


Epidemiology Kept Simple



Chapter 16: From Association to Causation

GerstmanChapter 161


Background

- **Causal inference** \equiv the process of deriving cause-and-effect conclusions by reasoning from knowledge and factual evidence
- “Proof” is impossible in empirical sciences.
- However, statements can be made strong or even overwhelming
- Inference differs for:
 1. Scientific inference
 2. When action is required and waiting for proof might be imprudent

GerstmanChapter 162

Causal Mechanism

“For want of knowledge, efforts which have been made to oppose [cholera] have often had contrary effect.”
– John Snow
(e.g., in reference to miasmas theory)



GerstmanChapter 163

Discovery of Preventive Measure May Predate Identification of Definitive Cause


Wynder (1994, p. 548)

Disease	Discoverer of preventive measure*	Year of discovery of preventive measure	Year of discovery of agent	Causative or preventive agent*	Discoverer of agent*
Scurvy	J. Lind	1753	1928	(Ascorbic acid)	A. Szent-Gyorgi
Pellagra	G. Casal	1755	1924	(Nicotin)	Goldberger et al.
Scrotal Cancer	P. Pott	1755	1933	Benzo(a)pyrene	J. W. Cook et al.
Smallpox	E. Jenner	1798	1958	Orthopoxvirus	F. Fenner
Postperal fever	I. Semmelweis	1847	1879	Streptococcus	L. Pasteur
Cholera	J. Snow	1849	1893	<i>Vibrio cholerae</i>	R. Koch
Bladder cancer [†]	L. Rehn	1895	1938	2-Naphtylamine	Hesper et al.
Yellow fever	W. Reed et al.	1901	1928	Flavivirus	A. Stokes et al.
Oral cancer [‡]	R. Abbe	1915	1974	N-nitrosomethylamine	Hoffmann et al.

What if we waited until the mechanism was known before employing citrus?

GerstmanChapter 164

§16.2 Surgeon General’s Report on Smoking




- Epi data must be coupled with clinical, pathological, and experimental data
- Epi data must consider multiple variables
- Multiple studies must be considered
- Statistical methods alone *cannot* establish proof
- Criteria for inference include consistency, strength of association, etc. (See Hill’s Framework)

[\[Link to Surgeon General’s report\]](#)

GerstmanChapter 165

Hill’s Inferential Framework

1. Consistency
2. Specificity
3. Temporality
4. Biological gradient
5. Plausibility
6. Coherence
7. Experimentation
8. Analogy




A. Bradford Hill
(1897–1991)

* Hill, A. B. (1965). The environment and disease: association or causation? *Proceedings of the Royal Society of Medicine*, 58, 295-300. [full text](#)

GerstmanChapter 166

Criterion 1: Strength

- **Stronger associations** are less easily explained by confounding than weak associations
- Ratio measures (e.g., RR, SMR, OR) are used to quantify the strength of an association
- **Example:** An RR of 10 is much less easily explain by confounding than an RR of 2




ARISTOTLE
Photograph of Aristotle by Raphael, 1504. The original painting is in the Vatican Museums, Rome. The photograph is by the artist's studio.

GerstmanChapter 167

Criterion 2: Consistency

- Consistency: studies using diverse methods in different populations under a variety of circumstances lead to similar conclusions
- **Example:** Ecological, cohort, and case-control done by independent researchers studying different populations all showed a strong association between smoking and lung cancer.




CONSISTENCY
It's Over a World in Yours. Not a Science.

GerstmanChapter 168

Criterion 3: Specificity

- The factor is linked to a **specific causal mechanism**
- **Example:** Smoking is linked to physical and chemical carcinogenesis of epithelial cells
- **Comment:** Mechanisms are difficult to establish when there is a vacuum of knowledge



Aristotle
(384 – 322 BCE)

GerstmanChapter 169

Criterion 4: Temporality

Exposure precedes disease by a reasonable amount of time

Lead encephalopathy scenario 1:

```

    graph LR
      A[deficient care giving environment] --> B[pica]
      B --> C[Exposure]
      D[paint chips] --> C
      C --> E[behavioral and Intellectual Impairments]
    
```

Lead encephalopathy scenario 2:

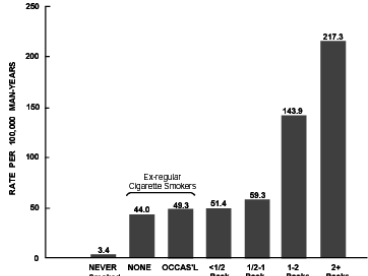
```

    graph LR
      A[deficient care giving environment] --> B[behavioral and Intellectual Impairments]
      B --> C[pica]
      C --> D[Lead Exposure]
      E[paint chips] --> D
    
```

GerstmanChapter 1610

Criterion 5: Biological Gradient

Increases in exposure dose ⇒ dose-response in risk




Smoking Status	Rate per 100,000 Man-Years	No. Deaths	No. Men
NEVER Smoked	3.4	4	32,392
NONE	44.0	15	10,095
OCCASL	49.3	2	1,322
<1/2 Pack	51.4	13	7,647
1/2-1 Pack	56.3	50	26,370
1-2 Packs	143.9	60	14,292
2+ Packs	217.3	22	5,100

GerstmanChapter 1611

Criterion 6: Plausibility

- Plausible mechanism in face of known biological facts
- Plausibility (defined): appearing worthy of belief
- **Comment:** All that is plausible is not always true



GerstmanChapter 1612

Criterion 7: Coherence

- All facts stick together to form a coherent whole.
- **Example:** Epidemiologic, pharmacokinetic, laboratory, clinical, and biological data create a cohesive picture about the smoking and lung cancer.



Gerstman

Chapter 16

13

Criterion 8: Experimentation

- Experimental evidence in conjunction with epidemiologic evidence
- *In vitro* and *in vivo* experimentation
- Experimentation is often not possible in humans
- Animal models of human disease



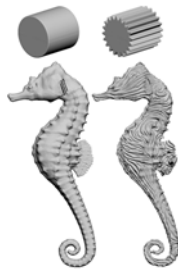
Gerstman

Chapter 16

14

Criterion 9: Analogy

- Similarities among things that are otherwise different
- Weak form of evidence
- **Example:** Before the HIV was discovered, epidemiologists noticed that AIDS and Hepatitis B had analogous risk groups, suggesting similar types of agents and transmission



Gerstman

Chapter 16

15