

Weight of Evidence: a review of concept and methods

Douglas L. Weed

Introduction

- an important issue in the process of scientific discovery: *experimentum crucis* (crucial experiment)
- it is a determinative experiment that creates new scientific consensus
 - examples can be found in physics, chemistry and engineering

Introduction

- public health: difficult to find crucial experiments
 - in general a single experiment cannot resolve a critical issue on the cause of a disease
 - several forms of evidence indirectly connected to the phenomenon
- some form of aggregation is needed in order to evaluate a causal hypothesis

Introduction

- WOE: taking into account different kinds of scientific evidence in assessing the validity of a causal hypothesis
- it is the method currently used by regulatory agencies in risk assessment process

Objective and Methodology

- **objective:** find out the meaning of the expression "weight of evidence" in recent literature
- computerized PubMed library search
- selection of 92 papers in which the term "weight of evidence" (WOE) was used
 - all publications from 2003 to 2004 (71)
 - all publications from 1994 to 2002 in which "WOE" appeared in the title (21)

Uses of WOE

- three different uses of the expression WOE have been identified:
 - metaphorical
 - methodological (with several subcategories)
 - theoretical

Metaphorical Use

- most common use
- not referred to any particular method
- the method is supposed to be implicitly understandable
- lack of transparency

Methodological Use

- *strenght of evidence*
 - use only a subset of the evidence
(only positive and statistically relevant or only unbiased evidence)
- *weight of evidence*
 - take into account all the evidence

Methodological Use

- use of established interpretative methodologies
 - systematic narrative reviews
 - quality criteria for toxicologic studies
 - epidemiology's causal criteria
 - meta-analysis
 - mixed epidemiology-toxicology methods

Systematic Narrative Reviews

- describing the state of science
- making research recommendations to authors about how to do claims of causality, or preventive recommendations
- fostering clarity and transparency in describing interpretative methods

Quality Criteria for Toxicologic Studies

- classifying toxicological studies according to their quality:
 - reliable without restriction
 - reliable with restriction
 - not reliable
 - not assignable
- first 2 classes: included in the interpretative process
- last 2 classes: included with a case-by-case approach

Epidemiology's Causal Criteria

- given a body of statistically significant epidemiological evidence and some lab based evidence
 - ➔ how can be established if there is a causal relationship?
 - ➔ *"Our observations reveal an association between two variables [...] What aspects of that association should we especially consider before deciding that the most likely interpretation of it is causation?"*
(A. B. Hill, 1965)

Epidemiology's Causal Criteria

Hill identified 9 different criteria:

- strength
- consistency
- specificity
- temporality
- biological gradient
(dose-response)
- plausibility
- coherence
- experimentation
- analogy

Meta-Analysis

- quantitative approach for summarizing evidence from several human population studies
- contribution of each study is weighted by the inverse variance of the effect estimate
- not sufficient to state causality but it can provide a reproducible weighted average of the estimate of effect across several studies

Mixed Epidemiology-Toxicology methods

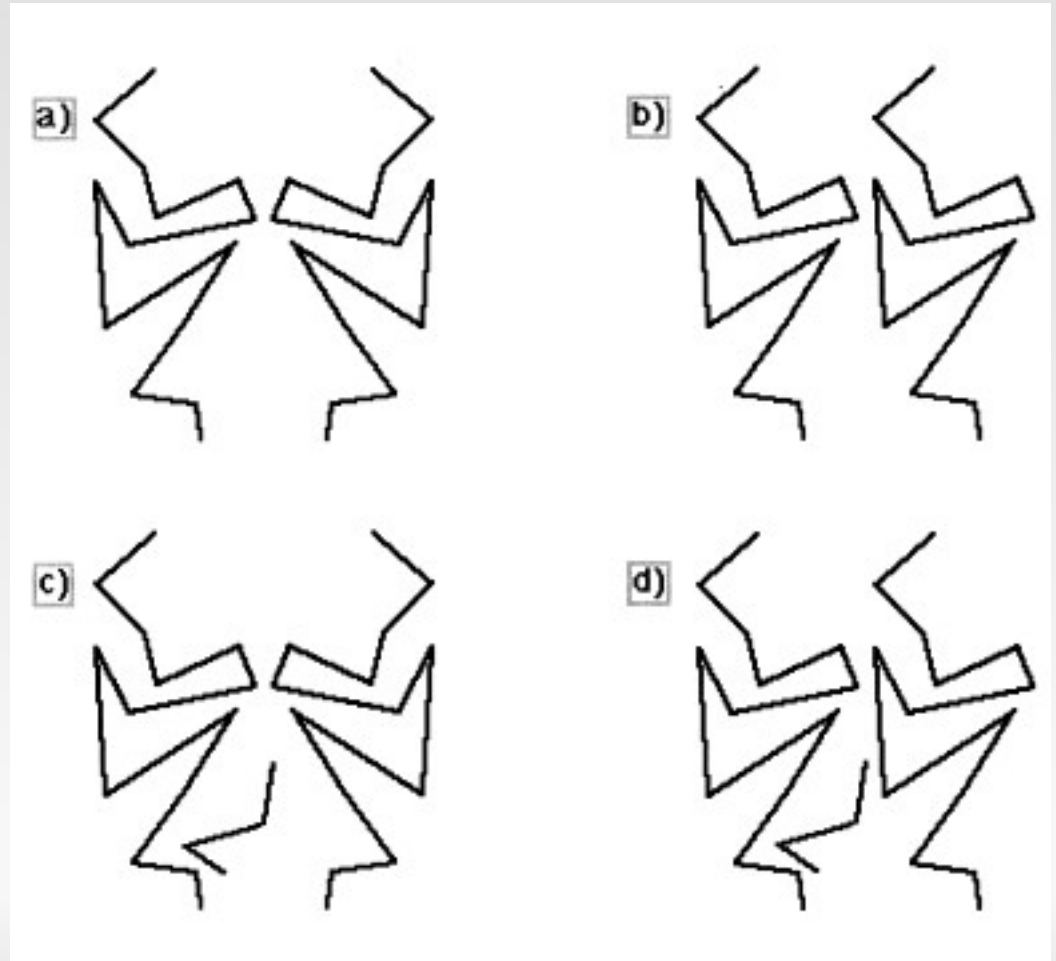
- methods to combine epidemiological evidence and lab-based biological evidence
- what is biological plausibility?
 1. a mechanism can be hypothesized
 2. some lab-based evidence supports the hypothesis
 3. sufficient evidence to show how the factor influences a known disease mechanism

Quantitative Weighting Schemes

- identifying a set of attributes
 - examples: strenght of association, temporal representativeness, use of a standard
- assigning a weight to each element according to the overall satisfaction of attributes

WOE in theory

- in cognitive science (E. M. Pothos, R. Ward, 2000)
 - WOE approach used to understand how visual patterns are perceived as regular phenomena (*figural goodness*)
 - can explain why symmetry is perceived as more regular than repetition



WOE in theory

- in law -> a conceptual framework for a WOE approach based on 4 features:
 - relevance
 - reliability
 - sufficiency
 - standard of proof

Problems & Proposals

- lack of transparency
 - encourage authors to describe the interpretative methods used
- multiplicity of WOE definitions and applications
 - moving towards the use of familiar and well-established criteria
- different kinds of weights
- role of judgement

Critics

- Letter to the editor (I. Linkov, 2006)
 - noncomprehensive search methods
 - lack of critical appraisal of included studies
 - missed studies
 - not a state-of-the-science review

Discussion

- clear formalization of the methodologies can help to reduce the subjectivity of judgements
- we must be aware that no interpretative method can abstract from expert judgement
 - further investigation is needed in order to clearly define the role of expert judgement in the different uses of WOE methodologies

References

- Weed, D. L. (2005). Weight of evidence: a review of concept and method. *Risk Analysis*, 25(6), 1545-1557.
- Hill, A. B. (1965). The environment and disease: Association or causation? *Journal of the Royal Society of Medicine*, 58, 295-30
- Pothos, E. M., and Ward, R. (2000). Symmetry, repetition, and figural goodness: An investigation of the weight of evidence theory. *Cognition*, 75(3), B65-78
- Krimsky, S. (2005). The weight of scientific evidence in policy and law. *American Journal of Public Health*, 95 (S1), S129-S136
- Linkov, I. (2006). Weight of Evidence: What is the State of the Science? *Risk Analysis*, 26(3), 573-575.
- Weed, D. L. (1997). On the use of causal criteria. *International Journal of Epidemiology*, 26, 1137-1141.
- Weed, D. L., and Hursting, S. D. (1998). Biologic plausibility in causal inference: Current method and practice. *American Journal of Epidemiology*, 147, 415-425.