Why the Father of Modern Statistics Didn't Believe Smoking Caused Cancer
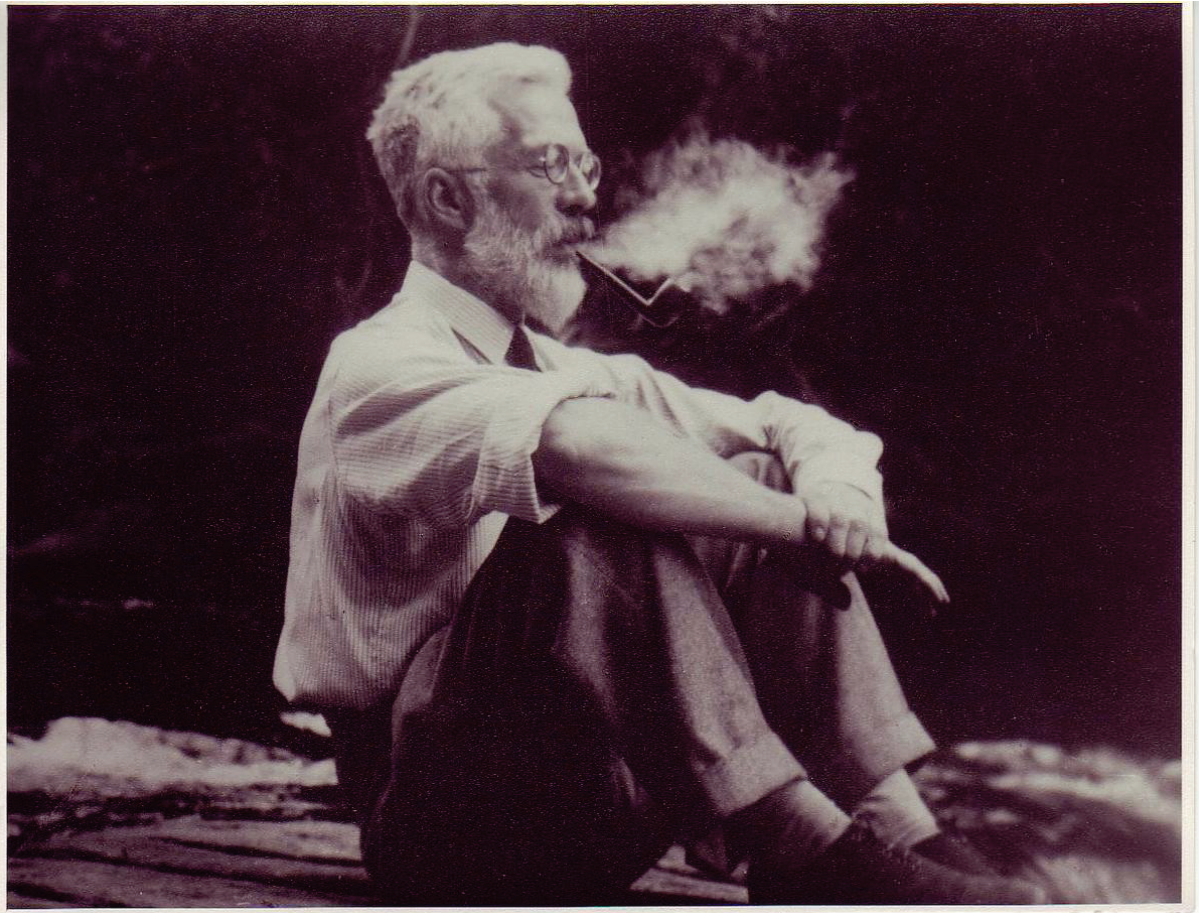
By **Ben Christopher** · **49,340 views** · **More stats**

Ronald A. Fisher, father of modern statistics, enjoying his pipe.

***

In the summer of 1957, Ronald Fisher, one of the fathers of modern statistics, sat down to write a strongly worded letter in defense of tobacco.

The letter was addressed to the *British Medical Journal*, which, just a few weeks earlier, had taken the editorial position that smoking cigarettes causes lung cancer. According to the journal's editorial board, the time for amassing evidence and analyzing data was over. Now, they wrote, "all the modern devices of publicity" should be used to inform the public about the perils of tobacco.

According to Fisher, this was nothing short of statistically illiterate fear mongering. Surely the danger posed to the smoking masses was "not the mild and soothing weed," he wrote, "but the organized creation of states of frantic alarm."

Fisher was a well-known hothead (and an inveterate pipe smoker), but the letter and the resulting debate, which lasted until his death in 1962, was taken as a serious critique by the scientific community. After all, Ronald A. Fisher had spent much of his career devising ways to mathematically evaluate causal claims—claims exactly like the one that the *British Medical Journal* was making about smoking and cancer. Along the way, he had revolutionized the way that biological scientists conduct experiments and analyze data.

And yet we know how this debate ends. On one of the most significant public health questions of the 20th century, Fisher got it wrong.

But while Fisher was wrong about the specifics, there is still debate over whether he was wrong about the *statistics*. Fisher never objected to the possibility that smoking caused cancer, only the certainty with which public health advocates asserted this conclusion.

"None think that the matter is already settled," he insisted in his letter to the *British Medical Journal*. "Is not the matter serious enough to require more serious treatment?"

The debate over smoking is now over. But on issues ranging from public health, to education, to economics, to climate change, researchers and policymakers still don't always agree on what sufficiently "serious treatment" looks like.

How can we ever confidently claim that A *causes* B? How do we weigh the costs of acting too early against the costs of acting too late? And when are we willing to set aside any lingering doubts, stop arguing, and act?

Big Ideas and Bad Blood

Ronald Fisher was known for being remarkably bright and remarkably difficult. The two qualities were probably not unrelated.

According to the writer and mathematician David Salsburg, who describes the history of 20th century statistics in his book, *The Lady Tasting Tea*, Fisher was often frustrated by those who failed to see the world as he did.

Few did.

At the age of seven, Fisher, sickly, shortsighted, and with few friends, began attending academic astronomy lectures. As an undergraduate at Cambridge, he published his first scientific paper, in which he introduced a novel technique to estimate unknown characteristics of a population. The concept, called maximum likelihood estimation, has been called "one of the most important developments in 20th century statistics."

A few years later, he set to work on a statistical problem that Karl Pearson, then the most prodigious statistician in England, had been trying to solve for decades. The question related to the challenge of estimating how different variables (for example, rainfall and crop yields) are related to one another when a researcher only has a small sample of data to work with. Pearson's

inquiry focused on how estimates of this correlation might differ from the actual correlation, but because the mathematics was so complex, he focused on only a handful of cases. After a week of work, Fisher solved the problem for *all* cases. Pearson initially refused to publish the paper in his statistical journal, *Biometrika,* because he did not fully understand it.

"The insights were so obvious to [Fisher] that he often failed to make them understandable to others," writes Salsburg. "Other mathematicians would spend months or even years trying to prove something that Fisher claimed was obvious."

It is no surprise that Fisher was not always popular with his colleagues.

Though Pearson eventually agreed to publish Fisher's paper, he printed it as a footnote to a much longer work that Pearson himself had written. Thus began a feud between the two men that would last beyond Pearson's death. When Pearson's son, Egon, became a prominent statistician, the Fisher-Pearson rivalry continued.

The Pearson family, including Karl and Egon. Source: National Portrait Gallery.

As one observer put it, Fisher had "a remarkable talent for polemic," and his professional debates often got personal. When the Polish mathematician, Jerzy Neyman, presented on his work to the Royal Statistical Society, Fisher opened up the post-lecture discussion with a dig: while he had hoped, he said, that Neyman would present "on a subject with which the author was fully acquainted, and on which he could speak with authority," he was sadly mistaken.

Though Fisher's cantankerous style "effectively froze [him] out of the mainstream of mathematical and statistical research," writes Salsburg, he nevertheless made his mark on the field.

After his falling out with the older Pearson, Fisher took a position at the Rothamsted Agricultural Experiment Station north of London in 1919. It was here that he helped introduce the concept of randomization as one of the most important tools in scientific experiments.

Up until then, the research station had been studying the effectiveness of different fertilizers by applying various chemical agents to separate tracts of land on a field-by-field basis. Field A would receive Fertilizer 1, Field B would receive Fertilizer 2, and so on.

But as Fisher pointed out, this type of experimentation was doomed to produce meaningless results. If the crops in Field A grew better than those in Field B, was that because Fertilizer 1 was better than Fertilizer 2? Or did Field A just happen to have richer soil?

As Fisher put it, the effect of the fertilizer was "confounded" by the effect of the field itself. Confoundedness made it impossible to say precisely what caused what.

The way around the problem, Fisher concluded, was to apply different fertilizers to different small plots *at random*. While Fertilizer 1 will sometimes be applied to a better plot than Fertilizer 2, if both are applied to enough plots in an arbitrary fashion, the opposite should happen just as often. On the whole, one would expect these differences to cancel out. On average, the soil under Fertilizer 1 ought to look exactly like the soil under Fertilizer 2.

This was big news. By randomizing the experimental treatment, the researcher could more confidently conclude that it really was Fertilizer 1, and not some confounding variable like soil quality, that *caused* the crops to grow taller.

But even if the researcher randomized and found that the various fertilizers seemed to produce different yields, how could she know that those differences were not just the product of random variation? Fisher devised a statistical answer to this question. He called the method "analysis of variance," which has since been shorthanded ANOVA. According to Salsburg, it is "probably the single most important tool of biological science."

The entrance to Rothamsted Experimental Farm, where Ronald Fisher introduced the principle of randomization. Source: [Jack Hill](#).

Fisher published these insights on research techniques in a series of books in the 1920s and 30s, and it had a profound effect on scientific research. From agriculture, to biology, to medicine, researchers suddenly had a mathematically rigorous way to answer one of the most important questions in science: what causes what?

The Case Against Smoking

At around the same time, British public health authorities were becoming increasingly concerned about one causal question in particular.

While most of the afflictions that had been killing British citizens for centuries were trending downward, the result of advances in medicine and sanitation, one disease was killing more and more people each year: carcinoma of the lung.

The figures were staggering. Between 1922 and 1947, the number of deaths attributed to lung cancer increased 15-fold across England and Wales. Similar trends were documented around the world. Everywhere, the primary target of the disease seemed to be men.

What was the cause? Theories abounded. More people than ever were living in large, polluted cities. Cars filled the nation's causeways, belching noxious fumes. Those causeways were increasingly being covered in tar. Advances in X-ray technology allowed for more accurate diagnoses. And, of course, more and more people were smoking cigarettes.

Which of these factors was to blame? All of them? None of them? British society had changed so dramatically and in so many ways since the First World War, it was impossible to identify a single cause. As Fisher would say, there were just too many confounding variables.

In 1947, the British Medical Research Council hired Austin Bradford Hill and Richard Doll to look into the question.

Though Doll was not well known at the time, Hill was an obvious choice. A few years earlier, he had made a name for himself with a pioneering study on the use of antibiotics to treat tuberculosis. Just as Fisher had randomly distributed fertilizer across the fields at Rothamsted, Hill had given out streptomycin to tubercular patients at random while prescribing bed rest to others. Once again, the goal was to make sure that the patients who received one treatment were, on average, identical to those who received the other. Any large difference in outcomes between the two groups had to be the result of the drug. It was medicine's first published randomized control trial.

Despite Hill's groundbreaking work with randomization, the question of whether smoking (or anything else) causes cancer was not one you could ask with a randomized control trial. Not ethically, anyway.

"That would involve taking a group of say 6,000 people, selecting 3,000 at random and forcing them to smoke for 5 years, while forcing the other 3,000 not to smoke for 5 years, and then comparing the incidence of lung cancer in the two groups," says Donald Gillies, an emeritus professor of philosophy of science and mathematics at University College London. "Clearly this could not be done, so, in this example, one has to rely on other types of evidence."

Hill and Doll tried to find that evidence in the hospitals of London. They tracked down over 1,400 patients, half of whom were suffering from lung cancer, the other half of whom had been hospitalized for other reasons. Then, as Doll later told the BBC, "we asked them every question we could think of."

These questions covered their medical and family histories, their jobs, their hobbies, where they lived, what they ate, and any other factor that might possibly be related to lung cancer. The two epidemiologists were shooting in the dark. The hope was that one of the many questions would touch on a trait or behavior that was common among the lung cancer patients and rare among those in the control group.

At the beginning of the study, Doll had his own theory.

"I personally thought it was tarring of the roads," Doll said. But as the results began to come in, a different pattern emerged. "I gave up smoking two-thirds of the way though the study."

Hill and Doll published their results in the *British Medical Journal* in September of 1950. The findings were alarming, but not conclusive. Though the study found that smokers were more likely than non-smokers to have lung cancer, and that the prevalence of the disease rose with the quantity smoked, the design of the study still left room for Fisher's dreaded "confounding" problem.

The problem was in the selection of the control. Hill and Doll had picked a comparison group that resembled the lung cancer patients in age, sex, approximate residence, and social class. But did this cover the entire list of possible confounders? Was there some other trait, forgotten or invisible, that the two researchers had failed to ask about?

Richard Doll, who quit smoking in 1950. Source: General Motors Cancer Research Foundation.

To get around this problem, Hill and Doll designed a study where they wouldn't have to choose a control group at all. Instead, the two researchers surveyed over 30,000 doctors across England. These doctors were asked about their smoking habits and medical histories. And then Hill and Doll waited to see which doctors would die first.

By 1954, a familiar pattern began to emerge. Among the British doctors, 36 had died of lung cancer. All of them had been smokers. Once again, the death rate increased with the rate of smoking.

The "British Doctor Study" had a distinct advantage over the earlier survey of patients. Here, the researchers could show a clear "this then that" relationship (what medical researchers call a "dose-response"). Some doctors smoked more than others in 1951. By 1954, more of those doctors were dead.

The back-to-back Doll and Hill studies were notable for their scope, but they were not the only ones to find a consistent connection between smoking and lung cancer. Around the same time, the American epidemiologists, E. C. Hammond and Daniel Horn conducted a study very similar to the Hill and Doll survey of British doctors.

Their results were remarkably consistent. In 1957, the Medical Research Council and the *British Medical Journal* decided that enough evidence had been gathered. Citing Doll and Hill, the journal declared that "the most reasonable interpretation of this evidence is that the relationship is one of direct cause and effect."

Ronald Fisher begged to differ.

Just Asking Questions

In some ways, the timing was perfect. In 1957, Fisher had just retired and was looking for a place to direct his considerable intellect and condescension.

Neither the first nor the last retiree to start a flame war, Fisher launched his opening salvo by questioning the certainty with which the *British Medical Journal* had declared the argument over.

"A good *prima facie* case had been made for further investigation," he wrote. "The further investigation seems, however, to have degenerated into the making of more confident exclamations."

The first letter was followed by a second and then a third. In 1959, Fisher amassed these missives into a book. He denounced his colleagues for manufacturing anti-smoking "propaganda." He accused Hill and Doll of suppressing contrary evidence. He hit the lecture circuit, relishing the opportunity to once again hold forth before the statistical establishment and to be, in the words of his daughter, "deliberately provocative."

Provocation aside, Fisher's critique came down to the same statistical problem that he had been tackling since his days at Rothamsted: confounding variables. He did not dispute that smoking and lung cancer tended to rise and fall together—that is, that they were correlated. But Hill and Doll and the entire British medical establishment had committed "an error…of an old kind, in arguing from correlation to causation," he wrote in a letter to *Nature.*

Most researchers had evaluated the association between smoking and cancer and concluded that the former caused the latter. But what if the opposite were true?

What if the development of acute lung cancer was preceded by an undiagnosed "chronic inflammation," he wrote. And what if this inflammation led to a mild discomfort, but no conscious pain? If that were the case, wrote Fisher, then one would expect those suffering from pre-diagnosed lung cancer to turn to cigarettes for relief. And here was the *British Medical Journal* suggesting that smoking be banned in movie theaters!

"To take the poor chap's cigarettes away from him," he wrote, "would be rather like taking away [the] white stick from a blind man."

The soothing properties of cigarettes was a common feature of mid-century tobacco advertising. This ad was from 1930. Source: Silberio77.

If that particular explanation seems like a stretch, Fisher offered another. If smoking doesn't cause cancer and cancer doesn't cause smoking, then perhaps a third factor causes both. Genetics struck him as a possibility.

To make this case, Fisher gathered data on identical twins in Germany and showed that twin siblings were more likely to mimic one another's smoking habits. Perhaps, Fisher speculated, certain people were genetically predisposed to crave of cigarettes.

Was there a similar familial pattern for lung cancer? Did these two predispositions come from the same hereditary trait? At the very least, researchers ought to look into this possibility before advising people to toss out their cigarettes.

And yet nobody was.

"Unfortunately, considerable propaganda is now being developed to convince the public that cigarette smoking is dangerous," he wrote. "It is perhaps natural that efforts should be made to discredit evidence which suggests a different view."

Though Fisher was in the minority, he was not alone in taking this "different view." Joseph Berkson, the chief statistician at the Mayo Clinic throughout the 1940s and 50s, was also a prominent skeptic on the smoking-cancer question, as was Charles Cameron, president of the American Cancer Society. For a time, many of Fisher's peers in academic statistics, including Jerzy Neyman, questioned the validity of a causal claim. But before long, the majority buckled under the weight of mounting evidence and overwhelming consensus.

But not Fisher. He died in 1962 (of cancer, though not of the lung). He never conceded the point.

Ulterior Motives

Today, not everyone takes Fisher's position on tobacco at face value.

In his review of the debate, the epidemiologist Paul Stolley lambasts Fisher for being "unwilling to seriously examine the data and to review all the evidence before him to try to reach a judicious conclusion." According to Stolley, Fisher undermined Hill and Doll's conclusions by cherry picking irregular findings and blowing them out of proportion. His use of the German twin data was either misguided or willfully misleading. Fisher, he writes, "sounds like a man with 'an axe to grind.'"

Others have offered less charitable interpretations.

In 1958, Fisher approached the British hematologist and geneticist Arthur Mourant, and proposed a joint research project to evaluate possible genetic differences between smokers and non-smokers. Mourant turned Fisher down and later speculated that the statistician's "obsession" with the topic "was the first sign of deterioration of an exceptionally brilliant mind."

Even worse was the suggestion that his skepticism had been bought. The Tobacco Manufacturers' Committee had agreed to fund Fisher's research into possible genetic causes of both smoking and lung cancer. And though it seems unlikely that a man who routinely insulted his peers and jeopardized his career in order to prove that he was right would sell his professional opinion at such an old age, some still regard him as doing so.

If Fisher wasn't swayed by money, it seems more likely that he was influenced by politics.

Throughout his life, Fisher was an unflinching reactionary. In 1911, while studying at Cambridge, he helped found the university's Eugenics Society. Though many well-educated English men of the day embraced this ideology, Fisher took to the issue with an unusual fervency. Throughout his career, he intermittently wrote papers on the subject. A particular

concern of Fisher's was that upper class families were having fewer children than their poorer and less educated counterparts. At one point, he suggested that the government pay "intelligent" couples to procreate. Fisher and his wife had eight children.

These political leanings may have colored his views on smoking.

"Fisher was a political conservative and an elitist," writes Paul Stolley. "Fisher was upset by the public health response to the dangers of smoking not only because he felt that the supporting data were weak, but also due to his holding certain ideological objections to mass public health campaigns."

If he were alive today, Ronald Fisher would have one hell of a Twitter account.

When Does Correlation Imply Causation?

Whatever his actual motives, it's hardly surprising that Fisher was drawn to this particular fight.
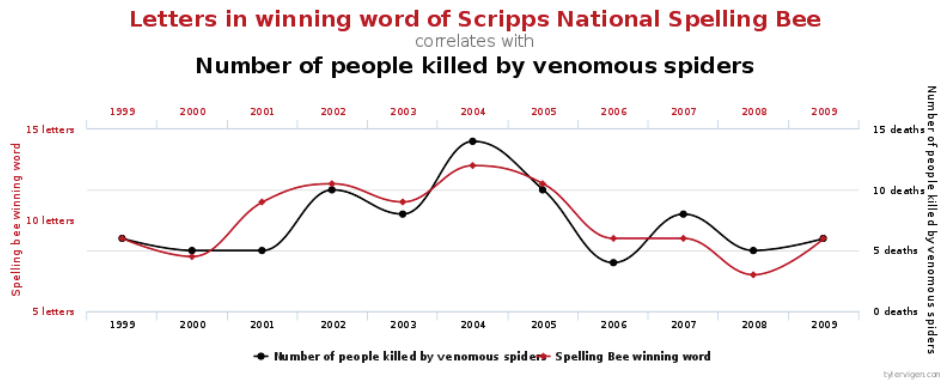
This was the man who had built his career on properly conducting research to avoid the pitfalls of confoundedness and to show, with mathematical precision, where correlation did and did *not* imply causation.

The fact that a younger generation of public health investigators (and much of the press) had come to a conclusion of such import without following the rules of causal inference that Fisher himself had established must have been infuriating. Fisher himself acknowledge that it would be impossible to conduct a randomized control trial on smoking. "It is not the fault of Hill or Doll or Hammond that they cannot produce evidence in which a thousand children of teen age have been laid under a ban that they shall never smoke," wrote Fisher. "And a thousand more chosen at random from the same age group have been under compulsion to smoke at least thirty cigarettes a day." But if scientists must depart from this gold standard of experimental design, he argued, they should consider every last explanation.

To a certain extent, this debate is ongoing.

"Nearly everyone now accepts that Fisher was wrong, but there are still contemporary issues of a similar character about which there is a great deal of debate," says Donald Gillies from University College London. "What causes obesity? What dietary factors, if any, cause coronary heart disease and diabetes?"

Add to that list countless debates over education (Does more spending on schools result in better education?), climate change (Do higher emissions cause global warming?), crime and punishment (Do harsher penalties lead to fewer crimes?), and some of the most basic facts of everyday life (Is flossing good for your teeth? Does coffee cause cancer? Does coffee *prevent* cancer?).

Letters in winning word of Scripps National Spelling Bee
correlates with
Number of people killed by venomous spiders

Correlation does not always imply causation. Source: Tyler Viglen.

Though the randomized control trial is seen as the gold standard for untangling correlation from causation, practicality and ethics often mean we have to make do, says Dennis Cook, a professor of statistics at the University of Minnesota. We make judgment calls.

"There has to be a balance," he says.

Cook recalls a popular study which made headlines a few years ago that found a statistically significant association between cranberry consumption and cancer. Should we as a society have implemented a cranberry ban?

"Fisher's point was you can't make that decision as a knee-jerk reaction," says Cook. "Some of the decisions, if you make them based on a knee-jerk reaction, are going to be right. Smoking would have been right. But others, like cranberries, are going to be wrong."

One of Ronald Fisher's most important contributions to modern statistics is the concept of the "null hypothesis." This is the philosophical starting point in any statistical test, the presumption that, in the absence of better evidence, you should not change your mind. When in doubt, assume that the fertilizer hasn't worked, that the antibiotic has had no effect, and that smoking does not cause cancer. This reluctance to "reject the null" imbues in science an inherent conservatism which keeps established knowledge from gyrating wildly with each new study about cranberries.

But it can also come at a steep cost.

In 1965, three years after Fisher's death, Austin Bradford Hill, by then a knight and a professor emeritus, gave a speech to the Royal Society of Medicine. In it he outlined a list of criteria that should be considered before establishing that one thing is the cause of another.

But more importantly, he wrote, none of these criteria should be considered sacrosanct. "Hard-and-fast" statistical rules do not clear away all uncertainty. They just help informed, well intentioned people make the best decisions that they can.

"All scientific work is incomplete," he said. "All scientific work is liable to be upset or modified by advancing knowledge. That does not confer upon us a freedom to ignore the knowledge we already have, or to postpone the action that it appears to demand at a given time."

Ronald Fisher had devised an ingenious way to separate correlation from cause. But waiting for absolute proof always comes at a price.

Source: https://priceonomics.com/why-the-father-of-modern-statistics-didnt-believe/