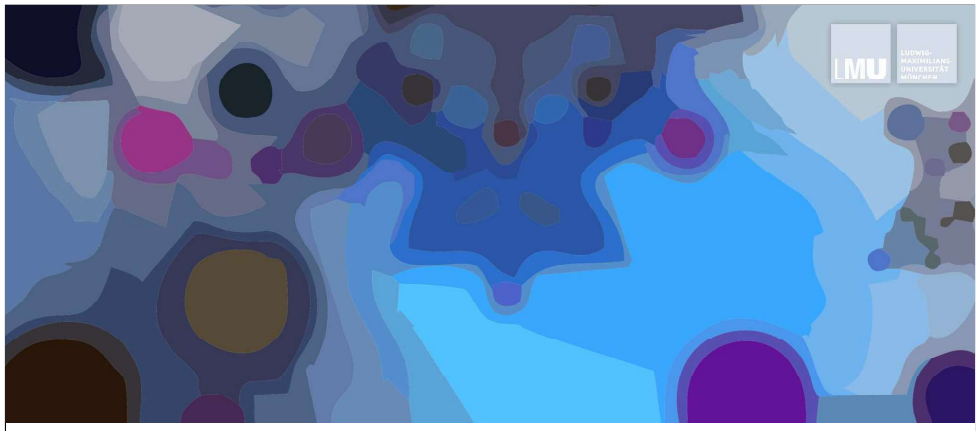


Probabilistic Causal Inference from Heterogeneous Evidence

Mechanisms in Medicine | University of Kent, Canterbury | VII/17

Roland Poellinger | r.poellinger@lmu.de
Munich Center for Mathematical Philosophy
Ludwig-Maximilians-Universität Munich

Evidence in medicine

The best-evidence paradigm selects evidence in a lexicographic hierarchy:

- Systematic reviews
- (meta-analyses of) randomized clinical trials or observational studies
- comparative non-randomized studies (cohort or case-control studies)
- evidence of mechanisms from basic science (lab experiments etc.)
- expert judgment

And: single case reports, case series, and animal studies

So, to generate reliable data in pharmacology, RCTs are the means of choice for confirming a drug's efficacy – but what about risk assessment?

Example: paracetamol and asthma

A case in pediatrics

1. Asthma epidemic calls for explanation
2. Suspicion fell on paracetamol after important study (Varner et al. '98)
3. Animal studies are conducted to test the effect of paracetamol on respiratory airways, show depletion of glutathione (antioxidant)
4. Observational studies on humans confirm the association
5. Accumulation of biological evidence leads to strong suspicion
6. Objections and alternative theories
   - increased exposure to outdoor and indoor pollutants
   - the "hygiene hypothesis" (decreased exposure to bacteria and childhood illnesses during infancy)
   - diet change and oxidant intake
   - increased obesity
   - and others ...

$\Longrightarrow$ Case unsettled: Further research needed!

Example: paracetamol and asthma, cont'd

The current debate

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 aged 6–7 years: analysis from phase three of the ISAAC programme. Lancet, 372 (9643):1039–48.

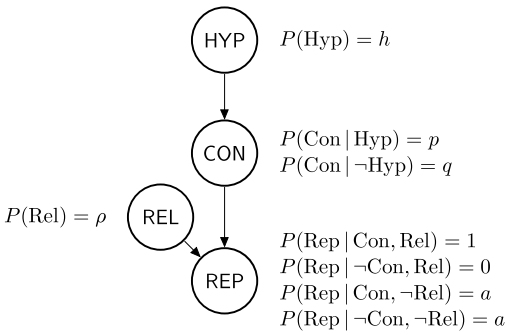
When RCTs are no option

Desiderata for an evidence-amalgamation framework for causal assessment:

1. It must allow for probabilistic hypothesis confirmation:
   We do not have to be certain (especially when high expected disutility suggests a low confidence threshold).
2. It must be able to incorporate heterogeneous kinds of data.
3. It must be able to integrate diverse types of inferential patterns (tracing the epistemic import of available evidence).

A Bayesian reconstruction of scientific hypothesis testing

Bovens & Hartmann (2003)



Confirmatory support, formally

(Qualitative) Bayesian (dis)confirmation

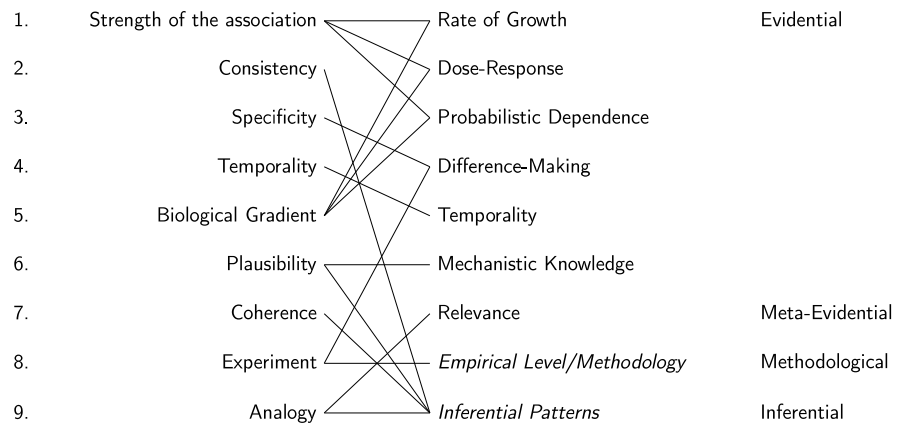
Evidence E confirms (or would confirm) hypothesis H just in case the prior probability of H conditional on E is greater than the prior unconditional probability of H:

$$P(H \mid E) > P(H).$$

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.

What are theoretical consequences of a causal hypothesis?

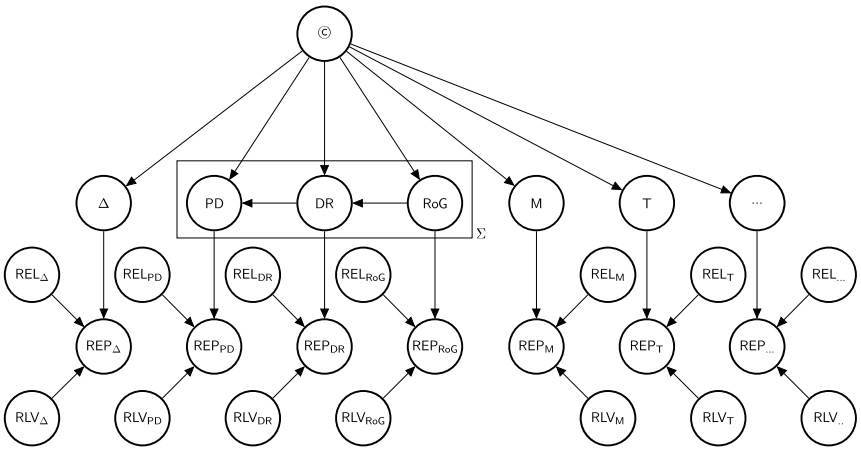
Cf. Hill's viewpoints (1965)



Hill on causal assessment

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?

Tracing epistemological dynamics in drug safety



Jürgen Landes, Barbara Osimani, and Roland Poellinger: Epistemology of Causal Inference in Pharmacology – Towards a Framework for the Assessment of Harms. European Journal for Philosophy of Science, 2017.

Probabilistic dynamics

The intended meaning of the variables yields the following inequalities:

$$P(©) < P(©|Ind) \quad \text{for all indicators } Ind \tag{9}$$

$$P(©) > P(©|\neg Ind) \quad \text{for all indicators } Ind \tag{10}$$

$$P(©|Ind_i \& Ind_k) \geq P(©|Ind_i) \quad \text{for all } i \neq k \tag{11}$$

$$P(©|Ind_i \& \neg Ind_k) \leq P(©|Ind_i) \quad \text{for all } i \neq k \tag{12}$$

$$P(©|\Delta),\ P(©|DR) > P(©|PD) \tag{13}$$

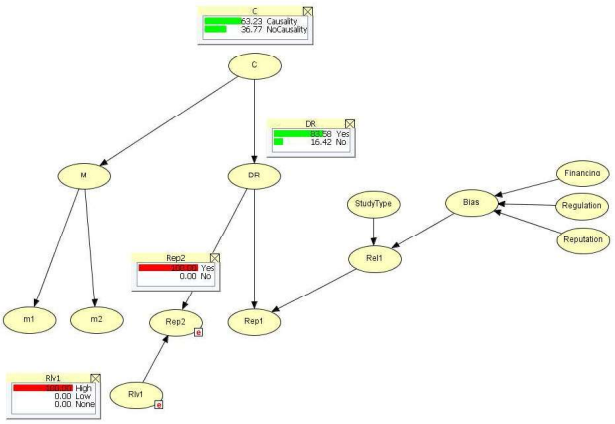
$$P(©|\neg PD) < P(©|\neg DR),\ P(©|\neg RoG) \tag{14}$$

$$\bigstar\ P(Ind_i|Rep_i \& Rlv_i \& Rel_i) > P(Ind_i) \quad \text{for all } i\ . \tag{15}$$

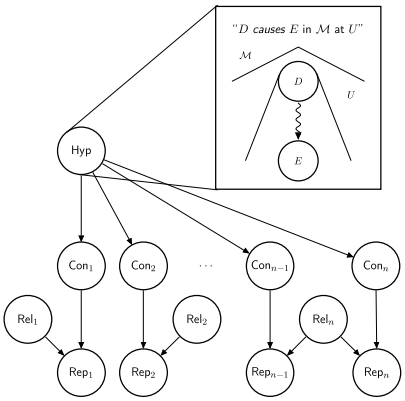
And, in addition, for a specific Ind variable and its associated report, reliability, and relevance variables:

$$\bigstar\ P(\neg Rep|\neg(Rel \& Rlv) \& Ind) > P(\neg Rep|Ind)\ . \tag{16}$$

Learning from relevant evidence



Dimensions of relevance (as confirmation criteria)



The causal hypothesis, unfolded: D ⓒ E in model $\mathcal{M}$ with context U.

Encoding dis/confirmation criteria in causal graphs

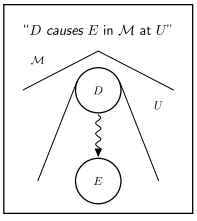
Adopt Bayes net causal modelling lingo:

The ingredients:
D, E: variables in V
U: situation, context, world, population, cet. par.
$\mathcal{M}$: conditional independence structure

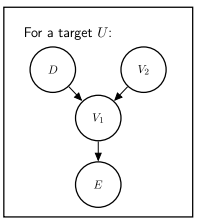
ⓒ **as contrast:** Choosing/defining/relating the variables defines how ⓒ can be disconfirmed.

Interpreting structures:
potential interventions for the difference-making theorist, counterfactual relations for the dispositionalist, connected parts for the process theorist, belief propagation for the epistemic causal theorist, closed systems for the holist, etc.

Scheme:

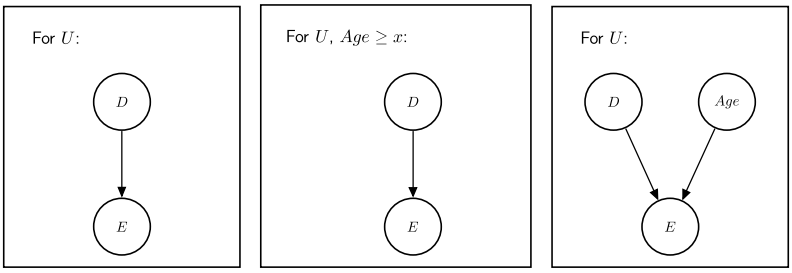


Example:



Shaping the causal hypothesis with relevant evidence

Simple example: evidence for ADR's (*only*) in elderly people.



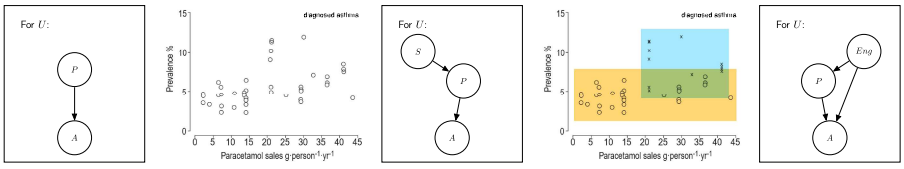
Unfolding © to $\langle D, E, \mathcal{M}, U \rangle$ allows for tracing

1. the interplay between population characteristics and causal structure;
2. how incoming evidence (if relevant) shapes © to carve out robust causal relations that remain stable under intervention and changing context – in other words: the *web of mechanisms*.

Example: interaction between $\mathcal{M}$ and U

Case study: paracetamol (sales) and asthma (Newson et al. 2000)

An ecological study in order to investigate whether the prevalences of asthma and other atopic diseases in children and adults are higher in countries with higher sales of paracetamol (data from up to 155 centers).



1. Cumulating relevant evidence leads to suspected causal association between paracetamol sales (S) and asthma (A)
2. Insert node S into structure $\mathcal{M}$
3. A coarser partition of population characteristics emerges: "It was [...] decided, *post hoc*, to carry out further regression analyses, controlling additionally for the *anglophone effect*." $\Longrightarrow$ P's effects nonsignificant

Discussion 1: the methodological debate

- **Standards for causal assessment:**
  - We argue that $P(\Delta \mid RCT) < 1$ for concrete *non-ideal* cases
  - ... to allow for additional confirmatory support from other pieces of evidence through indicators (of different weight)
  - ... locating EBM (where RCTs are trumping, perfect evidence)
  - ... locating RWT (where knowledge of difference-making plus mechanisms constitute necessary and sufficient conditions).
- **Modularity and holistic causation:**
  The structure keeps © and Δ apart to accommodate intuitions about holistic causation.
- **Reliability and relevance:**
  Distinguishing REL and RLV allows for locating and contrasting different criticisms about RCTs – issues of reliability (Worrall) vs. external validity (Cartwright & Stegenga).

Discussion 2: acquiring knowledge about mechanisms

- In our net, the M-indicator encodes the **existence of a mechanism**;
- in many cases, concrete pieces of evidence support © via M by way of contributing information about **sub-mechanisms** (integration?);
- Depending on
  (i) which concept of causation is chosen and
  (ii) how the concept of mechanism is spelled out,
  the **epistemological topology** (HYP and $INDs$) might have to be adjusted: e.g., flattening causation and mechanisms to a *thin* reading might mean linking or even merging ©, Δ, and M.

Summary: evidence synthesis and causal inference

- We needed a framework for causal inference when RCTs cannot be conducted, and when evidence is sparse and heterogenous
- B&H's scientific reconstruction separates the conceptual levels and makes epistemic dynamics transparent
- Our structure disentangles philosophical questions about causation from questions about diagnosis and methodology
- The network models cumulative learning and how evidence of different quality can be synthesized
- Different kinds of uncertainty are made explicit probabilistically and may point to gaps in justificatory arguments:
  - reliability of evidence (bias and random error);
  - relevance of evidence (potentially downgrading reliable studies);
  - confidence in the causal hypothesis (prior and posterior/s).
- Different inferential patterns can be expressed within our belief propagation structure: coherence, consistency, reasoning by analogy

Thank you!

Alexander von Humboldt Stiftung/Foundation
erc European Research Council
supported by

http://logic.rforge.com
r.poellinger@lmu.de