

# A coherentist approach to probabilistic causal assessment

The Bridges, Rutgers University, September 2015

Barbara Osimani & Roland Poellinger  
Munich Center for Mathematical Philosophy  
Ludwig-Maximilians-Universität Munich

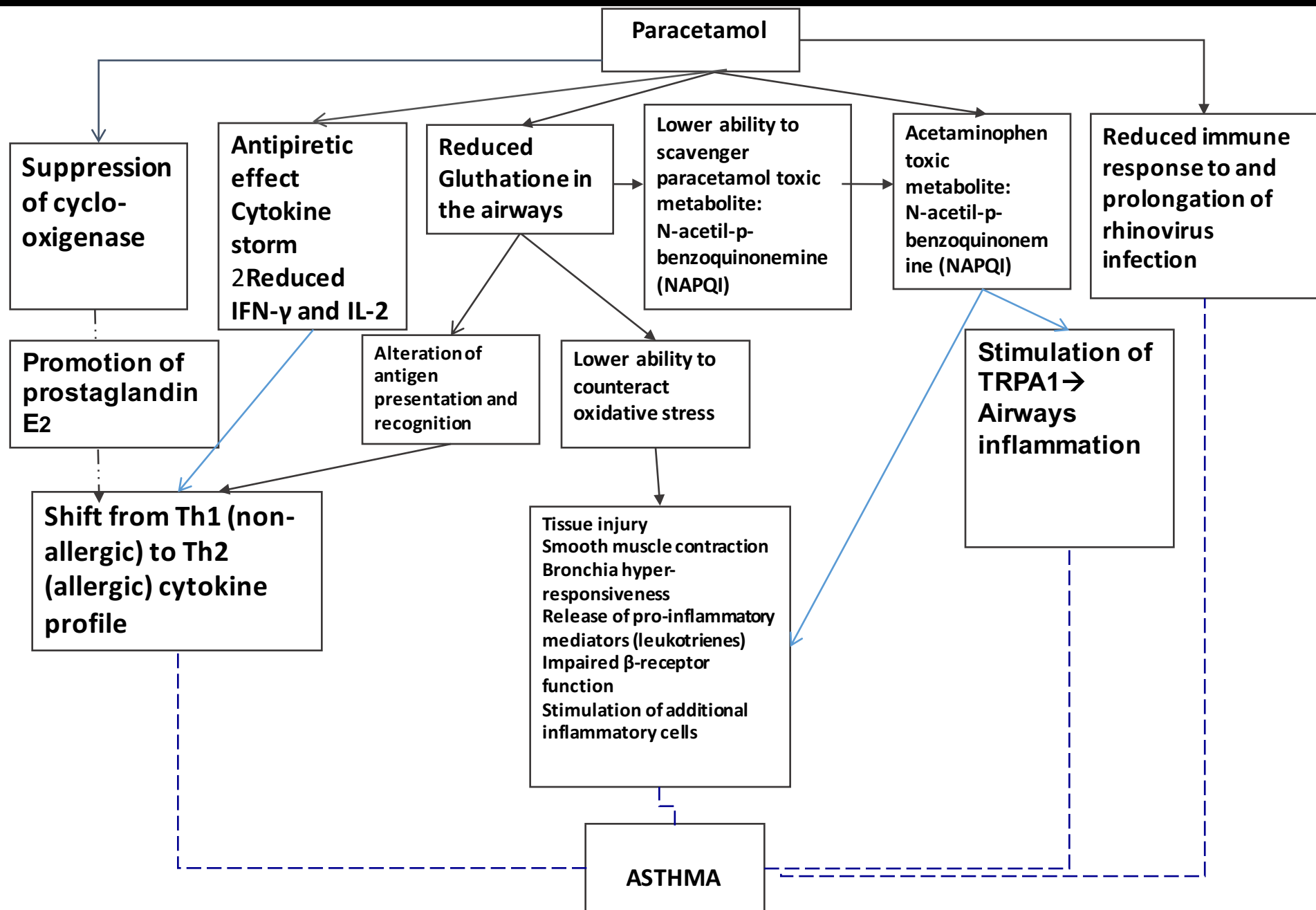
# Causal inference in pharmacology

# Outline

- Case study: does paracetamol cause asthma?
- Causal assessment of harms: the two paradigms
- (Towards) a coherentist approach to probabilistic causal assessment
- Conclusions and outlook

# Case study: does Paracetamol cause asthma?

- 1) Strength of the association: observational and experimental studies (RCTs without placebo); (PD;  $\Delta$ )
- 2) Robustness of association across geography, culture and age (R);
- 3) Dose-response relationship between acetaminophen exposure and asthma (D-R)
- 3) Relationship between asthma epidemic and per-capita sales of acetaminophen across countries (ecologic studies);
- 5) Plausible molecular mechanism (M)
- 6) Coincidence of time trends in acetaminophen use and asthma increase (C)
- 5) Lack of other equally strong causal explanations (NAA)



**Two interrelated questions:**

- 1. Does paracetamol cause asthma?**
- 2. What is the better candidate to explain the asthma epidemic?**



VS





Eneli et al. 2005, Allmers et al. 2009,  
Johnson and Ownby, 2011;  
Karimi et al., 2006, Wickens et al. 2011,  
Chang et al. 2011  
paracetamol-asthma relationship may be  
explained by

- 1) reverse causation,
- 2) confounding by indication or
- 3) preference for acetaminophen rather  
than ibuprofen in children at risk for  
asthma

Other authors are less sceptical but  
nevertheless equally require placebo-  
controlled trials to establish causation  
(Holgate, 2011; Henderson and Shaheen,  
2013).



Martinez-Gimeno and García-Marcos 2013,:  
“apart from tobacco smoke exposure, no other genetic or environmental factors, including genes, allergens, infections and bacterial substances, has shown the stubborn and consistent association with wheezing disorders prevalence as acetaminophen has done”

McBride (2011): **burden of proof reversal**  
“At present I need further studies not to prove that acetaminophen is dangerous but, rather, that it is safe.



# *The Debated association between Paracetamol and Asthma*

Patients with asthma and acetaminophen users might differ from corresponding reference groups in many major aspects. A randomized placebo-controlled trial is required to address the above controversy.

Chang K.C., C.C. Leung, C.M. Tam, F.Y. Kong (2011) Acetaminophen and Asthma: Spurious Association? Am. J. Respir. Crit. Care Med, 183: 1570-1571.

In response to Kwok Chiu Chang and colleagues, we reiterate that causality cannot be established from the ISAAC findings, owing to several potential biases that might confound the association, including but not limited to recall bias, misclassification bias, and confounding by indication, as discussed in detail in the article (1). However, when the study findings are considered together with other available data, there is substantive evidence that acetaminophen use in childhood may be an important risk factor for the development and/or maintenance of asthma, and that its widespread increasing use over the last 30 years may have contributed to the rising prevalence of asthma in different countries worldwide (2, 3).

Beasley R.W., T.O. Clayton, J. Crane et al. (2008) ISAAC Phase Three Study Group. Association between acetaminophen use in infancy and childhood, and risk of asthma, rhinoconjunctivitis, and eczema in children

Such methodological dissent concerning the best course of action among scholars hides differing epistemic views.

How to model this?

What view fares better when dealing with harms? Why?

Evidence hierarchies :

Best evidence → lexicographic decision rule (internal validity)

1. Meta-analyses of Randomized Clinical Trials/  
Systematic Reviews
2. Single Randomized Clinical Trials
3. Meta-analyses of observational studies
4. Comparative studies which are not randomized  
(e.g. cohort or case-control studies)
5. Pathophysiologic mechanisms  
(Basic science)
6. Expert judgment

# Hypothesis testing and *modus tollens*

1. Conjecture:  $H$  =

“Vitamin C has some effect on Flu”

→ Experimental hypothesis:  $H_0 \rightarrow \neg\Delta$

2. Test and observe result:  $\Delta$

3. Infer:  $\neg H_0$  (reject  $H_0$ )

→ Fisher disjunction & abduction

For the result to be at all meaningful, it is essential that the observed difference between groups is due to the treatment **and only to it**.

Which in turn explains the insistence on the exclusion of confounders.

Confounding by indication, confounding by contraindication, selection and self-selection bias...

→ the more likely a method is to be able to exclude confounders and systematic/random errors the more reliable is the inference we base on it

→→ the higher is the method ranked in the hierarchy (the better the evidence);

Evidence hierarchies are grounded on the assumption that if you have a study which has the capacity to eliminate more confounders than others, then the former should trump the latter.

# Putative roles of randomization

## 1) **SYSTEMATIC ERROR**

### a. **Confounders** (ontological)

Control would do (although obviously not for unknown confounders)

### b. **(self) selection bias** (due to experiment itself);

Blinding does the job

## 2) **RANDOM ERROR**

Single randomization = insufficient;

→ repeated randomization is needed = unpractical and unethical

→ Larger samples and meta-analyses

Quality level	Current definition (Balshem et al. 2011)	Previous definition (Guyatt et al. 2008)
High	We are <b>very confident</b> that the true effect lies close to that of the estimate of the effect	Further research is very unlikely to <b><u>change</u> our confidence in the estimate of effect</b>
Moderate	We are <b>moderately confident</b> in the effect estimate: The true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different	Further research is <b>likely</b> to have an important impact on our confidence in the estimate of effect and may change the estimate
Low	Our <b>confidence</b> in the effect estimate is <b>limited</b> : The <b>true effect may be substantially different</b> from the estimate of the effect	Further research is <b>very likely</b> to have an important impact on our confidence in the estimate of effect and is likely to change the estimate
Very low	We have <b>very little confidence</b> in the effect estimate: The true effect is likely to be substantially different from the estimate of effect	Any estimate of effect is <b>very uncertain</b>



# Bradford Hill criteria for causal assessment

1. Consistency of data within population / across populations;
2. Strength of the association;
3. Relationship in time;
4. Biological gradient;
5. Specificity;
6. Coherence of evidence;
7. Biological plausibility;
8. Reasoning by analogy;
9. Experimental evidence.

# Bradford Hill criteria for causal assessment

1. “None of my nine viewpoints can bring indisputable evidence for or against the cause-and-effect hypothesis and none can be required as a *sine qua non*. What they can do, with greater or less strength, is to help us make up our minds in the fundamental question – *is there **any other way of explaining the set of facts before us, is there any other equally, or more, likely than cause and effect?***”
2. “No formal tests of significance can answer those questions. Such tests can, and should, remind us of the effects that the play of chance can create, and they will instruct us on the likely magnitude of those effects. Beyond that, they contribute nothing to the proof of our hypothesis”.

# Efficacy vs. safety assessment

why standards should not be the same

1. *Integration of prior knowledge (theory, historical data, knowledge of same-class molecules) Price et al. 2014, Osimani 2013a);*
2. *High default prior for an undefined risk (Osimani 2013a);*
3. *Higher risk for false negatives than for false positives in the case of harm;*
4. *Cumulative learning and the virtues of probabilistic vs. categorical causal assessment (Osimani 2013b;*
5. *Risk-benefit balance and the precautionary principle (Rudén & Hansson, 2008; Osimani 2012, 2007);*
6. *Impartiality (conflicting interests among parties) (Teira, 2011);*
7. *Causal structure (Thompson, 2011, Joffe, 2011);*

# Bayesian methods for design and analysis of safety trials<sup>†</sup>

Karen L. Price,<sup>a\*</sup> H. Amy Xia,<sup>b</sup> Mani Lakshminarayanan,<sup>c</sup> David Madigan,<sup>d</sup> David Manner,<sup>a</sup> John Scott,<sup>e</sup> James D. Stamey,<sup>f</sup> and Laura Thompson<sup>g</sup>

Safety assessment is essential throughout medical product development. There has been increased awareness of the importance of safety trials recently, in part due to recent US Food and Drug Administration guidance related to thorough assessment of cardiovascular risk in the treatment of type 2 diabetes. Bayesian methods provide great promise for improving the conduct of safety trials. In this paper, the safety subteam of the Drug Information Association Bayesian Scientific Working Group evaluates challenges associated with current methods for designing and analyzing safety trials and provides an overview of several suggested Bayesian opportunities that may increase efficiency of safety trials along with relevant case examples. Copyright © 2013 John Wiley & Sons, Ltd.

*Pharmaceut. Statist.* **2014**, *13* 13–24

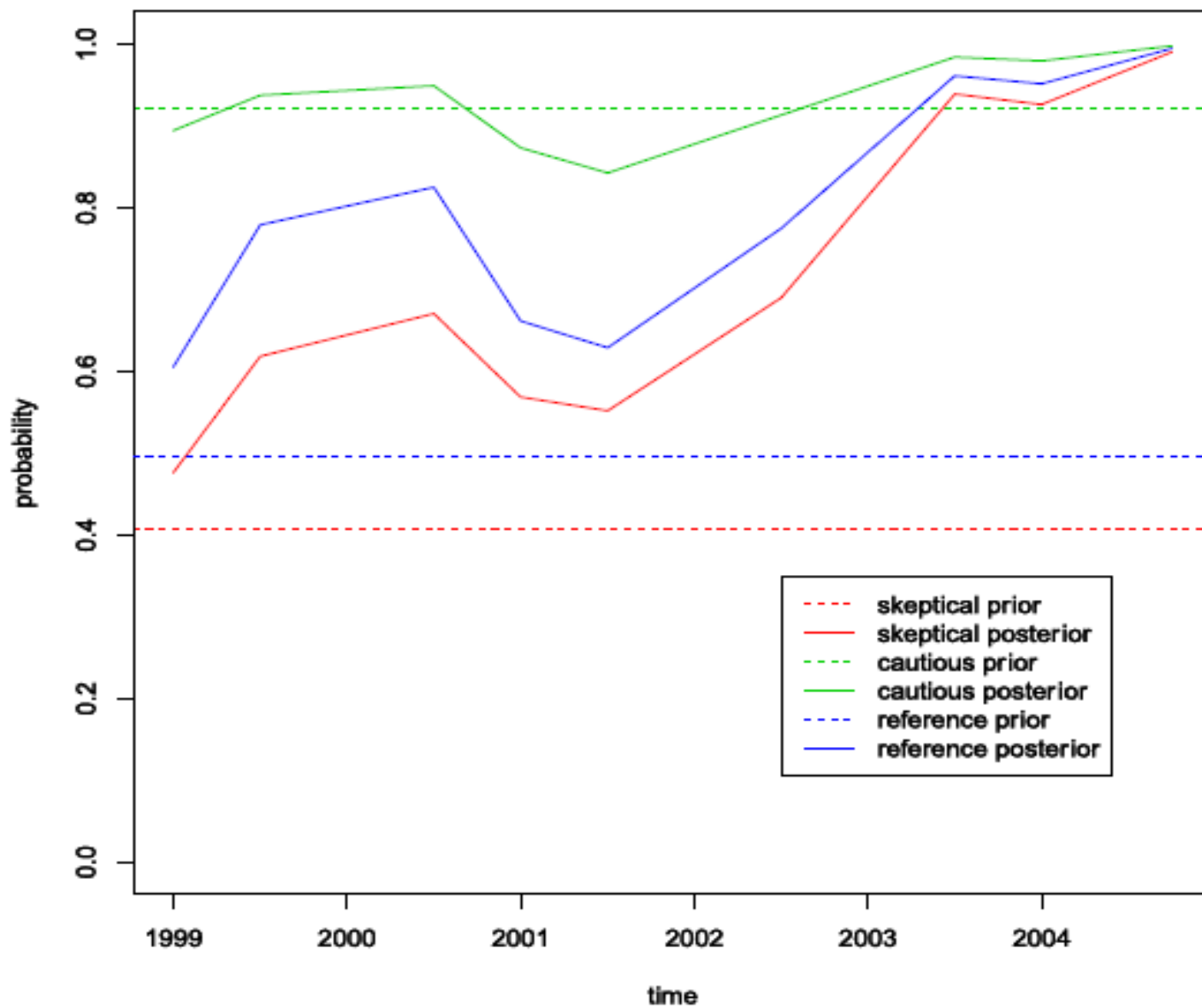
### 3.6. Continuous monitoring of events

Statistical issues posed by monitoring safety in clinical trials are considerably different from monitoring efficacy. In the presence of a safety concern, timely assessment of new data from an ongoing trial and establishment of a fast response system are important to protect patients participating in the trial. In this con-

#### 4.6. Surveillance case example

*Pharmaceut. Statist.* 2014, 13 13–24

Some safety issues from medical products are only observed long term, perhaps in a postmarket study. In postmarket surveillance of medical products, often the goal of a study is to estimate a quantity of interest (such as an adverse event rate or survival percentage) with a certain precision, and to show with high probability that it is above or below a prespecified threshold. Often, formal hypothesis testing is not used, and a type 1 error of a false safety issue is arguably less important than not detecting a real safety issue (type 2 error). Murray *et al.* [48] illustrate details of planning a postmarket surveillance study by using a Bayesian adaptive design. Their goal is to estimate the survival percentile for a med-



**Figure 1.** Posterior probability that the true relative risk exceeds 1.1 over time.

Coherentism (within a bayesian framework) as a valid alternative to the classical hypothesis testing approach in various respects:

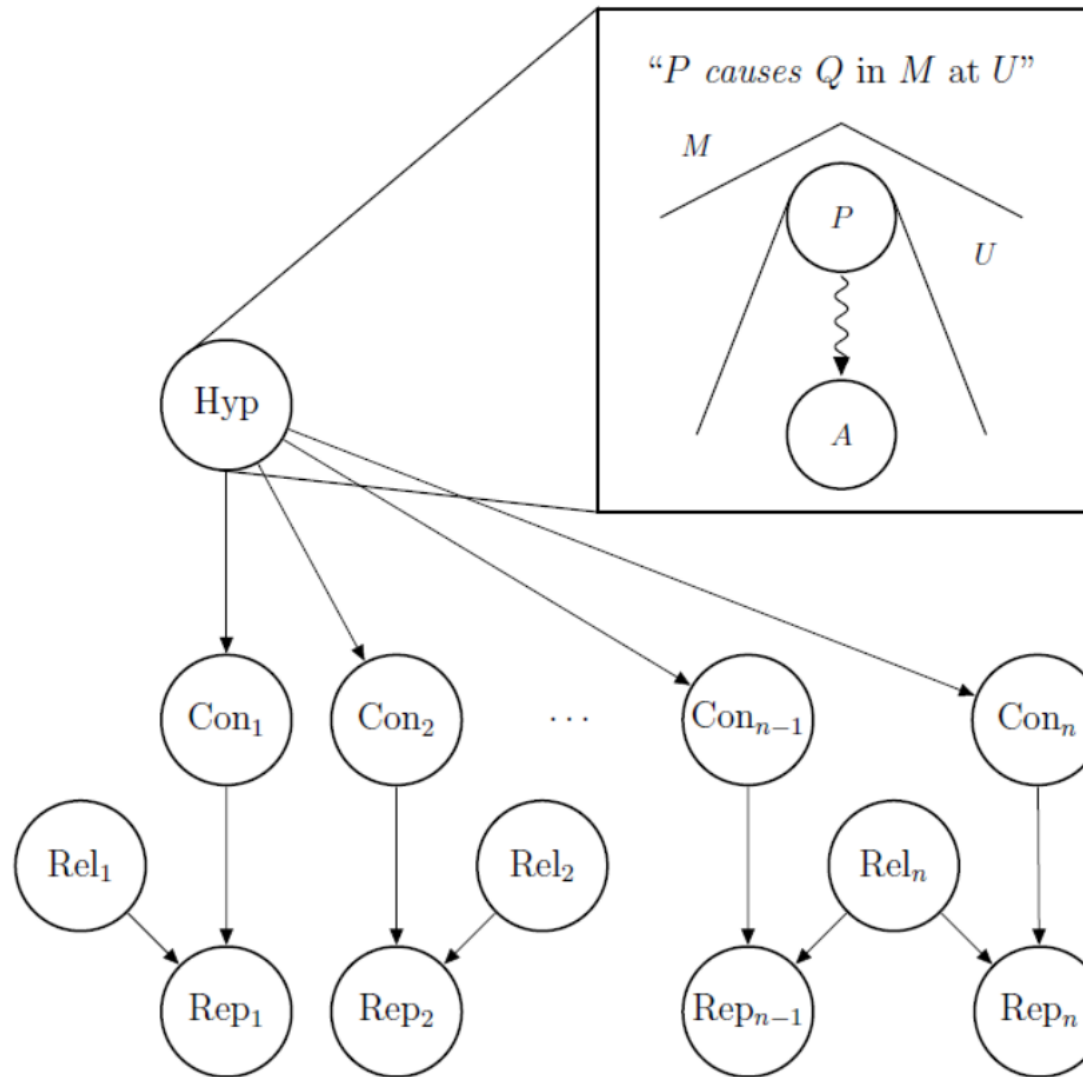
1. heuristic: → illustrate the structure of the problem;
2. “Pedagogical” → help professional to articulate their intuitions on this kind of situations  
→ raise awareness among authorities and guideline compilers about alternative epistemic paradigms;
3. Foundational: → justification of causal inference

→ evidential nets as unifying “inference engines” which flesh out the structure of the inference problem and allows to incorporate different kinds of (inconclusive) evidence (also for the purpose of “interim justification”)



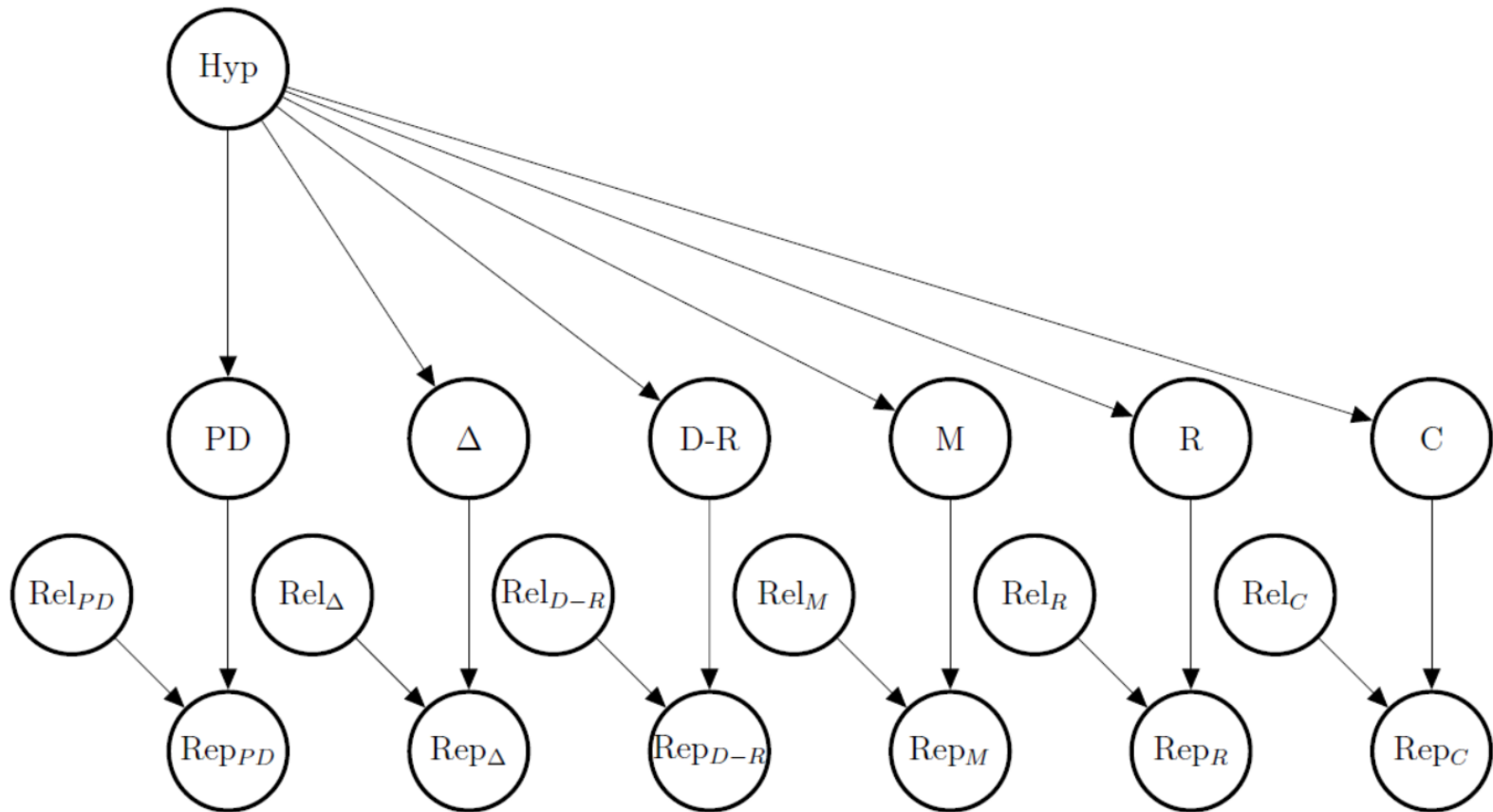
# Probabilistic causal assessment through networks of belief propagation

Relating the hypothesis and its observable consequences

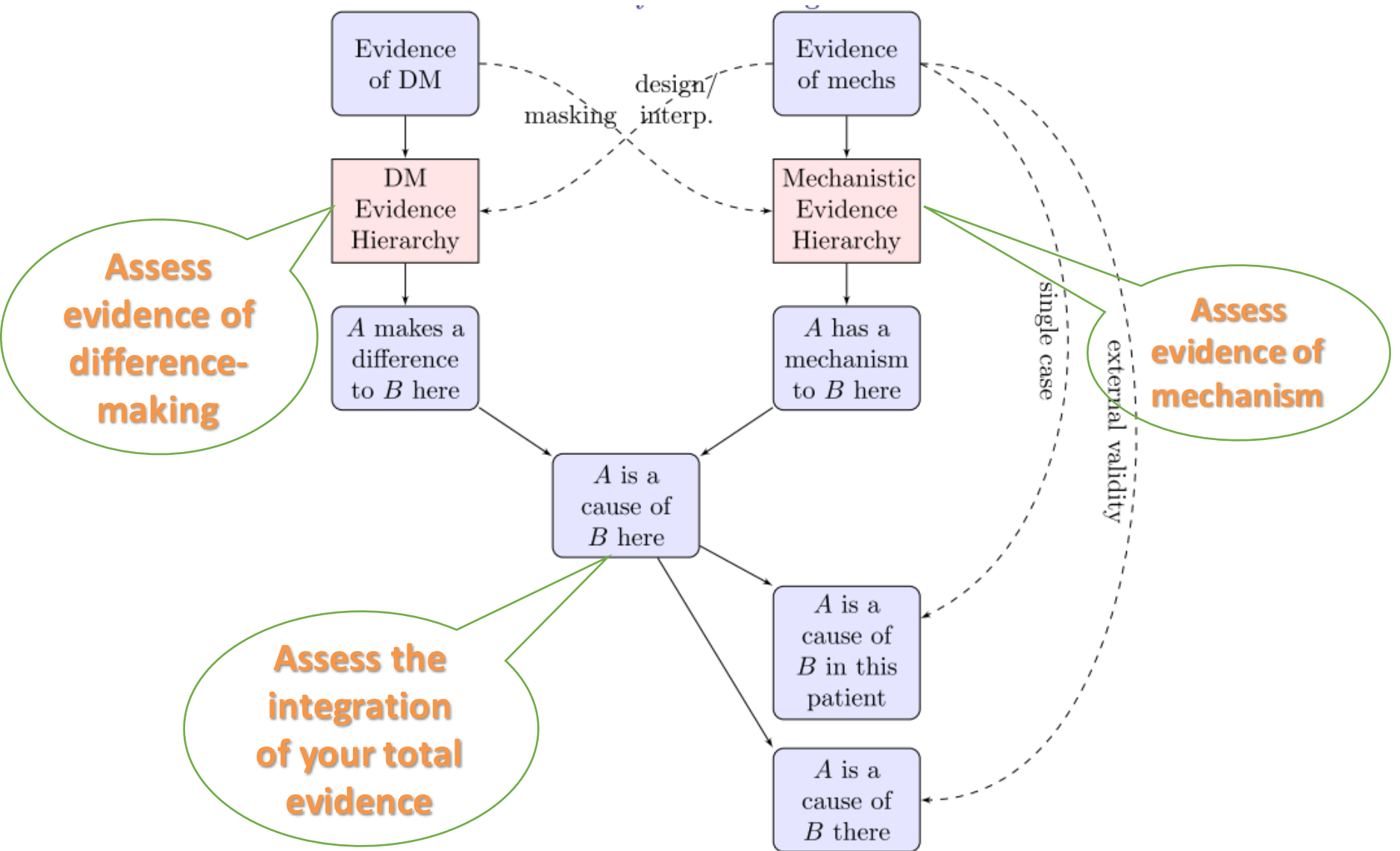


# Probabilistic causal assessment through networks of belief propagation

The case study, revisited



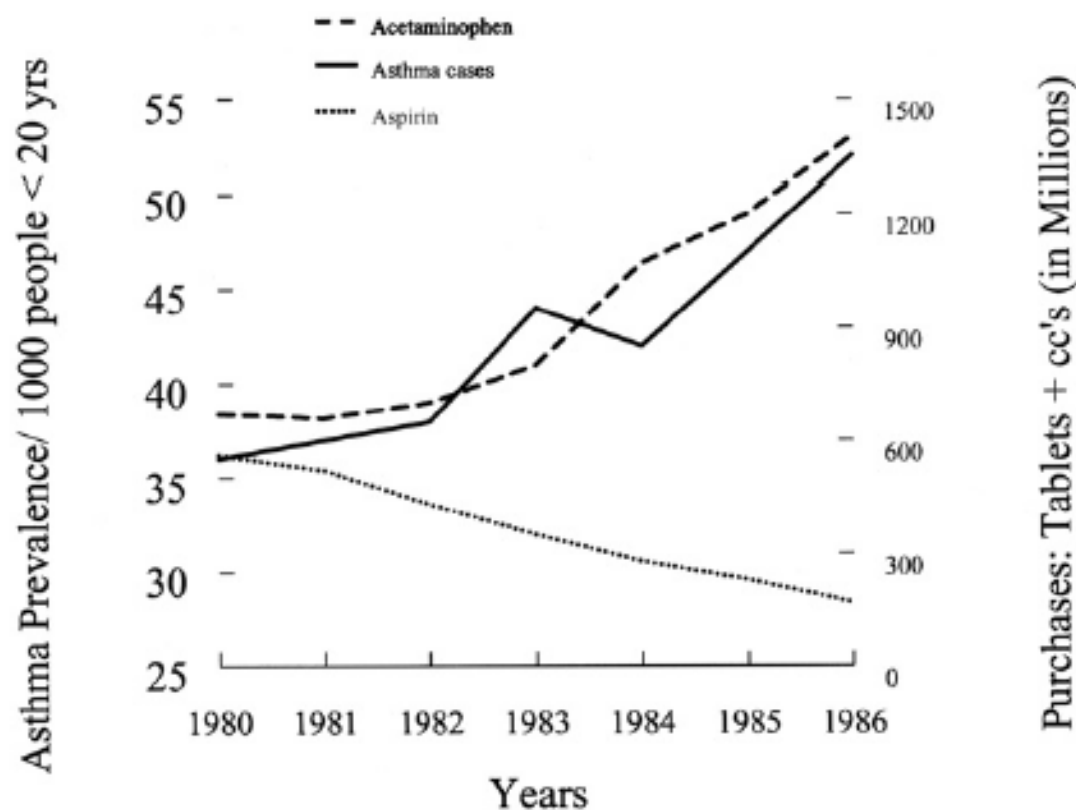
- **Epistemological value** (explain nature signals)
- **Methodological value** (best vs total evidence)
- **Heuristic value** (philosophical debate: Worral 2010, Cartwright 2007, 2010, Howick 2011, 2013, Williamson&Russo 2010, Illari et al. 2013)



Clarke B., D. Gillies, P. Illari, F. Russo, J. Williamson (2013) Mechanisms and the Evidence Hierarchy. *Topoi* (Special Issue: Evidence and Causality in the Sciences). October 2014, Volume 33, (2), pp 339-360

# Hypothesis: decreased use of pediatric aspirin has contributed to the increasing prevalence of childhood asthma

Arthur E Varner, MD\*,†; William W Busse, MD\*; and Robert F Lemanske, Jr, MD†,‡



*Ann Allergy Asthma Immunol* 1998;81:347–351.

# Explanatory hypotheses for asthma epidemic

- 1) increased exposure to outdoor and indoor **pollutants**;
- 2) decreased exposure to bacteria and childhood illnesses during infancy (the “**hygiene hypothesis**”);
- 3) cytokine imbalance as a reaction to **environmental allergens** in early childhood leading to lifelong T-helper type 2 (allergic) dominance over T-helper type 1 (nonallergic) reactions, thus increasing the risk for atopic disease;
- 4) changes in **diet** and oxidant intake;
- 5) increased **obesity** incidence and prevalence;

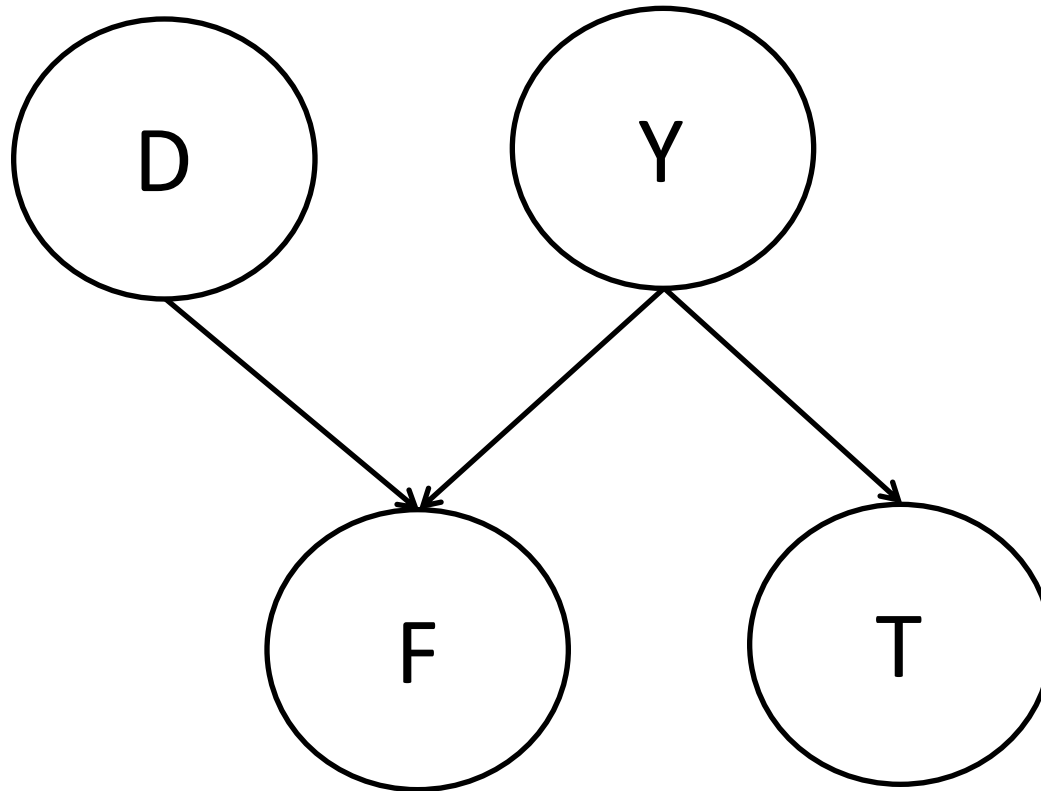
Eneli et al., 2005; Seaton et al. 1994, Shaheen et al. 2000. ETC

### **Diet** hypothesis:

- very complex one to prove because diet is difficult to measure; particularly it is difficult to identify the combined and independent effects of the different nutrients (Eder et al. 2006).
- the same element may have contrasting effects on the same outcome, e.g. selenium which is an antioxidant but may also upregulate immune responses typical of allergic asthma.
- aggregate measures of food consumption constitute an indirect index (or proxy) of actual nutritional antioxidant intake.

### **Hygiene** hypothesis:

- scarce consistency between the time trends of other allergic diseases (such as hay fever) and asthma (Platts-Mills et al. 2005).
- If true than one should see birth order effect: children born later on in the sequence of births should be less exposed to risk of asthma, but such tendency has not been observed.

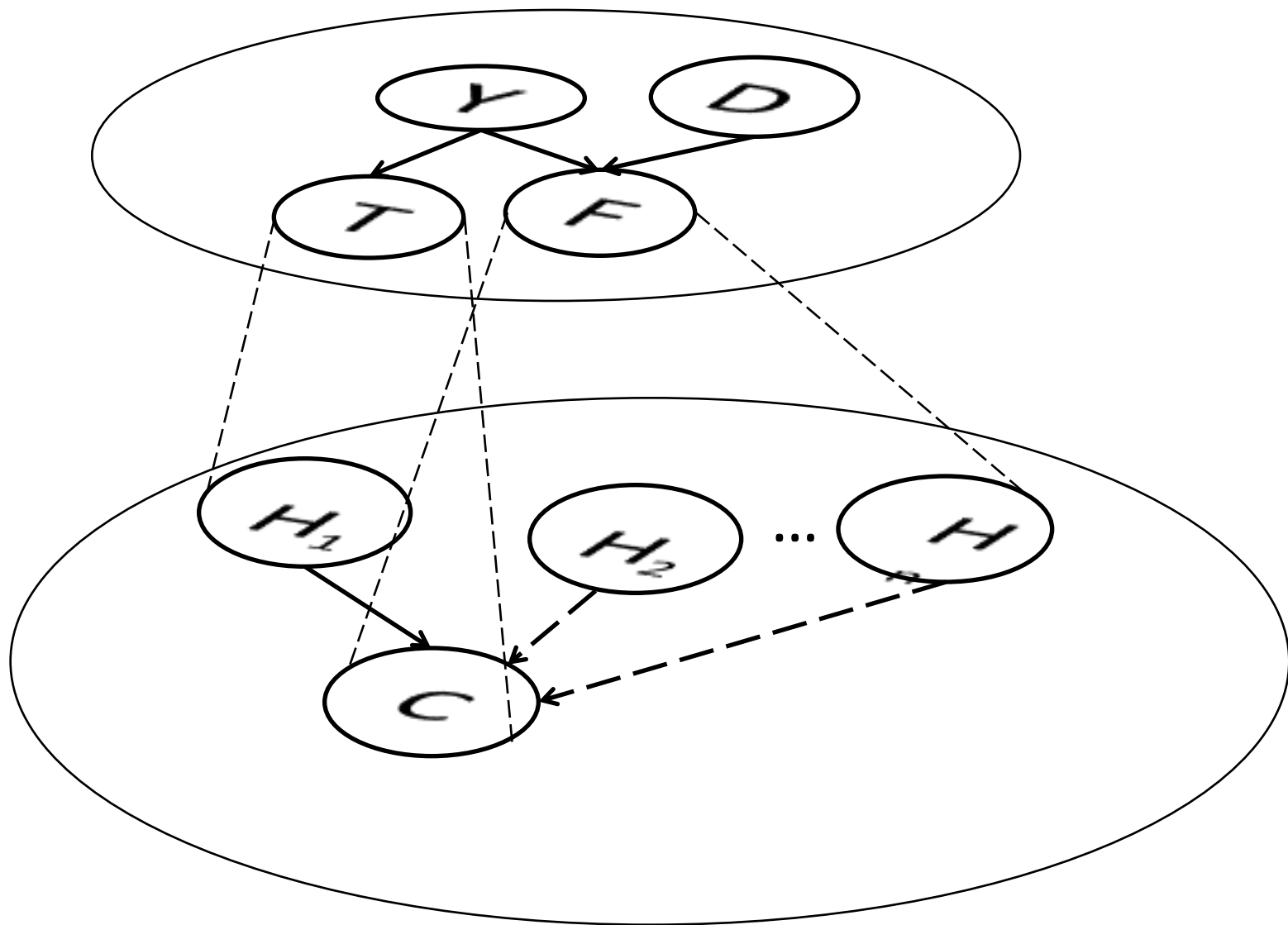


Dawid, Hartmann, Sprenger (2015)



## Meta-induction the no-alternatives argument

1. Coincidence of time-trends between asthma epidemic and paracetamol use → possible candidate for explanation of epidemic;
2. No other equally strong candidates → NAA
3. → support for causal claim on a meta-level



# Outlook

- How to use various kinds of evidential relationships (explanation, relevance, logical, semantic ...) in the same network?
- How to determine independencies between qualitatively diverse consequences of the hypothesis under consideration?
- How to mathematically relate coherence to evidential support?
- How should causal and evidential graphs be nested?

At the lower level:

- How to measure study reliability?
- How to measure evidential relevance (of different kinds)?
- ...

# Thankyou!

---



**Alexander von Humboldt**  
Stiftung/Foundation

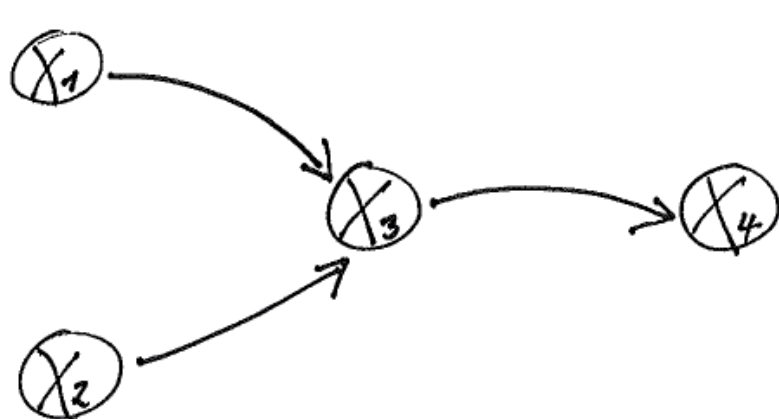


European Research Council

supported by 

---

## In Bayes net terms: parameters, graphs, and causes



$$\begin{aligned}
 X_1 &= f_1(u_1) \\
 X_2 &= f_2(u_2) \\
 X_3 &= f_3(X_1, X_2, u_3) \\
 X_4 &= f_4(X_3, u_4)
 \end{aligned}$$

causal interpretation	parameters ( $P$ )	graphs ( $G$ )
event types	random variables $V$	nodes
causal mechanisms	det. functions	family of edges
exceptions/cet. par. cond.	error terms/world $U$	–
Causal Markov Condition	Markov compatibility ( $P$ and $G$ )	
intervening (hypothetically) [to check for causal efficacy]	setting parameters	deleting edges

# Sticking point 1 (of 3): Modeling feedback

Cf. Clarke, Leuridan, and Williamson (2013)

*Cycles are everywhere in the sciences. They are particularly prevalent in the biomedical and biological sciences. Examples include metabolic cycles (such as Krebs' cycle), organismal life cycles (such as the malaria-causing organisms of the genus Plasmodium), homeostatic pathways (such as blood glucose regulation) and pathological processes. (§ 3)*

Idea:

Go from cycles to time-indexed Bayes net lattices (transition networks)

## Sticking point 1 (of 3): Modeling feedback

Cf. Clarke, Leuridan, and Williamson (2013)

*Cycles are everywhere in the sciences. They are particularly prevalent in the biomedical and biological sciences. Examples include metabolic cycles (such as Krebs' cycle), organismal life cycles (such as the malaria-causing organisms of the genus Plasmodium), homeostatic pathways (such as blood glucose regulation) and pathological processes. (§ 3)*

### Idea:

Go from cycles to time-indexed Bayes net lattices (transition networks)

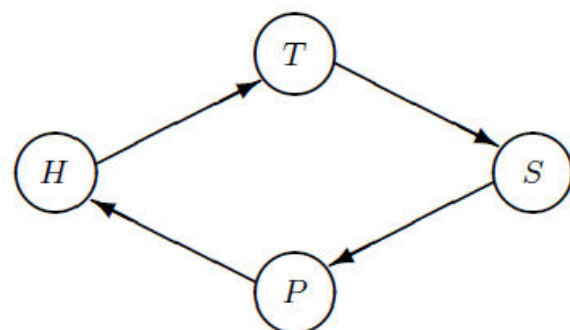
# Sticking point 1 (of 3): Modeling feedback

Cf. Clarke, Leuridan, and Williamson (2013)

*Cycles are everywhere in the sciences. They are particularly prevalent in the biomedical and biological sciences. Examples include metabolic cycles (such as Krebs' cycle), organismal life cycles (such as the malaria-causing organisms of the genus Plasmodium), homeostatic pathways (such as blood glucose regulation) and pathological processes. (§ 3)*

## Idea:

Go from cycles to time-indexed Bayes net lattices (transition networks)





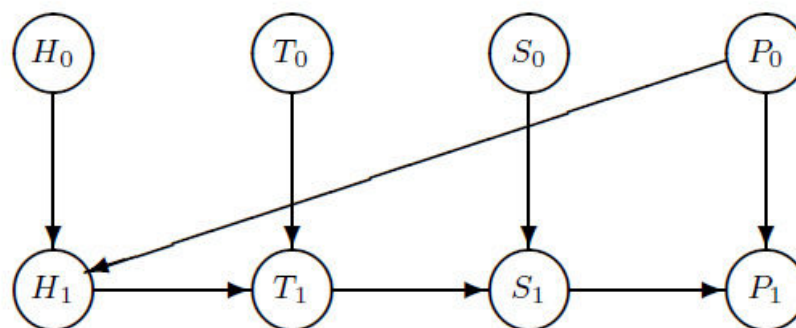
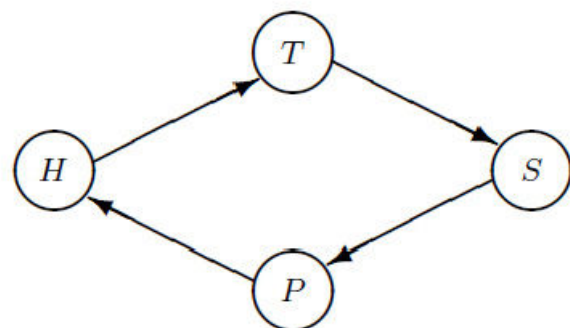
# Sticking point 1 (of 3): Modeling feedback

Cf. Clarke, Leuridan, and Williamson (2013)

*Cycles are everywhere in the sciences. They are particularly prevalent in the biomedical and biological sciences. Examples include metabolic cycles (such as Krebs' cycle), organismal life cycles (such as the malaria-causing organisms of the genus Plasmodium), homeostatic pathways (such as blood glucose regulation) and pathological processes. (§ 3)*

## Idea:

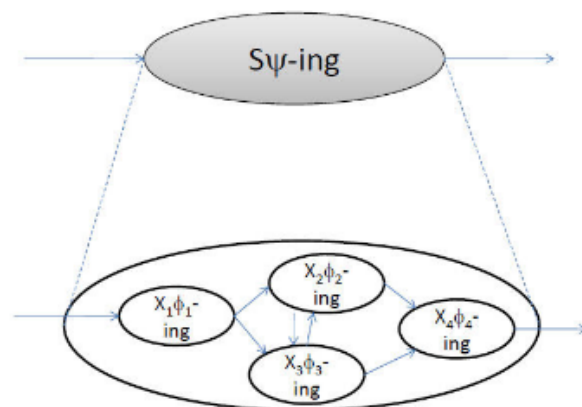
Go from cycles to time-indexed Bayes net lattices (transition networks)



## Sticking point 2 (of 3): Reference to mechanisms

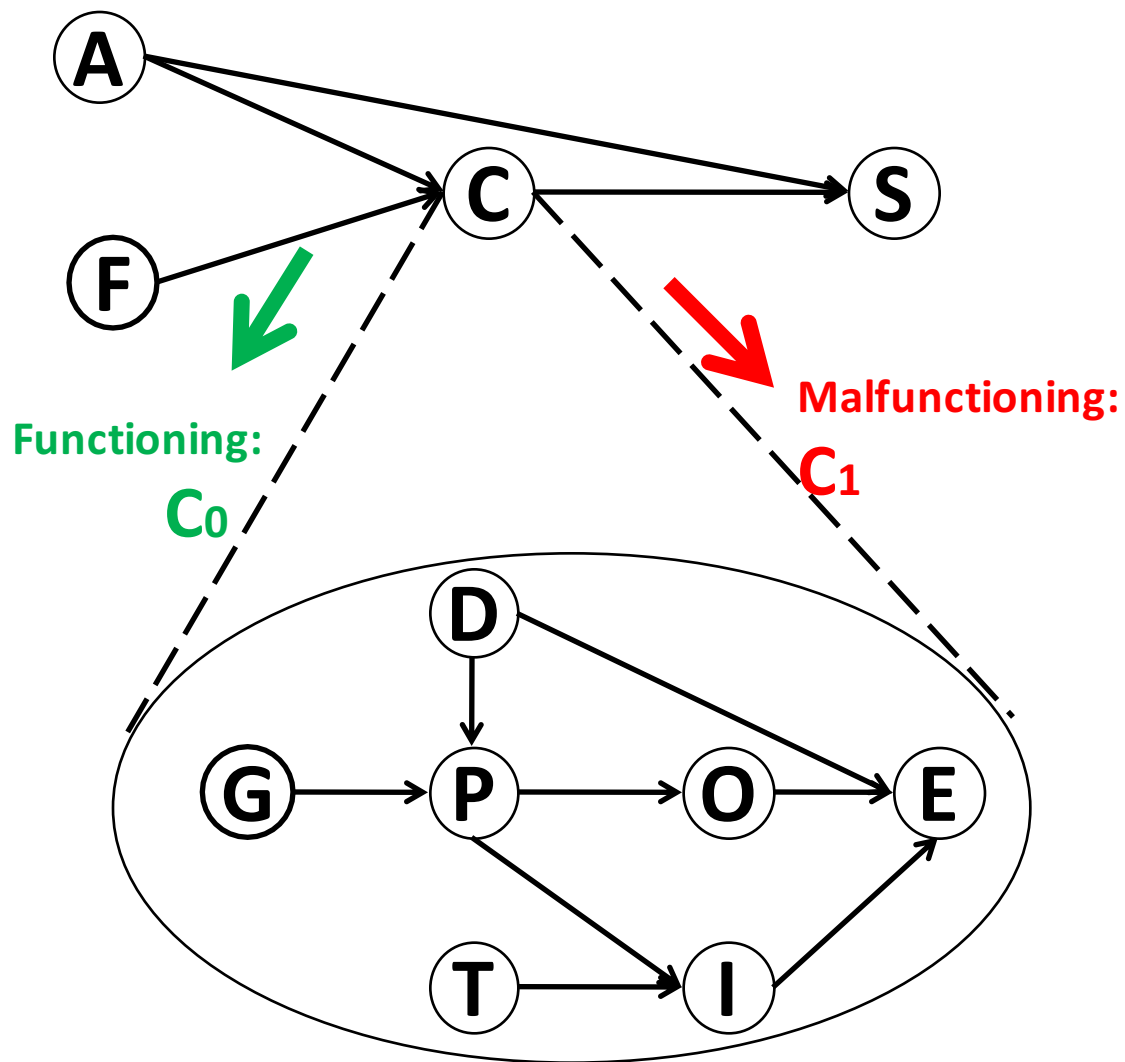
Cf. Clarke, Leuridan, and Williamson (2013) and Williamson (2009)

One might strive for (i) causal explanation by reference to mechanisms and for (ii) separation of top-level phenomena and bottom-level mechanisms:



But: Causal talk goes across levels (and disciplines), it works even without mechanisms (causation by omission).

⇒ Bayes nets are capable of expressing cross-level mechanisms (if variables are kept distinct, e.g., by flattening nets), integrating mechanistic and probabilistic knowledge, and expressing omissions.



*Adapted from: Clarke B., Leuridan B., Williamson J. (Forthcoming) Modelling Mechanisms with causal cycles.*

## Sticking point 3 (of 3): Expressing causal interaction

Weinberg (2007) is dissatisfied with the expressive power of causal graphs:

*The theory of directed acyclic graphs (DAGs), as extensively developed by Pearl [...] is producing growing pains for the field, even as it clarifies how we think about sampling biases and confounder adjustment in statistical models for causal relationships. [...] Most diseases are caused by multiple factors acting together and often through distinct pathways that can lead to a common final phenotype. Teasing apart the causal choreography will remain a prize worth the struggle; and the prize seems more attainable than ever, thanks to the rich array of molecular tools that are newly available to us.*

**Her idea:** Define structures to express different types of causal interaction and effect modification (where contributions are not additive).

## Sticking point 3 (of 3): Expressing causal interaction

Weinberg (2007) is dissatisfied with the expressive power of causal graphs:

*The theory of directed acyclic graphs (DAGs), as extensively developed by Pearl [...] is producing growing pains for the field, even as it clarifies how we think about sampling biases and confounder adjustment in statistical models for causal relationships. [...] Most diseases are caused by multiple factors acting together and often through distinct pathways that can lead to a common final phenotype. Teasing apart the causal choreography will remain a prize worth the struggle; and the prize seems more attainable than ever, thanks to the rich array of molecular tools that are newly available to us.*

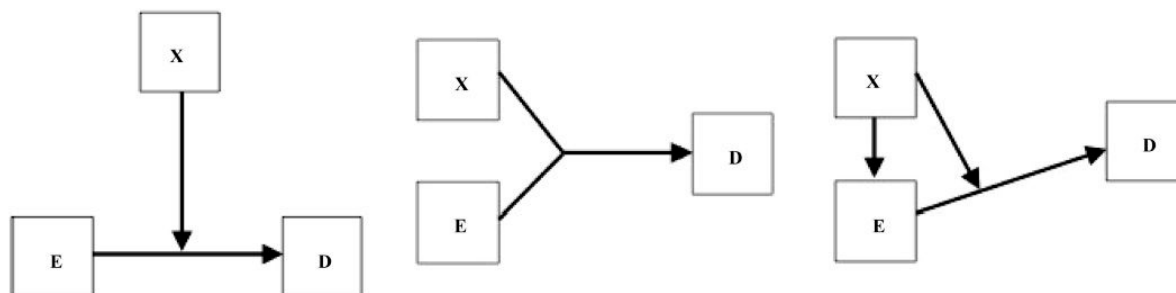
**Her idea:** Define structures to express different types of causal interaction and effect modification (where contributions are not additive).

## Sticking point 3 (of 3): Expressing causal interaction

Weinberg (2007) is dissatisfied with the expressive power of causal graphs:

*The theory of directed acyclic graphs (DAGs), as extensively developed by Pearl [...] is producing growing pains for the field, even as it clarifies how we think about sampling biases and confounder adjustment in statistical models for causal relationships. [...] Most diseases are caused by multiple factors acting together and often through distinct pathways that can lead to a common final phenotype. Teasing apart the causal choreography will remain a prize worth the struggle; and the prize seems more attainable than ever, thanks to the rich array of molecular tools that are newly available to us.*

**Her idea:** Define structures to express different types of causal interaction and effect modification (where contributions are not additive).



## Sticking point 3 (of 3): Expressing causal interaction

Weinberg (2007) is dissatisfied with the expressive power of causal graphs:

*The theory of directed acyclic graphs (DAGs), as extensively developed by Pearl [...] is producing growing pains for the field, even as it clarifies how we think about sampling biases and confounder adjustment in statistical models for causal relationships. [...] Most diseases are caused by multiple factors acting together and often through distinct pathways that can lead to a common final phenotype. Teasing apart the causal choreography will remain a prize worth the struggle; and the prize seems more attainable than ever, thanks to the rich array of molecular tools that are newly available to us.*

**Her idea:** Define structures to express different types of causal interaction and effect modification (where contributions are not additive).

**But:** This only addresses the graphical part of Bayes nets. Of course, Pearl's causal models are capable of encoding all kinds of interactions between independent factors —  $D$ 's value is defined as

$$D = f_D(X, E),$$

where the mechanism  $f_D$  is any nonlinear fct. (incl. case analysis etc.).

# Probabilistic causal assessment through networks of belief propagation

## Hypotheses and evidence in *networks of belief propagation*

In purely evidential networks, nodes of different types (representing theories, hypotheses, evidence) are structured according to how those nodes influence belief in other nodes, e.g.:

$$\textcircled{H} \xrightarrow{*} \textcircled{E} \quad (\text{with } *: \textit{entails, implies, predicts, necessitates, ...})$$

## (Qualitative) Bayesian (dis)confirmation

Evidence  $E$  confirms (or would confirm) hypothesis  $H$  just in case the prior probability of  $H$  conditional on  $E$  is greater than the prior unconditional probability of  $H$ :

$$P(H | E) > P(H).$$

Conversely,  $E$  disconfirms (or would disconfirm)  $H$  if the prior probability of  $H$  conditional on  $E$  is less than the prior unconditional probability of  $H$ .