

## On types of scientific inquiry: The role of qualitative reasoning

David A. Freedman  
Statistics Department  
UC Berkeley, CA 94720-3860

One type of scientific inquiry involves the analysis of large data sets, often using statistical models and formal tests of hypotheses. A moment's thought, however, shows that there must be other types of scientific inquiry. For instance, something has to be done to answer questions like the following. How should a study be designed? What sorts of data should be collected? What kind of a model is needed? Which hypotheses should be formulated in terms of the model and then tested against the data?

The answers to these questions frequently turn on observations, qualitative or quantitative, that give crucial insights into the causal processes of interest. Such observations generate a line of scientific inquiry, or markedly shift the direction of the inquiry by overturning prior hypotheses, or provide striking evidence to confirm hypotheses. They may well stand on their own rather than being subsumed under the systematic data collection and modeling activities mentioned above.

Such observations have come to be called "Causal Process Observations" (CPOs). These are contrasted with the "Data Set Observations" (DSOs) that are grist for statistical modeling (Brady and Collier 2004). My object in this essay is to illustrate the role played by CPOs, and qualitative reasoning more generally, in a series of well-known episodes drawn from the history of medicine.

Why is the history of medicine relevant to us today? For one thing, medical researchers frequently confront observational data that present familiar challenges to causal inference. For another, distance lends perspective, allowing gains and losses to be more sharply delineated. The examples show that an impressive degree of rigor can be obtained by combining qualitative reasoning, quantitative analysis, and experiments when those are feasible. The examples also show that great work can be done by spotting anomalies, and trying to understand them.

### Jenner and Vaccination

The setting is the English countryside in the 1790s. Cowpox, as will be clear from the name, is a disease of cows. The symptoms include sores on the teats. Those who milk the cows often became infected, with sores on their hands; by the standards of the time, the illness is rarely serious. In contrast, smallpox is one of the great killers of the 18th century.

In 1796, Edward Jenner took some matter from a cowpox sore on the hand of dairymaid Sarah Nelmes, and inserted it into the arm of an eight-year-old boy,

“by means of two superficial incisions, barely penetrating the cutis, each about half an inch long.” The boy was “perceptibly indisposed” on the ninth day, but recovered the following day. Six weeks later, Jenner inoculated him with matter taken from a smallpox pustule, “but no disease followed” (Jenner 1798, Case XVII).

Jenner published 23 case studies to demonstrate the safety and efficacy of “vaccination,” as his procedure came to be called: *vacca* is the Latin term for cow, and *vaccinia* is another term for cowpox. Despite initial opposition, vaccination became standard practice within a few years, and Jenner achieved international fame. By 1978, smallpox had been eradicated.

What led Jenner to try his experiment? The 18th century view of disease was quite different from ours. The great Scottish doctor of the time, William Cullen, taught that most diseases were “caused by external influences—climate, foodstuffs, effluvia, humidity, and so on—and . . . the same external factors could cause different diseases in different individuals, depending on the state of the nervous system” (Porter 1997, 262).

Despite such misconceptions, it was known that smallpox could somehow be communicated from one person to another; moreover a person who contracted smallpox and survived was generally immune to the disease from that point on. As a preventive measure, patients could be deliberately infected (through scratches on the skin) with minute quantities of material taken from smallpox pustules, the idea being to induce a mild case of the disease that would confer immunity later.

This procedure was called “inoculation” or “variolation.” It was not free of risk: serious disease was sometimes caused in the patient, and in people who came into contact with the patient (smallpox is highly contagious). On the other hand, failure to inoculate could easily lead to death from smallpox.

By the early part of the 18th century, variolation had reached England. Jenner was a country doctor who performed variolations. He paid attention to two crucial facts—although these facts were not explicable in terms of the medical knowledge of his time. (i) People who had the cowpox never seemed to contract smallpox afterwards, whether they had been inoculated or not. (ii) Some of his patients who had been ill with cowpox in the past still wanted to be inoculated: such patients reacted very little to inoculation—

“what renders the Cox-pox virus so extremely singular, is, that the person who has been thus affected is for ever after secure from the infection of the Small Pox; neither exposure to the variolous effluvia, nor the insertion of the matter into the skin, producing this distemper” (Jenner 1798, 6).

These two facts led him to a hypothesis: cowpox created immunity against smallpox. That is the hypothesis he tested, observationally and experimentally, as described above. In our terminology, Jenner vaccinated a boy (Case XVII) who showed no response to subsequent inoculation. Immunity to smallpox had been induced by the vaccination.

By “virus,” Jenner probably meant “contagious matter,” that being a standard usage in his time. Viruses in the modern sense were not to be discovered for another century. By a curious twist, smallpox and cowpox are viral diseases in our sense too.

### Semmelweis and Puerperal Fever

The time is 1844 and the place is Vienna. The discovery of microbes as the cause of infectious disease will not be made for some decades. Ignac Semmelweis is an obstetrician in the First Division of the Lying-in Hospital, where medical students are trained. (Midwives are trained in the Second Division.) Pregnant women are admitted to one division or the other, according to the day of the week that they come to the hospital, in strict alternation. Mortality from “puerperal fever” is much higher in the First Division (Semmelweis 1981, 356).

Eventually, Semmelweis discovers the cause. The medical students are doing autopsies, and then examining the “puerperae” (women who are giving birth, or who have just given birth). “Cadaveric particles” are thus transferred to the women, entering the bloodstream and causing infection. In 1847, Semmelweis institutes the practice of disinfection, and mortality plummets (Semmelweis 1981, 393–4).

But how did Semmelweis make his discovery? To begin with, he has to reject conventional explanations, including “epidemic influences,” which meant something different then:

“epidemic influences . . . are to be understood [as] certain hitherto inexplicable, atmospheric, cosmic, telluric changes, which sometimes disseminate themselves over whole countrysides, and produce childbed fever in individuals predisposed thereto by the puerperal state. [“Telluric” means earthly.] Now, if the atmospheric-cosmic-telluric conditions of the City of Vienna are so disposed that they cause puerperal fever in individuals susceptible thereto as puerperae, how does it happen that these atmospheric-cosmic-telluric conditions over such a long period of years have carried off individuals disposed thereto as puerperae in the First Clinic, while they have so strikingly spared others also in Vienna, even in the same building in the Second Division and similarly vulnerable as puerperae?” (Semmelweis 1981, 357).

The reasoning is qualitative; and similar qualitative arguments dispose of other theories—diet, ventilation, use of hospital linens, and so forth.

Now he has to discover the real cause. In 1847, his revered colleague Professor Kolletschka is accidentally cut with a knife used in a medico-legal autopsy. Kolletschka becomes ill, with symptoms remarkably similar to puerperal fever; then he dies. Again, qualitative analysis is crucial. Close attention to symptoms and their progression is used to identify Kolletschka’s illness with puerperal fever (Semmelweis 1981, 391). Tracing of causal processes comes into play as well:

“Day and night this picture of Kolletschka’s disease pursued me . . . I was obliged to acknowledge the identity of the disease, from which Kolletschka died, with that disease of which I saw so many puerperae die. . . . I must acknowledge, if Kolletschka’s disease and the disease from which I saw so many puerperae die, are identical, then in the puerperae it must be produced by the self-same engendering cause, which produced it in Kolletschka. In Kolletschka, the specific agent was cadaveric particles, which were introduced into his vascular system [the bloodstream]. I must ask myself the question: Did the cadaveric particles make their way into the vascular systems of the individuals, whom I had seen die of an identical disease? This question I answer in the affirmative” (Semmelweis 1981, 391–2).

The source of the infectious agent could also be a wound in a living person (Semmelweis 1981, 396). Once the cause is discovered, the remedy is not far away: eliminate the infectious particles from the hands that will examine the puerperae. Washing with soap and water is insufficient, but disinfection with chlorine compounds is sufficient (Semmelweis 1981, 392–6).

Semmelweis’ work was accepted by few of his contemporaries, due in part to his troubled and disputatious personality, although his picture of the disease was essentially correct. Puerperal fever is a generalized infection, typically caused by bacteria in the group *Streptococcus pyogenes*. These bacteria enter the bloodstream through wounds suffered during childbirth (for instance, at the site where the placenta was attached). Puerperal fever can be—and today it generally is—avoided by proper hygiene.

### Snow and Cholera

John Snow was a physician in Victorian London. In 1854, he demonstrated that cholera was an infectious disease, which could be prevented by cleaning up the water supply. The demonstration took advantage of a natural experiment. A large area of London was served by two water companies. The Southwark and Vauxhall company distributed contaminated water, and households served by it had a death rate “between eight and nine times as great as in the houses supplied by the Lambeth company,” which supplied relatively pure water (Snow 1965, 86, data in Table IX).

What led Snow to design the study and undertake the arduous task of data collection? To begin with, he had to reject the explanations of cholera epidemics that were conventional in his time. The predominant theory attributed cholera to “miasmas,” that is, noxious odors—especially odors generated by decaying organic material. Snow makes qualitative arguments against such explanations:

“[Cholera] travels along the great tracks of human intercourse, never going faster than people travel, and generally much more slowly. In extending to a fresh island or continent, it always appears first at a sea-port. It never attacks the crews of ships going from a country free from cholera, to one where the

disease is prevailing, till they have entered a port, or had intercourse with the shore. Its exact progress from town to town cannot always be traced; but it has never appeared except where there has been ample opportunity for it to be conveyed by human intercourse” (Snow 1965, 2).

These phenomena are easily understood if cholera is an infectious disease, but hard to explain on the miasma theory. Similarly,

“The first case of decided Asiatic cholera in London, in the autumn of 1848, was that of a seaman named John Harnold, who had newly arrived by the *Elbe* steamer from Hamburgh, where the disease was prevailing. . . . Now the next case of cholera, in London, occurred in the very room in which the above patient died” (Snow 1965, 3).

The first case was infected in Hamburgh; the second case was infected by contact with dejecta from the first case, on the bedding or other furnishings in that fatal room. The miasma theory, on the other hand, does not provide good explanations.

Careful observation of the disease led to the conclusion “that cholera invariably commences with the affection of the alimentary canal” (Snow 1965, 10). A living organism enters the body, as a contaminant of water or food, multiplies in the body, and creates the symptoms of the disease. Many copies of the organism are expelled from the body with the dejecta, contaminate water or food, then infect other victims. The task is now to prove this hypothesis.

According to Sir Benjamin Ward Richardson, who wrote the introduction to Snow’s book, the decisive proof came during the Broad Street epidemic of 1854:

“[Snow] had fixed his attention on the Broad Street pump as the source and centre of the calamity. He advised the removal of the pump-handle as the grand prescription. The vestry [in charge of the pump] was incredulous, but had the good sense to carry out the advice. The pump-handle was removed and the plague was stayed” (Snow 1965, xxxvi).

The pump-handle as the decisive test is a wonderful fable, which has beguiled many a commentator.

What are the facts? Contamination at the pump did cause the epidemic, Snow recommended closing the pump, his advice was followed, and the epidemic stopped. However, the epidemic was stopping anyway. Closing the pump had no discernible effect: the episode proves little. Snow explains this with great clarity (Snow 1965, 40–55, see esp. Table I on p. 49 and the conclusory paragraph on pp. 51–2). Richardson’s account is therefore a classic instance of post hoc, ergo propter hoc.

The reality is more interesting than the fable. Snow was intimately familiar with the Broad Street area, because of his medical practice. He says,

“As soon as I became acquainted with the situation and extent of this irruption of cholera, I suspected some contamination of the water of the much-frequented street-pump in Broad Street . . . but on examining the water, on

the evening of 3rd September, I found so little impurity in it of an organic nature, that I hesitated to come to a conclusion” (Snow 1965, 38–39).

Snow had access to the death certificates at the General Register Office, and drew up a list of the cholera fatalities registered shortly before his inspection of the pump. He then made a house-to-house canvass (the death certificate shows the address of the deceased), and discovered that the cases clustered around the pump, confirming his suspicion. Later, he made a more complete tally of cholera deaths in the area. His “spot map” displays the locations of cholera fatalities during the epidemic, and the clustering is apparent from the map (Snow 1965, 44–45; Cholera Inquiry Committee 1855, 106–9).

However, there were a number of exceptions that had to be explained. For example, there was a brewery near the pump; none of the workers contracted the disease: why not? First, the workers drank beer; second, if water was desired, there was a pump on the premises (Snow 1965, 42). For another example, a lady in Hampstead contracted cholera: why? As it turned out, she liked the taste of the water from the Broad Street pump, and had it brought to her house (Snow 1965, 44). Snow gives many other such examples.

Snow’s work on the Broad Street epidemic illustrates the power of case studies. His refutation of the usual explanations for cholera, and the development of his own explanation, are other indicators of the power of qualitative reasoning. The analysis of his natural experiment, referred to above, shows the power of simple quantitative methods and good research design. This was the great quantitative test of his theory that cholera was a waterborne infectious disease.

In designing the quantitative study, however, Snow made some key qualitative steps: (i) seeing that conventional theories were wrong, (ii) formulating the water hypothesis, and (iii) noticing that in 1852, the Lambeth company moved its intake pipe to obtain relatively pure water, while Southwark and Vauxhall continued to draw heavily contaminated water. It took real insight to see—a priori rather than a posteriori—that this difference between the companies allowed the crucial study to be done.

Snow’s ideas gained some circulation, especially in England. However, widespread acceptance was achieved only when Robert Koch isolated the causal agent (*Vibrio cholerae*, the comma-shaped bacillus) during the Indian epidemic of 1883. Even then, there were dissenters, with catastrophic results in the Hamburg epidemic of 1892 (Evans 1987).

Inspired by Koch and Louis Pasteur, there was a great burst of activity in microbiology during the 1870s and 1880s. The idea that microscopic life-forms could arise by spontaneous generation was cast aside, and the germ theory of disease was given solid experimental proof. Besides the cholera vibrio, the bacteria responsible for anthrax (*Bacillus anthracis*) and for tuberculosis (*Mycobacterium tuberculosis*) were isolated, and a vaccine was developed against rabies. However, as we shall see in a moment, these triumphs made it harder to solve the

riddle of beriberi. Beriberi is a deficiency disease, but the prestige of the new microbiology made investigators suspicious of any explanation that did not involve bacteria.

### Eijkman and Beriberi

Beriberi was endemic in Asia, from about 1750 until 1930 or so. Today, the cause is known. People need minute amounts (about one part per million in the diet) of a vitamin called “thiamin.” Many Asians eat a diet based on rice, and white rice is preferred to brown.

Thiamin in rice is concentrated in the bran—the skin that gives rice its color. White rice is obtained by polishing away the skin, and with it most of the thiamin; what is left is further degraded by cooking. The diet is then deficient in thiamin, unless supplemented by other foods rich in that substance. Beriberi is the sequel.

In 1888, knowledge about vitamins and deficiency diseases lay decades in the future. That year, Christiaan Eijkman—after studying microbiology with Koch in Berlin—was appointed director of the Dutch Laboratory for Bacteriology and Pathology in the colony of Java, near the city now called Jakarta. His research plan was to show that beriberi was an infectious disease, with Koch’s methods for the proof.

Eijkman tried to infect rabbits and then monkeys with blood drawn from beriberi patients. This was unsuccessful. He then turned to chickens. He tried to infect some of the birds, leaving others as controls. After a time, many of his chickens came down with polyneuritis, which he judged to be very similar to beriberi in humans. (“Polyneuritis” means inflammation of multiple nerves.)

However, the treated chickens and the controls were equally affected. Perhaps the infection spread from the treated chickens to the controls? To minimize cross infection, he housed the treated chickens and the controls separately. That had no effect. Perhaps his whole establishment had become infected? To eliminate this possibility, he started work on another, remote experimental station—at which point, the chickens began recovering from the disease.

“[Eijkman] wrote ‘something struck us that had escaped our attention so far.’ The chickens had been fed a different diet during the five months in which the disease had been developing. In that period (July through November 1889), the man in charge of the chickens had persuaded the cook at the military hospital, without Eijkman being aware of it, to provide him with leftover cooked [white] rice from the previous day, for feeding to the birds. A new cook, who started duty on 21 November, had refused to continue the practice. Thirty years later, Eijkman was to say that ‘[the new cook] had seen no reason to give military rice to civilian hens’ ” (Carpenter 2000, 38).

In short, the chickens became ill when fed cooked, polished rice; they recovered when fed uncooked, unpolished rice. This was an accidental experiment, arranged

by the cooks. One of Eijkman's great insights was paying attention to the results, because the cooks' experiment eventually changed the understanding of beriberi.

Eijkman's colleague Adolphe Vorderman undertook an observational study of prisons, to confirm the relevance to humans. Where prisoners were fed polished rice, beriberi was common; with a diet of unpolished rice, beriberi was uncommon. Beriberi is a deficiency disease, not an infectious disease.

The evidence may seem compelling, but that is because we know the answer. At the time, the picture was far from clear. Eijkman himself thought that white rice was poisonous, the bran containing the antidote. Later, he was to reverse himself: beriberi is an infectious disease, although a poor diet makes people (and chickens) more vulnerable to infection.

In 1896, Gerrit Grijns took over Eijkman's lab (Eijkman suffered from malaria, and had to return to Holland). Among other contributions, after a long series of careful experiments, Grijns concluded that beriberi was a deficiency disease, the missing element in the diet being concentrated in rice bran—and in other foods like mung beans.

In 1901, Grijn's colleague Hulshoff Pol ran a controlled experiment at a mental hospital, showing that mung beans prevented or cured beriberi. In three pavilions out of twelve, the patients were fed mung beans; in three pavilions, other green vegetables. In three pavilions, there was intensive disinfection, and three pavilions were used as controls. The incidence of beriberi was dramatically lower in the pavilions with mung beans.

Still, medical opinion remained divided. Some public health professionals accepted the deficiency hypothesis. Others continued to favor the germ theory, and still others thought the cause was an inanimate poison. It took another ten years or so to reach consensus that beriberi was a deficiency disease, which could be prevented by eating unpolished rice, or enriching the diet in other ways. From a public health perspective, the problem of beriberi might be solved, but the research effort turned to extracting the critical active ingredient in rice bran—no mean challenge, since there is about one teaspoon of thiamin in a ton of bran.

Around 1912, Casimir Funk coined the term "vitamines," later contracted to vitamins, as shorthand for "vital amines." The claim that he succeeded in purifying thiamin may be questionable. But he did guess that beriberi and pellagra were deficiency diseases, which could be prevented by supplying trace amounts of organic nutrients.

By 1926, B. C. P. Jansen and W. F. Donath had succeeded in extracting thiamin (vitamin B1) in pure crystal form. Ten years later, Robert R. Williams and his associates managed to synthesize the compound in the lab. In the 1930s, there were still beriberi cases in the East—and these could be cured by injecting a few milligrams of the new vitamin B1.



### Goldberger and Pellagra

“Pellagra was first observed in Europe in the eighteenth century by a Spanish physician, Gaspar Casal, who found that it was an important cause of ill-health, disability, and premature death among the very poor inhabitants of the Asturias. In the ensuing years, numerous. . . authors described the same condition in northern Italian peasants, particularly those from the plain of Lombardy. By the beginning of the nineteenth century, pellagra had spread across Europe, like a belt, causing the progressive physical and mental deterioration of thousands of people in southwestern France, in Austria, in Rumania, and in the domains of the Turkish Empire. Outside Europe, pellagra was recognized in Egypt and South Africa, and by the first decade of the twentieth century it was rampant in the United States, especially in the south. . .” (Roe 1973, 1).

Pellagra seemed to hit some villages much harder than others. Even within affected villages, many households were spared, but some had pellagra cases year after year. Sanitary conditions in diseased households were primitive: flies were everywhere. One blood-sucking fly (*Simulium*) had the same geographical range as pellagra, at least in Europe; and the fly was most active in the spring, just when most pellagra cases developed. Many epidemiologists concluded the disease was infectious, and—like malaria or yellow fever—was transmitted from one person to another by insects.

Joseph Goldberger was an epidemiologist working for the U. S. Public Health Service. In 1914, he was assigned to work on pellagra. Despite the climate of opinion described above, he designed a series of observational studies and experiments showing that pellagra was caused by a bad diet, and is not infectious. The disease could be prevented or cured by foods rich in what Goldberger called the P-P (pellagra-preventive) factor.

By 1926, he and his associates had tentatively identified the P-P factor as part of the vitamin B complex. By 1937, C. A. Elvehjem and his associates had identified the P-P factor as niacin, also called vitamin B3 (this compound had been discovered by C. Huber around 1870, but its significance had not been recognized). Since 1940, most of the flour sold in the United States has been enriched with niacin, among other vitamins.

Niacin occurs naturally in meat, milk, eggs, some vegetables, and certain grains. Corn, however, contains relatively little niacin. In the pellagra areas, the poor ate corn—and not much else. Some villages and some households were poorer than others, and had even more restricted diets. That is why they were harder hit by the disease. The flies were a marker of poverty, not a cause of pellagra.

What prompted Goldberger to think that pellagra was a deficiency disease rather than an infectious disease? In hospitals and asylums, the inmates frequently developed pellagra, the attendants almost never—which is unlikely if the disease is infectious, because the inmates could infect the attendants. This observation, although far from definitive, set Goldberger on the path to discovering the cause of pellagra and methods for prevention or cure. The qualitative thinking precedes the quantitative investigation. Pellagra is virtually unknown in the developed world today, although it remains prevalent in some particularly poor countries.

### McKay and Fluoridation

“Dental caries is an infectious, communicable, multifactorial disease in which bacteria dissolve the enamel surface of a tooth. . . . Soon after establishing his dental practice in Colorado Springs, Colorado, in 1901, Dr. Frederick S. McKay noted an unusual permanent stain or ‘mottled enamel’ (termed ‘Colorado brown stain’ by area residents) on the teeth of many of his patients. After years of personal field investigations, McKay concluded that an agent in the public water supply probably was responsible for mottled enamel. McKay also observed that teeth affected by this condition seemed less susceptible to dental caries” (Centers for Disease Control 1999, 933; internal citations omitted).

Mottling was caused by something in the drinking water: that was the main hypothesis at the time (McKay and Black 1916, 635). McKay and Black found that mottled teeth were endemic to specific areas. Mottling affected people born in the area, not people who moved to the area after their teeth had been formed. If mottling was prevalent in one area but not in a nearby area, the two areas had different water supplies. These observations supported the water hypothesis, the idea being that the causal agent affects the teeth as they are developing in the body.

McKay and Black could not identify the causal agent in the water, but explained that their chemical analyses—

“were made according to the standard quantitative form. There are present, however, in waters certain other elements of rarer varieties that exist only in traces, the determination of which requires much elaborate technique and spectroscopic and polariscopic tests, which are beyond the capacities of ordinary chemical laboratories” (McKay and Black 1916, 904).

As a consequence of mottling, two towns (Oakley in 1925 and Bauxite in 1928) changed the source of their water supply. After the change, newborn children in those towns developed normal teeth. This is, at least in retrospect, striking confirmation of the water hypothesis (McClure 1970, Chaps. 2–3).

Bauxite was a company town (Aluminum Company of America). H. V. Churchill, an ALCOA chemist, discovered in 1931 that fluorides were naturally

present in the original source—a deep well—at relatively high concentrations: he had a spectrograph at the company lab. McKay and Churchill also found high levels of fluorides in the water at several other towns where mottling was endemic, which suggested that fluorides might cause mottling and prevent tooth decay.

H. Trendley Dean, along with others in the U. S. Public Health Service, collected more systematic data on fluorides in the water, mottling, and tooth decay. The data confirmed the associations noted by McKay and Churchill. Moreover, the data indicated that, at lower doses, fluorides in the water could prevent decay without mottling the teeth. (Mottling was unsightly, and carried risks of its own.) Starting in 1945, community experiments strengthened these conclusions about the role of fluorides, although some controversy remained. Fluoridation of drinking water followed within a few years, and tooth decay in childhood declined precipitously.

### Fleming and Penicillin

Alexander Fleming was working at St. Mary's Hospital in London, under the direction of Sir Almroth Wright, studying the life cycle of staphylococcus (bacteria that grow in clusters, looking under the microscope like clusters of grapes). Fleming had a number of plates on which he was growing staphylococcus colonies. He left the plates in a corner of his office for some weeks while he was on holiday. When he returned, one of the plates had been contaminated by mold. So far, this is unremarkable. He noticed, however, "that around a large colony of a contaminating mould the staphylococcus colonies became transparent and were obviously undergoing lysis" (Fleming 1929, 226).

Bacteria "lyse" when their cell walls collapse. What caused the lysis? Rather than discarding the plate—the normal thing to do—Fleming thought that the lysis was worth investigating. He did so by growing the mold in broth, watching its behavior, and trying filtered broth on various kinds of bacteria. The mold, a species of *Penicillium*, generated a substance that "to avoid the repetition of the rather cumbersome phrase 'mould broth filtrate' [will be named] 'penicillin'" (Fleming 1929, 227). It was the penicillin that caused the bacteria to lyse. Fleming showed that penicillin destroyed—or at least inhibited the growth of—many kinds of bacteria besides staphylococcus.

Penicillin's therapeutic potential went unrealized until Howard Florey and his associates at Oxford took up the research in 1938 and found processes for purification and larger-scale production. Due to the exigencies of World War II, much of the work was done in the U. S., where a strain of *Penicillium* that gave high yields was found on a moldy cantaloupe at a market in Peoria. (Industrial-scale development was being done at a nearby Department of Agriculture laboratory under the direction of Kenneth Raper, and people were encouraged to bring in moldy fruit for analysis.)

Penicillin was widely used to treat battlefield injuries, largely preventing gangrene, for example. Along with the sulfa drugs (prontosil was discovered by Gerhard Domagk in 1932) and streptomycin (discovered by Selman Waksman in 1944), penicillin was among the first of the modern antibiotics.

### Gregg and German Measles

Norman Gregg was a pediatric ophthalmologist in Australia. In 1941, he noticed in his practice an unusually large number of infants with cataracts and heart defects. (“Cataracts” make the lens of the eye opaque.) On investigation, he found that many of his colleagues were also treating such cases. The similarity of the cases, and their widespread geographic distribution, led him to guess that the cause must have been exposure to some infectious agent early in the mother’s pregnancy, rather than genetics—which was the conventional explanation at the time for birth defects. But what was the infectious agent? This is how Gregg explained his thought process:

“The question arose whether [the cause] could have been some disease or infection occurring in the mother during pregnancy which had then interfered with the developing cells of the lens. By a calculation from the date of the birth of the baby, it was estimated that the early period of pregnancy corresponded with the period of maximum intensity of the very widespread and severe epidemic in 1940 of the so-called German measles” (Gregg 1941).

Detailed epidemiological research showed that exposure of the mother to German measles in the first or second month of pregnancy markedly increases the risk of birth defects in the baby. The association is generally viewed as causal. Today, there is a vaccine that prevents German measles, and cataracts at birth are exceedingly rare.

### Herbst and DES

Herbst and Scully described seven cases of adenocarcinoma of the vagina in adolescent girls. This is an unusual kind of cancer, especially in adolescence. What was the cause? The mother of one patient suggested diethylstilbestrol (DES), an artificial hormone often prescribed in those days to prevent miscarriage. Arthur Herbst and his associates were intrigued, but skeptical. They did a case control study, and established a highly significant association, confirmed by a number of other studies, and now accepted as causal.

Two key insights precede any statistical analysis: (i) this is a cluster of cancers worth investigating, and (ii) the cause might have been exposure of the mother during pregnancy—not the daughter after birth—to some toxic substance. *A priori*, neither point could have been obvious.

## Conclusions

In the health sciences, there have been enormous gains since the time of Jenner, many of which are due to statistics. Snow's analysis of his natural experiment shows the power of quantitative methods and good research design. Semmelweis' argument depends on statistics; so too with Goldberger, Dean, Gregg, and Herbst et al. On the other hand, as the examples demonstrate, substantial progress also derives from informal reasoning and qualitative insights. Recognizing anomalies is important; so is the ability to capitalize on accidents. Progress depends on refuting conventional ideas if they are wrong, developing new ideas that are better, and testing the new ideas as well as the old ones. The examples show that qualitative methods can play a key role in all three tasks.

In Fleming's lab, chance circumstances generated an anomalous observation. Fleming resolved the anomaly and discovered penicillin. Semmelweis used qualitative reasoning to reject older theories about the cause of puerperal fever, to develop a new theory from observations on a tragic accident, and to design an intervention that would prevent the disease. The other examples lead to similar conclusions.

What are the lessons for methodologists in the 21st century? Causal inference from observational data presents many difficulties, especially when underlying mechanisms are poorly understood. There is a natural desire to substitute intellectual capital for labor, and an equally-natural preference for system and rigor over methods that seem more haphazard. These are possible explanations for the current popularity of statistical models.

Indeed, far-reaching claims have been made for the superiority of a quantitative template that depends on modeling—by those who manage to ignore the far-reaching assumptions behind the models. However, the assumptions often turn out to be unsupported by the data (Duncan 1984; Berk 2004; Brady and Collier 2004; Freedman 2005). If so, the rigor of advanced quantitative methods is a matter of appearance rather than substance.

The historical examples therefore have another important lesson to teach us. Scientific inquiry is a long and tortuous process, with many false starts and blind alleys. Combining qualitative insights and quantitative analysis—and a healthy dose of skepticism—may provide the most secure results.

## Further Reading

Brady, Collier and Seawright (2004) compare qualitative and quantitative methods for causal inference in the social sciences. As they point out,

“it is difficult to make causal inferences from observational data, especially when research focuses on complex political processes. Behind the appar-

ent precision of quantitative findings lie many potential problems concerning equivalence of cases, conceptualization and measurement, assumptions about the data, and choices about model specification. . .” (Brady, Collier and Seawright 2004, 9–10).

These authors recommend using a diverse mix of qualitative and quantitative techniques in order to exploit the available information; no particular set of tools is universally best. Causal process observations (including anomalies and results of accidental experiments, even experiments with  $N = 1$ ) can be extremely helpful, as they were in the epidemiological examples discussed here.

The role of anomalies in political science is also discussed by Rogowski (2004). He suggests that scholars in that field may be excessively concerned with hypothesis testing based on statistical models. Scholars may underestimate the degree to which the discovery of anomalies can overturn prior hypotheses and open new avenues of investigation. Anomalies that matter have been discovered in case studies—even when the cases have been selected in ways that do considerable violence to large- $N$  canons for case selection. He also suggests that failure to search for anomalies can lead to a kind of sterility in research programs.

Scientific progress often begins with inspired guesswork. On the other hand, if guesses cannot be verified, progress may be illusory. For example, Snow (1965, 125–33) theorized that—by analogy with cholera—plague, yellow fever, dysentery, typhoid fever, and malaria (which he calls “ague” or “intermittent fever”) were infectious waterborne diseases. His supporting arguments were thin. As it turns out, these diseases are infectious; however, only dysentery and typhoid fever are waterborne.

Proof for dysentery and typhoid fever, and disproof for the other diseases, was not to come in Snow’s lifetime. Although William Budd (1873) made a strong case on typhoid fever, reputable authors of the late 19th century still denied that such diseases were infectious (Bristowe and Hutchinson 1876, 211, 629; Bristowe et al. 1879, 102–3). In the following decades, evidence from epidemiology and microbiology settled the issue.

Plague is mainly spread by fleas, although transmission by coughing is also possible. The causal agent is the bacterium *Yersinia pestis*. Yellow fever and malaria are spread by mosquitoes. Yellow fever is caused by a virus. Malaria is caused by several species of *Plasmodium*, one-celled organisms with a nucleus and an extravagantly complicated life cycle spent partly in humans and partly in mosquitoes. The medieval Black Death is usually identified with modern plague, but this is still contested by some scholars (Nutton 2008).

Buck et al. (1989) reprints many of the classic papers in epidemiology; some classic errors are included too. Porter (1997) is a standard reference on history of medicine. Jenner’s papers are reprinted in Eliot (1910). Bazin (2000) discusses the history of smallpox, Jenner’s work, and later developments, including the eradication of smallpox; the last recorded cases were in 1977–78. There is

a wealth of additional information on the disease and its history in Fenner et al. (1988).

Inoculation was recorded in England by 1721 (Bazin 2000, 13; Fenner et al. 1988, 214–6). However, the practice was described in the journals some years before that (Timonius and Woodward 1714). It was a common opinion in Jenner’s time that cowpox created immunity to smallpox (Jenner 1801; Baron 1838, 122). Over the period 1798–1978, techniques for producing and administering the vaccine were elaborated. As life spans became longer, it became clear that—contrary to Jenner’s teachings—the efficacy of vaccination gradually wore off. Revaccination was introduced. By 1939, the virus in the vaccines was a little different from naturally-occurring cowpox virus. The virus in the vaccines is called “vaccinia” (Bazin 2000, Chap. 11; Fenner et al. 1988, Chaps. 6–7, esp. p. 278).

Bulloch (1938) reviews the history of bacteriology. Bacteria were observed by Hooke and Leeuwenhoek before 1700. Otto Friderich Müller in Denmark developed a workable classification before 1800, improved about 50 years later by Ferdinand Cohn in Germany.

Some of Koch’s work on anthrax was anticipated by Pierre François Rayer and Casimir-Joseph Davaine in France. Likewise, Pasteur’s experiments disproving spontaneous generation built on previous work by others, including Lazzaro Spallanzani; contemporaneous research by John Tyndall should also be mentioned.

Freedman (2005, 6–9) reports on Snow and cholera. For detailed information on Snow’s work, see Vinten-Johansen et al. (2003). Evans (1987) gives a historical analysis of the cholera years in Europe. Koch’s discovery of the vibrio was anticipated by Filippo Pacini in 1854, but the implications of Pacini’s work were not recognized by his contemporaries.

Henry Whitehead was a clergyman in the Soho area. He did not believe that the Broad Street pump—famous for the purity of its water—was responsible for the epidemic. He saw a gap in Snow’s argument: the fatalities cluster around the pump, but what about the population in general?

Whitehead made his own house-to-house canvass to determine attack rates among those who drank water from the pump and those who did not. Then he drew up a  $2 \times 2$  table to summarize the results. The data convinced him that Snow was correct (Cholera Inquiry Committee 1855, 121–33). Snow made this kind of analysis only for his natural experiment.

William Farr, statistical superintendent of the General Register Office, was a leading medical statistician in Victorian England and a “sanitarian,” committed to eliminating air pollution and its sources. He claimed that the force of mortality from cholera in an area was inversely related to its elevation. More specifically, if  $y$  is the death rate rate from cholera in an area and  $x$  is its elevation, Farr proposed the equation

$$y = \frac{a}{b + x}.$$

The constants  $a$  and  $b$  were estimated from the data. For 1848–49, the fit was excellent.

Farr held the relationship to be causal, explained by atmospheric changes, including attenuation of noxious exhalations from the Thames, changes in vegetation, and changes in the soil. After the London epidemic of 1866, however, he came to accept substantial parts of Snow's theory—without abandoning his own views about miasmas and elevation (Humphreys 1885, 341–84; Eyles 1979, 114–22; Vinten-Johansen et al. 2003, 394).

For better or worse, Farr's belief in mathematical symbolism had considerable influence on the development of research methods in medicine and social science. Furthermore, the tension between the pursuit of social reform and the pursuit of truth, so evident in the work of the sanitarians, is still with us.

There are two informative web sites on Snow, Whitehead, and other major figures of the era (these sites were active on August 12, 2008):

<http://www.ph.ucla.edu/epi/snow.html>

<http://matrix.msu.edu/~johnsnow/>

Loudon (2000) is highly recommended on puerperal fever; but also see Nuland (1979) for a more sympathetic account of Semmelweis' life. Hare (1970, Chap. 7) discusses efforts to control puerperal fever in a London maternity hospital in the 1930s. The strain of *Staphylococcus pyogenes* causing the disease turned out to be a common inhabitant of the human nose and throat (Loudon 2000, 201–4).

A definitive source on beriberi, only paraphrased here, is Carpenter (2000). He gives a vivid picture of a major scientific advance, including discussion of work done before Eijkman arrived in Java.

The discussion of pellagra is based on Freedman et al. (2007, 15–16). Goldberger's papers are collected in Terris (1964). Goldberger (1914) explains the reasoning that led him to the deficiency-disease hypothesis; Goldberger et al. (1926) identifies the P-P factor as part of the vitamin B complex. Carpenter (1981) reprints papers by many pellagra researchers, with invaluable commentary. He explains why in Mexico a corn-based diet does not lead to pellagra, discusses the role of tryptophan (an amino acid that can be converted to niacin in the body), and points out the gaps in our knowledge of the disease and the reasons for its disappearance.

The primary papers on fluoridation are McKay and Black (1916), Churchill (1931), and Dean (1938). There is a considerable secondary literature; see, for instance, McClure (1970) and Centers for Disease Control (1999). McKay (1928) is often cited, but seems mainly about another topic: whether enamel in teeth is living tissue.

An excellent source on Fleming is Hare (1970), with Goldsmith (1946) adding useful background. Today, "penicillin" refers to the active ingredient in Fleming's mold broth filtrate. What is the cell-killing mechanism? In brief, cell



walls of most bacteria include a scaffolding constructed from sugars and amino acids. Components of the scaffolding have to be manufactured and assembled when the cells are dividing to form daughter cells. In many species of bacteria, penicillin interferes with the assembly process, eventually causing the cell wall to collapse (Walsh 2003).

Some species of bacteria manufacture an enzyme (“penicillinase”) that disables penicillin—before the penicillin can disable the cell. There are other bacterial defense systems too, which explain the limits to the efficacy of penicillin. Penicillin inhibits cell wall synthesis by a process that is reasonably well understood, but how does inhibition cause lysis? That is still something of a mystery, although much has been learned (Walsh 2003, 41; Bayles 2000; Giesbrecht et al. 1998).

Penicillin only causes lysis when bacteria are dividing. For this reason among others, a rather unusual combination of circumstances was needed to produce the effect that Fleming noticed on his Petri dish (Hare 1970, Chap. 3). Was Fleming merely lucky? Pasteur’s epigram is worth remembering: “Dans les champs de l’observation, le hasard ne favorise que les esprits préparés.”

Almroth Wright, Fleming’s mentor, was one of the founders of modern immunology (Dunnill 2001). Among other accomplishments, he developed a vaccine that prevented typhoid fever. Wright was a close friend of George Bernard Shaw’s, and was the basis for one of the characters in *The Doctor’s Dilemma*.

Material on Gregg may be hard to find, but see Gregg (1941), Lancaster (1996), and Webster (1998). Gregg (1944) discusses infant deafness following maternal rubella.

On DES, the basic papers are Herbst and Scully (1970) and Herbst et al. (1971), with a useful summary by Colton and Greenberg (1982). Also see Freedman et al. (2007, 9–10). DES was an unnecessary tragedy. Doctors who prescribed DES were paying attention to observational studies that showed a positive effect in preventing miscarriage. However, clinical trials showed there was no such effect. DES was banned in 1971 for use in pregnant women.

## References

- Baron, J. 1838. *The Life of Edward Jenner*. Vol. I. London: Henry Colburn. Available on Google Books, August 12, 2008.
- Bayles, K. W. 2000. The bactericidal action of penicillin: New clues to an unsolved mystery. *Trends in Microbiology* 8: 274–78.
- Bazin, H. 2000. *The Eradication of Smallpox*, translated into English by A. and G. Morgan, London: Academic Press.
- Berk, R. A. 2004. *Regression Analysis: A Constructive Critique*. Thousand Oaks: Sage Publications.

- Brady, H. E. and Collier, D., eds. 2004. *Rethinking Social Inquiry: Diverse Tools, Shared Standards*. Lanham, Maryland: Rowman & Littlefield Publishers, Inc.
- Brady, H. E., Collier, D., and Seawright, J. 2004. Refocusing the discussion of methodology. In Brady and Collier 2004: 3–20.
- Bristowe, J. S., Hutchinson, J. S. 1876. *A Treatise on the Theory and Practice of Medicine*. Philadelphia: Henry C. Lea. Available on Google Books, August 12, 2008.
- Bristowe, J. S., Wardell, J. R., Begbie, J. W., et al. 1879. *Diseases of the Intestines and Peritoneum*. New York: William Wood and Company. Available on Google Books, August 12, 2008.
- Buck, C., Llopis, A., Nájera, E., and Terris, M., eds. 1989. *The Challenge of Epidemiology: Issues and Selected Readings*. Geneva: World Health Organization.
- Budd, W. (1873). *Typhoid Fever: Its Nature, Mode of Spreading, and Prevention*. London: Longmans, Green, and Co.  
<http://www.deltaomega.org/typhoid.pdf>
- Bulloch, W. 1938. *The History of Bacteriology*. London: Oxford University Press.
- Carpenter, K. J., ed. 1981. *Pellagra*. Stroudsburg, PA.: Hutchinson Ross Pub. Co.
- Carpenter, K. J. 2000. *Beriberi, White Rice, and Vitamin B*. Berkeley: University of California Press.
- Centers for Disease Control. 1999. Fluoridation of drinking water to prevent dental caries. *Morbidity and Mortality Weekly Report*, October 22, Vol. 48, No. 41, 933–40. U. S. Department of Health and Human Services.
- Cholera Inquiry Committee. 1855. *Report on the Cholera Outbreak in the Parish of St. James, Westminster during the Autumn of 1854*. London: Churchill.
- Churchill, H. V. 1931. Occurrence of fluorides in some waters of the United States. *Journal of Industrial and Engineering Chemistry*, 23: 996–8.
- Colton, T. and Greenberg, E. R. 1982. Epidemiologic evidence for adverse effects of DES exposure during pregnancy. *The American Statistician*, 36: 268–72.
- Dean, H. T. 1938. Endemic fluorosis and its relation to dental caries. *Public Health Reports*, 53: 1443–52. Reprinted in Buck et al. 1989: 271–8.
- Duncan, O. D. 1984. *Notes on Social Measurement*. New York: Russell Sage.
- Dunnill, M. S. 2001. *The Plato of Praed Street: The Life and Times of Almroth Wright*. London: Royal Society of Medicine Press.
- Eliot, C. W., ed., 1910 edn., *Scientific Papers: Physiology, Medicine, Surgery, Geology*. Vol. 38 in *The Harvard Classics*. New York: P. F. Collier & Son; originally published 1897. Available on Google Books, August 12, 2008.
- Evans, R. J. 1987. *Death in Hamburg: Society and Politics in the Cholera Years, 1830–1910*. Oxford: Oxford University Press.

Eyler, J. M. 1979. *Victorian Social Medicine: The Ideas and Methods of William Farr*. Baltimore: Johns Hopkins University Press.

Fenner, F., Henderson, D. A., Arita, I., Jezek, Z., and Ladnyi, I. D. 1988. *Smallpox and its Eradication*. Geneva: World Health Organization.

<http://whqlibdoc.who.int/smallpox/9241561106.pdf>

Fleming, A. 1929. On the antibacterial action of cultures of a penicillium, with special reference to their use in the isolation of *B. influenzae*. *British Journal of Experimental Pathology*, 10: 226–36.

Freedman, D. A. 2005. *Statistical Models: Theory and Practice*. New York: Cambridge University Press.

Freedman, D. A., Pisani, R., and Purves, R. A. 2007. *Statistics*. 4th edn. New York: W. W. Norton & Company.

Giesbrecht, P., Kersten, T., Maidhof, H., and Wecke, J. 1998. Staphylococcal cell wall: Morphogenesis and fatal variations in the presence of penicillin. *Microbiology and Molecular Biology Reviews*, 62: 1371–1414.

Goldberger, J. 1914. The etiology of pellagra. *Public Health Reports*, 29: 1683–86. Reprinted in Buck et al. 1989: 99–102, and in Terris 1964: 19–22.

Goldberger, J., Wheeler, G. A., Lillie, R. D., and Rogers, L. M. 1926. A further study of butter, fresh beef, and yeast as pellagra preventives, with consideration of the relation of factor P-P of pellegra (and black tongue of dogs) to vitamin B1. *Public Health Reports*, 41: 297–318. Reprinted in Terris 1964: 351–70.

Goldsmith, M. 1946. *The Road to Penicillin*. London: Lindsay Drummond.

Gregg, N. M. 1941. Congenital cataract following German measles in the mother. *Transactions of the Ophthalmological Society of Australia*, 3: 35–46. Reprinted in Buck et al. 1989: 426–34.

Gregg, N. M. 1944. Further observations on congenital defects in infants following maternal rubella. *Transactions of the Ophthalmological Society of Australia*, 4: 119–31.

Hare, R. 1970. *The Birth of Penicillin and the Disarming of Microbes*. London: Allen & Unwin.

Herbst, A. L. and Scully, R. E. 1970. Adenocarcinoma of the vagina in adolescence: A report of 7 cases including 6 clear cell carcinomas. *Cancer* 25: 745–57.

Herbst, A. L., Ulfelder, H., and Poskanzer, D. C. 1971. Adenocarcinoma of the vagina: Association of maternal stilbestrol therapy with tumor appearance in young women. *New England Journal of Medicine*, 284: 878–81. Reprinted in Buck et al. 1989: 446–50.

Humphreys, N. A., ed. 1885. *Vital Statistics: A Memorial Volume of Selections from the Reports and Writings of William Farr*. London: Edward Stanford. Available on Google Books, August 12, 2008.

- Jenner, E. 1798. *An Inquiry into the Causes and Effects of the Variolae Vaccinae, a Disease Discovered in Some of the Western Counties of England, Particularly Gloucestershire, and Known by the Name of the Cow Pox*. London: printed for the author by Sampson Low. Reprinted in Eliot 1910: 151–80.
- Jenner, E. 1801. *The Origin of the Vaccine Inoculation*. London: D. N. Shury. Reprinted in Fenner et al. 1988: 258–61.
- Lancaster, P. A. L. 1996. Gregg, Sir Norman McAlister (1892–1966), Ophthalmologist. Pp. 325–327 in *Australian Dictionary of Biography*, vol. 14, ed. John Ritchie. Melbourne: Melbourne University Press.  
<http://www.adb.online.anu.edu.au/biogs/A140370b.htm>
- Loudon, I. 2000. *The Tragedy of Childbed Fever*. Oxford: Oxford University Press.
- McKay, F. S. 1928. Relation of mottled enamel to caries. *Journal of the American Dental Association*, 15: 1429–37.
- McKay, F. S. and Black, G. V. 1916. An investigation of mottled teeth: An endemic developmental imperfection of the enamel of the teeth, heretofore unknown in the literature of dentistry. *Dental Cosmos*, 58: 477–84, 627–44, 781–92, 894–904.
- McClure, F. J. 1970. *Water Fluoridation*. Bethesda, MD.: National Institute of Dental Research.
- Nuland, S. 1979. The enigma of Semmelweis—An interpretation. *Journal of the History of Medicine and Allied Sciences*, 34: 255–72.
- Nutton, V., ed. 2008. *Pestilential Complexities*. London: Wellcome Trust.
- Porter, R. 1997. *The Greatest Benefit to Mankind*. New York: W. W. Norton & Company.
- Roe, D. A. 1973. *A Plague of Corn*. Cornell University Press.
- Rogowski, R. (2004). How inference in the social (but not the physical) sciences neglects theoretical anomaly. In Brady and Collier 2004: 75–82.
- Semmelweis, I. 1981 edn. *The Etiology, Concept, and Prophylaxis of Childbed Fever*. English translation by F. P. Murphy. Birmingham: Classics of Medicine Library; originally published in German 1860.
- Snow, J. 1965 edn. *On the Mode of Communication of Cholera*. New York: Hafner; originally published 1855.
- Terris, M., ed. 1964. *Goldberger on Pellagra*. Baton Rouge: Louisiana State University Press.
- Timonius, E. and Woodward, J. 1714. An account, or history, of the procuring the small pox by incision, or inoculation; as it has for some time been practised at Constantinople. *Philosophical Transactions* 29: 72–82.
- Vinten-Johansen, P. Brody, H., Paneth, N., and Rachman, S. 2003. *Cholera, Chloroform, and the Science of Medicine*. New York: Oxford University Press.

Walsh, C. 2003. *Antibiotics: Actions, Origins, Resistance*. Washington, D. C.: ASM Press.

Webster, W. S. 1998. Teratogen update: Congenital rubella. *Teratology*, 58: 13–23.

Acknowledgments. David Collier (Berkeley), Thad Dunning (Yale), Paul Humphreys (University of Virginia), Erich Lehmann (Berkeley), and Janet Macher (Berkeley) made many helpful comments.

Revision of Chapter 12 in *The Oxford Handbook of Political Methodology*, edited by Janet M. Box-Steffensmeier, Henry E. Brady and David Collier.

Freedman, David A. "On Types of Scientific Inquiry: The Role of Qualitative Reasoning." Chapter 12 in *The Oxford Handbook of Political Methodology*. Edited by Janet M. Box-Steffensmeier, Henry E. Brady, and David Collier. Oxford University Press, 2010. (PDF - 8.4MB)

American Political Science Review 97, no. 1 (2003): 75-90.

Brady, Henry E. "Data-Set Observations versus Causal-Process Observations: The 2000 U.S. Presidential Election." Chapter 12 in *Rethinking Social Inquiry: Diverse Tools, Shared Standards*. 2nd edition. Edited by Henry E. Brady and David Collier.