

How do medical researchers make causal inferences?

Olaf Dammann, Ted Poston, and Paul Thagard¹

ABSTRACT: Bradford Hill (1965) highlighted nine aspects of the complex evidential situation a medical researcher faces when determining whether a causal relation exists between a disease and various conditions associated with it. These aspects are widely cited in the literature on epidemiological inference as justifying an inference to a causal claim, but the epistemological basis of the Hill aspects is not understood. We offer an explanatory coherentist interpretation, explicated by Thagard's ECHO model of explanatory coherence. The ECHO model captures the complexity of epidemiological inference and provides a tractable model for inferring disease causation. We apply this model to three cases: the inference of a causal connection between the Zika virus and birth defects, the classic inference that smoking causes cancer, and John Snow's inference about the cause of cholera.

Introduction

Bradford Hill asked “In what circumstances can we pass from ... [an] observed *association* to a verdict of *causation*? Upon what basis should we proceed to do so?” (Hill 1965, p. 295) Hill's expertise lay in the relationship between work conditions and illness. He often

¹ **Acknowledgments:** Thanks to Mike Bishop, Kostos Kampouratis, Kevin McCain, and Chase Wrenn for comments on an earlier draft.

had information that revealed associations among many factors, and he had to determine which factors, if any, cause which others. He aimed to provide guidelines (what he called “viewpoints”) for justifying a particular causal inference.

The Hill aspects are widely discussed and used in epidemiological inference, yet how they justify causal inference is poorly understood. Morabia (2013, p. 1526) remarked that “Hill’s viewpoints may be philosophically novel, *sui generis*, still waiting to be validated and justified.”

We advance Hill’s contribution by interpreting his viewpoints as contributions to inference to the best explanation. We first introduce the Hill aspects, and then discuss explanatory coherentism based on the principles of explanatory coherence. We then apply these principles to three cases of epidemiological inference using the ECHO model of computing explanatory coherence: the recent case of inferring a causal relationship between the Zika virus and birth defects, the classic case of inferring that smoking causes cancer, and the historical case of Snow’s inference to the cause of cholera. Each case illustrates the central coherentist theme that justified inferences require balancing various lines of evidence with various competing theoretical claims. Moreover, the cases illustrate the utility of the ECHO program for modeling epidemiological inference. Finally, we provide a general interpretation of Hill’s aspects in terms of principles of explanatory coherence, and reply to objections to our approach.

Hill’s Viewpoints

Epidemiological inference is complex. It is rarely obvious what statistically correlated factors are causally responsible for others. It is typical for multiple possible causes to be

viable explanations. Partial evidence is usually misleading. Some of the identified associations normally conflict with expected theory. In such a complex evidential situation, it is difficult to justifiably infer any causal relationship. Even so, the need to improve public health makes it imperative to discern causal relationships.

Hill's viewpoints address this complex situation and provide particular questions that a medical researcher should attempt to answer. Reordering his list, we group his nine aspects as follows:

1. Temporality – does the putative cause precede the effect?
2. Strength of association – is the association strong?
3. Consistency of association – is the association consistent across a variety of conditions?
4. Specificity of association – how specific is the association?
5. Biological gradient – is there a strong dose-response curve (i.e., the curve of an independent and dependent variable)?
6. Experiment – is the association supported by experimental study?
7. Plausibility – how plausible is the causal claim given existing biological knowledge?
8. Coherence – does the causal claim cohere with the existing history and biology of the disease?
9. Analogy – how similar is the potential causal claim with other accepted causal claims?

The first aspect, *temporality*, suggests that one should determine the beginning point of each factor and then formulate causal hypotheses guided by the rule that causes

come before their effects. Hill observed that the onset of certain factors is not always evident. Illnesses often have a long incubation period, and an illness may cause a particular factor rather than vice versa. For example, Hill asked: “Does a particular diet lead to disease or do the early stages of the disease lead to those peculiar dietetic habits?” (Hill, 1965, p. 297).

The features of association are the *strength of association*, the *consistency of association*, the *specificity of association*, the *biological gradient*, and *experiment*. The strength of association between a possible causal condition C and an effect E should examine the ratios of (i) C&E to C&~E and of (ii) C&E to ~C&E. The first ratio compares the number of cases in which the putative cause and effect are present to the cases in which the putative cause but not the effect is present. The second ratio compares the number of cases in which the putative cause and effect are present to the number of cases in which the effect is present without the putative cause. Causal relations are consistent with a low ratio (i). For instance, smoking causes lung cancer even though few smokers develop lung cancer. The key to detecting this causal relation is that lung cancer is rare in non-smokers so that there is a strong ratio (ii) of smoking and cancer to not-smoking and cancer.

The *consistency* of an association concerns whether it has been observed by different persons in different places at different times. This aspect is aimed against alternative explanations of an association such as chance and bias. Similarly, *experiment* looks for cases where removing a possible cause decreases an effect, also making less plausible alternative explanations such as chance and confounding factors. Consistency looks at existing studies in diverse circumstances, whereas experiment looks at interventional studies.

The *specificity* of association favors more precise causal paths over more general ones. Workers at several chemical plants may develop an illness, suggesting that working at chemical plants cause illness. But the suggestion is stronger if the association is limited to specific workers, sites, and diseases, and when there is no association between the work and other diseases.

The last aspect of association that Hill mentioned is *biological gradient*, which corresponds to John Stuart Mill's (1970; original 1843) method of concomitant variation. More of a cause is associated with more of an effect, and less of a cause produces less of an effect. Evidence for the causal connection between smoking and lung cancer is enhanced by the fact that the number of cigarettes smoked per day is proportional to the rate of lung cancer.

The guiding aspect of temporality together with the five aspects of association are alone inadequate to infer a causal relationship. Causal inference should also be guided by theory, captured by Hill's aspects of plausibility, coherence, and analogy. *Plausibility* assesses how the potential causal relationship fits with general biological knowledge. *Coherence* assesses how the potential causal relationship fits with the history and biology of the disease. Finally, *analogy* assesses whether the potential causal relationship is similar to other established causal relationships.

The importance of background theory in causal inference is illustrated by the history of the practice of bloodletting. Based on the theory that disease involved an imbalance of the four humors (blood, black bile, yellow bile, and phlegm), bloodletting evacuated 'bad blood' from the body to restore the proper balance of the humors. This practice was supported both by the association between bleeding patients and fever

reduction, and by the theory of disease as humoral imbalance. The germ theory of disease introduced by Pasteur dramatically changed the biological background and led to the abandonment of bloodletting. A strength of Hill's perspective is his sensitivity to the theoretical dynamics in causal inference.

Explanatory Coherence

The complexity of epidemiological inference suggests a coherentist interpretation. Evidence for a claim *emerges* from a body of information in which the relations of support between claims are bi-directional and may involve rejecting some of the originally accepted claims. On a coherentist picture of inference each claim in a body of information may contribute to the justification of any other (See, for example, Poston 2014, ch. 3).

Medical researchers highlight the emergence of conclusions from evidence. Rasmussen et. al note the emergence of a causal relation in the case of the Zika virus. They write, "As is typically the case in epidemiology and medicine, no 'smoking gun' (a single definitive piece of evidence that confirms Zika virus as a cause of congenital defects) should have been anticipated. Instead, the determination of a causal relationship would be expected to emerge from various lines of evidence, each of which suggests, but does not on its own prove, that prenatal Zika virus infection can cause adverse outcomes." (2016, p. 1982)

Dammann (2018) proposed that epidemiological inferences concerning the causes of disease can be understood in terms of explanatory coherence through Poston's (2014) development of explanatory coherentism. Furthermore, Dammann conjectured that Thagard's (1989) ECHO model of coherence computation could provide a rigorous

account of such inferences. We now develop Dammann's proposal both specifically and generally. We show how Thagard's principles of explanatory coherence apply to three important cases of epidemiological reasoning, all of which can be simulated using ECHO. We then describe more generally how these principles connect with Hill's viewpoints and similar attempts to characterize inference in epidemiology. Our results confirm and deepen the remark of Broadbent (2017, p. 104) that Hill's reasoning is a good example of inference to the best explanation.

Philosophers such as Wilfred Sellars (1973), Gilbert Harman (1973), and Ted Poston (2014) have argued that knowledge is justified by explanatory coherence: you are justified in believing that P if P is part of the best explanation of the evidence as determined by coherence with everything that you know. Thagard (1989) proposed a precise theory of explanatory coherence accompanied by a computational model, ECHO, which has been used to simulate numerous examples of scientific, medical, legal, and everyday inference (Thagard, 1992, 1999, 2000, 2004, 2012; Eliasmith and Thagard 1997; Nowak and Thagard 1992a, 1992b).

Box 1 presents principles of explanatory coherence. In the Zika case, the data (Principle E4) are the results of observations, for instance the Brazilian finding of a strong association between Zika virus infection and cases of microcephaly. The hypotheses are conjectures about what might be causing the data, for example that Zika virus causes microcephaly and other birth defects. Principle E2 says that hypotheses cohere with what they explain, so the hypothesis that Zika virus causes birth defects coheres with the evidence concerning increased microcephaly in Brazil. Hypotheses can be stacked up in complex causal networks, for example Zika virus causes birth defects because of biological

mechanisms of infection disrupting cell growth. In accord with Principle E1, the coherence relation is symmetrical: hypothesis and data cohere with each other. In contrast, the probability of a hypothesis given data is usually very different from the probability of data given evidence.

Principle E1. Symmetry. Explanatory coherence is a symmetric relation, unlike, say, conditional probability. That is, two propositions p and q cohere with each other equally.

Principle E2. Explanation. (a) A hypothesis coheres with what it explains, which can either be evidence or another hypothesis; (b) hypotheses that together explain some other proposition cohere with each other; and (c) the more hypotheses it takes to explain something, the lower the degree of coherence.

Principle E3. Analogy. Similar hypotheses that explain similar pieces of evidence cohere.

Principle E4. Data priority. Propositions that describe the results of observations have a degree of acceptability on their own.

Principle E5. Contradiction. Contradictory propositions are incoherent with each other.

Principle E6. Competition. If P and Q both explain a proposition, and if P and Q are not explanatorily connected, then P and Q are incoherent with each other. (P and Q are explanatorily connected if one explains the other or if together they explain something.)

Principle E7. Acceptance. The acceptability of a proposition in a system of propositions depends on its coherence with them.

Box 1. Principles of explanatory coherence, from Thagard (2006).

Principle E3 recognizes that analogy can contribute to coherence, for example when Darwin (1859) argued that one of the supports for his theory of evolution by natural selection was the analogy with artificial selection carried out by breeders (Thagard 1978). In the Zika case, epidemiologists note analogous explanations such as the causation of birth defects by the rubella virus.

Principles E5 and E6 establish incoherence relations between hypotheses that are flat-out contradictory or merely competing to explain the same data. The alternatives to the hypothesis that the Zika virus causes birth defects are that something else causes birth defects, or that the defects occur randomly.

Principles E1-E6 establish complex networks of data, explanations, and competing hypotheses at different levels. Principle E7 directs how to determine what to believe and what not to believe, based on how well a proposition (hypothesis or piece of evidence) fits with everything else. For example, the hypothesis that Zika virus causes birth defects should fit with all the data and outcompete alternative hypotheses. In a complex evidential situation it is difficult to determine the best fit of all the explanatory constraints. The computer program ECHO shows how to best satisfy these constraints.

ECHO Simulations

To determine overall coherence, the computer program ECHO uses a neural network algorithm for approximately maximizing coherence. ECHO represents each proposition by a unit, a simplified artificial neuron that is connected to other units by excitatory and

inhibitory links. As in real neurons, an excitatory link is one that enables one neuron to increase the firing of another, whereas an inhibitory link decreases firing. After cycles of excitation and inhibition, the firing rates (activations) of the units settle into stable patterns.

Zika Simulation

In the Zika example, we can represent the hypothesis that the virus causes defects by the unit ZIKA-CAUSES-DEFECTS, and the Brazilian association between virus and defects by a unit BRAZIL-ASSOCIATION. Then whenever principles E2 and E3 establish relations of coherence between two propositions, the units that represent the propositions get excitatory links between them. So ZIKA-CAUSES-DEFECTS and BRAZIL-ASSOCIATION have an excitatory link between them that is symmetric in accord with principle E1. Principle E4 is implemented by making an excitatory link between a special unit EVIDENCE and any unit such as BRAZIL-ASSOCIATION that represents a proposition based on observation. Principles E5 and E6, which establish incoherence between competing hypotheses, are implemented by inhibitory links between units: when two hypotheses are incoherent, e.g. ZIKA-CAUSES-DEFECTS vs. OTHER-CAUSE, then the units that represent them get an inhibitory link between them.

The acceptability of a unit is represented by its activation, corresponding roughly to the firing rate of a real neuron. Just as firing rates of neurons are determined by their excitatory and inhibitory neurons, the activation of units in ECHO are determined by their excitation and inhibition and the activation of the units to which they are connected. When the network settles (i.e. activations stabilize), the resulting activations (positive or negative) indicate whether the hypotheses and data represented by the units are accepted or rejected.

A test of the theory of explanatory coherence is whether examples such as the Zika virus can be plausibly modeled using the program ECHO.

The neural networks used by ECHO are not biologically plausible because single neuron-like units represent complex propositions such as that Zika virus causes birth defects, and because the excitatory and inhibitory links between units are symmetric. Thagard and Aubie (2008) showed how to translate ECHO networks into more biologically realistic networks with one-directional links between neurons in groups that collectively represent propositions. Techniques are now available for translating complex symbolic propositions into neural networks (Eliasmith and Thagard, 2001).

The input to ECHO for the Zika virus simulation consists of statements of what explains what, analogies, and evidence, shown in box 2. Fleshed out, the main evidence and hypotheses are:

E1. Infection is present during prenatal development.

E2. Rare microcephaly is associated with Zika. Reports of fetuses and infants with microcephaly who are born to women with brief periods of travel to countries with active Zika virus transmission are consistent with Zika virus being a rare exposure. The defect, congenital microcephaly, is rare, with a birth prevalence of approximately 6 cases per 10,000 liveborn infants, according to data from birth-defects surveillance systems in the United States.

E3. Zika virus is in brain tissue.

E4. A study during the outbreak in Brazil found a significant association between Zika virus infection and microcephaly. Eighty-eight pregnant women who had had an onset of rash in the previous 5 days were tested for Zika virus RNA. Among the 72

women who had positive tests, 42 underwent prenatal ultrasonography, and fetal abnormalities were observed in 12 (29%); none of the 16 women with negative tests had fetal abnormalities. The abnormalities that were observed on ultrasonography varied widely, and some findings lacked postnatal confirmation because the pregnancies were ongoing.

E5. A study on subjects in French Polynesia found a significant association between Zika virus infection and microcephaly.

E6. No results of an animal model with Zika virus infection during pregnancy and fetal effects have yet been published.

E7. Birth defects are associated with rubella virus.

E8. Animal models have shown that Zika virus is neurotropic, supporting biologic plausibility.

E9. Zika virus infects neural progenitor cells and produces cell death and abnormal growth.

H1. The Zika virus causes microcephaly.

H2. There is some other cause of microcephaly.

H3. Rubella causes birth defects.

H4. The Zika virus is neurotropic.

; EVIDENCE

(proposition 'E1 "infection-during-prenatal-development")

(proposition 'E2 "rare-microencephaly-with-zika")

(proposition 'E3 "zika-virus-in-brain-tissue")

(proposition 'E4 "Brazil-more-microencephaly-after-infection")

(proposition 'E5 "Polynesia-more-microencephaly-after-infection")

(proposition 'E6 "no-animal-models")

(proposition 'E7 "birth-defects-rubella")

(proposition 'E8 "animal-models-neurotropic")

(proposition 'E9 "Zika-produces-abnormal-growth")

(data '(E1 E2 E3 E4 E5 E6 E7 E8 E9))

; HYPOTHESES

(proposition 'H1 "zika-virus-causes-microencephaly")

(proposition 'H2 "other-cause")

(proposition 'H3 "rubella-causes-defects")

(proposition 'H4 "Zika-virus-is-neurotropic")

; EXPLANATIONS

(explain '(H1) 'E2)

(explain '(H1 E1 E3) 'E4)

(explain '(H1 E1 E3) 'E5)

(explain '(H2) 'E6)

(explain '(H3) 'E7)

(explain '(H4 E9) 'H1)

```
; ANALOGY
(analogous '(H1 H3) '(E4 E7))

; CONTRADICTION
(contradict 'H1 'H2)
```

Box 2. Input to ECHO for Zika simulation. The parentheses and quotation marks are artifacts of the implementation of ECHO in the programming language LISP.

Figure 1 provides a simplified picture of the causal network that ECHO turns into a neural network. When ECHO is run, all units begin with activation 0, and after 118 cycles of activation adjustment activations stabilize. ECHO accepts the hypothesis that the Zika virus causes birth defects while rejecting the alternative hypothesis of some other cause. The specific numbers for activation are not significant: what matters is whether the final activation is above 0, indicating acceptance, or below 0, indicating rejection. Hence explanatory coherence and the ECHO model explain how the conclusion that the Zika virus causes brain defects arises by inference to the best explanation.

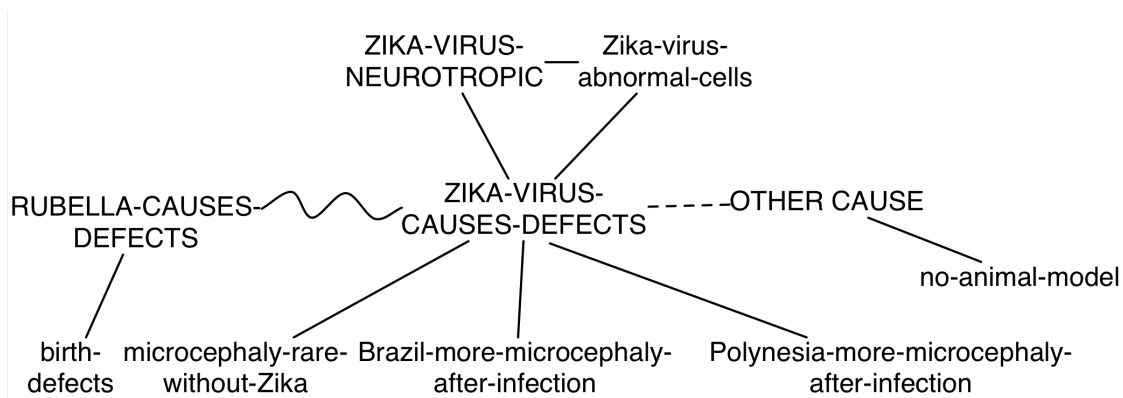


Figure 1. Causal and neural network for the Zika simulation, simplified. Solid lines indicate excitatory links based on coherence from explanation. The wavy line indicates an excitatory link based on coherence from analogy. The dotted line indicates an inhibitory link based on incoherence from contradiction. Hypotheses are shown in capital letters, and evidence in lower case.

Smoking/Cancer Simulations

One of the great public health accomplishments of epidemiology is the demonstration that tobacco smoking causes cancer and other diseases. Hill was one of the earlier researchers who found a statistical association between smoking and cancer (Doll and Hill, 1950), but the overall case that smoking causes cancer was made by the American Surgeon General (1964). This study used five “criteria” for establishing causal relationships based on statistical associations: consistency, strength, specificity, temporal relationship, and coherence. All of these are included in Hill’s viewpoints, and the four viewpoints not included in the report (biological gradient, experiment, plausibility, analogy) might be absorbed into coherence. Subsequent reports to the Surgeon General (e.g. 2010, 2014)

made an even stronger case that smoking causes many diseases, including various forms of cancer, cardiovascular and pulmonary problems, and reproductive effects.

The Proctor Simulation

Proctor (2013) reviewed the history of the discovery of the connection between cigarettes and lung cancer. He says that four lines of evidence converged to establish cigarette smoking as the leading cause of lung cancer: population studies, animal experimentation, cellular pathology, and cancer-causing chemicals in cigarette smoke. Population studies repeatedly found that smokers of cigarettes were far more likely to contract lung cancer than non-smokers. Animal experiments found that applying tobacco products to rabbits and mice led to cancer. Cellular pathology research showed that smokers experienced damage to lung cells. Finally, chemical research determined that cigarette smoke contains many carcinogens.

How these lines of evidence converged to back the conclusion that smoking causes cancer is a matter of explanatory coherence, as shown in figure 2. The hypothesis that smoking causes cancer explains why smokers and tobacco-applied animals are more likely to get cancer. The studies about cellular pathology and cancer-causing chemicals sketch the mechanisms which explain how smoking causes cancer, through the effects of carcinogenic chemicals on lung cells.

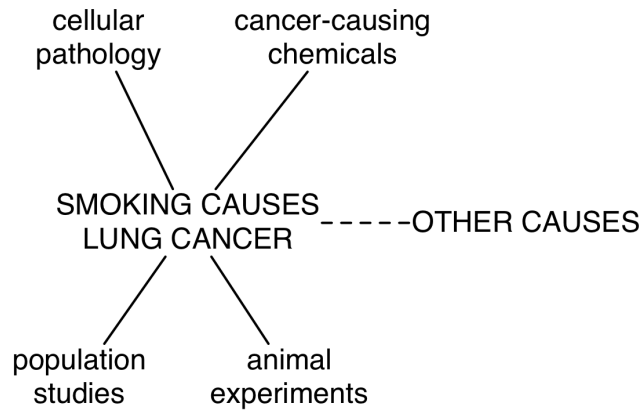


Figure 2. Explanatory coherence of the conclusion that smoking causes cancer, following Proctor (2013). Conventions are the same as in figure 1.

The input for the ECHO simulation of this case is shown in box 3. After 95 cycles of activation adjustment, the neural network produced by this input settles, with positive activation of the unit smoking-causes-cancer indicating acceptance of this hypothesis, rejecting other-causes.

```

; EVIDENCE

(proposition 'population "population studies associate smoking and lung cancer")

(proposition 'animal "animal experimentation associate tobacco and cancer")

(proposition 'cellular "cellular pathology finds that smoking damages cells")

(proposition 'chemicals "there are cancer causing chemicals in smoke")

(data '(population animal cellular chemicals))

; HYPOTHESES

(proposition 'smoking-causes-cancer "tobacco smoking causes cancer")

(proposition 'other-cause "cancer has other causes")

```

; EXPLANATIONS

(explain '(smoking-causes-cancer) 'population)

(explain '(smoking-causes-cancer) 'animal)

(explain '(chemicals cellular) 'smoking-causes-cancer)

; CONTRADICTION

(contradict 'smoking-causes-cancer 'other-cause)

Box 3. Input to ECHO for Proctor simulation.

John Snow's Communication Theory of Cholera

John Snow (1855) is considered one of the originators of epidemiology because of his arguments in the 1840s and 1850s that cholera is caused by communication via excremental evacuations. Tulodziecki (2011) has provided a thorough analysis of Snow's arguments showing the explanatory power of his theory compared to the prevalent view that cholera results from miasma - bad air due to decaying matter.

This analysis translates into explanatory coherence as shown in figure 3. The superior explanatory power of the communication theory comes primarily from its ability to explain numerous phenomena that the miasma theory cannot. For example, communication of "cholera poison" via evacuation explains why cholera usually starts with digestive problems and why people with bad hygiene got cholera more often than people with good hygiene. In addition, the communication theory explains why physicians (who were careful about washing hands and not eating while visiting the sick) were less likely to get cholera than ordinary people. In contrast, on the miasma theory physicians would be

more likely to get the disease via miasmatic effluvia from the sick people they visited (Tulodziecki, 2011, p. 312). Figure 3 displays the superior explanatory power of the communication theory.

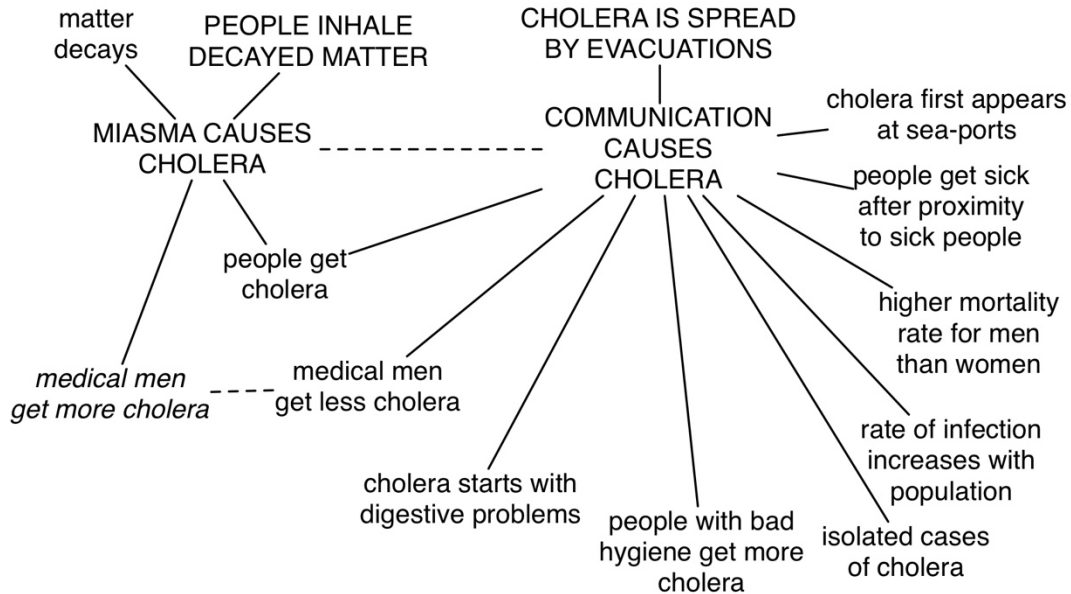


Figure 3. Explanatory coherence of the conclusion that smoking causes cancer, following Tulodziecki (2011). Conventions are the same as in figure 1.

The relations between propositions shown in figure 3 generate the input to ECHO shown in box 4. In less than a second, with 142 cycles of activation updating, ECHO settles with the acceptance of H2 (communication causes cholera) and the rejection of H1 (miasma causes cholera).

; EVIDENCE

(proposition 'E1 "people get cholera")

(proposition 'E2 "medical men get less cholera")

(proposition 'NE2 "medical men get more cholera") ; negative evidence

(proposition 'E3 "cholera starts with digestive problems")

(proposition 'E4 "people with bad hygiene get more cholera")

(proposition 'E5 "there are isolated cases of cholera")

(proposition 'E6 "the rate of infection increases with population")

(proposition 'E7 "higher mortality rater for men than women")

(proposition 'E8 "people get sick after proximity to sick people")

(proposition 'E9 "cholera first appears at sea-ports")

(proposition 'E10 "matter decays")

(data '(E1 E2 E3 E4 E5 E6 E7 E8 E9 E10))

; HYPOTHESES

(proposition 'H1 "miasma causes cholera")

(proposition 'H2 "communication causes cholera")

(proposition 'H3 "people inhale decayed matter")

(proposition 'H4 "cholera is spread by evacuations");

CONTRADICTIONS

(contradict 'E2 'NE2)

; EXPLANATIONS

(explain '(H1) 'E1)

(explain '(H1) 'NE2)

(explain '(H3 E10) 'H1)

(explain '(H2) 'E1)

(explain '(H2) 'E2)

(explain '(H2) 'E3)

(explain '(H2) 'E4)

(explain '(H2) 'E5)

(explain '(H2) 'E6)

(explain '(H2) 'E7)

(explain '(H2) 'E8)

(explain '(H2) 'E9)

(explain '(H4) 'H2)

Box 4. Input to ECHO for Snow simulation.

Connections with Epidemiological Standards for Causality

Thagard (1998, 1999) described how an explanatory coherence account of reasoning concerning the causation of stomach ulcers by *H. pylori* bacteria fit well with the criteria of causality advocated by Evans (1993). More generally, table 1 maps the relation between principles of explanatory coherence and additional ways that epidemiologists have characterized determination of causality, due to Hill (1965) and Shepard et al. (1994). Rasmussen et al. (2016) connected their inference that Zika virus causes birth defects with Shepard's criteria, which map onto Hill's viewpoints and Evans's criteria as shown by table

1. These viewpoints and criteria are not necessary and sufficient conditions for causality, but serve as standards of evaluation.

<i>Hill's viewpoints</i>	<i>Evans's criteria</i>	<i>Shepard's criteria</i>	<i>Principles of explanatory coherence</i>
1. temporality	4. temporally, disease follows exposure	1. exposure to the agent	E2. explanation
2. strength of association	1. prevalence 2. exposure 3. incidence	2. epidemiology findings	E2. explanation E4. data priority
3. consistency of association		2. consistent findings 3. delineation of cases	E4. data priority E5. contradiction E7. acceptance
4. specificity of association		4. rarity of exposure and defect	E2. explanation E5. contradiction
5. biological gradient	5. spectrum of host responses 6. measurable host response		E2. explanation

6. experiment	7. experimental reproduction 8. elimination -> reduction 9. prevention of host response	5. experimental animals 7. experimental system for agent	E2, explanation
7. plausibility	10. biological sense	6. biological sense	E2. explanation (higher order) E3. analogy
8. coherence	10. biological sense		E5. contradiction E7. acceptance
9. analogy		6. biological sense	E3. analogy

Table 1. Mapping of standards for causation onto explanatory coherence principles.

According to principle E2, the hypothesis that an environmental condition causes a disease coheres with the evidence that it explains. But if the disease precedes the condition, then there is no causation, hence no explanation, and hence no coherence. The rule that causes come before effects is not true *a priori*, because it is conceivable that time travel could enable a future event such as getting into a time machine in the year 3000 to cause an earlier event such as arriving in a place in the year 1000. But there has never been an observed case of the future causing the past, so the temporality rule is a reasonable way of

dismissing cases of backward causation. Thus, failure of temporality blocks some applications of E2 where an explanatory cause happens after the event explained.

Principle E2 is also key to understanding Hill's aspects 2-6. The hypothesis that an environmental condition causes a disease can explain why there are associations between the condition and the disease that are strong, consistent, and specific. These explanations therefore enhance the coherence and acceptability of the causal hypothesis. Alternative hypotheses such as chance and the occurrence of some unknown factor cannot furnish comparable explanations.

Similarly, the hypothesis that a condition causes a disease explains why there is a biological gradient such that more of the condition causes more disease. When experimental evidence of a successful intervention is available, it also increases explanatory coherence because the hypothesis that a condition causes a disease explains why changing the condition changes the disease. Alternative hypotheses concerning chance and unknown factors cannot mount similar explanations, and hence gain no support from E2.

Principles E4, E5, E6, and E7 are also relevant to understanding why aspects 2-6 help to indicate causality. E4 (data priority) ensures that evidence collected by observations and experiments gets a degree of priority over hypotheses. E4 does not imply that data are always taken at face value, because observations and experimental results can be mistaken; but it does help to ensure that evidence will have a greater contribution to coherence than hypotheses that may be fanciful. Aspects E5 and E6 set up a battle between the hypothesis that a condition causes a disease and its alternatives, either because the alternative is flat-

out contradictory (e.g. cause vs. chance) or merely competitive in cases where multiple causes might be operating.

One weakness of Hill's method is that he gave no indication of how all the aspects can be combined into an overall inference that a condition causes a disease. Principle E7 asserts that maximizing coherence is the key to evaluating a causal hypothesis and other beliefs. E7 does not say how to maximize coherence, but the construal of coherence as constraint satisfaction and the availability of various algorithms for approximately maximizing coherence (including the neural network algorithm used by ECHO) takes care of this problem. From the explanatory coherence perspective, medical researchers should use Hill's aspects to establish non-rigid constraints that need to be coherently satisfied. The ECHO program then determines best overall fit.

Hill's aspects 7-8 of plausibility and coherence also fall under the theory of explanatory coherence. They urge that a causal hypothesis should fit with general biological and medical knowledge. Principle E5 (contradiction) handles the most extreme case where a new hypothesis contradicts what is generally believed. As Hill noted, contradicting orthodoxy does not always provide the grounds for rejecting a hypothesis because the orthodoxy may be wrong.

Another source of fit with biological and medical knowledge comes from the availability of higher-order explanations. The hypothesis that smoking causes cancer became more plausible once mechanisms were understood for how the ingredients in smoke irritate tissues and encourage the development of mutations that lead to growth of tumors. Also relevant is E7 (coherence) which encourages an overall fit with all knowledge, not just the narrow domain in which the causal hypothesis operates.

Hill's ninth aspect, analogy, encourages that a causal hypothesis be analogous to other kinds of explanations used in biology and medicine. Analogy is taken care of by Principle E3 of explanatory coherence, and analogical reasoning can also be understood as a kind of parallel constraint satisfaction (Holyoak and Thagard, 1995).

Objections and Replies

Our explanatory coherence account of epidemiological reasoning generates worries.

1. *Explanation.* The theory of explanatory coherence is empty without an account of the nature of explanation.

Reply. Explanation follows different patterns in different fields (see Poston (2014), pp 70-80). For example, the hypothesis that the Zika virus causes birth defects explains the evidence that in Brazil the virus is associated with microcephaly because of a partially understood mechanism where the parts are viruses and neurons, the main interaction is infection, and the regular changes are defective neurons and brains. In contrast, Snow lacked a detailed understanding of the mechanism of cholera infection, which needed the germ theory developed by Pasteur in the 1860s.

2. *Causality.* The hypothesis that a condition causes a disease is meaningless without an understanding of causality.

Reply: Hill acknowledged the difficulty of analyzing causality, and no definition has ever survived for long. But causality can be characterized using the method of 3-analysis, which is based on a new theory of concepts that describes how they combine exemplars (typical examples), typical features, and explanations (Blouw, Solodkin, Thagard, and Eliasmith 2016). There are many familiar exemplars of causes, such as

pushes, pulls, motions, collisions, actions, and diseases whose effects are symptoms (Thagard, forthcoming).

The typical features of causality include:

1. Causes happen before effects.
2. Causes operate in sensory-motor-sensory patterns, e.g. when you see and feel a bike not moving, move the pedals, then see and feel the bike move.
3. Cause and effects sometimes yield regularities, for example that hitting your finger with a hammer always hurts.
4. Statistical dependencies occur, with causes increasing the probabilities of effects.
5. Manipulations and interventions lead from causes to effects.

None of these typical features is a necessary or sufficient condition of causality, but matching a lot of them suggests that cause/effect relations have been identified. There are obvious relations between these 5 typical features of causality and the epidemiological criteria in table 3.

Such relations provide explanations of why things happen and how they can be changed. Causality in particular cases is explained by the presence of underlying mechanisms connecting cause and effect. Before concluding that C causes E, you need to consider alternative explanations such as that E has a different cause, or that C and E are both caused by something else, or that E occurs randomly. We cannot directly observe causal relations, but can infer that they exist as part of the best explanation of systematic observations, in accord with explanatory coherence.

3. Inferences against causality. Epidemiology sometimes leads to the rejection of causal hypotheses, not just their acceptance.

Reply. Explanatory coherence understands the rejection of causal hypotheses as resulting from the acceptance of alternatives concerning other causes, chance, bias, or confounding. For example, the popular hypothesis that stomach ulcers are caused by excess acidity was rejected because of the explanatory coherence of the new hypothesis that bacteria cause ulcers (Thagard 1999). More recently, the hypothesis that multiple sclerosis is caused by compromised flow of blood in veins to the head has been largely rejected for many reasons. Factors include the shoddiness of initial studies used to support the hypothesis, the conflicts of interest of the investigators who proposed it, the failure of more careful studies to find that balloon angioplasty reduces the symptoms of multiple sclerosis, and the finding that the correlation between venous insufficiency and multiple sclerosis is dubious (Traboulee et al., 2014; Kruger, Patel, and Lee, 2015). All of these factors could be incorporated into an explanatory coherence analysis and ECHO model of rejection of the hypothesis that venous insufficiency causes multiple sclerosis.

Conclusion

Our paper addresses causal reasoning in epidemiology, but explanatory coherence extends to other kinds of medical inference. Thagard and Larocque (2018) model mental health assessment as inference to the best explanation performed by ECHO. Other forms of diagnosis can also be construed as abductive inference, i.e. inference to the best explanation (Josephson and Josephson, 1994; Peng and Reggia, 1990), in ways that naturally translate into explanatory coherence. For example, physicians who diagnose lung cancer in patients can take into account all of (1) evidence explained by the diagnosis such as coughing and test results, (2) history of heavy smoking which explains why the patient got sick, and (3)

alternative explanations such as emphysema. Finally, reasoning in evidence-based medicine concerning the effectiveness of medical treatments can be understood as inference to the best explanation (Thagard, 2010), but detailed analysis in terms of explanatory coherence remains to be developed.

More narrowly, we have provided an epistemological interpretation and justification for Bradford Hill's influential recommendations about how to infer causality in epidemiology. Our interpretation is based on the epistemology of explanatory coherentism, fleshed out using a detailed theory of explanatory coherence. We have shown the applicability of this approach by applying the ECHO computational model for calculating explanatory coherence to three important cases of epidemiological reasoning, concerning the Zika virus, smoking, and cholera. The result is a deeper understanding of the nature of medical inference concerning the causes of disease.

References

- Bechtel, W. (2008). *Mental mechanisms: Philosophical perspectives on cognitive neuroscience*. New York: Routledge.
- Blouw, P., Solodkin, E., Thagard, P., & Eliasmith, C. (2016). Concepts as semantic pointers: A framework and computational model. *Cognitive Science*, 40, 1128-1162.
- Broadbent, A. (2017). Philosophy of epidemiology. In J. A. Marcum (Ed.), *The Bloomsbury companion to contemporary philosophy of medicine* (pp. 93-112). London: Bloomsbury.

- Craver, C. F., & Darden, L. (2013). *In search of mechanisms: Discoveries across the life sciences*. Chicago: University of Chicago Press.
- Dammann, O. (2018). Hill's heuristics and explanatory coherentism in epidemiology. *American Journal of Epidemiology*, 187(1), 1-6.
- Darwin, C. (1859). *The origin of species*. London: Murray.
- Doll, R., & Hill, A. B. (1950). Smoking and carcinoma of the lung. *British Medical Journal*, 2(4682), 739-748.
- Eliasmith, C., & Thagard, P. (1997). Waves, particles, and explanatory coherence. *British Journal for the Philosophy of Science*, 48, 1-19.
- Eliasmith, C., & Thagard, P. (2001). Integrating structure and meaning: A distributed model of analogical mapping. *Cognitive Science*, 25, 245-286.
- Evans, A. S. (1993). *Causation and disease: A chronological journey*. New York: Plenum.
- Harman, G. (1973). *Thought*. Princeton: Princeton University Press.
- Hennekens, C. H., & Buring, J. E. (1987). *Epidemiology in medicine*. Boston: Little, Brown.
- Hill, A. B. (1965). The environment and disease: Association or causation? *Proceedings of the Royal Society of Medicine*, 58, 295-300.
- Holyoak, K. J., & Thagard, P. (1995). *Mental leaps: Analogy in creative thought*. Cambridge, MA: MIT Press/Bradford Books.
- Josephson, J. R., & Josephson, S. G. (Eds.). (1994). *Abductive inference: Computation, philosophy, technology*. Cambridge: Cambridge University Press.
- Kugler, N., Patel, P. J., & Lee, C. J. (2015). Chronic cerebrospinal venous insufficiency in multiple sclerosis: a failed concept. *Vascular Specialist International*, 31(1), 11-14.

- Kuhn, T. S. (1970). *The structure of scientific revolutions* (2nd ed.). Chicago: University of Chicago Press.
- Lakatos, I. (1970). Falsification and the methodology of scientific research programs. In I. Lakatos & A. Musgrave (Eds.), *Criticism and the growth of knowledge* (pp. 91-195). Cambridge: Cambridge University Press.
- Mill, J. S. (1970). *A system of logic* (8 ed.). London: Longman.
- Morabia, A. (2013). Hume, Mill, Hill, and the sui generis epidemiologic approach to causal inference. *American Journal of Epidemiology*, *178*(10), 1526-1532.
- Nowak, G., & Thagard, P. (1992a). Copernicus, Ptolemy, and explanatory coherence. In R. Giere (Ed.), *Cognitive models of science*, (Vol. 15, pp. 274-309). Minneapolis: University of Minnesota Press.
- Nowak, G., & Thagard, P. (1992b). Newton, Descartes, and explanatory coherence. In R. Duschl & R. Hamilton (Eds.), *Philosophy of Science, Cognitive Psychology and Educational Theory and Practice*. (pp. 69-115). Albany: SUNY Press.
- Peng, Y., & Reggia, J. (1990). *Abductive inference models for diagnostic problem solving*. New York: Springer-Verlag.
- Poston, T. (2014). *Reason and explanation: A defense of explanatory coherentism*. London: Palgrave Macmillan.
- Proctor, R. N. (2012). The history of the discovery of the cigarette–lung cancer link: Evidentiary traditions, corporate denial, global toll. *Tobacco Control*, *21*(2), 87-91.
- Rasmussen, S. A., Jamieson, D. J., Honein, M. A., & Petersen, L. R. (2016). Zika virus and birth defects—reviewing the evidence for causality. *New England Journal of Medicine*, *2016*(374), 1981-1987.

- Sellars, W. (1973). Givenness and explanatory coherence. *Journal of Philosophy*, 70, 612-624.
- Shepard, T. H. (1994). "Proof" of human teratogenicity. *Teratology*, 50(2), 97-98.
- Surgeon General. (1964). *Smoking and Health: Report of the Advisory Committee to the Surgeon General of the Public Health Service*. Washington, DC.
- Surgeon General. (2010). *How Tobacco Smoke Causes Disease: The Biology and Behavioral Basis for Smoking-Attributable Disease*. Washington, DC.
- Surgeon General. (2014). *The Health Consequences of Smoking - 50 Years of Progress*. Washington, DC.
- Snow, J. (1855). *On the mode of communication of cholera*. London: John Churchill.
- Thagard, P. (1978). The best explanation: criteria for theory choice. *Journal of Philosophy*, 75, 76-92.
- Thagard, P. (1989). Explanatory coherence. *Behavioral and Brain Sciences*, 12, 435-467.
- Thagard, P. (1992). *Conceptual revolutions*. Princeton: Princeton University Press.
- Thagard, P. (1998). Ulcers and bacteria I: Discovery and acceptance. *Studies in History and Philosophy of Science. Part C. Studies in History and Philosophy of Biological and Biomedical Sciences*, 29, 107-136.
- Thagard, P. (1999). *How scientists explain disease*. Princeton: Princeton University Press.
- Thagard, P. (2000). *Coherence in thought and action*. Cambridge, MA: MIT Press.
- Thagard, P. (2004). Causal inference in legal decision making: Explanatory coherence vs. Bayesian networks. *Applied Artificial Intelligence*, 18, 231-249.
- Thagard, P. (2006). Evaluating explanations in science, law, and everyday life. *Current directions in psychological science*, 15, 141-145.

- Thagard, P. (2010). *The brain and the meaning of life*. Princeton, NJ: Princeton University Press.
- Thagard, P. (2012). *The cognitive science of science: Explanation, discovery, and conceptual change*. Cambridge, MA: MIT Press.
- Thagard, P. (forthcoming). *Natural philosophy: From social brains to knowledge, reality, morality, and beauty*. Oxford: Oxford University Press.
- Thagard, P., & Aubie, B. (2008). Emotional consciousness: A neural model of how cognitive appraisal and somatic perception interact to produce qualitative experience. *Consciousness and Cognition*, 17, 811-834.
- Thagard, P., & Larocque, L. (2018). Mental health assessment: Inference, explanation, and coherence. *Journal of Evaluation in Clinical Practice*, 24, 649-654.
- Traboulee, A. L., Knox, K. B., Machan, L., Zhao, Y., Yee, I., Rauscher, A., . . . Kopriva, D. (2014). Prevalence of extracranial venous narrowing on catheter venography in people with multiple sclerosis, their siblings, and unrelated healthy controls: A blinded, case-control study. *The Lancet*, 383(9912), 138-145.
- Tulodziecki, D. (2011). A case study in explanatory power: John Snow's conclusions about the pathology and transmission of cholera. *Studies in History and Philosophy of Biological and Biomedical Sciences*, 42(3), 306-316.