AGAINST ALL ODDS
EPISODE 14 – "THE QUESTION OF CAUSATION"
TRANSCRIPT

FUNDER CREDITS

Funding for this program is provided by Annenberg Learner.

INTRO

Pardis Sabeti
Hello, I'm Pardis Sabeti and this is *Against All Odds*, where we make statistics count.

Causation in statistics can be a tricky thing because appearances can often be deceiving. Although we might observe a strong association between two variables, that doesn't automatically mean there's a cause and effect relationship.

One of the biggest challenges in attempting to prove causation comes from hidden factors. They might not immediately be apparent; they just lurk in the background. And that's exactly what we call them: lurking variables.

Take, for example, a study that found people who owned two cars lived longer than people who owned only one car. Those who owned three cars were even more likely to live a longer life, and so on. Does this mean buying more cars will allow you to avoid death? No, of course not. In this case, the lurking variable is the car buyer's affluence: richer individuals own more cars AND tend to live longer, probably because they have access to better medical care and healthier food. The cars have nothing to do with it.

Of course, sometimes causation seems to be the only reasonable explanation of the relationship between an explanatory variable and a response variable. We can see this in the case of smoking and lung cancer.

> **Winston Commercial**
> Brought to you by, Winston!

Pardis Sabeti
Believe it or not, there was a time when cartoons sold cigarettes on TV…

> **Fred Flintstone**
> Winston tastes good, like a cigarette should!

Pardis Sabeti
People smoked in hospitals, and smokers didn't give a second thought to the health risks they might be taking. A far cry from today, when anyone can tell you smoking causes lung cancer; it's even printed on cigarette packs.

But how can we say for sure that smoking causes lung cancer? Yes, there's a strong association between the two, but how do we know there are no lurking variables affecting our conclusions?

The best way to make a case for causation is to conduct an experiment. An experiment imposes a treatment in order to observe its effects. In the case of smoking, we could get a local hospital and randomly assign newborn babies to one of two groups: those we force to smoke, and those we prevent from smoking. We could then keep our subjects in isolation and, as they get older, compare the cancer rates between the two groups. The only difference between the groups would be their smoking habits.

Of course, this particular experiment is highly unethical and logistically impossible. So, if we can't conduct an experiment, how do we prove a causal relationship between smoking and lung cancer? The answer to that question was long in coming, but offers a fascinating look at biostatistical research.

Cigarette smoking became increasingly popular in America after WWI when cigarettes were handed out to soldiers to boost morale. Per capita consumption rose from 49 cigarettes in 1900, to 611 cigarettes in 1920. Along with a rise in smoking came a disturbing rise in lung cancer rates.

A handful of doctors began to raise early warnings about the dangers of smoking. One was Richard Overholt, a young lung surgeon who noticed that his nonsmoking patients had quicker surgical recoveries and a higher survival rates than the smokers. He tried to get other doctors to encourage their patients to quit.

> **Richard Overholt**
> A good share of the doctors were smokers, they didn't believe this, and it was very difficult to convince people that this was doing damage to them because they felt that they were getting pleasure from smoking.

Pardis Sabeti
During the 1930s and '40s the popularity of cigarette smoking soared as a new image of the sophisticated, independent, and fun-loving smoker took hold. Little thought was given to the risk. But in the early 1940s, new studies sounded a louder alarm on the dangers of smoking.

One of the earliest and most compelling was a retrospective study conducted by Ernst Wynder and his teacher Evarts Graham.

Subjects were separated into two groups: those with cancer and those without. Their health status was treated as a categorical variable, with one of two possible values: cancer patient or non-cancer patient. Both groups were asked the same questions about their past habits. Responses were compared to see which habits distinguished people with cancer from those without. Smoking stood out.

Dwight Harken
Dr. Wynder, then a student, in St. Louis under Dr. Evarts Graham, came to me and he said, “Camel cigarettes cause cancer of the lung.” I couldn’t believe it. And you know, you see what you look for and you look for what you know, and it never occurred to me that cigarettes caused cancer. So we went to see my patients, and at that time I had quite a large practice. We discovered to our amazement that patients who had cancer of the lung were 17 times to 1 as apt to be two-pack-a-day smokers. So here was a fact trying to tell us something.

Pardis Sabeti
Despite the remarkable discrepancy in smoking habits between the two groups of patients, this so-called retrospective study was not good enough. Because the study looked at past behavior, behavior it could not control, it’s possible that the lung cancer was due to any number of lurking variables.

What if, for example, something in the smokers’ genetics predisposed them to smoke and also made them more inclined to develop lung cancer? In this case, it wouldn’t be causation, but rather a common cause: the smokers’ DNA.

Or, what if the smokers lived in heavily polluted environments that might cause lung cancer? Their smoking habits would be confounded, or tangled up, with pollution, making it impossible to discern the true cause of the disease.

Or, perhaps it was a coincidence that their cancer patients just happened to be smokers. After all, it was a popular habit at the time.

Common cause, confounding factors, coincidence – these are all alternative explanations for a strong association between two factors that don’t have a causal relationship.

The next step in solving this epidemiological mystery was setting up prospective studies. Doctors Hammond and Horn of the American Cancer Society gave about 200,000 people a smoking questionnaire and followed them for four years. Unlike a retrospective study, which begins with sick people – cancer patients – and works backwards to examine their habits; a prospective study looks ahead, following healthy people – both smokers and non-smokers – forward through time to see which ones develop lung cancer. It’s still not quite an experiment, since people make their own choices about whether or not to smoke, rather than being randomly assigned to take up the habit– but it brings science one step closer to nailing down evidence of causation between smoking and lung cancer.

Lawrence Garfinkel
The preliminary study published in 1954 caused quite a sensation. It was the largest study on smoking that had been done, it was the first prospective study that had been published, and it showed that people who

smoke cigarettes have a lung cancer rate ten times as high as people who never smoke. And that the risk goes up proportionally according to the amount that—the number of cigarettes that people smoke per day.

Pardis Sabeti
This high correlation was consistent with the findings of the retrospective studies. But there was still disagreement about how to explain the correlation. Remember those pesky lurking variables could be present, making the association only appear strong.

To tackle this problem, doctors designed a follow-up study of one million people. The smokers and nonsmokers were matched on such variables as age, geography and occupation. The goal was to make smoking habits the only major difference between the two groups.

Lawrence Garfinkel
And we matched them on 19 different variables so they were as alike as we could make it. And when we finished this analysis, the death rate of the smokers was higher than the death rate of the nonsmokers.

Pardis Sabeti
Even though the two groups were as similar as possible, they were not identical. Sir Ronald Fisher, known by many as the “Father of Modern Statistics,” argued that, without a controlled experiment, there was no way to know for sure whether the lung cancer was due to smoking or to another unknown variable. So despite the accumulating evidence from dozens of studies, for many scientists and statisticians, the labeling of causation was still premature.

So, they turned to the lab for help. Throughout the ‘50s and ‘60s, biochemical and animal studies provided compelling supporting evidence.

Researchers painted tar on mouse skin and rabbit ears, causing cancerous tumors to develop. They also identified carcinogens in tobacco smoke. And lab work confirmed another aspect of the prospective studies: the dose-response relationship. By exposing hamsters to cigarette smoke, it was possible to determine how higher numbers of cigarettes smoked is related to higher incidences of lung cancer.

Ernst Wynder
So you have carcinogen tobacco smoke, it is proved to be carcinogenic to animal tissue, it is highly correlated in retrospective and prospective studies in respect to cancer of the lung in humans, and that was more than sufficient to establish tobacco smoke as a cause of lung cancer and subsequently other cancers in humans.

Pardis Sabeti
For some, the evidence was sufficient. But the struggle to prove causation to a nation of smokers was not over. Rich, powerful tobacco companies did not ignore the growing evidence that cigarettes posed a serious risk to health. Instead, they set out to reassure the American public that their brands were safe through strategic advertising that featured doctors' endorsements and claims of being less irritating to consumers' throats.

In 1962, the Surgeon General assembled an advisory committee on smoking and health. The goal was to take on the question of causation. Its task was formidable: to review more than 6,000 studies on smoking.

Based on pre-determined criteria, the committee concluded that the association between cigarette smoking and lung cancer was found in many studies with many different groups of people.

The association was very strong.

Cancer regularly followed smoking in time.

Smoke did contain cancer-causing substances.

And no explanation other than causation was plausible.

> **Luther Terry**
> The committee states on page 61 of the report, and I quote, in view of the continuing and mounting evidence from many sources, it is the judgment of the committee that cigarette smoking contributes substantially to mortality from certain specific diseases and to the overall death rate.

Pardis Sabeti
In the wake of the surgeon general's report, millions of Americans quit smoking. Congress passed a law requiring warnings on cigarette packs. And the antismoking movement as we know it was born. The scientific consensus was in: Smoking does cause lung cancer.

Now you can see just how hard it is to prove causation when direct experiments aren't possible. Here, the non-experimental evidence is about as strong as it gets. Remember, not all associations can be explained by cause and effect. But some can. With enough evidence from enough different sources – and the help of statistics – we can be confident in our conclusions. The public seems to be getting the message, too. Under 20% of Americans smoke these days – that's less than half the proportion who did at smoking's peak back in the 20th century.

I'm Pardis Sabeti for *Against All Odds*. Stay tuned!

PRODUCTION CREDITS

Host – Dr. Pardis Sabeti

Writer/Producer/Director – Maggie Villiger

Associate Producer – Katharine Duffy

Editors – Brian Truglio
-- Jared Morris
-- Seth Bender

Director of Photography – Dan Lyons

Sound Mix – Richard Bock

Animation – Jason Tierney

Title Animation – Jeremy Angier

Web + Interactive Developer – Matt Denault / Azility, Inc.

Website Designer – Dana Busch

Teleprompter – Sue Willard-Kiess

Hair/Makeup – Emily Damron

Additional Footage and Stills

- iStock/mrod
- iStock/philpell
- Kashif Mardani
- CDC, Office on Smoking & Health
- National Archives & Records Administration; Historical Films, 111-H-1518
- Library of Congress; Chronicling America: Historic American Newspapers Collection
- National Library of Medicine
- T3Media
- Library of Congress:
    - Harris & Ewing Collection, LC-DIG-hec-20150
    - Prints & Photographs; LC-USZ62-113367, LC-USZ62-73952
    - Tobacco Collection, D0308
    - National Photo Company Collection, LC-USZ62-106964
    - Cartoon Drawings: Swann Collection of Caricature & Cartoon; LC-DIG-ppmsca-01583, LC-DIG-ppmsca-01589
    - National Photo Company Collection, LC-DIG-npcc-07666

- Prelinger Archives
- UCSF Legacy Tobacco Documents Library
- American Lung Association

Music
DeWolfe Music Library

Based on the original Annenberg/CPB series *Against All Odds*,
Executive Producer Joe Blatt

Annenberg Learner Program Officer – Michele McLeod

Project Manager – Dr. Sol Garfunkel

Chief Content Advisor – Dr. Marsha Davis

Executive Producer – Graham Chedd

FUNDER CREDITS

Funding for this program is provided by Annenberg Learner.

For information about this, and other Annenberg Learner programs, call 1-800-LEARNER, and visit us at www.learner.org.