Causal Inference: New data for an old problem

Jonathan M. Samet, MD, MS
Distinguished Professor and Flora L. Thornton Chair
Department of Preventive Medicine
Keck School of Medicine of USC

NASEM Workshop
Advances in Causal Understanding for Human Health
Risk Based Decision Making

March 6, 2017
Objectives

• To describe the role of causal inference within the risk assessment framework
• To address the evolution of causal inference in public health
• To address the challenges posed by new data streams
The “Red Book”

Four Components

- Causal Inference
- Dose-response
- Exposure assessment
- Risk characterization
Morton Levin’s Attributable Risk Formula

\[ PAR = \frac{P \times (RR - 1)}{1 + P(RR - 1)} \]

- Estimate the Relative Risk (RR)
- Estimate the prevalence (P) of each risk factor.

Johns Hopkins School of Hygiene and Public Health, Department of Epidemiology, c. 1935-36
Causal Inference for What?

• To identify determinants of disease etiology as a basis for intervention
• As an element of quantitative risk assessment
• As a basis for decision-making
  – What level of evidence is needed for decision-making?
• To identify critical uncertainties
The Importance of Causation in Public Health

“A causal conclusion conveys the inference that changing a given factor will actually reduce a population’s burden of disease, either by reducing the overall number of cases or by making disease occur later than it would have.”

CAUSATION
GETTING TO THE 21ST CENTURY
What is a cause?

• Long-standing and ongoing discussion among philosophers and others

• Hume’s problem—Is causation unknowable since we only know that putative cause precedes putative effect in observations made?

• Differing frames of reference across fields of science and among people in general
David Hume (1711-1776)

“Rules by which to judge of causes and effects”
Hume’s Rules 1 and 2

"The cause must be prior to the effect"

Temporality
Hume and the counterfactual

Counterfactual = condition contrary-to-fact

“We may define a cause to be an object, followed by another, and where all objects similar to the first are followed by objects similar to the second” (...)

“and in other words where, if the first object had not been, the second never had existed”

(Enquiry VII, pt II, SB p, 76)
“But make the three field goals they missed and get a few other breaks and the Terps would have been sailing into the fourth Quarter with a safe lead.”

8/31/08, p4 Sports
Robert Koch (1843 – 1910)

- Considered one of the founders of bacteriology

- 1870s – Koch develops a sequence of experimental steps for directly relating a specific microbe to a specific disease (Henle-Koch postulates).
  - Discovers: *Bacillus anthracis*, *Mycobacterium tuberculosis*, *Vibrio cholera*

- Contributed to the germ theory of disease
Henle-Koch postulates

All four must be fulfilled in order to establish a causal relationship between an organism and a disease:

- The organism must be found in all animals suffering from the disease, but not in healthy animals.
- The organism must be isolated from a diseased animal and grown in pure culture.
- The cultured organism should cause disease when introduced into a healthy animal.
- The organism must be re-isolated from the experimentally infected animal.
Selected mortality rates in US, 1900-2005

The Environment and Disease: Association or Causation?

by Sir Austin Bradford Hill CBE DSc FRCP(hon) FRS (Professor Emeritus of Medical Statistics, University of London)

Amongst the objects of this newly-founded Section of Occupational Medicine are firstly 'to provide a means, not readily afforded elsewhere, whereby physicians and surgeons with a special knowledge of the relationship between sickness and injury and conditions of work may discuss their problems, not only with each other, but also with...

Meeting January 14 1965

President’s Address

observed association to a verdict of causation? Upon what basis should we proceed to do so?

I have no wish, nor the skill, to embark upon a philosophical discussion of the meaning of 'causation'. The 'cause' of illness may be immediate and direct, it may be remote and indirect underlying the observed association. But with the aims of occupational, and almost synonymously preventive, medicine in mind the decisive question is whether the frequency of the undesirable event B will be influenced by a change in the environmental feature A. How such a change exerts that influence may call for a great deal of research. However, before deducing...
Table 1  Guidelines for causal inference. Data from the 1964 *Smoking and Health: Report of the Advisory Committee to the Surgeon General* (71) and from Hill 1965 (33)

<table>
<thead>
<tr>
<th>US Surgeon General Report’s criteria</th>
<th>Hill’s criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>Consistency of association</td>
<td>Strength</td>
</tr>
<tr>
<td>Strength of association</td>
<td>Consistency</td>
</tr>
<tr>
<td>Specificity of association</td>
<td>Specificity</td>
</tr>
<tr>
<td>Temporal relationship of association</td>
<td>Temporality</td>
</tr>
<tr>
<td>Coherence of association</td>
<td>Biological gradient</td>
</tr>
<tr>
<td></td>
<td>Coherence</td>
</tr>
<tr>
<td></td>
<td>Experiment</td>
</tr>
</tbody>
</table>

Source: Google Scholar search for “The environment and disease: association or causation? 1965.”
These Guidelines are Embedded in:

- US EPA Cancer Guidelines
- US EPA IRIS Preamble
- IARC Preamble
- And more!
The Process of Causal Inference
The Process of Causal Inference

Develop evidence:
- Observations of people
  - individuals
  - groups
- Human experiments
- Other research

Synthesize:
- Systematic reviews
- Meta-analysis
- Other evidence

Evaluate:
- Expert judgment
- Causal criteria

Domain of research
Domain of Systematic Review
Domain of Expert Judgment

CAUSAL INFERENCE
Some Guidelines for Judging Whether an Association is Causal

- Temporal relationship
  - Exposure should precede effect
Some Guidelines for Judging Whether an Association is Casual

• Temporal relationship

• \textit{Strength of the association} — Stronger associations more likely to be causal
# Odds Ratios (OR) for Lung Cancer in Males

<table>
<thead>
<tr>
<th>Study</th>
<th>Years</th>
<th>OR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wynder/Graham</td>
<td>&lt; 1950</td>
<td>12.8</td>
</tr>
<tr>
<td>Doll/Hill</td>
<td>1948-51</td>
<td>9.1</td>
</tr>
<tr>
<td>Wynder et al</td>
<td>1966-69</td>
<td>9.7</td>
</tr>
<tr>
<td>Higgins/Wynder</td>
<td>1979-84</td>
<td>11.2</td>
</tr>
</tbody>
</table>
Some Guidelines for Judging Whether an Association is Causal

- Temporal relationship
- Strength of the association
- Dose–response relationship
  - Increasing risk with increasing dose implies causality
DEATH RATES FROM LUNG CANCER
BY NUMBER OF CIGARETTES SMOKED
(Age-Adjusted, per 100,000 Man-Years)
Relative Risk of Lung Cancer by Cigarettes Smoked per Day: Current Smokers vs. Never Smokers

Men

Relative Risk vs. Cigarettes per Day

TRIAL EXHIBIT 30,092
Some Guidelines for Judging Whether an Association is Causal

• Temporal relationship
• Strength of the association
• Dose–response relationship
• *Consistency of the association*

- Consistency on replication implies causality
Relative Risk of Lung Cancer by Cigarettes Smoked per Day: Current Smokers vs. Never Smokers

Men

[Graph showing the relationship between Cigarettes per Day and Relative Risk]

TRIAL EXHIBIT 30,092
EPIDEMIOLOGIC STUDIES OF HUSBAND'S SMOKING & LUNG CANCER

Relative Odds of Lung Cancer

10

1

0.1
The 1986 Surgeon General’s Report

“Involuntary smoking is a cause of disease, including lung cancer, in healthy nonsmokers.”

C. Everett Koop, MD, DSc
Surgeon General, 1982-89
Some Guidelines for Judging Whether an Association is Causal

- Temporal relationship
- Strength of the association
- Dose–response relationship
- Consistency of the association
- Effect of removing the exposure
  - Removing exposure should lower risk, if association is causal
Relative Risk of Lung Cancer by Number of Years Quit Smoking: Former Smokers vs. Never Smokers

Men
Some Guidelines for Judging Whether an Association is Causal

- Temporal relationship
- Strength of the association
- Dose–response relationship
- Consistency of the association
- Effect of removing the exposure
- Biologic plausibility
  - Causal association should be biologically plausible
These studies are inconclusive—so far we’ve only succeeded in giving cancer and heart disease to laboratory humans.
STRENGTH OF EVIDENCE
The Evidence Scale

Act

Evidence

Uncertainty

Not Act

Politics
Costs
Activists
Advocates
Equipoise and Evidence

What is it?

The balance point for strength of evidence on causation
Terminology of Conclusions and the Causal Claims

A four-level hierarchy for classifying the strength of causal inferences based on available evidence:

A. Evidence is **sufficient** to infer a causal relationship.

B. Evidence is **suggestive but not sufficient** to infer a causal relationship.

C. Evidence is **inadequate** to infer the presence or absence of a causal relationship (which encompasses evidence that is sparse, of poor quality, or conflicting).

D. Evidence is **suggestive of no causal relationship**.
NEW DATA-OLD PROBLEM
What are the new data streams?

• Exposomics

• New, non-apical endpoints
  – Markers of early disease, e.g., imaging
  – Other biomarkers
  – Pathway perturbation

• Population stratification by phenotype
  – New phenotyping approaches
  – Genomics
One Person: Too Much Data
Using 21st Century Science to Improve Risk-Related Evaluations

Committee on Incorporating 21st Century Science into Risk-Based Evaluations

Board on Environmental Studies and Toxicology

Division on Earth and Life Studies

The National Academies of
SCIENCES • ENGINEERING • MEDICINE
Tox21 and ES21 Reports

- *Toxicity Testing in the 21st Century: A Vision and a Strategy* was published in 2007 and envisioned a future in which toxicology relied primarily on high-throughput in vitro assays and computational models based on human biology to evaluate potential adverse effects of chemical exposure.

- *Exposure Science in the 21st Century: A Vision and a Strategy* was published in 2012 and provided a vision that was hoped to inspire transformational changes in the breadth and depth of exposure assessment that would improve integration with and responsiveness to toxicology and epidemiology.
The advances in exposure science, toxicology, and epidemiology described in this report support the new direction for risk assessment,

- one based on biological pathways and processes rather than on observation of apical responses and

- one incorporating the more comprehensive exposure information emerging from new tools and approaches in exposure science.
Questions to Address

- Can an identified pathway, alone or in combination with other pathways, when sufficiently perturbed, increase the risk of an adverse outcome or disease in humans, particularly in sensitive or vulnerable individuals?

- Do the available support the judgment that the chemical or agent perturbs one or more pathways linked to an adverse outcome?

- How does the response or pathway activation change with exposure? By how much does a chemical or agent exposure increase the risk of outcomes of interest?

- Which populations are likely to be the most affected? Are some more susceptible because of co-exposures, pre-existing disease, or genetic susceptibility? Are exposures of the young or elderly of greater concern?
With the shift from observing apical responses to understanding biological processes or pathways comes the recognition that a single adverse outcome might result from multiple mechanisms, which can have multiple components.
The committee found that Bradford-Hill causal guidelines could be extended to help to answer such questions as whether specific pathways, components, or mechanisms contribute to a disease or outcome and whether a particular agent is linked to pathway perturbation or mechanism activation.

Although the committee considered various methods for data integration, it concluded that guided expert judgment should be used in the near term for integrating diverse data streams for drawing causal conclusions.
The 21st century science with its diverse, complex, and very large datasets, however, poses challenges related to analysis, interpretation, and integration of data and evidence for risk assessment.

The committee emphasizes that insufficient attention has been given to analysis, interpretation, and integration of various data streams from exposure science, toxicology, and epidemiology.
Components of a Research Agenda

- Develop case studies that reflect various scenarios of decision-making and data availability.

- Test case studies with multidisciplinary panels.

- Catalogue evidence evaluations and decisions that have been made on various agents so that expert judgments can be tracked and evaluated, and expert processes calibrated.

- Determine how statistically based tools for combining and integrating evidence, such as Bayesian approaches, can be used for incorporating 21st century science into all elements of risk assessment.
The data that are being generated today can be used to help to address many of the risk-related tasks that agencies face.

Although the challenges to achieving the visions of the earlier reports often seem daunting, 21st century science holds great promise for advancing risk assessment and ultimately for improving public health and the environment.
Looking forward and backwards

"It's like deja-vu, all over again."
Yogi Berra