled out but are manifestations of a general underdetermination problem. In the second part, I consider Pietsch's claim that the practical benefit of big data approaches indicates that variational induction implemented by machine learning algorithms generates causal knowledge. I provide a critical assessment of his notion of causal knowledge, and I argue that his conclusion relies on an inaccurate depiction of scientific practice.

In developing his views on variational induction, Pietsch relies on a difference-making account of causation. More specifically, he defines causal relevance and causal irrelevance as three-place relations between two variables and a given context or background:

In a context  $B$ , in which a condition  $A$  and a phenomenon  $C$  occur,  $A$  is causally relevant to  $C$ , in short  $A \mathcal{R} C \mid B$ , iff the following counterfactual holds: if  $A$  had not occurred,  $C$  would also not have occurred.

In a context  $B$ , in which a condition  $A$  and a phenomenon  $C$  occur,  $A$  is causally irrelevant to  $C$ , in short  $A \mathcal{J} C \mid B$ , iff the following

counterfactual holds: if A had not occurred, C would still have occurred. (Pietsch 2016b, 5)

The truth value of the defining counterfactual statement is assessed in terms of difference-making, taking into account instances that are or were realized in our world (Pietsch 2016a, 11).<sup>1</sup> Methodologically, this assessment rests on the framework of variational induction, which stands in the tradition of Mill's (1889, 253ff) methods and comprises two key methods, namely, the method of difference and the strict method of agreement. To determine if a circumstance A is causally relevant for a phenomenon C with respect to background B, the *method of difference* must be employed:

If two instances with the same background B are observed, one instance, in which circumstance A is present and phenomenon C is present, and another instance, in which circumstance A is absent and phenomenon C is absent, then A is causally relevant to C with respect to background B, iff B guarantees homogeneity. (Pietsch 2021, 33)<sup>2</sup>

Simply put, the homogeneity of the background ensures that all the circumstances that are potentially causally relevant for phenomenon C are held fixed, except for circumstance A, whose influence on phenomenon C is explicitly studied.<sup>3</sup>

In contrast, the *strict method of agreement* allows us to identify relations of causal irrelevance:

If two instances with the same background B are observed, one instance, in which circumstance A is present and phenomenon C is present, and another instance, in which circumstance A is absent and phenomenon C is still present, then A is causally irrelevant to C with respect to background B, iff B guarantees homogeneity. (Pietsch 2021, 33)

- 
- 1 This conception of counterfactual statements differs fundamentally from traditional counterfactual approaches to causation, such as those advanced by Lewis, who analyzes the truth conditions of counterfactual statements by referring to possible worlds (1973, 560–561).
  - 2 In principle, causal relationships between continuous variables can be established likewise by extending the framework of variational induction by the method of concomitant variation (Pietsch 2021, 34). In the following, I will be concerned with binary variables exclusively.
  - 3 I examine the homogeneity condition more closely in the context of epiphenomena. For a detailed discussion, cf. Pietsch (2021, 33–34; 2016b, 11–13).

For variational induction to yield reliable results, several conditions must be fulfilled. Most importantly, (i) all variables that are potentially causally relevant for the phenomenon of interest must be known, and (ii) the dataset must contain a sufficiently large number of observations covering all relevant constellations of the variable values. Pietsch acknowledges that, due to the fact that he makes these presuppositions, his account is (what he calls) *externally theory-laden*. However, he contends that his account avoids *internal theory-ladenness*, i.e., assumptions about causal connections between the variables considered. In other words, he claims to avoid the kind of theory-ladenness that is distinctive of hypothesis-driven approaches.<sup>4</sup> With the framework of variational induction, by contrast, the causal structure of the phenomenon of interest is supposed to be elaborated from the data alone.<sup>5</sup>

## 1 Causal Inference by Variational Induction and the Underdetermination Problem<sup>6</sup>

To point out the fundamental difficulty of inferring causal relationships by variational induction, I first focus on more complex causal structures, namely, (i) *symmetric overdetermination and preemption* and (ii) *epiphenomena*. Then, I evaluate whether the (iii) *directionality* of the relation of causal relevance can be established or not. For this assessment, I take for granted that the above-mentioned conditions for variational induction are met. In particular, I assume that every possible constellation of variable values that could have been generated by the causal structure in question is indeed observed and, moreover, that the set of observations involves neither measurement errors nor accidentally correlating variables.

(i) Let us consider the following dataset consisting of observations 1–4, which all share the same background B:

---

4 A prominent advocate of such an approach is Pearl, who maintains that “causal questions can never be answered from data alone” and that answering those questions “require[s] us to formulate a model of the process that generates the data, or at least some aspects of that process,” also in the context of big data (Pearl and Mackenzie 2018, 351).

5 If the requirements for variational induction are met, “then there are enough data to avoid spurious correlations and to map the causal structure of the phenomenon without further internal theoretical assumptions about the phenomenon” (Pietsch 2015, 910–911). See also Pietsch (2021, 65–66).

6 Woodward uses the term *underdetermination problem* to refer to the circumstance that, given a set of variables, different causal structures encompassing these same variables can generate an identical pattern of correlations and conditional correlations (2003, 106–107).

- Observation 1:  $A_1 \& A_2 \& C$
- Observation 2:  $A_1 \& \neg A_2 \& C$
- Observation 3:  $\neg A_1 \& A_2 \& C$
- Observation 4:  $\neg A_1 \& \neg A_2 \& \neg C$

Further, let us suppose that the causal relationships that do, in fact, underlie these observations are those depicted by the model in figure 1a. In this model, we have two potential causes of C,  $A_1$  and  $A_2$ , with  $A_1$  preempting  $A_2$  when both obtain. How does Pietsch propose to deal with this dataset? *Preempted* or *alternative causes* such as  $A_2$  require other circumstances, in this case  $A_1$ , to be absent in order to have an impact on the phenomenon of interest. He claims that such alternative causes can be singled out based on the method of difference and the method of strict agreement too, but the background B must be specified by an additional condition X, the *preempting cause*:

A is an ‘alternative cause’ to C with respect to background B, iff there exists an X such that A is causally relevant to C with respect to a background  $B \& \neg X$ , but causally irrelevant to C with respect to a background  $B \& X$  (i.e., C is always present in  $B \& X$ ). (Pietsch 2021, 34)

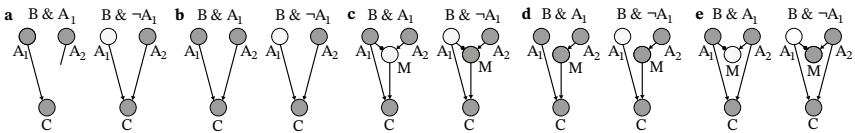


Figure 1: Symmetric overdetermination and preemption. In scenario (a), the exact mechanism of preemption is not specified and, therefore, symbolized by the prematurely terminated line originating from  $A_2$ .

Since  $A_2$  is a preempted cause of C, it is causally relevant to C with respect to B only in the absence of  $A_1$ , which can be deduced by contrasting observations 3 and 4:  $A_2 \mathcal{R} C \mid B \& \neg A_1$ . Comparing observations 1 and 2,  $A_2$  turns out to be causally irrelevant for C with respect to  $B \& A_1$ , relying on the strict method of agreement ( $A_2 \mathcal{J} C \mid B \& A_1$ ). Yet, two problems arise from Pietsch’s approach for the identification of alternative causes: First,  $A_1$ , which is the *preempting cause* of C, is determined to be causally irrelevant for C when applying the strict method of agreement to observations 1 and 3. As seen before, it is only

in the absence of  $A_2$  that the causal relevance between  $A_1$  and  $C$  with regard to background  $B$  can be established ( $A_1 \mathcal{R} C \mid B \ \& \ \neg A_2$ , observations 2 and 4). However, by definition, the impact of  $A_1$  on  $C$  does not depend on the value of the variable  $A_2$ . Thus, variational induction fails to depict the asymmetry between the preempting and the preempted cause. Second, the above-listed observations 1–4 are compatible with another underlying causal structure, namely a model of symmetric overdetermination, as displayed in figure 1b. Hence, without any prior knowledge about the causal connections between the variables, it is impossible to decide which causal structure really underlies the observed constellations of variable values.

Then, a common feature of big data is its high dimensionality, meaning that each observation includes numerous different variables. So, it could be put forward that this problem only arises because the dataset is not sufficiently complex and not enough variables were regarded. For example, introducing the variable  $M$ , which mediates the causal effect of  $A_2$  on  $C$  and whose instantiation is prevented in the presence of  $A_1$ , would definitively allow to distinguish between the case of preemption and the case of symmetric overdetermination, as depicted in figures 1c and 1d.<sup>7</sup>

Observation 5:  $A_1 \ \& \ A_2 \ \& \ \neg M \ \& \ C$

Observation 6:  $A_1 \ \& \ \neg A_2 \ \& \ \neg M \ \& \ C$

Observation 7:  $\neg A_1 \ \& \ A_2 \ \& \ M \ \& \ C$

Observation 8:  $\neg A_1 \ \& \ \neg A_2 \ \& \ \neg M \ \& \ \neg C$

Observation 5 is indeed not compatible with the model of symmetric overdetermination in figure 1d because, according to that model, the variable  $A_1$  has no impact on the other cause  $A_2$  or its mediating variable  $M$ . Yet, these four observations are consistent with another model of symmetric overdetermination, where the instantiation of  $M$  depends both on the presence of  $A_2$  and the absence of  $A_1$ , but  $M$  does not mediate the causal effect of  $A_2$  on  $C$ , as illustrated in figure 1e. Hence, including more variables does not solve, but, at best, deflects the above-mentioned difficulties.

(ii) In connection with *epiphenomena*, similar problems arise. Epiphenomena, such as  $A_4$  in figure 2c, have a common cause with the phenomenon of interest but have no causal impact on it. Let us consider another dataset con-

<sup>7</sup> According to Lewis' terminology, the causal model displayed in figure 1c is an example of early preemption (1986b, 200).

sisting of observations 9 and 10, which, in turn, share the same background B:

Observation 9:  $A_3 \ \& \ A_4 \ \& \ C$

Observation 10:  $\neg A_3 \ \& \ \neg A_4 \ \& \ \neg C$

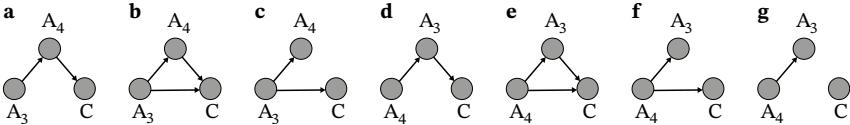


Figure 2: Epiphenomena and directionality of causal connections.

Then, let us suppose that these observations, which are compatible with all the models depicted in figure 2a–f, were generated by the causal model in 2c. Since  $A_3$  and  $A_4$  covary, neither the method of difference nor the strict method of agreement can be applied to determine whether  $A_4$  stands in a relation of causal relevance or irrelevance to  $C$  with respect to background  $B$ .<sup>8</sup> By contrast,  $A_3$  proves to be causally relevant to  $C$  with respect to  $B$ , which, in this case, does guarantee homogeneity.<sup>9</sup> On behalf of Pietsch, it could be put forward that this problem can be circumvented by considering the combination of the variables  $A_3$  and  $A_4$  instead of examining them separately. Following this approach, the method of difference establishes either the conjunction or the disjunction of  $A_3$  and  $A_4$  to be causally relevant for  $C$  with respect to background  $B$  ( $A_3 \ \& \ A_4 \ \mathcal{R} \ C \mid B$  or  $A_3 \ \vee \ A_4 \ \mathcal{R} \ C \mid B$ ).<sup>10</sup> Still, in all cases, variational induction fails to establish that  $A_4$  is causally irrelevant to  $C$  with respect to  $B$ .

Pietsch acknowledges the issue that algorithms employing variational induction may mistakenly single out epiphenomena as causally relevant for a

8  $A_3$ , which is potentially causally relevant for  $C$  as well, cannot be held fixed, as it strictly covaries with  $A_4$ . Therefore,  $B$  does not guarantee homogeneity.

9  $B$  guarantees homogeneity with respect to the relationship between  $A_3$  and  $C$  if “only circumstances that are causally irrelevant to  $C$  can change” or that “lie on a causal chain through  $A_3$ ] to  $C$  or that are effects of circumstances that lie on this causal chain” (Pietsch 2021, 33–34). Since  $A_4$  is an effect of  $A_3$ ,  $B$  does guarantee homogeneity, although  $A_4$  cannot be held fixed. However, presuming that  $A_4$  is connected to  $A_3$  in this way contradicts Pietsch’s claim that his account avoids assumptions about causal connections between the variables considered.

10 Such Boolean expressions are, as Pietsch maintains, a possible result of variational induction (2021, 50).

phenomenon of interest, although they lack the difference-making character of a cause (Pietsch 2021, 55; 2016a, 153–154).<sup>11</sup> However, he attributes it to the fact that either the dataset is incomplete or the algorithm does not fully implement variational induction. By contrast, this example demonstrates that this erroneous conclusion is not due to missing observations because it is drawn despite considering all observations compatible with a given causal model. On the other hand, it cannot be ascribed to the algorithmic implementation either, as the manual, non-algorithmic application of variational induction does not satisfyingly deal with epiphenomena either.<sup>12</sup>

(iii) Finally, the fact that observations 9–10 could have been generated by a causal structure with  $A_3$  being a cause of  $A_4$  (figures 2a–c) or a causal structure with  $A_4$  being a cause of  $A_3$  (figures 2d–f) demonstrates that the *direction of the relation of causal relevance* cannot be established by variational induction. Besides, the same holds true for phenomenon C, which could as well be a cause and not an effect of variable  $A_3$  or  $A_4$  if not predefined as the phenomenon of interest. To solve this problem, Pietsch has suggested introducing a time index for the phenomenon of interest and the variables examined (2014, 424). Yet, from a conceptual point of view, this seems like an ad hoc solution since the truth condition he specifies for the counterfactual defining causal relevance fails to capture the asymmetry in the relation of causal relevance. Additionally, this solution is not practicable for cross-sectional data, where all the variables are recorded at the same time and, accordingly, the timepoint of their occurrence remains unknown.

In light of these difficulties, Pietsch’s assertion that it is possible “to determine the true causal relationships by means of variational induction” seems to be unwarranted (2021, 61). In the causal discovery literature, it is well established that, given the causal Markov condition and the causal faithfulness condition, certain features of the underlying causal structure can be deduced from the probability distribution in the data. But, aside from special cases, it

---

<sup>11</sup> Strictly speaking, he refers to *proxies*, which I take to be the equivalent of epiphenomena.

<sup>12</sup> Needless to say, some of such wrong conclusions can be traced back to issues regarding data acquisition. For instance, a sampling error can result in an accidental correlation between variables that is not present in the population from which the sample was drawn. Let us suppose that observations 9–10 were generated by the causal structure displayed in figure 2g. In this case, the observed correlation between  $A_3$  and  $A_4$ , on one side, and C, on the other side, must have occurred by chance. Yet, by employing variational induction, the conjunction or disjunction of  $A_3$  and  $A_4$  is mistakenly identified as causally relevant for C with respect to B, and this misattribution can be recognized as such and corrected only when analyzing another, possibly larger dataset devoid of this accidental correlation.



is not possible to uniquely determine the true causal structure.<sup>13</sup> In my view, variational induction similarly faces the problem of underdetermination as it rests ultimately on the analysis of patterns of (conditional) dependencies and independencies in the data. That is to say, variational induction aims at identifying the constellation of variables  $V$  that has difference-making character with respect to background  $B$ : If this exact configuration of variables is present,  $C$  is always present as well (i.e.,  $P(C | V) = 1$ ); in its absence,  $C$  is always absent as well (i.e.,  $P(C | \neg V) = 0$ ). From this dependence between  $C$  and  $V$ , a relation of causal relevance is inferred ( $V \mathcal{R} C | B$ ). Thus, the procedure of variational induction can be viewed as the comparison of the conditional probabilities of  $C$  rather than the comparison of individual observations. From the pattern of conditional probabilities based on observations 9 and 10, for example, it can be deduced that it is either the conjunction or the disjunction of  $A_3$  and  $A_4$  that makes a difference for the value of the variable  $C$  with respect to background  $B$  (i.e.,  $V = (A_3 \& A_4)$  or  $V = (A_3 \vee A_4)$ ).<sup>14</sup> While these two Boolean expressions are highly useful for predicting the value of  $C$ , the pattern of dependencies is, as already stated for the direct comparison of individual observations, compatible with all the models depicted in figures 2a–f. Accordingly, Pietsch’s claim that “the difference-making circumstances identified by variational induction are exactly the circumstances that need to be manipulated or intervened upon in order to change a phenomenon” does not seem justified. Although a given configuration of circumstances might unequivocally determine the value of the phenomenon of interest  $C$  in an observational setting, the exact connection between these circumstances and  $C$  remains elusive. Therefore, successful intervention strategies cannot be deduced from the Boolean expression of these circumstances. While in 2a, 2d, 2e, and 2f a single intervention on  $A_4$  is an effective way of manipulating the value of  $C$ , in 2c this is clearly not the case. The Boolean expression may

---

13 These two conditions are so-called *bridge principles*, which are required to connect the observations of a given set of variables to the underlying causal model that generated these observations. More specifically, the causal Markov condition allows the inference from a probabilistic dependence between two variables to a causal connection, whereas the causal faithfulness condition allows the inference from a probabilistic independence to causal separation. Cf. Eberhardt (2017, 82–85). For a discussion of underdetermination in causal inference in relation to different success criteria and background assumptions, see Zhang (2009).

14 Since the configurations ( $A_3 \& \neg A_4$ ) and ( $\neg A_3 \& A_4$ ) do not occur in a purely observational setting, these two possibilities cannot be distinguished.

encompass the necessary and together sufficient circumstances for *observing* phenomenon C but not for *producing* it.<sup>15</sup>

## 2 Objectives of Big Data Analysis and Causal Knowledge<sup>16</sup>

Pietsch distinguishes two central functions of big data approaches, namely, *prediction* and *intervention*, and claims that the exertion of both requires some access to causal knowledge.<sup>17</sup> Arguably, the view that causal knowledge is indispensable for effectively manipulating a phenomenon of interest is hardly contested. However, causal knowledge is usually not considered a prerequisite for predictive success.<sup>18</sup> In that regard, it is useful to touch upon Pietsch's notion of causal knowledge, which bears on his distinction of direct and indirect causal connections: If a certain variable is causally relevant for another variable in a given context, as it is the case for  $A_3$  and C in figure 2c, the relationship between those two variables constitutes a *direct causal connection*, as Pietsch suggests. If, by contrast, two variables are not causally relevant for one another but are related via a common cause, then there exists an *indirect causal connection* between these two variables, as it is the case for  $A_4$  and C in

- 
- 15 As an anonymous reviewer pointed out, a promising way of dealing with this problem of underdetermination is the appeal to theoretical virtues such as parsimony. For example, Forster et al. introduce the principle of frugality that favors those causal structures with the fewest causal connections (2018). I fully agree that, technically, the procedure of variational induction could be combined with an algorithm that ranks the possible causal structures in terms of simplicity. Yet, this constraint regarding the total number of causal connections involves an assumption about the causal connections between the variables since causal models with more numerous connections, such as 2b or 2e, are dismissed in favor of models with fewer connections, although perfectly compatible with the data. Therefore, this strategy is *internally theory-laden* and not reconcilable with the concept of variational induction as a purely data-driven approach. An alternative strategy to determine the true causal structure is experimentation. For a detailed discussion of experimentation as a means for resolving underdetermination, cf. Eberhardt (2013).
- 16 In this section, which is concerned with variational induction as a means of causal inference from big data specifically, I acknowledge Pietsch's claim that the most successful algorithms are based on variational induction without further examination.
- 17 Rather than distinguishing between different functions, I would propose to differentiate between two questions that are to be answered by big data analysis. To specify *intervention* as a function of big data approaches presupposes what is under consideration. Besides, it remains unclear how to discern which function, prediction or intervention, is exerted in a given case.
- 18 For example, Woodward maintains that accurate predictions can be made based on correlations solely; furthermore, he points out that "inferences from effect to cause are often more reliable than inferences from cause to effect" (2003, 31–32).

figure 2c, which are both effects of  $A_3$  (Pietsch 2021, 55).<sup>19</sup> Certainly, only for successful interventions upon a phenomenon of interest must it be known whether there is a direct causal connection between that same phenomenon and the variable that is to be manipulated or not. But, as he argues, for accurate predictions, either such a direct causal connection or an indirect causal connection between the phenomenon of interest and a potential predictor variable is required. Thus, when a machine learning algorithm singles out a variable as a promising predictor variable for a given phenomenon of interest, the algorithm thereby generates causal knowledge to a certain degree. In my opinion, this broad notion of causal knowledge allowing for different degrees is particularly problematic in three respects:

(i) First of all, as a cause (usually) correlates with the phenomenon of interest, so does an epiphenomenon of this cause. The first correlation is indicative of a (direct) causal connection, whereas the second is indicative of a common cause. As discussed for epiphenomena, variational induction does not allow us to distinguish between a correlation ascribable to a direct causal connection and a correlation ascribable to a common cause. It follows that not only variational induction but also the analysis of (conditional) correlations yields causal knowledge in this wide sense. Accordingly, it does not seem consistent to specify correlation as a contrasting notion for causation. Furthermore, since the procedure of variational induction makes use of the pattern of dependencies in the data, it does not even allow for a distinction between correlations that are indicative of some sort of causal connection and purely accidental correlations. Therefore, it remains unclear in what sense big data algorithms are capable of delimiting causation from correlation, as Pietsch maintains.<sup>20</sup>

(ii) This broad notion of causal knowledge stands in tension with Pietsch's claim that the primary function of causal knowledge is to guide us on how to effectively intervene in the world (2021, 54). If the knowledge of an indirect causal connection between two variables is regarded as causal knowledge as well, having access to causal knowledge in this wide sense does not help to discriminate between effective and ineffective strategies to manipulate a phenomenon of interest.

19 Pietsch's distinction of direct and indirect causal connections differs from the conventional view, according to which the difference between direct and indirect causal connection results from the absence or presence of a mediating variable. See, for example, Woodward (2003, 55).

20 "By relying on variational induction, big data approaches are to some extent able to distinguish causation from correlation" (Pietsch 2021, 57).

(iii) And, finally, it risks obscuring the distinction between questions of prediction and questions of intervention, which are addressed in scientific practice: Causal knowledge in the strict sense, that is to say, knowledge about relations of causal relevance and irrelevance in a given set of variables, is no precondition for predictions. Thus, the fact that an algorithm implementing variational induction yields accurate predictions cannot be cited in support of the view that variational induction is capable of establishing causal relationships. Conversely, interventions indeed depend on causal knowledge in the strict sense. If such an algorithm truly did enable us to efficiently intervene in the world, this could speak in favor of Pietsch's view that variational induction is capable of inferring direct causal connections. As an example, he refers to "algorithms [that] are designed to determine the best medicine to cure a certain cancer" (Pietsch 2021, 54).<sup>21</sup> In fact, there are a number of studies that relied on machine learning in order to predict the response to a given drug. In a recently published work, the tumor tissue of patients with breast cancer was analyzed with different methods at diagnosis (Sammut et al. 2022). Patients were subsequently treated with chemotherapy, and treatment response was evaluated. Using a machine learning algorithm, the authors built a model to predict the response to chemotherapy, which was based on the molecular profile of the tumor as well as clinicopathological features, and model performance was successfully validated on a different dataset. Amongst other things, they drew the conclusion that patients predicted to show a poor response to standard-of-care chemotherapy should be enrolled in clinical trials investigating novel therapies. Therefore, the results of this big data approach may allow for better stratification of patients that will or will not benefit from conventional chemotherapy and are inasmuch action-guiding. However, Pietsch maintains that these algorithms are, moreover, designed to *determine the best treatment* for a given cancer, in this way allowing us to effectively intervene upon the phenomenon of interest, namely, tumor growth. For the sake of argument, let us suppose the algorithm revealed that three signaling pathways are hyperactive in tumors poorly responding to chemotherapy compared to tumors displaying a good treatment response. But, as outlined above, it is impossible to determine if (or which of) these three pathways are indeed driving tumor growth and which are rather an epiphenomenon of the cause of excessive tumor growth or even a consequence thereof. Accordingly, the question whether one of these hyperactive signaling

---

21 He does not cite any specific publication to underpin his assertion.

pathways truly constitutes a promising therapeutic target or not cannot be answered based on the observational data alone but requires experimentation. Besides, in order to successfully intervene in the world, it is essential not only to identify the causes of the phenomenon but also to understand how these causes can be manipulated, specifically which drug effectively targets a given pathway (compare figure 3). This kind of causal knowledge may be generated in randomized controlled trials or *in vitro* studies but, again, cannot be derived from observational data. To trace back the practical benefit of big data approaches to the generation of causal knowledge by the algorithms used does not accurately reflect the scientific practice, which builds upon different sources of knowledge to determine effective interventions.

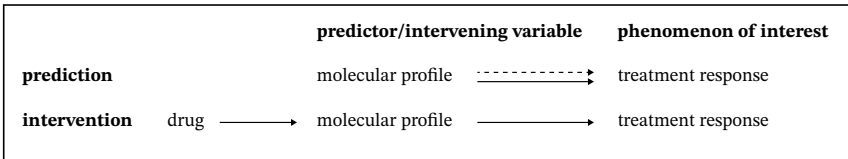


Figure 3: Prediction of and intervention upon a phenomenon of interest. Indirect causal connections are represented by dashed arrow lines, direct causal connections by solid arrow lines.

### 3 Conclusions

In my view, variational induction fails to elucidate causal structures involving preemption, symmetric overdetermination, or epiphenomena, establishing causal relationships between variables that actually are conjoined in a relation of causal irrelevance and vice versa. Furthermore, the direction of the relation of causal relevance cannot be specified by variational induction, which poses a problem for even the simplest causal models possible. These shortcomings are neither specific to the method of variational induction nor ascribable to an imperfect dataset with missing observations, an insufficient number of observed variables, or any measurement errors. Rather, the attempt to infer causal relationships from observational data (including big data) itself faces important limitations: Since for a given set of observations multiple underlying causal structures are usually conceivable, generally it is impossible to uniquely determine the true causal model from this set of observations alone without

endorsing any background assumptions about or having any prior knowledge of the causal relationships between the variables involved.

Pietsch's notion of causal knowledge explains, at least partially, why he reaches another assessment of variational induction as a method for generating causal knowledge. Supposing a broader notion of causal knowledge, he seems to have in mind a less strict success criterion: It suffices for variational induction to approximate causal relationships, namely, to determine if there is any causal connection, direct or indirect, between two variables. This could create the appearance that the conflicting assessment of variational induction as a means to infer causal relationships is merely due to two divergent, equally plausible notions of causal knowledge. However, in my opinion, Pietsch's broad notion of causal knowledge is problematic because it blurs the distinction between causation and correlation and between the prerequisites for prediction and for intervention.

If the practical benefit of big data approaches cannot be attributed to the elucidation of causal relationships, an alternative explanation is needed. The identification of predictive markers may indeed improve patient care by sparing patients who are unlikely to respond to the adverse reactions of an ineffective treatment. Randomized controlled trials can yield negative results only because patients are not selected appropriately. This could be obviated by a more adequate patient stratification based on reliable predictor variables. The analysis of the molecular profile of a tumor can generate promising hypotheses about chemoresistance in a relatively unbiased way, which may prove to be true in experimental assays. Undoubtedly, the results of machine learning algorithms contain very valuable information, which, in conjunction with knowledge derived from other sources, provide reasons to act in a certain way. In this sense, Pietsch is right in stating that precisely data-rich sciences such as medicine are fundamentally concerned with difference-making relationships and that the correlations unveiled by machine learning algorithms certainly do not replace causation. But, although such results of big data analysis can be action-guiding and aid in singling out potentially effective interventions, this should not be taken as a confirmation of the claim that machine learning algorithms indeed elucidate the causal structure underlying the observational dataset.\*

Serena Galli

---

\* I thank Peter Schulte and two anonymous reviewers for their helpful comments on an earlier version of this paper.

## References

- EBERHARDT, Frederick. 2013. "Experimental Indistinguishability of Causal Structures." *Philosophy of Science* 80(5): 684–696, doi:10.1086/673865.
- . 2017. "Introduction to the Foundations of Causal Discovery." *International Journal of Data Science and Analytics* 3: 81–91, doi:10.1007/s41060-016-0038-6.
- FORSTER, Malcolm, RASKUTTI, Garvesh, STERN, Reuben and WEINBERGER, Naftali. 2018. "The Frugal Inference of Causal Relation." *The British Journal for the Philosophy of Science* 69(3): 821–848, doi:10.1093/bjps/axw033.
- LEWIS, David. 1973. "Causation." *The Journal of Philosophy* 70(17): 556–567. Reprinted, with a postscript (Lewis 1986b), in Lewis (1986a, 159–213), doi:10.2307/2025310.
- . 1986a. *Philosophical Papers, Volume 2*. Oxford: Oxford University Press, doi:10.1093/0195036468.001.0001.
- . 1986b. "Postscript to Lewis (1973)." in *Philosophical Papers, Volume 2*, pp. 172–213. Oxford: Oxford University Press, doi:10.1093/0195036468.001.0001.
- MILL, John Stuart. 1843. *A System of Logic, ratiocinative and inductive*. London: Longmans, Green & Co.
- . 1889. *A System of Logic, Ratiocinative and Inductive, Being a Connected View of the Principles of Evidence and the Methods of Scientific Investigation*. London: Longmans, Green & Co. First edition: Mill (1843).
- PEARL, Judea and MACKENZIE, Dana. 2018. *The Book of Why: the New Science of Cause and Effect*. New York: Basic Books, <http://bayes.cs.ucla.edu/WHY/>.
- PIETSCH, Wolfgang. 2014. "The Structure of Causal Evidence Based on Eliminative Induction." *Topoi* 33(2): 421–435, doi:10.1007/s11245-013-9190-y.
- . 2015. "Aspects of Theory-Ladeness in Data-Intensive Science." *Philosophy of Science* 82(5): 905–916, doi:10.1086/683328.
- . 2016a. "The Causal Nature of Modeling with Big Data." *Philosophy & Technology* 29(2): 137–171, doi:10.1007/s13347-015-0202-2.
- . 2016b. "A Difference-Making Account of Causation." Unpublished manuscript, <https://philsci-archive.pitt.edu/id/eprint/11913>.
- . 2021. *Big Data*. Cambridge: Cambridge University Press, doi:10.1017/9781108588676.
- SAMMUT, Stephen-John, CRISPIN-ORTUZAR, Mireia, CHIN, Suet-Feung, PROVENZANO, Elena, BARDWELL, Helen A., MA, Wenxin, COPE, Wei, et al. 2022.

“Multi-Omic Machine Learning Predictor of Breast Cancer Therapy Response .”  
*Nature* 601(January): 623–629, doi:[10.1038/s41586-021-04278-5](https://doi.org/10.1038/s41586-021-04278-5).

WOODWARD, James F. 2003. *Making Things Happen. A Theory of Causal Explanation*.  
Oxford Studies in the Philosophy of Science. Oxford: Oxford University Press,  
doi:[10.1093/0195155270.001.0001](https://doi.org/10.1093/0195155270.001.0001).

ZHANG, Jiji. 2009. “Underdetermination in Causal Inference.” *Studies in Logic* 2(4):  
16–47.