Seeking the Holy Grail:
The Search for Causality in Air Pollution and Health Data

Dan Greenbaum, President
Health Effects Institute

PM Causality Symposium
Rizzo Conference Center
Chapel Hill, NC
October 4, 2018
The Holy Grail
The Seekers....

MONTY PYTHON AND THE HOLY GRAIL 1975
Paths to Causality Enlightenment....

There are several ways forward...

• **Systematic Literature Review and Synthesis**
  – Drawing on all lines of inquiry

• **Accountability Research**
  – Testing whether, after an intervention reduces exposure, there is evidence of health improvement

• **Causal Inference Statistical Methods**
  – Rigorous *a priori* specification of assumptions and analytic approach
The Health Effects Institute
Trusted Science – Cleaner Air – Better Health
www.healtheffects.org

• Independent Non-profit Research Institute since 1980

• Balanced Core Support
  – Also partnerships with other governments, industries, foundations

• Independent Board and Expert Science Committees
  – Board agreed to by EPA Administrator and core industry sponsors
  – Research Committee selects all research competitively
  – Separate Review Committee intensively peer reviews all results

• Full Transparency
  – All Results – positive and negative – published
  – Works to make all data accessible to others

• Does not take policy positions
HEI Science
All available at www.healtheffects.org

• **Targeted Research and Reanalysis**
  – Over 350 Studies on a wide variety of air pollutants: PM, ozone, diesel, air toxics, others
  – Exposure, Toxicology, Epidemiology
  – Accountability Research

• **Rapid Scientific Review**
  – E.g. The Health Effects of Exposure to Traffic

• **Global Health**
  – North America, Europe
  – Developing Asia

• **NEW Energy Research Program**
  – Potential Exposures and effects from unconventional oil and gas development
Paths to Causality Enlightenment....

There are several ways forward...

• *Systematic Literature Review and Synthesis*
  – Drawing on all lines of inquiry

• *Accountability Research*
  – Testing whether, after an intervention reduces exposure, there is evidence of health improvement

• *Causal Inference Statistical Methods*
  – *Rigorous a priori specification of assumptions and analytic approach*
Systematic Review and Synthesis

- Applied by many agencies (WHO, EPA, NIEHS....)
Drawing on All Lines of Evidence...

- **Animal and toxicology studies:** can contribute to the understanding of potential biological mechanisms of how air pollutants can cause effects.

- **Human Epidemiology Studies:** can take many forms, from small “panel” studies of carefully selected subjects to large population studies of selected population cohorts ranging to very large administrative data sets (e.g. Medicare recipients).

- **Controlled Human Exposure Studies:** can provide direct evidence of whether there are effects in humans from specific pollutants, and can do so at levels relevant to actual ambient levels.
Bradford-Hill considerations
(J Roy Soc Med 1965:58:295-300)

1. Strength of association: Stronger associations are more likely causal.
2. Consistency of findings across populations, study designs, times
3. Specificity of association: one cause, one effect
4. Temporality: Causes precede their effects
5. Biological gradient: Dose-response
6. Biological plausibility: Plausible biological mechanism
7. Coherence: Agrees with knowledge of disease biology
8. Experiment: Reducing exposure reduces effect

https://www.healthknowledge.org.uk/e-learning/epidemiology/practitioners/causation-epidemiology-association-causation
Can be adapted for application in specific contexts:

EPA ISA for Oxides of Nitrogen 2016

<table>
<thead>
<tr>
<th>Table I</th>
<th>Aspects to aid in judging causality.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aspect</td>
<td>Description</td>
</tr>
<tr>
<td>Consistency</td>
<td>An inference of causality is strengthened when a pattern of elevated risks is observed across several independent studies. The reproducibility of findings constitutes one of the strongest arguments for causality. Statistical significance is not the sole criterion by which the presence or absence of an effect is determined. If there are discordant results among investigations, possible reasons such as differences in exposure, confounding factors, and the power of the study are considered.</td>
</tr>
<tr>
<td>Coherence</td>
<td>An inference of causality from one line of evidence (e.g., epidemiologic, controlled human exposure, animal, welfare studies) may be strengthened by other lines of evidence that support a cause-and-effect interpretation of the association. There may be coherence in demonstrating effects from evidence across various fields and/or across multiple study designs or related health endpoints within one scientific line of evidence. For example, evidence on welfare effects may be drawn from a variety of experimental approaches (e.g., greenhouse, laboratory, field) and subdisciplines of ecology (e.g., community ecology, biogeochemistry, paleontological/historical reconstructions).</td>
</tr>
<tr>
<td>Biological plausibility</td>
<td>An inference of causality is strengthened by results from experimental studies or other sources demonstrating biologically plausible mechanisms. A proposed mechanism, which is based on experimental evidence and which links exposure to an agent to a given effect, is an important source of support for causality.</td>
</tr>
<tr>
<td>Biological gradient (exposure-response relationship)</td>
<td>A well-characterized exposure-response relationship (e.g., increasing effects associated with greater exposure) strongly suggests cause and effect, especially when such relationships are also observed for duration of exposure (e.g., increasing effects observed following longer exposure times).</td>
</tr>
<tr>
<td>Strength of the observed association</td>
<td>The finding of large, precise risks increases confidence that the association is not likely due to chance, bias, or other factors. However, it is noted that a small magnitude in an effect estimate may or may not represent a substantial effect in a population.</td>
</tr>
<tr>
<td>Experimental evidence</td>
<td>Strong evidence for causality can be provided through “natural experiments” when a change in exposure is found to result in a change in occurrence or frequency of health or welfare effects.</td>
</tr>
<tr>
<td>Temporality of the observed association</td>
<td>Evidence of a temporal sequence between the introduction of an agent and appearance of the effect constitutes another argument in favor of causality.</td>
</tr>
<tr>
<td>Specificity of the observed association</td>
<td>Evidence linking a specific outcome to an exposure can provide a strong argument for causation. However, it must be recognized that rarely, if ever, does exposure to a pollutant invariably predict the occurrence of an outcome, and that a given outcome may have multiple causes.</td>
</tr>
<tr>
<td>Analogy</td>
<td>Structure activity relationships and information on the agent’s structural analogs can provide insight into whether an association is causal. Similarly, information on mode of action for a chemical, as one of many structural analogs, can inform decisions regarding likely causality.</td>
</tr>
</tbody>
</table>
Explicit Weight of Evidence to determine Causality

- Not Likely
- Inadequate
- Suggestive
- Likely
- Causal

<table>
<thead>
<tr>
<th>Table II</th>
<th>Weight of evidence for causal determination.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Health Effects</strong></td>
<td><strong>Ecological and Welfare Effects</strong></td>
</tr>
<tr>
<td>Causal relationship</td>
<td>Evidence is sufficient to conclude that there is a causal relationship with relevant pollutant exposures. Evidence is sufficient to conclude that there is a causal relationship with relevant pollutant exposures. That is, the pollutant has been shown to result in health effects in studies where the effects are not explained by other factors. Generally, the determination is based on multiple high-quality studies conducted by multiple research groups.</td>
</tr>
<tr>
<td>Likely to be a causal relationship</td>
<td>Evidence is sufficient to conclude that there is a likely causal relationship with relevant pollutant exposures. That is, an association has been observed between the pollutant and the outcome in studies where the effects are not explained by other factors. Generally, the determination is based on multiple high-quality studies.</td>
</tr>
<tr>
<td>Suggestive of, but not sufficient to infer, a causal relationship</td>
<td>Evidence is suggestive of a causal relationship with relevant pollutant exposures but is limited, and other biases cannot be ruled out. For example, when the body of evidence is not consistent, at least one high-quality epidemiologic study shows an association with a given health outcome and/or at least one high-quality toxicologic study shows evidence of causation. If the data from multiple studies are consistent, there may be coherence across lines of evidence. Generally, the determination is based on multiple high-quality studies.</td>
</tr>
<tr>
<td>Inadequate to infer a causal relationship</td>
<td>Evidence is inadequate to determine that a causal relationship exists with relevant pollutant exposures. The available studies are of insufficient quantity, quality, consistency, or statistical power to permit a conclusion regarding the presence or absence of an effect.</td>
</tr>
<tr>
<td>Not likely to be a causal relationship</td>
<td>Evidence indicates there is no causal relationship with relevant pollutant exposures. Several studies examining relationships with relevant exposures are consistent in not showing an effect at any level of exposure.</td>
</tr>
</tbody>
</table>
Challenges...

• There are of course questions about the robustness of these steps:
  – There are limitations in each type of research
  – Epidemiology has it limitations
    • Exposure assessment, control for confounders, model selection
  – How does one
    • Select which studies to review?
    • Weigh the strength of individual studies?
    • Test for inherent biases
  – What tools are applied to assess overall weight of evidence?
Opportunities for improvement in *Systematic Review and Synthesis*

- **Better Systematic Review** – New Techniques for
  - Identifying Literature
  - Weighing Quality and Risk of Bias
  - Using Machine Learning to Extract and analyze data

- **Increased use of publicly available data bases**
  - Government funded studies (e.g. NCI/NIOSH DEMS, MESA/AIR, Others)
  - Government Data Sets (e.g. Medicare, National Health Interview Survey, Others)

- **Continued enhancements to the already relatively high level of transparency**
Paths to Causality Enlightenment....

There are several ways forward...

- **Systematic Literature Review and Synthesis**
  - Drawing on all lines of inquiry

- **Accountability Research**
  - Testing whether, after an intervention reduces exposure, there is evidence of health improvement

- **Causal Inference Statistical Methods**
  - Rigorous a priori specification of assumptions and analytic approach
Accountability Research

• How can we learn if air quality changes can improve health?

• *Accountability Research*: Testing the extent to which air quality interventions improve air quality and public health

• Part of a broader effort to assess the performance of environmental regulatory policies
One way of answering:
By following the **Chain of Accountability**

(HEI 2016)
HEI Accountability Studies

• Over a dozen short and longer term intervention studies to date...

• Today, a sample of those results
  – Shorter term:
    • Atlanta Olympics
    • Ireland Coal Ban
  – Longer term
    • AQ Actions in the LA Basin
    • AQ Actions in the Atlanta Region
Impact of Improved Air Quality During the 1996 Atlanta Olympic Games

HEI Research Report 198

Jennifer Peel and colleagues; Colorado State University

- Study to assess impact of traffic reduction measures in downtown Atlanta during the Olympic Games and the effect on air quality and health

- An earlier study by the CDC reported a decrease in ozone and childhood asthma hospital admissions (Friedman et al, JAMA 2001)
Atlanta Olympics – Earlier Study
Michael Friedman et al, JAMA 285 (2001) 897-905

Reported significant reduction in morning traffic (-22%), reduced peak daily ozone levels (-28%), and reduced asthma acute care events (-41%)
Objectives of HEI Analysis

• Test what happened and explore alternate explanations more rigorously:

  – Did the measures to control traffic actually reduce traffic?
  – Did traffic changes reduce Ozone, PM and other pollutants?
  – What happened to Emergency Department Visits for multiple cardiovascular and respiratory cases?
  – And what happened in the same periods in the years before and after the Olympics? (i.e. were their seasonal effects?)
Results – Traffic Counts
~10-15% decline in morning rush hour peak
*BUT* overall daily count unchanged
Results – Ozone (1-hour max)

~30% Reduction Pre-During-Post

**BUT**, Similar Reductions Throughout the Southeast
Emergency Department Visits (all ages)

Little evidence of reduced visits when the same trends in other years are included
So Did This Confirm Cause and Effect?

• NO – but there wasn’t really an air pollution change in the first place:
  – The traffic “controls” were voluntary – and basically seemed to just shift the times of traffic
  – Ozone levels were down, but likely due to weather patterns across the Southeast

• And, there was a seasonal confounder not considered earlier

• **Bottom Line:** It is possible that traffic changes could have had benefits, but they would have had to be much more comprehensive and effective
A second key element of such studies:
Making sure you have a control group that was not affected by the action

My stomach hurts

Maybe it was the bad tuna I had for lunch

Too bad I don't have a proper control group, now I'll never know
A 2002 Study in Ireland:
A ban on home coal use in Dublin reduced air pollution

Seasonal mean black smoke and SO$_2$ concentrations, September 1984-96

Clancy et al. (2002)
Lancet
360:1210-14
Also, the authors reported that heart and lung mortality declined over the next 6 years.

<table>
<thead>
<tr>
<th></th>
<th>Unadjusted % change (95% CI)</th>
<th>p</th>
<th>Adjusted % change* (95% CI)</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Total</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-trauma</td>
<td>-8.0 (-9.8 to -6.2)</td>
<td>&lt;0.0001</td>
<td>-5.7 (-7.2 to -4.1)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td><strong>Cause-specific</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cardiovascular</td>
<td>-13.4 (-15.9 to -10.8)</td>
<td>&lt;0.0001</td>
<td>-10.3 (-12.6 to -8.0)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Respiratory</td>
<td>-16.1 (-20.4 to -11.6)</td>
<td>&lt;0.0001</td>
<td>-15.5 (-19.1 to -11.6)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Other</td>
<td>1.4 (-1.6 to 4.6)</td>
<td>0.36</td>
<td>1.7 (-0.7 to 4.2)</td>
<td>0.17</td>
</tr>
<tr>
<td><strong>Age-specific</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Younger than age 60 years</td>
<td>-8.1 (-12.3 to -3.7)</td>
<td>&lt;0.0001</td>
<td>-7.9 (-12.0 to -3.6)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Age 60–74 years</td>
<td>-8.6 (-12.3 to -9.6)</td>
<td>&lt;0.0001</td>
<td>-6.2 (-8.8 to -3.5)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Age 75 years or older</td>
<td>-7.6 (-8.1 to -7.0)</td>
<td>&lt;0.0001</td>
<td>-4.5 (-6.7 to -2.3)</td>
<td>&lt;0.0001</td>
</tr>
</tbody>
</table>

*Adjusted in robust Poisson regression for temperature, relative humidity, day of week, respiratory epidemics, and standardised cause-specific death rates in rest of Ireland.

Table 3: Change in age-standardised total, cause-specific, and age-specific mortality rates for Dublin County Borough for 72 months before and after ban of sale of coal in Dublin

A Deeper Look Funded by HEI*

- Clancy et al., focused on changes in Dublin only
- Dockery et al extended the original study:
  - Irish govt. extended coal ban to 11 other cities in 1995, 1998 and 2000
  - Added a “comparison” population: residents of the Midland counties where coal ban was not instituted

= Comparison County
Confirmed that air pollution went down...

Mean BS and TGA concentrations 5 years before and 5 years after the coal bans
Key result: Comparison counties saw **same** changes in heart deaths as those that had reduced air pollution (likely because of improving heart health overall)

Key result: There did appear to be reductions in respiratory deaths

So did this confirm cause and effect?

- There was an improvement in air pollution from this action.
- And YES there was evidence of improved lung health.
- But also NO: the original study probably overestimated the effect of the Dublin coal ban on total and cardiovascular mortality.
- “Detecting changes in public health indicators associated ... with clear improvements in air quality, as in this case, remains difficult when there are simultaneous secular improvements in the same health indicators.” (Dockery et al.)
An HEI LA Accountability Study: Los Angeles Then and Now

Source: New York Times
After substantial numbers of mobile source and other action, air quality improved (though not ozone).

HEI Report 190: The Effects of Policy-Driven Air Quality Improvements on Children’s Respiratory Health
By Frank Gilliland et al. (January 2017)
Cleaner Air and Improved Lung Health

- Tracked growth in Lung Function in 3 “cohorts” (2,100 children total) in Southern California 1994 – 2011
- Reported notable improvement in lung function, symptoms in the most recent cohort (who grew up 2007 – 2011 in cleaner air)
- Still some questions about other differences in the 3 cohorts (e.g. more Hispanics in the latest one)
  - And not possible to isolate a specific action that had an effect
  - But overall strong relationship
Did the study confirm cause and effect?

• There *were* measurable and substantial improvements in air quality
  – Though difficult to isolate specific actions and their affects on air quality

• And there were carefully-measured improvements in children’s lung health for those growing up in the cleaner air
Air Quality and Health Benefits of Regulations:  
A Detailed Accounting

Lucas Henneman, Joe Abrams, Cong Liu, Mitch Klein, Jim Mulholland, Yongtao Hu, Talat Odman, Huizhong Shen, Abiola Lawal, Stefanie Sarnat, Howard Chang, Matt Strickland, Paige Tolbert, and Armistead (Ted) Russell

Georgia Institute of Technology and Emory University
A key study component: Estimating a counterfactual of what would have happened without the actions

They first found substantial before-after emissions reductions from mobile and EGU sources

EGU emissions (measured)
90% decreases in NO\textsubscript{X} & SO\textsubscript{2} in southeastern U.S.

On-road mobile emissions (modeled)
60-90% decreases in multiple pollutants (NO\textsubscript{X}, SO\textsubscript{2}, PM\textsubscript{2.5}, CO, VOC, EC, OC) in Atlanta, GA

EPA Air Markets Program Data (2016)
Henneman et al. (2017), *Air Quality, Atmosphere, and Health*
But what would have happened *without* the controls?
Source-Specific Counterfactual PM$_{2.5}$ concentrations

- May be able to provide source and regulation-specific emission impacts

All tested regulations counterfactual time series

Russell et al. (2018) HEI report
Increased Avoided Emergency Department Visits across Time (especially for asthma)
Did the study confirm cause and effect?

- There *were* measurable and substantial improvements in air quality
  - Some indication of different effects from different source controls, but difficult to isolate specific actions and their affects on air quality and health
- And there were estimates of “avoided” emergency department visits as a result
  - Though results depend to some extent on the accuracy of the counterfactual estimates
So what do Accountability/Intervention Studies Add?

- They offer real world opportunities to test whether changes in air pollution result in changes in health.
- If well-designed – control populations, well-defined health outcomes, etc. – they can offer significant insight on cause and effect.
- You need to find the “sweet spot” where an action has caused a significant change in air quality; coincident with a good health data set.
- More to be done!
Paths to Enlightenment....

There are several ways forward...

- **Systematic Literature Review and Synthesis**
  - Drawing on all lines of inquiry

- **Accountability Research**
  - Testing whether, after an intervention reduces exposure, there is evidence of health improvement

- **Causal Inference Statistical Methods**
  - Rigorous *a priori* specification of assumptions and analytic approach
“Causal Inference” Methods

- Using data to learn about consequences of specific actions
- A “cause” is an action that either occurs or does not
- An “effect” is a consequence of the explicitly defined cause
- They bring an analytic perspective to:
  - frame scientific questions to be answered with observed data
  - clarify common threats to validity of observational (and other) studies
  - increase transparency of assumptions required to infer causal relationships

Slides Courtesy Zigler, Dominici 2015
A causal inference approach to accountability

HEI report examining what happened in western state nonattainment areas – and how that compares to what would have happened without EPA designating nonattainment

**Framing As Hypothetical Experiment**

- **“Treatment:”** Initial nonattainment designations for PM\(_{10}\) NAAQS following 1990 Clean Air Act Amendments.
- **“Control:”** EPA takes no nonattainment action (hypothetical).

**Question:** What is the causal effect (on health outcomes, pollution, etc. ...) of the initial PM\(_{10}\) nonattainment designations vs. what would have happened if the designations had not occurred?
KEY LESSONS:  
Design Is More Important Than Analysis!

Judging a study’s potential for estimating a causal effect should pertain primarily to the “design” and how well it approximates a randomized study

- What are the “treatment” groups being compared?
- What is the effect of interest?
- How is a “control” group constructed?

More Important Than

- What statistical model was used to estimate, provide standard errors, etc.?

Slides Courtesy Zigler, Dominici 2015
Key Lesson

Constructing an Adequate Control Group

Identify observations that can accurately reflect what *would have happened* under the alternative action

- Based on space
  - E.g., areas with low pollution as “controls” for areas with high pollution
  - E.g., areas where an intervention did not occur as “controls” for areas where one did
- Based on time
  - E.g., Observations before an intervention as “controls” for observations after
- Based on time and space

Slides Courtesy Zigler, Dominici 2015
KEY LESSON

Good Control Group = Adjusting for Confounding

• The most common threat to validity in observational studies is confounding
• Poor “design” can lead to confounding IF
  – the “Control group” actually differs on some important factor
  – outcome differences may be partially due to the confounder, not the “treatment”
  – it doesn’t accurately reflect what would have happened without the action

Slides Courtesy Zigler, Dominici 2015
So What Can Causal Inference Add?

• There’s no question that the explicit consideration of, and transparency about, major assumptions is a valuable step forward for epidemiology-type studies.

• Well-designed studies can add insight that can contribute to causality determinations needed for public agency health hazard assessment.

• However, at the end of the day, they are still epidemiology studies, with:
  – the attendant strengths (these are people in real world settings) and
  – limitations (don’t need to list those all here...)
Paths to Causality Enlightenment....

There are several ways forward...

- **Systematic Literature Review and Synthesis**
  - Drawing on all lines of inquiry
- **Accountability Research**
  - Testing whether, after an intervention reduces exposure, there is evidence of health improvement
- **Causal Inference Statistical Methods**
  - Rigorous *a priori* specification of assumptions and analytic approach
So can we link cause and effect?

• The “Holy Grail” is elusive, but...

• Systematic Review, can draw on all lines of scientific evidence
  – With opportunities to enhance how it is carried out

• Accountability studies can help ask:
  – Did the actions “cause” the targeted improvements in AQ and health?
  – Could the actions have been designed better, or better implemented?
  – Regulatory changes often overlap with (many) other changes and trends
    (e.g. changes in health care and status) which have to be separated

• Causal Inference still in early stages
  – But has the opportunity to foster much more rigorous and transparent
    epidemiology

• And we at HEI hope to further add to this developing
  field with our newest RFA...
Thank you

Questions? Contact:

Dan Greenbaum
dgreenbaum@healtheffects.org