Causality in the Time of Cholera: John Snow and the Process of Scientific Inquiry

Thomas S. Coleman*
Julia Koschinsky†
Dan Black‡

Draft June 14, 2022

Abstract

John Snow in 1849 proposed the intestinal fecal-oral theory for cholera and provided substantial evidence throughout the 1850s supporting a water-borne explanation of how the disease was transmitted. Nevertheless, Snow's theory was not widely accepted among airborne theorists until many years later. Our goal in this paper is two fold. First, to provide a rational reconstruction (in the words of Imre Lakatos) for the history of cholera explanations, showing how alternative theories (rationally) accommodated water as a cause and thus were protected from refutation by evidence on water alone. We use Lakatos' sophisticated falsification criteria to argue that, in spite of slow acceptance by contemporaries, the wide range of evidence that was available in the 1850s was sufficient to demonstrate that the fecal-oral theory was superior. Our second goal is to use this rational reconstruction of cholera theories as a practical guide to demonstrating a causal effect. Here we use the competition and the comparison across theories in the 1850s as a practical example of the process for scientific inquiry – a template for modern teaching and research of quantitative methods. We build on ideas of Charles Sanders Peirce and Imre Lakatos to illustrate how causal explanation and falsification works in an iterative process of scientific inquiry – and how testing for specific causal effects as is often done in empirical economics, statistics, and computer science was insufficient in differentiating theories. (231 words)

Acknowledgements: We would like to thank Steven Durlauf, Marcos Falcone, Luc Anselin, Robert Michael, Xi Song, Karina Acosta, Austin Wright, Emlen Metz, Michelle Si, members of the CEHD Lifecycle Working Group and the Center for Spatial Data Science's Study Group at the University of Chicago. Errors are our own.

Contents

*Harris School of Public Policy, University of Chicago, tscoleman@uchicago.edu
†Center for Spatial Data Science, University of Chicago, jkoschinsky@uchicago.edu
‡Harris School of Public Policy, University of Chicago, danblack@uchicago.edu

List of Tables

List of Figures

1 Introduction

Cholera is a horrible and often deadly disease. Victims experience diarrhea, vomiting, and rapid dehydration; without treatment the mortality rate is roughly 50 percent. It was a scourge of the 19^{th} century – it visited London in 1831, 1849, 1854, and 1866 and Tumbe (2020) estimates that between 1817 and 1920 it killed 50 million people in a world of about 1.2 billion. Today we know cholera is caused by a bacterium, *Vibrio cholerae*, but in the 19^{th} century the underlying cause was uncertain and discovering it was of the highest priority and represented one of the outstanding achievements of 19^{th}-century science. Debate in England over cholera was robust, generally reasoned, and in many ways quite modern. Participants used evidence to develop and refine their ideas, and also to advance their theories within the medical and public sphere. The goal, sometimes stated explicitly, was to use evidence to understand the causes of cholera, to develop theories of the disease that would lead to policies to ameliorate its impact. The evidence was often anecdotal (particularly prior to the 1850s) but there was an understanding that evidence was crucial to understanding the disease, discriminating between competing theories, and developing effective public policies.

John Snow (1813-1858) was a London doctor who in 1849 proposed the intestinal (fecal-oral) theory.[1] In *On the mode of communication of cholera* (Snow, 1849a) Snow proposed his theory, provided extensive supporting evidence particularly on the water-borne nature of disease transmission, and argued persuasively why the evidence supported the fecal-oral theory. Nonetheless, despite substantial evidence, Snow's theory was not widely accepted during the 1850s, and the intestinal fecal-oral theory was not universally accepted until the end of the 19^{th} century. The lack of early recognition of Snow's theory – in spite of substantial supporting evidence – has puzzled many (including the present authors), as seemingly a case of "smart people cling[ing] to an outlandishly incorrect idea despite substantial evidence to the contrary" (Johnson, 2007, 125).

[1] Snow and the story of cholera has long exercised a fascination over both the popular imagination and academic inquiry. A very limited sampling of recent books and papers devoted to the topic is Johnson (2007); Freedman (1999, 1991); Hempel (2007); Vinten-Johansen et al. (2003); Vinten-Johansen (2020); McLeod (2000); Rothman (2002).

The puzzle arises, however, from a too-narrow conception of causal and scientific inquiry, a view that might be caricatured as the lone genius undertaking a single crucial experiment that in one fell swoop overturns accepted doctrine and sets the world on the path of true knowledge. And the story of cholera carries the outlines of such a caricature: John Snow *did*, on his own, develop the correct theory of cholera; the demonstration of cholera as water-borne *does* look like a crucial experiment, fits neatly in a treatment-control framework for causality, and Snow was the first to use experimental designs we continue to use today for causal analysis (difference-in-differences and randomization as an instrumental variable); and Snow *did* advocate for many years for his theory in the face of a skeptical medical and public health establishment.

But in fact the story of cholera *contradicts* popular understandings of science as a one-off experiment or test of theory that decisively demonstrates a causal effect. There was no single, definitive experiment or test that proved cholera was an intestinal disease transmitted by fecal contamination and oral ingestion. Even an *experimentum crucis* such as the South London "Grand Experiment" of 1854 only occurred after several years of analyses at individual and neighborhood scales (including the famous Broad Street pump outbreak) and even then did not on its own decide between the competing theories. "Identification" of the underlying causal theory required, not one crucial experiment, but consideration of a wide range of evidence across alternative models and theories.

Developing and demonstrating a causal explanation or *theory* such as for cholera requires a broader view of the dynamic interplay between data and theory, moving beyond a treatment-control view of causality to embrace an iterative process of scientific inquiry. Central to this process is choosing between competing theories and the criterion for choosing. Karl Popper Popper (1985) demonstrated that scientific theories can never be *proven*, only *falsified*, arguing against inductivism and also Rudolf Carnap's Leitgeb and Carus (2021) confirmationism. Naive falsification, however, is not itself a suitable criterion. Theories and hypotheses can be adjusted in response to contradictory evidence – the essence of the Duhem-Quine thesis (Quine 1953, chapter 11; Lakatos 1980, 93-101). As a practical matter replacing a theory or *research programme* is a long war of attrition (quoting Lakatos, 1980, 133) requiring a wide range of evidence comparing across multiple predictions from alternate theories. From a theoretical perspective, Imre Lakatos's *sophisticated falsification* (Lakatos, 1980) provides a workable criterion.

We have two goals in retelling the classic story of cholera and bringing to life this logic of scientific inquiry. The first is a (following Lakatos, 1980, 52-53) *rational reconstruction* of the history of cholera research during the 1850s to better understand why Snow's theory, although correct, was not quickly or widely accepted. The second is to use Snow's work as a template for best practice in demonstrating a causal effect. It is a practical example of sophisticated falsification, providing a compelling argument for the superiority of the fecal-oral theory even when limited to the data available in the 1850s. The methods, data, and analysis used by Snow and contemporaries illustrate the dynamic interplay between theory, data, and testing – what we call the iterative process of scientific inquiry – and why a narrow focus on counterfactuals, single causal effects, and experiments was insufficient in this context.

Our focus on the iterative process of scientific inquiry fits with other recent work that seeks to add stronger conceptual framing to today's statistical methods curriculum for social scientists to counter mechanical applications of statistical techniques (e.g. Fowler and Bueno de Mesquita (2021)). Our

aim with this article is to encourage teaching counterfactual frameworks to social scientists within the broader context of scientific inquiry. In this context, we are also making teaching materials (Vinten-Johansen and Koschinsky (2022)), a re-analysis with modern statistical methods and tools (Coleman (2020)) and data (Coleman (2020), Falcone et al. (2021)) available to replicate some of Snow's and colleague's work. With these materials, we hope to help reconnect students with the thrill of social scientific discovery and with the larger logic of scientific inquiry that statistical methods are embedded in.

We aim to illustrate the process of falsification (and corroboration), through four steps: (1) laying out the theories, being honest about both the preferred theory *and* alternatives; (2) developing predictions for all theories; (3) comparing predictions against evidence, comparisons that may involve both formal statistical testing and less formal comparisons; and (4) weighing the balance of evidence for and against the alternatives.

Section 2 starts with a brief overview of cholera theories (Snow's and alternatives) in mid-century Britain. We also emphasize our focus on predictions and evidence about the patterns of mortality, what we might call statistical or observational "epidemiological" evidence, to distinguish from "biological" evidence. In Section **??** we turn to the process for choosing or discriminating between theories or causal explanations: the criterion of *sophisticated falsification* and the iterative process of scientific inquiry. We build on ideas from Imre Lakatos – sophisticated falsification and scientific research programmes – and Charles Sanders Peirce – stages of scientific inquiry. Section 5.1 presents the development of and debate over Snow's fecal-oral theory in the 1850s as a case study for the application of sophisticated falsification and the iterative process of scientific inquiry.

In Section 3.1 we discuss the treatment-control or potential outcome approach to causal analysis (Neyman, 1923; Cox, 1958; Rubin, 1974; Holland, 1986) that is widely used today and show why this framework was insufficient in deciding between theoretical alternatives.

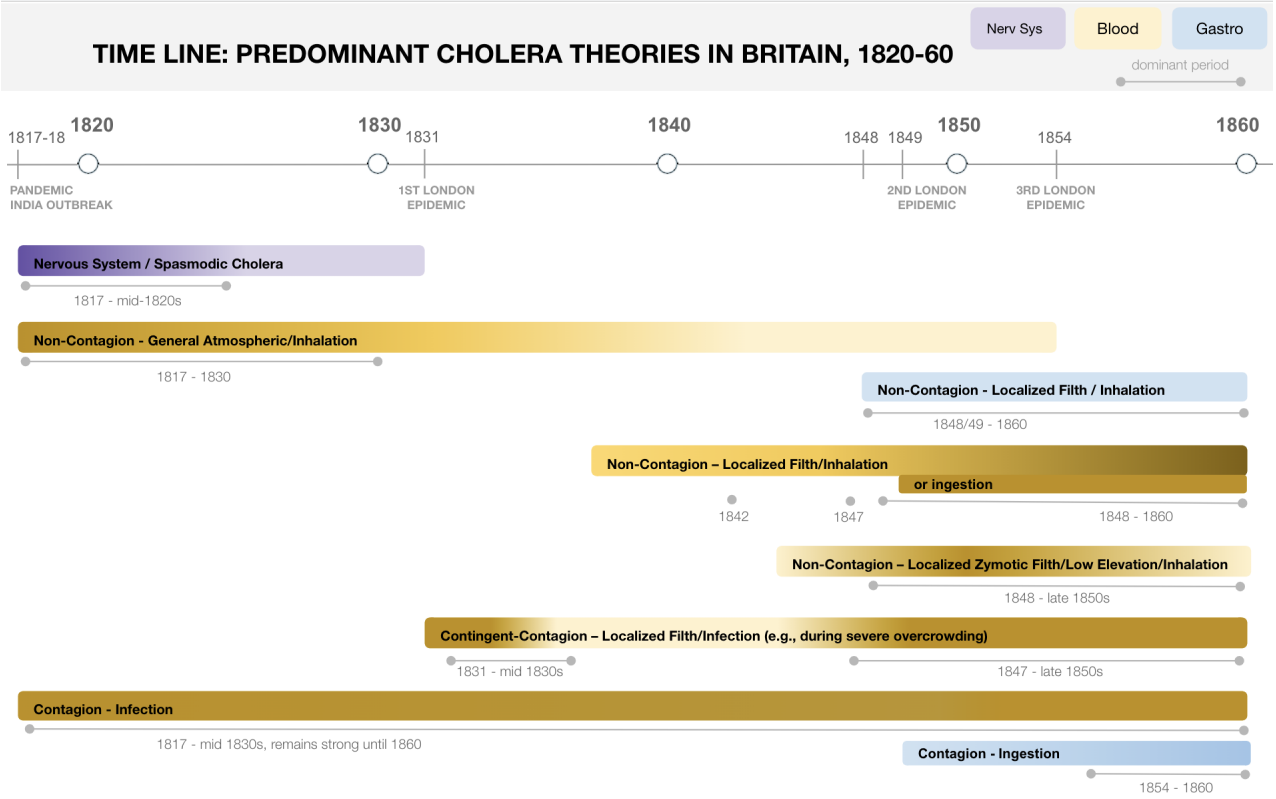
In Section 5.3 we examine, as a case study of sophisticated falsification, Snow's efforts in 1855-56 (primarily Snow, 1855, 1856b) to establish the fecal-oral theory versus alternatives. In the Conclusion, Section 6, we speculate on why the fecal-oral theory was accepted only slowly and discuss lessons learned from cholera in the 1850s.

Snow's efforts to demonstrate the fecal-oral theory in the 1850s were only partially successful, but the episode remains both an instructive historical episode and a template for modern investigators for scientific inquiry and falsification of alternative theories. It is highly relevant for social scientists today, an outstanding case of social science research based on observational and statistical evidence.

2 Cholera Theories and Predictions

In this section we summarize the theories of cholera, which are discussed in more detail in Appendix A. For present purposes we can collapse the non-Snow alternative theories into one broad category – inhaled blood disease – which then split into two, depending on whether the airborne cholera poison is contagious (generated in and transmitted from one person to another) or non-contagious (generally created by environmental conditions but not passed person-to-person). The non-contagious theory,

Figure 1: Timeline for Theories of Cholera in Nineteenth Century England



in response to contradictory observations, was modified in the 1830s by the addition of a "contingent-contagious" auxiliary hypothesis stating that cholera could under some conditions and for a limited period become contagious.[2] This is a substantial simplification of the actual theories, but captures the important component shared by all, that the cholera poison is airborne.

The important second step is making explicit the predictions from *all* theories being considered. This was not done in the 1850s although the predictions are implicit throughout the writings of Snow and contemporaries.

2.1 Cholera Theories in 1849 and After

Figure 1 shows a timeline summarizing cholera theories in the 19^{th} century (Vinten-Johansen and Koschinsky, 2022). In late 1848 Snow, in line with most of his colleagues, believed cholera was an inhaled disease of the blood.[3] By late 1849, Snow "had been led, contrary to the usual opinion, to consider cholera as being ... a local affection of the mucous membrane of the alimentary canal" (Snow, 1849b, 431). In other words, between late 1848 and late 1849 Snow had developed and adopted a new theory of cholera as a gastrointestinal disease, the theory that we recognize today as

[2]"certain conditions may favour its spread from person to person, as when great numbers of the sick are crowded together in close unventilated apartments, yet this is not to be considered as affecting the general principal of its non-contagious nature" General Board of Health (1848)

[3]We can sse this in Snow (1848, 507-508), where he discusses treatments, presuming that cholera affects the lungs, "in some points resembling [asphyxia], as far as the internal congestion is concerned."

correct. From Snow's writings we can trace his thinking and create a "rational reconstruction" of the development of his theory (to use Lakatos's phrase).[4]

The prominent theories in 1848 assumed transmission was via airborne inhalation, but diverged on whether cholera was contagious or non-contagious. The divergence reflected contradictory evidence, with some observations supporting airborne contagion and others supporting non-contagion. In a formal sense, all airborne inhaled theories were refuted by the evidence and had been from the early days.[5] In fact, the hypothesis of contingent contagion had been introduced in the 1830s specifically to adjust the non-contagious theories to (or protect them from) the evidence supporting airborne contagion.

Combining his medical knowledge, examination of cholera cases, and evidence on transmission (which was contradictory assuming airborne transmission), Snow proposed the fecal-oral theory as a plausible theory and mechanism for cholera. Snow explicitly laid out the theory for cholera as an intestinal disease transmitted from person-to-person via the inadvertent swallowing of contaminated matter: "That a portion of the ejections or dejections must often be swallowed by healthy persons is, however, a matter of necessity" Snow (1849a, 9). Alternatives were generally not so explicit as to the disease or mechanism of transmission. In Appendix A we attempt, from a wealth of sometimes contradictory contemporary accounts, to construct historically accurate but cogent descriptions of alternative theories.

The three important dimensions along which we can categorize theories of the 1850s are: 1) the mode of transmission, either inhaled or ingested; 2) whether contagious or non-contagious; 3) the "seat of the disease" – the nervous system, blood, or the gastrointestinal tract; . The major alternatives to Snow were all inhaled disease theories. A major point of debate (prior to but continuing after 1849) was whether cholera was contagious or not, with important implications for predictions. By 1849 it was widely assumed that cholera was a blood disease, but for observations on mortality the distinction between blood and nervous system disease is not especially relevant – airborne and contagion were the important characteristics.

There are important distinctions – for example between general atmospheric non-contagious and localized non-contagious – but these do not have any large impact on the predictions from the alternate theories for the epidemiological data on observed mortality.

2.2 Predictions

We focus here on predictions concerning patterns of mortality across time, space, and groups. We might call this "epidemiological" or statistical evidence, which aggregates deaths across specific time periods, geographic locations, and groups of people – e.g. in epidemic curves, maps, and frequency tables of getting sick or not. In this context, predictions are formulated and evidence is collected across different geographic scales, from person-to-person transmission within homes, to neighborhood

[4]Snow was not the only one to develop an intestinal theory of cholera. Writing shortly after Snow, William Budd, a Bristol surgeon, independently proposed an intestinal theory (Budd, 1849).

[5]Blane and Corbyn (1821) cited evidence on both sides: For example pro-contagious was an episode from Mauritius where cholera seemed to be imported by the frigate Topaz from Cylon (Blane and Corbyn, 1821, 148-150, 152-153); anti-contagious was a case of three Sepoys in Seroor hospital who escaped disease despite being in close proximity to patients (Blane and Corbyn, 1821, 143)

outbreaks around pumps, municipal outbreaks due to contaminated water supply in London, and global transmission along trade routes.

Fecal-Oral Predictions Snow's core theory posited cholera as a disease of the intestine, transmitted by some unknown material that passed out of one individual (vomit or more commonly diarrhea) and was then inadvertently ingested by others. This core theory or hypothesis implied three modes of communication or transmission: (1) person-to-person ordinary or normal transmission; (2) neighborhood transmission when a local water (or food) source becomes contaminated; and (3) municipal transmission when a municipal water source becomes contaminated. These predict what we should observe about patterns of mortality.

- **Person-to-Person** or **Normal** Transmission: People in the same room and sharing air would not tend to get cholera unless they are in close physical contact or share food and drink but not otherwise (for example, doctors attending patients would tend not to become ill); Once a case of cholera occurs in a house or neighborhood, it will tend to transmit to others in close physical contact; Cholera "may be conveyed to a distance, and into quarters having apparently no communication with the sick" when "the patient, or those waiting on him, are occupied in the preparation or vending of provisions" Snow (1849a, 10).
- **Neighborhood** Transmission: Explosive growth in a local area when a localized source (say a pump-well) becomes contaminated. Cases will rise very quickly, but often tail off quickly – either because the source of the local contamination is removed or people move away. (Both factors came into play with the Broad Street outbreak in 1854.)
- **Municipal** Transmission: Widespread cholera in a city or area when a common water-source such as a river or a water company is contaminated. Large differential in mortality rate at the beginning of the outbreak, diminished differential as cholera spreads more widely through normal transmission.

Airborne Predictions The airborne theories were not well-articulated, coherent, and complete theories, in contrast to Snow's fecal-oral theory. We have attempted a rational reconstruction of the extant theories. A more detailed (but still cursory) discussion is in the appendix, but here we focus on the core attribute that is important for our predictions, that the cholera poison was airborne and so cholera would be transmitted via air. Auxiliary hypotheses stated that local or environmental factors would cause or amplify cholera disease, for example local dampness or filth generating or amplifying the cholera poison. The split between contagious and non-contagious could also by treated as an auxiliary hypothesis.

- **Person-to-Person** Transmission:
  - **Non-Contagious** & **Contingent Contagious**: No – People in the same room will not necessarily become sick. In the 1830s, an auxiliary hypothesis of *Contingent Contagion* was introduced to account for contrary observations, that cholera person-to-person

airborne contagion was sometimes observed, sometimes not.[6]

- **Contagious**: Yes – People in the same room will likely become sick.

- **Transmitted via Air**: People sharing the same air should suffer the same mortality, those not sharing the same air should experience different mortality.

- **Factors Increasing Mortality**: Predisposing causes and susceptibilities increase cholera mortality, and these factors will be associated with higher mortality. The list is extensive and includes: decomposing vegetation, soil contamination, overcrowding, dampness, filth, poor ventilation, proximity to graveyards, bad water, bad sanitation, poverty, unhealthy & moist living conditions. (Predisposing causes may generate or amplify the lethality of the cholera poison, and susceptibilities increase susceptibility to cholera illness.) Some of these factors were direct predictions of the theories – for example that crowded housing and poor ventilation, by forcing people into closer contact and sharing air, would increase transmission and mortality. Others were some combination of predictions plus auxiliary hypotheses adopted to account for observations – for example "filth" or sewage was presumed to produce air-borne cholera poison, but was also adopted in response to observations. It is important to recognize, however, that the airborne theories and predictions were *not* stated explicitly (as was Snow's intestinal fecal-oral theory) and that what we present is a rational reconstruction of those theories and predictions.

- **Water** (post-1849): Added as an auxiliary hypothesis – Water as a, possibly the most important, predisposing cause. Additionally, the airborne cholera poison may dissolve in water and be ingested (thus adding ingestion and water as a mode of transmission).

Appendix B (Table 4) lays out the detailed predictions that correspond to the evidence and mortality observations available at the time.

3 Testing for Specific Causal Effects: Counterfactuals & Causal Water 1850s

Single hypothesis testing within a treatment-control or potential outcome framework is the most commonly applied statistical framework for causal inference in social science research. Snow's analysis of water-borne transmission of cholera is a good example of demonstrating a causal effect, and Snow is credited with the first use of two research designs widely used today for demonstrating causal treatment effects – randomization as an instrumental variable and difference-in-differences.[7] Demonstrating a causal effect, however, was not and we argue should not be sufficient to convince skeptics. Demonstrating the superiority of Snow's theory requires examining a broad range of evidence and

[6] This is a rather obvious *ad hoc* hypothesis under Lakatos's standards. Such hypotheses are allowed but only when they produce new empirical content – when they are "progressive" and produce predictions beyond the facts or observations that motivated their introduction. This auxiliary hypothesis could account for the "old" observation that transmission sometimes appeared to be person-to-person airborne, sometimes not. But it did not produce any new theoretical or empirical predictions.

[7] See Greene (2018, 228), and also Deaton (1997); Grootendorst (2007), for randomization, and Angrist and Pischke (2014, 205) and Angrist and Pischke (2008, 227) for differences-in-differences. Arguably, Simon (1856, Table 1) is an even cleaner early example of difference-in-differences.

testing the plausibility of alternative potential explanations. This requires a shift in attention from treatment effects to explanations – from the *effects of causes* (Neyman (1923); Cox (1958); Rubin (1974, 1978); and Holland (1986)) to the *causes of effects* (Pinto and Heckman (2021)).

3.1 Treatment-Control Framework

Testing for causal effects within a treatment-control or Neyman-Rubin potential outcome framework (Neyman, 1923; Cox, 1958; Rubin, 1974; Holland, 1986) is the focus for much causal analysis in social, behavioral, and medical sciences. The basic question for any treatment, say consuming clean water in the case of cholera, is whether the outcome (mortality) is different when the treatment is applied (clean water) versus when it is not (contaminated water). This is asking a hypothetical or counterfactual question – would an individual's chance of dying from cholera be lower when drinking clean water? This counterfactual is the important question but impossible to measure empirically, because the individual drinks either clean water (treated) or contaminated (control). One cannot measure the same person both treated and not.

The idea is to explicitly introduce the counterfactual world (potential outcome) where the individual *does* consume clean water and we *do* observe the mortality outcome, contrary to the real world:

$$mortality\,rate = Y = \begin{cases} Y_1\,if\,treated\,(clean) & counter\text{-}factual \\ Y_0\,if\,control\,(contaminated) & observed \end{cases}$$

We cannot measure that counterfactual and thus the effect of treatment for any individual, but we can hope to measure some average treatment effect across a population: $ate = E\left[Y_1 - Y_0\right]$. The potential outcome approach focuses our attention on the statistics and mechanics of correctly estimating the effect of treatment (most basically an average treatment effect). The primary challenge is confounding by observable or unobservable factors – spurious association (correlation). The goal is to control for (eliminate the effect of) those confounding factors.

Tools for controlling for confounders are available. Regression and matching on observables control for observable characteristics. The real challenge is unobservable factors. Instrumental variables, of which randomization is the most intuitive and commonly used, controls for unobservable confounders. Experimental design, such as difference-in-differences (DiD), provide another set of tools. Both of these show up (possibly for the first time) in Snow's work.

3.2 Single Tests of 'Causal Water' in 1855

Many components of Snow's cholera analysis fall squarely within a treatment-control framework. Although Snow and his contemporaries did not have the statistical tools to conclusively demonstrate that water has a causal effect, they did have an understanding of many of the issues, and they did provide strong evidence to support the claim that water had a causal effect on cholera mortality. We need to reinforce, however, that single hypothesis tests of water within a treatment-control or potential outcome framework were *not* sufficient for separating waterborne and airborne research programmes.

Snow (1855) undertook an analysis of the large (400,000+) population in South London, exploiting two research designs which are in widespread use today – randomization as an instrumental variable and difference-in-differences (DiD). Although Snow's analyses suffered from flaws as discussed in the appendix, it nonetheless provided valuable evidence supporting water as a causal factor.[8] There was other, arguably stronger, evidence provided in the 1850s which showed the causal effect of water within a counterfactual framework. We highlight two instances: First, Sutherland (1851, 15-16) reported an outbreak in Hope Street, Manchester, in 1849 that showed very strongly a well as the source of disease. Second, Simon (1856) discussed at some length the same 1854 South London outbreak and population that Snow analyzed, but with the benefit of population-by-supplier detailed by subdistricts (which was not available when Snow published in 1855). Simon (1856, 9) concludes, justifiably, that

> The above conclusions [that higher mortality was caused by drinking contaminated water] rest on so large a basis of facts, that I venture to believe they will be accepted as the final solution of any existing uncertainty as to the dangerousness of putrefiable drinking-water during visitations of epidemic cholera.

In summary, the evidence available in 1855 strongly supported water as a causal factor in cholera mortality, even by today's standards. But in the end this result contributed little to deciding between theories, for two reasons. First, as will be shown in Section 5.1, existing airborne theories were modified so that water was incorporated as a causal factor, thus protecting them from refutation by "causal water". The second, more general, reason is that testing for and demonstrating an isolated causal effect is rarely sufficient to refute a theory within the overall process of scientific inquiry – Lakatos argues there are no "crucial experiments". In Section 5.3 we expand on this theme and show that "causal water" was only one piece of the overall evidence for comparing between theories. We argue that the broad range of evidence should have been sufficient to justify replacing existing theories with the fecal-oral theory, but single hypothesis tests of water alone were not sufficient.

4 Framework for Falsification and the Iterative Process of Scientific Inquiry

4.1 Sophisticated Falsification

The limited reception of Snow's theory in the 1850s presents something of a puzzle if one starts (as did the current authors, see Coleman (2018)) from a simple treatment-control approach to testing and discriminating between Snow's and alternative theories. The puzzle of strong empirical support but scant acceptance of Snow's theory is resolved by recognizing that choosing between theories (or research programmes) is rarely resolved quickly with a definitive experiment, but generally involves a long competition or war of attrition with multiple strands of evidence and repeated tests and comparisons. Lakatos's *sophisticated falsification* plays the central role as the criterion for comparison.

[8] Re-analysis using current statistical tools in (Coleman, 2020) shows that Snow's evidence does demonstrate water as a causal factor.

We use a framework that combines Peirce's idea for the stages of scientific inquiry together with Lakatos's methodology of *scientific research programmes*. Figure 2 provides an schematic overview, which is discussed in this section.

The process of scientific inquiry is an on-going interplay between data and theory, with theories modified in response to new evidence, and new facts in turn predicted by theory – what Heckman and Singer (2017) call the "abductive approach". This iterative process of inquiry, which goes back and forth between theory and data, stands in contrast to common paradigms in both economics and computer science. In empirical economics the conceptual framework (as opposed to actual practice) is often static, with theory development and testing separated: "Many analysts operating in this framework ... define empirical economics to be the activity of testing (falsifying or corroborating) *a priori* models" (Heckman and Singer, 2017, 299). Revising theories after analyzing data is seen with suspicion as cheating or data mining. In contrast, computer science is largely data-driven without reliance on prior theories, expressed in the view that "The Data Deluge Makes the Scientific Method Obsolete" (Anderson, 2008).

Before moving ahead we should say that we are not ignoring other substantive contributions to causal inference, but rather focusing on Peirce and Lakatos because they provide a concise and practical model for scientific inquiry. A very incomplete list of additional work includes discussions of causal inference in history of science (Laudan, 1968), philosophy of science (Salmon, 2006), statistics (Pearl, 2009) and other social sciences (Gangl, 2010; HedstrÃ¶m and Ylikoski, 2010); methodology in economics (Friedman, 1953),[9] and recent work in econometrics (Pinto and Heckman, 2021).

Lakatos (1980, 32) proposes *sophisticated falsification* as the specific criterion for separating between theories:[10]

> a scientific theory T is falsified if and only if another theory T' has been proposed with the following characteristics: (1) T' has excess empirical content over T: that is, it predicts novel facts, that is, facts improbable in the light of, or even forbidden by, T; (2) T' explains the previous success of T, that is, all the unrefuted content of T is included (within the limits of observational error) in the content of T'; and (3) some of the excess content of T' is corroborated.

An important implication (maybe not obvious from the way that Lakatos stated the criterion) is the importance of *both* corroborations of the new T' (3) *and* contradictions of the original T (1). In other words, sophisticated falsification requires contradiction of the existing (or parts of the existing) theory, which looks something like a form of naive or even dogmatic falsification. But with two important additions. First, additional empirical content that is corroborated is also required, and second "falsification" is always provisional and cannot immediately or automatically lead to rejection of the original theory T.

[9] Friedman (1953) lays out a methodological framework with many similarities to Lakatos's research programmes: "Such a theory [for a positive science] is, in general, a complex intermixture of two elements ... a 'language' ... [and] a body of substantive hypotheses." See also Friedman's footnote 11 for a discussion of "identification" in economics and references to the economics literature.

[10] Lakatos distinguishes between three forms of falsification: *dogmatic*, *naive* or methodological, and *sophisticated* – see Lakatos (1980, 12-47). Both dogmatic and naive falsification are open to the Duhem-Quine critique (Quine 1953, chapter 11; Lakatos 1980, 93-101). Lakatos discusses rules for the falsification of research programmes (as opposed to theories within a programme) on page 69 and page 179, quoted above.

This last condition, that adjustments must predict new facts (increase the empirical content of the theory) manages the problem that any theory may be adjusted to accommodate contradictory evidence (the Duhem-Quine thesis). Lakatos's framework of research programmes recognizes and embraces the fact that theories – particularly the protective belt of auxiliary hypotheses – may and indeed *should* be adjusted to incorporate and subsume "anamolies" and new evidence. The challenge is to distinguish between *ad hoc* adjustments that adjust only to accommodate inconvenient observations and should be disallowed, versus "legitimate" adjustments that refine and enhance the overall programme. Rather than provide prior conditions to distinguish between disallowed versus allowed adjustments, here the focus is on the result and whether an adjustment improves the theory by predicting new facts (increasing empirical content). Failure to predict new facts does not immediately disqualify a theory or programme, but does separate between progressive programmes that increase scientific knowledge, versus degenerating programmes that, ultimately, do not.

In this context, we need to consider the *working unit* for inquiry. Rather than a single hypothesis or theory, we use Lakatos's *research programme*, a structured set of theories and hypotheses: "my unit of appraisal is not an isolated hypothesis (or a conjunction of hypotheses): a research programme ... consists of a developing series of theories [with] ... a structure. It has a tenacious *hard core* ... a *heuristic*, [a set of rules for analyzing the theory and data] ... [and] a vast belt of auxiliary hypotheses [that translate from the core to the world of observations]" Lakatos (1980, 178).

4.2 Iterative Comparison Across Theories & Programmes

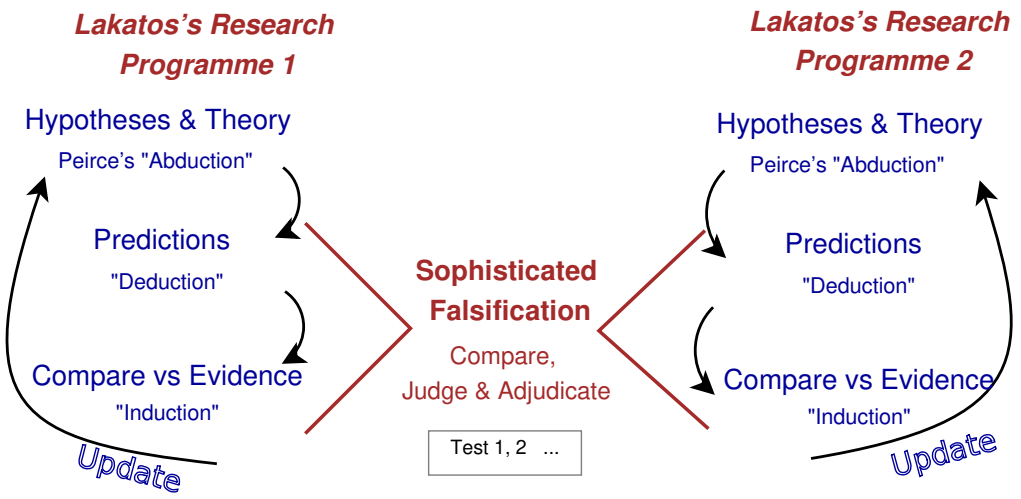
An important component of the inquiry process is the comparison across and adjudication between theories and programmes. The core of the programme is not subject to refutation or contradiction, until and unless the complete programme is replaced by another by the scientific community – to which we turn shortly when we discuss *sophisticated falsification* and the comparison across programmes. The belt of auxiliary hypotheses, in contrast, may be refuted, altered, and added to. This belt of auxiliary hypotheses and theories provides the translation from the core to the world of observed phenomena, while also protecting the core from refutation.

Comparison across programmes shows us how and why we should replace one programme (set of theories) with another. Our initial thought might be to a simple comparison of a theory's predictions versus observations or facts. If the predictions pass, the theory is (provisionally) accepted, if not then the theory is summarily rejected. This is what Lakatos calls dogmatic or naive falsification.[11] But the adjustment of theories to accommodate conflicting facts raises substantial problems for a naive approach to falsification (including Popper's). Lakatos emphasizes that "Popper's great negative crucial experiments disappear" Lakatos (1980, 111) and that, "Contrary to naive falsification, *no experiment, experimental report, observation statement or well-corroborated low-level falsifying hypothesis alone can lead to falsification*" Lakatos (1980, 35). Thus the necessity and importance of Lakatos's sophisticated falsification, discussed above.

Lakatos (1980, 179) provides the criterion for replacing one research programme (collection of core and auxiliary theories) by another:

[11] See Lakatos (1980, Section 1.2) and also page 93 ff for more detail on various flavors of falsification.

Figure 2: Schematic for the Iterative Process of Scientific Inquiry



We combine Peirce's three stages of scientific inquiry with Lakatos's *scientific research programmes* and sophisticated falsification. Scientific inquiry involves a dynamic back-and-forth of data and theory, from theory generation to predictions to comparing against data, and back again to theory generation and modification. Scientists work with research programmes, which are collections of theories & hypotheses. Sophisticated falsification is the criterion to appraise and compare across programmes, and may involve testing various hypotheses and predictions (*Test 1, 2, ...*) within a Neyman-Rubin potential outcome framework.

> Having specified that the unit of mature science is a research programme, I now lay down rules for appraising programmes. A research programme is either progressive or degenerating. It is *theoretically progressive* if each modification leads to new unexpected predictions and it is *empirically progressive* if at least some of these novel predictions are corroborated. It is always easy for a scientist to deal with a given anomaly by making suitable adjustments to his programme (e.g. by adding a new epicycle). Such manoeuvres are *ad hoc*, and the programme is *degenerating*, unless they not only explain the given facts they were intended to explain but also predict some new fact as well.
>
> ...
>
> One research programme *supersedes* another if it has excess truth content over its rival, in the sense that it predicts progressively all that its rival predicts and some more besides.

Research programmes are thus fundamentally dynamic, being a series of theories and hypotheses that progress (or degenerate) over time. A progressive research programme provides a good working definition of effective scientific inquiry. A degenerating programme does not produce excess content or predict new facts; it is not effective scientific inquiry. The distinction between progressive and degenerating research programmes resolves the problem (at least in the long run) for how to choose between theories or research programmes: (1) scientific progress requires discarding degenerating programmes, and (2) adjustments to auxiliary hypotheses that protect a theory "are *ad hoc*, and the programme is degenerating, unless they not only explain the given facts they were intended to explain but also predict some new fact as well" (Lakatos, 1980, 179).

Figure 2 shows a schematic for the iterative process of scientific inquiry, with sophisticated falsifi-

cation playing the central role for comparing across theories. Peirce proposed scientific inquiry as comprising three stages: generating theories or hypotheses (abduction), deriving predictions from theory (deduction), and comparison against evidence (induction) Misak (2004, 93-95). He saw the path of scientific inquiry as a dynamic interplay between data and theory, a progression from a settled belief that encounters a surprise, through doubt, inquiry, and back to settled (but altered) belief: "Peirce characterises the path of inquiry as follows: belief—surprise—doubt—inquiry—belief" (Misak, 2004, 50). This process is not linear, however, but dynamic and iterative – comparing predictions against evidence will inevitably produce anomalies, contradictions, and surprises, which entail further development and modifications of theory and hypotheses. Our sharp distinction between the three stages is a stylization and something of an exaggeration, but captures important aspects of the process. Finally, Figure 2 shows that hypothesis tests (1, 2, ...) as described in /refsec:causal-water form one component of the larger iterative process of scientific inquiry – crucial but only a component.

5 Historical Examples of Iterative Process and Sophisticated Falsification

5.1 The Iterative Process of Scientific Inquiry: Snow's Cholera Theory 1849-1866

This section provides an overview of the development and progress of Snow's intestinal fecal-oral theory from 1849 to 1866. It demonstrates the iterative process of inquiry, and particularly the use of auxiliary hypotheses in protecting the airborne theories.

Some time in 1849 Snow seems to have made a mental leap, an abductive leap, to recognize that airborne transmission was not necessary: "There is, however, no reason to conclude, *a priori*, that this [contagion via inhalation] must be the mode of communication of cholera" Snow (1849a, 6). Having broken from inhalation, and calling upon both his medical experience in treating patients and widespread observations of the symptoms of cholera ("it has always appeared, from what the writer could observe, that in cholera the alimentary canal is first affected" Snow 1849a, 7), Snow hypothesized that cholera was a disease of the intestines (the alimentary canal) and transmitted person-to-person via ingesting contaminated fecal matter because, unintentionally, "a portion of the ejections or dejections [from a patient with cholera] must often be swallowed by healthy persons" (Snow, 1849a, 9). Snow proceeded to develop a variety of implications or predictions for observed mortality and transmission of cholera, as discussed in Section 2 and Appendixes A and B.

Snow combined the fact of long-standing evidence contradictory to inhalation (evidence refuting those theories), observations and evidence concerning cholera, and intuition and knowledge of medicine to develop his theory. But this was only the first step in an on-going process of inquiry that went back-and-forth between data and theory. To supplant existing and widely-accepted theories, more was required than a good idea and some scattered evidence, as Snow himself recognized: "The opinions now made known [in late 1849] have been entertained by the author since the latter part of last year ... but he [the author, Snow] hesitated to publish them, thinking the evidence in their

favour of so scattered and general a nature as not to be likely to make a ready and easy impression" Snow (1849a, 12).

Snow had to move to the next stages of the process, deduction of predictions and induction versus evidence. The was the first attempt at falsification of alternatives, the attempt to use evidence to refute alternative theories while showing that the evidence supported his theory – corroborating his predictions and showing that his theory had "higher corroborated content". Snow had to collect and examine evidence, both existing and new. He had to marshal facts, logic, and explanation to make a case for his theory versus alternatives. We know that Snow found additional evidence in late 1849: "Within the last few days, however, some occurrences have come within his [the author, Snow's] knowledge which seem to offer more direct proof [of the theory], and have induced him [Snow] to take the present course [publishing]" (Snow, 1849a, 12).

The new evidence Snow discovered were the neighborhood outbreaks in Thomas Street (Horsleydown), and Albion Terrace (Wandsworth Road). Snow spent roughly 10 pages providing a detailed description of the two sets of buildings and specifically the way in which faulty plumbing allowed sewage-contaminated water to be shared among the residences. Snow was undertaking the falsification or comparison part of the inquiry process, comparing predictions against evidence to sort among alternative theories. For now we only outline the falsification effort, returning in Section 5.3 to use Snow (1855) and related publications to examine it in detail. Here we highlight that Snow (1849a) 1) developed predictions (deduction) – particularly how cholera would be transmitted via hand-to-hand contact and via contaminated water; 2) provided evidence (induction) – particularly on the importance of water in cholera transmission and mortality; 3) compared predictions versus evidence for his theory and (rather cursorily) for alternatives (falsification).

For present purposes the important result is that Snow provided strong evidence that sewage-contaminated water was an important factor in cholera mortality. Contemporaries acknowledged the importance of water, due to Snow's work and also evidence from others. William Farr of the General Registrar Office concluded that "Dr. Snow is unfortunately able to show that this excremental distribution ... is possible to a very considerable extent.... The experience of ... many of our own towns lends some countenance to Dr. Snow's theory" (General Register Office, 1852, lxxviii). During and shortly after the 1849 epidemic, local officials throughout England and Wales submitted reports to London that detailed similar associations between water quality and cholera mortality. In its final report, the General Board of Health noted that "much additional evidence has been elicited proving the influence of the use of impure water in predisposing to the disease" (General Board of Health, 1850, 50). John Sutherland, a supervising inspector, noted that "in nearly every city and town affected, ... a number of most severe and fatal outbursts of cholera were referable to no other cause except the state of the water supply" (Sutherland, 1851, 14). Richard Grainger, who summarized reports received from local inspectors in the London metropolis after the 1848-49 epidemic, cited "unquestionable evidence" of water "contaminated in various ways by decomposed organic and noxious matter" and then "taken up ... by the blood vessels of the alimentary canal ... as noxious aerial agents do by way of the respiratory apparatus" (Grainger, 1851, 91)

Remember that, prior to Snow, cholera was assumed to be airborne and propagated by inhalation. And yet by 1851 impure water – broadly defined as water containing decomposing organic matter – was widely recognized as at least a contributing cause for cholera; for some, water was elevated

to "perhaps the most fatal of all" predisposing causes (Sutherland, 1851, 14). It is natural for a modern reader to assume that evidence for water as a causal factor would serve to refute airborne theories and confirm the ingested fecal-oral theory, and furthermore that falsification of inhalation would be widely recognized. This was not the case and instead water was incorporated into the existing theories, making them in effect hybrid theories of inhalation with water as a secondary mechanism. Even though this was an incorrect and an unfortunate scientific response, it was a reasonable response.

The research programme framework tells us that any prediction is a combination of core and protective hypotheses. Evidence refuting a prediction refutes the joint hypotheses, not the core theories alone. It is perfectly appropriate to maintain the core and reject (and then adjust) the auxiliary hypotheses. A single anomaly, even an accumulation of contrary evidence, does not of necessity refute the core theories.

For pre-1849 inhalation theories the core would be that cholera is a disease of the blood with individuals inhaling a choleric poison which entered through the lungs, into the blood, and thus caused disease (by some unknown mechanism). The protective belt of hypotheses would detail how the choleric poison was produced and transmitted and thus the observed patterns of disease and mortality. One hypothesis would be that cholera is contagious with the choleric poison created in an individual and passed person-to-person via inhalation. If this hypothesis were contradicted by evidence, this would not (of necessity) contradict the core of an inhaled blood disease, but simply require adjustment of the auxiliary contagion hypothesis – adjusting by hypothesizing that cholera is non-contagious.

Now we come to evidence that water is strongly associated with cholera mortality. Rather than taking such evidence as refuting the underlying core of an inhaled blood disease, such observations were interpreted as positive evidence that the cholera poison was produced or transmitted through water and that a source of the cholera poison had been previously overlooked – water as a predisposing causal factor was introduced as a (protective) auxiliary hypothesis to the airborne theories.

Simplifying somewhat, the outcome of the Snow's scientific inquiry and abductive efforts from the 1849 outbreak, culminating in the publication of the 1849 first edition of *On the mode of communication of cholera*, was a productive stand-off: (1) the causal importance of water was widely (although not universally) recognized; (2) water as a causal factor was incorporated into inhaled blood disease theories through expansion of the protective belt of auxiliary hypotheses; (3) water was removed as a crucial distinguishing factor between theories, with inhaled blood diseases now "protected" against evidence of water as a causal factor; and most importantly (4) the fecal-oral theory was recognized although not widely adopted. In the following years both Snow and skeptics were impelled to collect new evidence, re-examine existing evidence, and refine predictions to compare against that evidence.

Cholera struck again in 1854, bringing new hardship but providing substantial new evidence on cholera. The Broad Street outbreak in Soho in early September 1854 is probably the most widely known episode in the history of epidemiology. The case of South London, which closely approximated a randomized trial with treatment by clean and contaminated water, provided quantitative evidence from a population of over 450,000. Snow published the 2^{nd} edition of *OMCC* in 1855 (Snow, 1855), and also papers in learned journals (e.g. Snow, 1856b) in a renewed attempt at falsification, which

we discuss in Section 5.3. The fecal-oral theory gained marginally wider acceptance but was not generally accepted.

In 1866, cholera arrived in London for the fourth (and final) time. In the years since 1855 (and Snow's death at the age of 45) the fecal-oral theory had gained wider although not universal acceptance. Reports to the Privy Council, to Parliament, and by the *Lancet* all recognized the fecal-oral theory, and Filippo Pacini's identification of the "germs of vibrions" (Farr, 1868, xv), as the correct theory for the cause of cholera.

5.2 Protective Auxiliary Hypotheses: Airborne Theories and Water 1849-1855

As mentioned in the prior section, water was introduced as a causal factor, an auxiliary hypothesis, for the airborne theories. It is important to recognize that this was not irrational, and was not contrary to accepted scientific knowledge of the time, or even current scientific knowledge. Possibly water and general dampness contributes to organic decomposition that amplifies production of airborne cholera poison that is then inhaled, thereby producing a strong (and causal) association between contaminated water and cholera mortality. Contemporary writers explicitly proposed such mechanisms, with Grainger (1851, 34) as one example:

> when the numerous sewers of a city reach the stream [rivers], ... [this] presents a physical condition favourable to its subsequent escape into the atmosphere in the form of mephetic gases

Another mechanism, possibly acting in parallel, is that the airborne cholera poison can be dissolved into water, ingested, and then make its way to the bloodstream. Exactly this mechanism was laid out by Grainger (1851, 91):

> water ... is contaminated in various ways by decomposed organic and other noxious matter; as by the absorption of deleterious gases ... these injurious substances get into the blood ... the most poisonous being soaked up by the blood-vessels of the alimentary canal as readily as the most harmless. Thus a second great inlet is established by which injurious substances in a liquid form as certainly find their way into the circulating blood, as noxious aerial agents do by the way of the respiratory apparatus.

Today we know these proposals are mistaken but they were not contrary to either medical knowledge or scientific method – they are *not* examples of smart people clinging to outlandishly incorrect ideas. They are examples of an unfortunate but still rational response to contrary evidence or anomalies, adjusting the protective belt of hypotheses that translate from the core theories (inhaled blood disease) to the world of observations (cholera mortality). Yes, strong evidence of the causal effect of water supports Snow's fecal-oral theory. And yes, the evidence is contrary to a simple (non-water) version of inhaled blood diseases. But instead of throwing out the existing theory, new auxiliary hypotheses were introduced concerning the role of water in the production and transmission of the cholera poison, and thus the previously contradictory evidence incorporated into the inhaled blood-disease theory (or research programme).

5.3 Sophisticated Falsification: Comparing Theories 1855

Choosing among competing theories is not easily formalized and does not neatly fit in a statistical testing framework – the process can be qualitative, fuzzy, and somewhat ambiguous. It is nonetheless one of the most important components of scientific inquiry, and it *can* be organized and structured. Choosing between competing theories is often a slow and complicated war of attrition (Lakatos, 1980, 133), producing neither the "instant rationality" of Popper's falsification (Popper, 1985) nor the "sudden, irrational changes in vision" of Kuhn (1962) (quoting in both instances Lakatos 1980, 6).

We rely on Lakatos's *sophisticated falsification* – the process of one theory (with higher corroborated content) superseding another (with lower content) – discussed in Section 4.1 above (and Lakatos, 1980, 32). Lakatos emphasizes that with sophisticated falsification crucial experiments disappear (see, e.g. Lakatos, 1980, 35, 111) and it is multiple sources of evidence across various predictions that is compelling. Snow's efforts in 1855 and 1856 are an exemplar and test case because it exhibits the slow and sometimes fitful process of developing scientific knowledge, and also the use of multiple strands of evidence – evidence and comparisons that sometimes includes qualitative in addition to quantitative or statistical evidence – for comparing across a wide range of predictions.

Section 2 introduces the theories (Snow's and alternatives) and predictions. The appendixes discuss in more detail the theories (Appendix A), predictions (Appendix B), and comparison versus evidence (Appendix D). Here, in Table 1, we focus on a subset of the predictions and comparison versus evidence: person-to-person transmission (2 a&b); airborne transmission (4); water and other factors (5b and 5c-f); and the time-pattern of mortality or the "epidemic curve" (8 and 9). These are predictions that are particularly valuable for discriminating between theories (2 a&b, 5c-f, 8, 9), or highlight the consequence of introducing the water-based auxiliary hypotheses (4 and 5b), showing why the airborne theories are degenerating according to Lakatos's criterion.

There are two levels on which we consider whether the airborne theories were "falsified" by the evidence in 1855-56. First, naive falsification, whether the evidence refutes or contradicts the airborne theories. We might be tempted to answer "yes" but close examination of Tables 1 and **??** show why this is a vain hope. The final two columns of Table **??** *do* show that the airborne theories are (naively) contradicted by some but only some of the evidence – the theories are protected on two crucial predictions regarding airborne and waterborne transmission (4 and 5b). For those predictions that *are* contradicted, suitable adjustments to the theories could no doubt be made: "It is always easy for a scientist to deal with a given anomaly by making suitable adjustments to his programme" Lakatos (1980, 179). Furthermore, the evidence for other factors not affecting mortality (5c-f) was rather weak. The fecal-oral theory is without doubt better supported by the evidence, but there is no single observation to which we can point as definitively refuting the airborne alternatives.

Turning to Lakato's three-part criterion for sophisticated falsification (Lakatos 1980, 32 and Section 5.3 above), the fecal-oral theory satisfies all three conditions:

1. the fecal-oral theory has excess empirical content, predicting new patterns of person-to-person transmission different from airborne theories (2b); the absence of factors other than water affecting mortality (5c-f); and time-patterns of mortality (epidemic curve, 8 & 9)

Table 1: Highlights of Comparing Predictions versus Evidence

	Prediction	Evidence / Observation	Value	Support? Fecal	Air
	PERSON-TO-PERSON CONTAGION				
2a & 2b	***Airborne theories*** predict that when contagion occurs it is between those sharing airspace in close proximity ***Fecal-oral theory*** predicts contagion occurs for those in close proximity, via sharing food, clothing	*Quality*=**Good**. Anecdotes, reports, and case studies. See Blane and Corbyn (1821, 143, 148-150, 152-153), Snow (1855, 2-9)). *Support*=**CONTRADICTORY** for Airborne theories since some reports show those sharing airspace get sick, others not. *Support*=**YES** for fecal-oral since reports are consistent with transmission via food, clothing, hand-to-hand contact.	**High**	Y	N
	GENERAL CHARACTERISTICS				
4	***Transmission by Air – Airborne theories*** predict transmission by air: Sharing same air (different water) ⇒ similar mortality; sharing different air (same water) ⇒ different mortality	*Quality*=**Good**. Broad Street data showing that transmission is *not* by air – evidence on those sharing air but not water, and water but not air (Snow 1855, 42, 44-45; Whitehead, Cholera Inquiry Committee, 1855, 128 ff). *Support*=**NO** for airborne theories but **Protected** by the auxiliary hypotheses concerning water as a predisposing cause and as a mode of transmission.	Medium	Y	P/N
5b	***Water – Airborne and fecal-oral*** theories (post-1849) all predict water is a causal factor for cholera mortality	*Quality*=**High**. Supported by multiple strands of high-quality evidence (narrative of neighborhood outbreaks (Albion Terr, Broad St, etc. Snow 1849a, 12 ff, 15 ff; Snow 1855, 23-56); quantitative analysis of municipal & neighborhood outbreaks (Snow 1855, 89, Table IX ; Sutherland 1851, 14 ff; Simon 1856, 9 & 13-14; Whitehead, Cholera Inquiry Committee, 1855, 128 ff)). *Support*=**YES** for fecal-oral, **NO** but **Protected** for airborne by auxiliary hypotheses (water as a predisposing cause, mode of transmission).	Low	Y	P/N
5c, d, e, f	***Other Factors – Airborne theories*** predict cholera is associated with variety of factors, independent of water ***Fecal-oral theory*** predicts water will be the overriding factor in a neighborhood or municipal outbreak and other factors will have no or very small effect	*Quality*=**High**. Snow (1856b) attempts to show the "overwhelming influence" of water – doing so would be strong evidence contradicting airborne theories. *Support*(1856)=Weak **YES** for fecal-oral and **NO** for air – without regression tools Snow could not convincingly demonstrate effect *Support*(2022)=Strong **YES** for fecal-oral, **NO** for air – Re-analyzing with modern statistical tools contradicts the airborne hypothesis – 1854 South London other factors had no significant effect.	Medium	Y	N?
	EPIDEMIC CURVE				
8	***Fecal-oral theory*** predicts explosive neighborhood outbreaks: Mortality grows quickly initially then falls off	*Quality*=**High**; *Support*=**YES** – Narratives of neighborhood outbreaks (Snow 1849a, 12 ff, 15 ff; Snow 1855, 23-56) showing explosive growth from low background rates	**High**	Y	N
9	***Fecal-oral theory*** predicts municipal outbreak mortality will differ early vs late: Early: high mortality for those exposed to contaminated water; Later: difference remains but diminishes as normal and neighborhood propagation contribute	*Quality*=**High**; *Support*(1855)=**YES**: Snow highlighted in multiple instances, (Snow 1855, 82, 86, 88; 1856b; 1856a, 243; 1857) *Quality*=**High**; *Support*(2021)=**YES**: DiD combined plus mixed population (quasi-randomization) split into first 4 vs next 3 vs final 8 weeks	**High**	Y	N

Table 2: Comparison of Predictions versus Evidence

	Prediction & Observation	Fecal-Oral, Intestinal (Snow)	Airborne, Blood Disease: Non-Cont Contingent	Airborne, Blood Disease: Contagious	Evidence / Observation	Value	Contradict? Fecal	Contradict? Air
PERSON-TO-PERSON CONTAGION		2	3	4	5	6	7	8
1	***Geographic Spread***: Cholera occurs along trade routes, at the speed of travel	yes Support	no Weak contradict	yes Support	*Quality*: **Moderate**; *Agree*: **YES** – Anecdotes and general observations that it spreads along trade routes – not specific or detailed	Low	N	M
2a	***Person-to-Person airborne contagion***: occurs between those sharing airspace in close proximity	na	yes/no Protected	yes Contradict	*Quality*: **Good**; *Agree*: **CONTRADICTORY** for Airborne – in some cases those sharing the same airspace are infected, sometimes not *Quality*: **Good**; *Agree*: **YES** for fecal-oral Anecdotes, reports, and case studies. See Blane and Corbyn (1821, 143, 148-150, 152-153), Snow (1855, 2-9)	**High**	na	M
2b	***Person-to-Person non-airborne contagion***: occurs for those in close proximity, via sharing food, clothing	yes SUPPORT	na	na			N	na
3	***Winds***: Cholera follows winds (not contagious)	no Weak support	no Weak support	yes Weak contradict	*Quality*: **Poor**; *Agree*: **NO** – Anecdotes and general observations that it does not follow winds – not specific or detailed	Low	N	M
GENERAL CHARACTERISTICS								
4	***General transmission by air***: Sharing same air (different water) ⇒ similar mortality; sharing different air (same water) ⇒ different mortality	no SUPPORT	yes Protected	yes Protected	*Quality*: **Good**; *Agree*: **NO** – Broad Street provides events and data on those sharing air but not water, and water but not air (Snow 1855, 42, 44-45; Whitehead, Cholera Inquiry Committee, 1855, 128 ff)	Medium	N	P
5a	***Water is causal***: pre-1849, cholera not associated with water (NB – prediction from airborne is reversed)	na	no Contradict	no Contradict	*Quality*: **High**; *Agree*: **YES** – Multiple strands: narrative of neighborhood outbreaks (Albion Terr, Broad St, etc. Snow 1849a, 12 ff, 15 ff; Snow 1855, 23-56); quantitative analysis of municipal & neighborhood outbreaks (Snow 1855, 89, Table IX ; Sutherland 1851, 14 ff; Simon 1856, 9 & 13-14; Whitehead, Cholera Inquiry Committee, 1855, 128 ff)	Low	na	Y
5b	***Water is causal***: post-1849, cholera associated with water – post-1849	yes SUPPORT	yes Protected	yes Protected			N	P
5c, d, e, f	***Other factors***: Cholera associated with variety of factors, independent of water: 5c) Overcrowding & poor ventilation; 5d) Hot weather & decomposing refuse; 5e) Poverty & unhealthy living conditions; 5f) Dampness, filth, bad sanitation	no 1856: weak 2021: strong support	yes 1856: Weak 2021: strong contradict	yes 1856: Weak 2021: strong contradict	*Quality*: **High**; 1856 *Agree*: **Weak NO**: Snow (1856b) attempts to show the "overwhelming influence" of water but without statistical tools of regression only modestly successful 2022 *Agree*: **Strong NO**: Strong evidence that other observed correlations are spurious	Medium	N	Y
6	***General characteristics***: Cholera erupts simultaneously in non-contiguous locations; Visits same areas in subsequent outbreaks	yes Support	yes Support	no Contradict	*Quality*: **Moderate**; *Agree*: **YES** – Anecdotes and general observations – not specific or detailed	Low	N	M
7	***Elevation***: Mortality falls with higher elevation and further from river	no 1856: weak 2021: strong support	yes 1856: Support 2021: Contradict	no Neutral	*Quality*: **Good**; *Agree*: **YES but spurious** – Quantitative analysis (Farr, General Register Office, 1852, lxi ff) but Farr recognized association with water and crowding. 2021 re-analysis shows relation is spurious due to association with water supply.	Low	N	M
EPIDEMIC CURVE								
8	***Explosive neighborhood outbreaks***: Mortality grows quickly initially then falls off	yes SUPPORT	no Contradict	no Contradict	*Quality*: **High**; *Agree*: **YES** – Narratives of neighborhood outbreaks (Snow 1849a, 12 ff, 15 ff; Snow 1855, 23-56) showing explosive growth from low background rates	**High**	N	Y
9	***Municipal outbreaks early vs late***: Early: high mortality for those exposed to contaminated water; Later: difference remains but diminishes as normal and neighborhood propagation contribute	yes SUPPORT	no Contradict	no Contradict	*Quality*: **High**; *Agree*: **1855 YES**: Snow highlighted in multiple instances, (Snow 1855, 82, 86, 88; 1856b; 1856a, 243; 1857) *Quality*: **High**; *Agree*: **2021 YES**: DiD combined plus mixed population (quasi-randomization) split into first 4 vs next 3 vs final 8 weeks	**High**	N	Y

Each row presents a prediction about observed patterns of mortality, as in Table **??**. The sixth column indicates the evidence regarding the prediction. In columns 3-5 the "yes" or "no" indicates whether the theory and evidence agree with the prediction, or the opposite. The color of each cell indicates whether the evidence supports (teal), contradicts (red), or weakly supports / contradicts (orange) the theory in the column. Protected (violet) indicates the theory is protected from refutation by adoption of a protective auxiliary hypothesis. The seventh column is a subjective assessment of the value of the prediction in separating between theories, and may be "Low" because either the predictions do not separate, or the evidence is low quality. The final two columns indicate whether the indicated theory is contradicted by the evidence: "Y"=yes, "N"=no, "M"'=mixed.

2. the fecal-oral theory explains the previous success, regarding geographic spread, transmission by water, and other observations

3. new predictions (2b, 5c-f, 8 & 9) are corroborated

Existing airborne theories could be and *were* adjusted to accommodate the evidence, but in a manner that was *ad hoc* and indicative of a degenerating programme. The contingent-contagious hypothesis was introduced to accommodate contradictory observations that cholera sometimes appeared to be airborne-contagious, sometimes not. Contingent-contagion was, however, neither theoretically nor empirically progressive, producing no new predictions concerning contagion or other aspects of cholera mortality. This is in contrast to the fecal-oral hypothesis for person-to-person transmission, that 1) predicted (and thus made consistent) the cases that seemed to be contradictory under an airborne theory; and 2) predicted measures such as hand-washing (see Snow, 1855, 133 ff) that would prevent person-to-person transmission and could in principle be observed and tested.

The adoption by airborne theories of water as a major predisposing cause, and even as a mode of transmission (airborne cholera poison being diffused into water) was also *ad hoc* – it accommodated contrary observations but produced no new predictions. Again, the fecal-oral theory *did* produce new predictions (5c-f that other factors would not be important in municipal outbreaks, and 8 & 9 concerning the time-pattern of neighborhood and municipal outbreaks) all of which were at least weakly corroborated in the 1850s (and strongly corroborated using modern statistical tools – see Coleman 2020).

In summary, while continued adherence to the airborne theories was not *per se* irrational or illogical, the theories could not avoid the conclusion that they were degenerating under Lakatos's sophisticated falsification criterion. The fecal-oral theory, in contrast, had attributes of a progressive programme, and we can conclude there was sufficient evidence for the fecal-oral theory to supersede the alternatives.

6 Conclusion

Snow's 1850s efforts at falsification of alternative theories (primarily Snow, 1849a, 1855, 1856b) had limited success in persuading the medical, public health, and political establishment that the fecal-oral theory was superior to alternatives and in fact correct. In Section 5.3 and Appendix D we argued that the evidence available in the 1850s was sufficient to separate the theories according to Lakatos's criteria of sophisticated falsification – the fecal-oral programme was progressive while the airborne programme was degenerating. In this section, we touch upon later stages of the story, i.e. what happened to the fecal-oral theory after 1856, before concluding the article.

6.1 The Ultimate Success of the Fecal-Oral Theory

One might speculate that had Snow presented his evidence structured as we do in Section 5.3 he might have been more successful, but this would be too facile. There was some reason for slow acceptance. As discussed in Section 2.2, evidence supporting the fecal-oral theory was statistical or

epidemiological, not biological. Although Filippo Pacini identified *vibrio cholerae* in 1854 (see Carboni, 2021), there was no substantive or reliable biological evidence in the 1850s. In some cases there was evidence on the specific mechanism of transmission of fecal material – for example the Broad Street index case of baby Lewis (Appendix D) – but such evidence was scarce. Without biological evidence on the causative agent and detailed mechanism of transmission, skeptics questioned the plausibility of the fecal-oral theory, in spite of the strong support from statistical or epidemiological evidence. Advocates of the fecal-oral theory had to rely on primarily epidemiological and statistical (circumstantial) evidence until the end of the 19^{th} century.

By the time cholera struck London for the last time in 1866, the fecal-oral theory, and particularly the importance of water in wide-spread municipal transmissions, had been accepted by important figures within the medical establishment. Snow's careful analysis, his attempt at falsification of alternative theories through accumulation of evidence and careful analysis – a template for good scientific inquiry – had indeed won converts. William Farr, in the report to Parliament from the Registrar-General, (Farr, 1868, 99), states the fecal-oral theory and water propagation as a seeming accepted matter and the *Lancet*, in reviewing Farr's report, agreed on the importance of the water supply (Lancet, 1868a).

The report to the Privy Council by J. Netten Radcliffe (Radcliffe, 1867) also identified waterborne transmission as the primary cause of the high and localized mortality (see also the review in the Lancet 1867). Further, Farr, the Lancet, and Radcliffe advocated a germ theory that we can recognize today:

> if the cholera germs are living molecules or organisms, the uniformity of their distribution in water would depend on a variety of conditions. Only one in ten or more persons may actually swallow these germs; and even if the whole ten did swallow them, it by no means follows that they would all be affected in an equal degree. (Lancet, 1867, 559, describing Radcliffe's views)

Farr discussed Fillippo Pacini's discovery and identification in 1854 of the cholera bacterium, what Farr called *cholera corpuscles*. Farr (also quoted in the Lancet 1868b) used estimates from Pacini of the volume of diarrhoea (dejections) and concentration of bacteria to account for how a single individual could contaminate a river (Farr, 1868, xv, lxiii). However, there do not appear to have been efforts to actually identify *v. cholerae* in water samples via microscopy – possibly because Farr, and through Farr other medical professionals, were not aware of Pacini's results until 1867 and the end of the outbreak (Farr apparently visited Pacini in 1867) or possibly Pacini's microscopy tools and skills were not available in London in 1866.

Despite some acceptance, the fecal-oral or water-only theory was not yet ascendant. Luckin (1977) provides some of the answers for why, discussing the commercial and political forces that shaped the social and medical environment. Luckin emphasizes that one crucial piece missing from the fecal-oral germ-based theory was a method for testing (and identifying) the cholera "poison" in contaminated water – what we have termed "biological evidence". Without such a means of clearly identifying the causal agent, skeptics could always cast doubt on the conclusions. It was only after Robert Koch re-identified and isolated *v. cholerae* in pure culture that the fecal-oral theory became widely accepted.

6.2 Lessons from Cholera in the 1850s

This study of cholera and John Snow's role in developing a causal explanation runs counter to popular understandings of the case and statistical practices. First, Snow was not the lone voice of cholera knowledge in a wilderness of miasma ignorance. Instead, many contributed to the accumulation and analysis of data and actively debated the causes of cholera, in the 1850s and in following years. Evidence and the case for the fecal-oral theory were improved through this competition between theories. The case illustrates how scientific discovery did not only occur in the minds of individual scientists but was facilitated through a community of competing stakeholders (see also **?**).

Second, evidence of water as a causal factor was not, and in the larger context of scientific inquiry could not be, the crucial piece of evidence that defeated alternatives and cemented Snow's theory as the explanation for cholera. The incorporation of water as a causal factor by alternative theories, through adopting auxiliary hypotheses in response to contradictory observations, demonstrates how theories (research programmes) can rationally adapt to adverse evidence. In the 1830s the contingent-contagion hypothesis (that cholera, while normally non-contagious, could under certain circumstances become contagious) was adopted by non-contagionists to reconcile the observation that cholera sometimes appeared to be passed person-to-person (airborne) and other times not. After 1849 water was recognized as an important, possibly the most important, predisposing cause. Adoption of these hypotheses was not per se irrational or anti-scientific (according to Lakatos). They served to protect the airborne theories against contrary observations but failed to produce any new facts and thus were degenerating (*ad hoc*), arguing against the airborne theories as a useful causal explanation.

Causal explanations, such as for cholera, generally require a broader approach to causation and evaluation of evidence than provided within even a well-specified treatment-control causal framework. The case of contingent-contagion and causal water serve as examples for why individual hypothesis tests and crucial experiments are often insufficient (on their own and taken individually) for deciding between alternative research programmes. Cholera in the 1850s provides a case study for the iterative process of scientific inquiry and the dynamic interplay theory and data. We close with, and echo, David Freedman's endorsement: "Snow's work on cholera is ... a success story for scientific reasoning based on nonexperimental data" (Freedman, 1991, 291).

References

Anderson, C. (2008, June). The End of Theory: The Data Deluge Makes the Scientific Method Obsolete. *Wired*. Section: tags.

Angrist, J. D. and J.-S. Pischke (2008, December). *Mostly Harmless Econometrics: An Empiricist's Companion* (1 edition ed.). Princeton University Press.

Angrist, J. D. and J.-S. Pischke (2014, December). *Mastering 'Metrics: The Path from Cause to Effect* (with French flaps edition ed.). Princeton ; Oxford: Princeton University Press.

Bacon, F. (1620). *The New Organon and Related Writings*. New York: Liberal Arts Press.

Blane, G. and F. Corbyn (1821). Account of the epidemic spasmodic cholera which has lately prevailed in India, ... *Medico-Chirurgical Transactions 11*(Pt 1), 110–56.

Budd, W. (1849). *Malignant Cholera: Its Mode of Propogation, and Its Prevention*. London: Churchill. reprint International Journal of Epidemiology, 42 #6 2013.

Carboni, G. P. . (2021). The enigma of Pacini's Vibrio cholerae discovery. *Journal of Medical Microbiology 70*(11), 001450. Publisher: Microbiology Society,.

Cholera Inquiry Committee (1855). *Report on the cholera outbreak in the parish of St. james, Westminster, during the autumn of 1854*. London: J. Churchill.

Coleman, T. S. (2018, October). Causality in the Time of Cholera: John Snow as a Prototype for Causal Inference. SSRN Scholarly Paper ID 3262234, Social Science Research Network, Rochester, NY.

Coleman, T. S. (2020, September). John Snow, Cholera, and South London Reconsidered. SSRN Scholarly Paper, Social Science Research Network, Rochester, NY.

Cox, D. R. (1958). *Planning of Experiments*. New York: Wiley.

Deaton, A. (1997). *The analysis of household surveys: a microeconometric approach to development policy*. Baltimore, MD: Published for the World Bank [by] Johns Hopkins University Press.

Deaton, A. (2013). *The great escape: health, wealth, and the origins of inequality*. Princeton: Princeton University Press.

Falcone, M., J. Koschinsky, P. Vinten-Johansen, T. Coleman, and L. Anselin (2021). John Snow & the Cholera Epidemic in Mid-19th Century London: 9 Datasets With Documentation for Use in GeoDa.

Farr, W. (1853, November). Supplement to the Weekly Return: Cholera and the London Water Supply. In *Weekly Return of Births and Deaths in London*, Volume XIV, pp. 401–406. General Register Office. https://wellcomecollection.org/works/ygtd8752.

Farr, W. (1868). Report on the Cholera Epidemic of 1866 in England: Supplement to the 29th Annual Report of the Registrar-General of Births, Deaths, and Marriages in England. In *U.K. Parliament Sessional Papers, 1867-1868*, Volume 37 of *Command Papers*, pp. 415. London. https://parlipapers-proquest-com.proxy.uchicago.edu/parlipapers/result/pqpdocumentview?accountid=14657&groupid=95684&pgId=0315a53f-b3f8-4072-a698-e427be699bed&rsId=17DB0024559.

Fowler, A. and E. Bueno de Mesquita (2021). *Thinking Clearly with Data: A Guide to Quantitative Reasoning and Analysis*. Princeton, NJ: Princeton University Press.

Freedman, D. (1991). Statistical Models and Shoe Leather. *Sociological Methodology 21*, 291–313.

Freedman, D. (1999, August). From association to causation: some remarks on the history of statistics. *Statistical Science 14*(3), 243–258.

Friedman, M. (1953). Methodology of Positive Economics. In *Essays in Positive Economics*. Chicago: University Of Chicago Press.

Gangl, M. (2010). Causal inference in sociological research. *Annual review of sociology 36*, 21–47.

General Board of Health (1848, October). The Public Health. *Times*, 1.

General Board of Health (1850). *Report on the Epidemic Cholera of 1848 and 1849*. London: Her Majesty's Stationary Office. http://kora.matrix.msu.edu/files/21/120/15-78-1E5-22-1850-GBoHReportCholera1848-49.pdf.

General Register Office (1852). *Report on the Mortality of Cholera in England, 1848-49*. London: Her Majesty's Stationary Office. http://johnsnow.matrix.msu.edu/work.php?id=15-78-12A.

Grainger, R. (1851). Appendix B. In *Report on the Epidemic Cholera of 1848 and 1849*. London: Her Majesty's Stationary Office.

Greene, W. H. (2018). *Econometric Analysis, 8th Edition* (8th ed.).

Grootendorst, P. (2007, December). A review of instrumental variables estimation of treatment effects in the applied health sciences. *Health Services and Outcomes Research Methodology 7*(3), 159–179.

Heckman, J. J. and B. Singer (2017, May). Abducting Economics. *American Economic Review 107*(5), 298–302.

HedstrÃ¶m, P. and P. Ylikoski (2010). Causal mechanisms in the social sciences. *Annual review of sociology 36*, 49–67.

Hempel, S. (2007, January). *The Strange Case of the Broad Street Pump: John Snow and the Mystery of Cholera* (First edition ed.). Berkeley: University of California Press.

Holland, P. W. (1986, December). Statistics and Causal Inference. *Journal of the American Statistical Association 81*(396), 945–960.

Johnson, S. (2007, October). *The Ghost Map: The Story of London's Most Terrifying Epidemic–and How It Changed Science, Cities, and the Modern World* (Reprint edition ed.). New York: Riverhead Books.

Kuhn, T. S. (1962). *The Structure of Scientific Revolutions: 50th Anniversary Edition* (4th edition, 2012 ed.). Chicago ; London: University of Chicago Press.

Lakatos, I. (1980, November). *The Methodology of Scientific Research Programmes: Volume 1: Philosophical Papers*. Cambridge: Cambridge University Press.

Lancet (1867, November). Mr. Radcliffe's Report on the London Cholera Epidemic. *Lancet 90*(2305), 558–559.

Lancet (1868a, August). Dr. Farr's Cholera Report. *The Lancet 92*(2346), 220–223.

Lancet (1868b, August). THE PROPAGATION OF CHOLERA. *The Lancet 92*(2346), 217–219.

Laudan, L. (1968). Theories of scientific method from Plato to Mach. *History of Science 7*(1), 1–63.

Leitgeb, H. and A. Carus (2021). Rudolf Carnap. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Summer 2021 ed.). Metaphysics Research Lab, Stanford University.

Luckin, W. (1977, January). The final catastrophe– cholera in London, 1866. *Medical History 21*(1), 32–42.

McLeod, K. S. (2000, April). Our sense of Snow: the myth of John Snow in medical geography. *Social Science & Medicine (1982) 50*(7-8), 923–935.

Misak, C. J. (2004). *Truth and the End of Inquiry: A Peircean Account of Truth*. Oxford Philosophical Monographs. Oxford: Oxford University Press.

Neyman, J. (1923). Statistical problems in agricultural experiments. *Journal of the Royal Statistical Society II (Supplement) 2*, 107–180.

Parkes, E. A. (1855, April). Review of Snow's Mode of Communication of Cholera. *British and Foreign Medical Review*, 449–63.

Parkin, J. (1832). Suggestions respecting the cause, nature, and treatment of cholera. *London Medical and Surgical Journal 2*, 151–153. https://babel.hathitrust.org/cgi/pt?id=mdp.39015057196241&view=1up&seq=165 https://www.amazon.com/London-Medical-Surgical-Journal-1833/dp/0266519105.

Pearl, J. (2009). Causal inference in statistics: An overview. *Statistics surveys 3*, 96–146.

Pinto, R. and J. J. Heckman (2021). The Econometric Model for Causal Policy Analysis. *Annual Review of Economics*.

Popper, K. R. (1985, February). *Popper Selections*. Princeton, N.J: Princeton University Press.

Quine, W. V. (1953). *From a logical point of view: 9 logico-philosophical essays* (2d ed. rev. 1961 ed.). Cambridge: Harvard University Press.

Radcliffe, J. N. (1867). Ninth report of the Medical Officer of the Privy Council. In *U.K. Parliament Sessional Papers, 1867*, Volume 37 of *Command Papers*. https://parlipapers-proquest-com.proxy.uchicago.edu/parlipapers/result/pqpdocumentview?accountid=14657&groupid=95684&pgId=8439fd51-d7cc-4463-a048-cd527c15f4fb&rsId=17DB6527783.

Rothman, K. J. (2002). *Epidemiology: an introduction*. New York, N.Y.: Oxford University Press.

Rubin, D. B. (1974). Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of Educational Psychology 66*(5), 688–701.

Rubin, D. B. (1978). Bayesian inference for causal effects: The role of randomization. *Annals of Statistics 6*(1), 34–58.

Salmon, W. C. (2006). *Four decades of scientific explanation.* Pittsburgh, PA: University of Pittsburgh Press.

Simon, J. (1856). *Report on the last two cholera-epidemics of London, as affected by the consumption of impure water*. London: printed by George E. Eyre and William Spottiswoode, for H.M.S.O. addressed to the Rt. Hon. the President of the General Board of Health.

Snow, J. (1848, November). WESTMINSTER MEDICAL SOCIETY.: OCTOBER 21, 1848.–J. WEBSTER, M.D., F.R.S., PRESIDENT. *The Lancet 52*(1314), 506–508.

Snow, J. (1849a). *On the mode of communication of cholera*. London: John Churchill. OCLC: 14550757.

Snow, J. (1849b, October). WESTMINSTER MEDICAL SOCIETY.: SATURDAY, OCTOBER 13, 1849.–MR. HIRD, PRESIDENT. *The Lancet 54*(1364), 431–433. Publisher: Elsevier.

Snow, J. (1854, August). Cholera in the Baltic fleet. *Medical Times and Gazette 9*.

Snow, J. (1855). *On the mode of communication of cholera* (2nd ed.). London: John Churchill.

Snow, J. (1856a, June). Cholera and the water supply. *Times*, 12, col. B.

Snow, J. (1856b, October). Cholera and the water supply in the south district of London in 1854. *Journal of Public Health and Sanitary Review 2*, 239–257.

Snow, J. (1857, October). Cholera and the water supply in the south district of London. *British Medical Journal 2*, 864–865.

Sutherland, J. (1851). Appendix A. In *Report on the Epidemic Cholera of 1848 and 1849*. London: Her Majesty's Stationary Office.

Tumbe, C. (2020, December). *Age Of Pandemics (1817-1920): How they shaped India and the World*. HarperCollins India.

Vinten-Johansen, P. (Ed.) (2020, May). *Investigating Cholera in Broad Street: A History in Documents:* (1st edition ed.). Broadview Press.

Vinten-Johansen, P., H. Brody, N. Paneth, S. Rachman, and M. R. Rip (2003, May). *Cholera, Chloroform and the Science of Medicine: A Life of John Snow* (1 edition ed.). Oxford ; New York: Oxford University Press.

Vinten-Johansen, P. and J. Koschinsky (2022). Teaching spatial data analysis as part of scientific inquiry. Cholera theories and evidence in 19th century Britain.

Wakley, T. (1832, April). Editorial. *Lancet 18*(452), 122–126.

Appendices

A Theories of Cholera in Mid-Century Britain

Theories of cholera in the 1850s can be categorized along three dimensions. These do not exactly correspond with modern understanding of medicine and physiology, but are close enough to be both understandable and useful. First, the mode of transmission, being (for cholera) either inhaled or ingested. Second, whether contagious or non-contagious. Third, the "seat of the disease", being the nervous system, the blood, or the gastrointestinal tract.

Figure 1 shows a timeline and Table 3 lists the alternative theories under consideration during the 1850s. The first detailed reports to medical practitioners in Britain (Blane and Corbyn, 1821) presumed cholera was a disease of the nervous system, although by the 1830s it was more widely thought to be a blood disease. For present purposes the distinction between nervous system versus blood disease is not important, because the mode of transmission for both was inhaled and through the lungs. Up until Snow's 1849 *On the mode of communication of cholera* (Snow, 1849a) inhalation was almost universally accepted.[12]

The major point of contention and debate prior to 1849 was whether cholera was contagious (infectious) or non-contagious. (The debate over contagion continued after ingestion was introduced as a method of communication.) For contagionists a cholera poison was produced in the bodies of the sick and communicated through the disease process known at the time as infection, what we would now call inhalation or respiratory infection. Non-contagionists considered that cholera was generated by the interaction of an imperceptible atmospheric influence and local concentrations of decomposing organic matter (filth). They classified cholera as a typical epidemic disease, but we need to recognize that in 1850 'epidemic' had a somewhat different meaning than today, meaning a localized outbreak of non-contagious nature.

In the 1830s a hybrid termed *contingent-contagion* was introduced to non-contagious theories to allow for cholera to become temporarily contagious under certain conditions, say overcrowding (see Wakley, 1832, 124-125). The hypothesis was necessary to account for conflicting evidence. Two instances from early in the period suffice to demonstrate the issue: Pro-contagious was a episode from Mauritius where cholera seemed to be imported by the frigate Topaz from Cylon (Blane and Corbyn, 1821, 148-150, 152-153); anti-contagious was a case of three Sepoys in Seroor hospital who escaped disease despite being in close proximity to patients (Blane and Corbyn, 1821, 143).

The contingent-contagion hypothesis is noteworthy for a number of reasons. First, it is an example of what Lakatos termed a protective auxiliary hypothesis introduced to protect a core theory, non-contagion in this case (Lakatos, 1980, particularly section 1.3). It made non-contagious theories immune from contradiction by evidence of contagion, a sign of a degenerative research programme in Lakatos's terms. Second, as a result the hypothesis weakened non-contagion relative to any alternative that could account for apparently conflicting evidence within the core theory. We will see that Snow's ingested fecal-oral theory could account for the evidence that appeared contradictory

[12] Parkin (1832) suggests that an airborne cholera poison can be diffused into water and then ingested, but ingestion and gastrointestinal disease were not widely considered prior to 1849.

Table 3: Cholera Theories in Mid-Nineteenth Century England

Theory		Comment
Blood Disease, Airborne (inhaled, + water post-1849)	**Infection** (contagious, airborne): Airborne morbid poison produced in victims' bodies. Transmitted via inhalation, with water possible post-1849.	Current in 1850s. **Predisposing Causes**: close proximity to victims; overcrowding, poor ventilation; Post-1849: impure water/food a major predisposing cause; **Susceptibility**: constitutional depression of vital powers; previous or current pulmonary diseases.
	General Atmospheric (generally non-contagious, airborne): An active morbid poison in atmosphere. Transmitted via inhalation, with water possible post-1849.	Less in favor after 1850. **Predisposing Causes**: Environmental: hot, moist, stagnant atmosphere. Associated with decomposing vegetation, and/or soil contamination. **Susceptibility**: constitutional depression of vital powers; poverty; unhealthy, moist living conditions
		Contingent-contagion – protective "auxiliary hypothesis" aligns theory with evidence of contagion
	Localization (generally non-contagious, airborne): Concentrations of the "cholera poison" from localized sources of organic decomposition (including sewer gases and "privy atmospheres"). Transmitted via inhalation, with water possible post-1849.	**Predisposing Causes**: Environmental: overcrowding, dampness, filth, poor ventilation, proximity to graveyards, bad water, bad sanitation; Post-1849: impure water/food a major predisposing cause; **Susceptibility**: living in unhealthy situations, constitutional weaknesses
		Contingent-contagion – protective "auxiliary hypothesis" aligns theory with evidence of contagion
	Elevation / Zymotic: A refinement of localization (above).	Evaporation & inhalation: sewage & organic matter in rivers evaporates, those at higher elevation (further distance) inhale less
Gastrointestinal tract, Ingested	**Alimentary Canal** (contagious, now termed fecal-oral transmission): "excretions of the sick ... being accidentally swallowed" Snow (1849a, 8)	**Normal (Ordinary) Propagation**: close communication with infected individual leading to inadvertently swallowing fecal matter. Predominant mode of transmission. **Neighborhood Propagation**: Choleraic contamination of a common food or (more frequently) water source **Municipal Propagation**: water contaminated by sewage containing cholera dejections is distributed through municipal water supply (either rivers or water supply companies)
	Intestinal Canal (non-contagious)	Introduced around 1850 to account for strong association with water

under inhaled non-contagious (or contagious) theory and thus had some support relative to theories that relied on contingent-contagion.

All these theories invoked predisposing causes or conditions, and susceptibility. Predisposing causes were environment factors that caused or contributed to cholera. For airborne contagious, overcrowding would be a predisposing cause because more people in close proximity would lead to more transmission and more disease. For non-contagious (localization in particular) local sources of decomposing organic matter, say pigsties or open sewers, would activate or amplify the airborne cholera poison. Susceptibility was invoked to explain why some but not all those exposed fell ill with cholera, and might include exposure to other diseases, poor food, or poor living conditions.

In 1849 John Snow published the first edition of his monograph *On the mode of communication of cholera*, where he proposed that cholera was gastrointestinal and transmitted by the inadvertent ingestion of contaminated fecal matter, either through food, water, or hand-mouth contact: "excretions of the sick ... being accidentally swallowed" Snow (1849a, 8). The fecal-oral theory predicts three modes of communication or propagation, all of which Snow described more or less explicitly: (1) ordinary or normal propagation, person-to-person transmission through contamination of food or water, or hand-mouth contact: "evacuations ... become attached unobserved to the hands of the person nursing the patient, and are unconsciously swallowed" (Snow, 1849a, 9-10); (2) neighborhood propagation, the contamination of a local water (or food) source accessed by multiple households in a neighborhood(Snow, 1855, 22-23); and (3) municipal propagation where sewage was mixed with municipal water supply: "the emptying of sewers into the drinking water of the community" (Snow, 1849a, 11)

Snow's was a contagionist theory, but broke from most others by proposing that cholera was not airborne, and affected the intestines rather than the blood. William Budd, a Bristol surgeon, independently proposed a non-contagious fecal-oral theory. Writing just a few weeks after the publication of Snow's pamphlet in early September 1849, Budd concluded from investigations undertaken during the epidemic in Bristol that "the germs of the cholera-fungus [first described in July 1849 by two colleagues at the Bristol Medico-Chirurgical Society] are diffused in three ways: 1, through the air; 2, in contact with articles of food; and 3, and chiefly, in the drinking-water of infected places" (Budd, 1849, 16-17). Morbid cholera matter must be "taken by the act of swallowing into the intestinal canal ... [and] develop[s] only in the human intestine ..." (Budd, 1849, 5). In other words, airborne cholera matter was only morbid if it settled on something that was subsequently swallowed. As such, Budd agreed with Snow that the sole mode of cholera communication was via the fecal-oral route.

B Predictions

Repeating from above, Snow's core theory posited cholera as a disease of the intestine, transmitted by some unknown material that passed out of one individual (vomit or more commonly diarrhea) and was then inadvertently ingested by others. The three modes of communication (person-to-person or normal, neighborhood, and municipal transmission) together predicted the

The core of the alternative theory (or theories) posited cholera as an airborne poison causing a

blood disease.[13] For predictions the important characteristic is that the cholera poison was airborne. various There were many and various factors that could produce or amplify the airborne poison ("predisposing causes"), or increase individuals' susceptibility – factors such as crowded hosing conditions, dampness, filth, bad sanitation. Prior to 1849 transmission was generally considered via inhalation. After 1849 auxiliary hypotheses added water as a mode of transmission – the airborne poison could diffuse into water and enter through the stomach – and also added water as a cause – contaminated water could produce or amplify the airborne poison that could be inhaled (predisposing cause), or could increase an individuals' susceptibility. These water-borne auxiliary hypotheses are not prima facie contrary to logic or the science of the 1850s (or even science today). Today we know they are wrong, but they were not and are not silly.

Table 4 details predictions from the alternative theories. Each row is a specific predicted observation – matching with evidence. The second column ("Predicted Observation") shows the prediction about the pattern of observed mortality. The third column indicates the theory and specific hypothesis generating that prediction. The third, fourth, and fifth columns show whether the listed theory generates the predicted observation or not.

For example, prediction (2a) concerns person-to-person transmission for those sharing airspace – the prediction is that increased cholera mortality occurs for those in close proximity via sharing airspace. The airborne contagious blood disease theory produces this prediction because it hypothesizes that transmission is person-to-person via inhalation, and so the cell in the fifth column is "Yes". The airborne non-contagious theory does not generate this prediction (so "No"), but the auxiliary contingent-contagious hypothesis specifies that cholera can become contagious under certain circumstances. The entry is therefore "No/Yes" – the theory predicts both airborne contagion and non-contagion. The fecal-oral theory predicts "No" because the theory predicts fecal-oral transmission, and thus sometimes those sharing airspace will become infected, sometimes not.

The table is valuable because it allows us to identify predictions and evidence (patterns of mortality) that distinguishing between theories. A row with mixed "Yes" and "No" distinguishes between theories, while a row with all "Yes" does not. A row is particularly valuable when predictions separate between theories *and* the quality of evidence is good. On this basis, the most important predictions are: (2a) and (2b) person-to-person contagion; (5) transmission via air; (6c-f) association with water versus other factors; (9) and (10) concerning the time-pattern of mortality, what is now termed the epidemic curve. The final column ("Overall Value") summarizes the overall value of the prediction for separating between theories.

Some predictions distinguish between theories (for example predictions (1), (4), or (7)) but the evidence is not good quality. Other predictions have strong evidence but do not allow one to distinguish between theories. Prediction (5a) – water as a causal factor for cholera mortality – is particularly notable, and helps us understand some puzzles about the competition among theories in the 1850s. There was strong evidence for water as a causal factor in cholera mortality – Snow (1855) is well-known and widely cited, but Simon (1856) provided even stronger evidence – and today we are puzzled that such evidence did not lead to the acceptance of Snow's theory (remember the comments of Johnson (2007, 125) about "smart people cling[ing] to an outlandishly incorrect

[13]Today, we would say "an airborne poison infecting the blood" but in the 1850s the word "infection" had a different meaning, reserved for contagious epidemic diseases, and we will try to avoid its use.

Table 4: Predictions for Alternative Theories as of 1855

	Predicted Observation	**Theory**: Hypothesis	Fecal-Oral, Intestinal (Snow)	Airborne, Blood Disease		Quality of Evidence	Overall Value
				Non-Contagious / Contingent	Contagious		
	PERSON-TO-PERSON CONTAGION						
1	***Geographic Spread***: Cholera occurs along trade routes, at the speed of travel	**Airborne and Fecal-oral**: Person-to-person transmission	YES	NO / neutral	YES	Moderate, anecdote	Low
2a	***Person-to-Person airborne contagion***: occurs between those sharing airspace in close proximity	**Airborne**: person-to-person transmission	NO	YES/NO	YES	Good, case studies	**High**
2b	***Person-to-Person non-airborne contagion***: Cholera occurs for those in close proximity, via sharing food, clothing	**Fecal-oral**: person-to-person transmission	YES	NO	NO	Good, case studies	
3	***Winds***: Cholera follows winds (not contagious)	**Airborne**: but no person-to-person transmission	NO	YES	NO	Poor	Low
	GENERAL CHARACTERISTICS						
4	***General transmission by air***: Sharing same air (different water) ⇒ similar mortality; sharing different air (same water) ⇒ different mortality	**Airborne**: transmitted via air	NO	YES	YES	Good, case studies	Medium
5a	***Water is causal***: pre-1849, cholera not associated with water (NB – prediction from airborne is reversed)	**Airborne** pre-1849: water not a predisposing cause or method of transmission	–	NO	NO	High	Low
5b	***Water is causal***: post-1849, cholera associated with water – post-1849	**Airborne** post-1849: water *is* a predisposing cause or method of transmission **Fecal-oral**: water is a cause through neighborhood and municipal transmission	YES	YES	YES	High, quantitative (statistical)	
5c, d, e, f	***Other factors***: Cholera associated with variety of factors, independent of water: 5c) Overcrowding & poor ventilation; 5d) Hot weather & decomposing refuse; 5e) Poverty & unhealthy living conditions; 5f) Dampness, filth, bad sanitation	**Airborne**: many other factors are predisposing causes and susceptibilities **Fecal-oral**: other factors may have some influence but will be negligible relative to water in a municipal outbreak	NO	YES	YES	1856: Good; 2022: High, quantitative (statistical)	Medium
6	***General characteristics***: Cholera erupts simultaneously in non-contiguous locations; Visits same areas in subsequent outbreaks	**Airborne non-contagious** (localized): dependent on local predisposing conditions **Fecal-oral**: dependent on local conditions that favor municipal outbreaks (e.g. water from the Thames)	YES	YES	NO	Moderate, anecdotal	Low
7	***Elevation***: Mortality falls with higher elevation and further from river	**Zymotic** (subset of airborne non-contagious): Mortality falls with higher elevation and further from river	YES - but confounding	YES	NO	Good, quantitative	Low
	EPIDEMIC CURVE						
8	***Explosive neighborhood outbreaks***: Mortality grows quickly initially then falls off	**Fecal-oral**, Neighborhood propagation: contamination of a point source, usually water	YES	NO	NO	High, quantitative	**High**
9	***Municipal outbreaks early vs late***: Early: high mortality for those exposed to contaminated water; Later: difference remains but diminishes as normal and neighborhood propagation contribute	**Fecal-oral**, Municipal propagation: contamination of a municipal water supply, usually a river	YES	NO	NO	High, quantitative	**High**

Each row presents a prediction about observed patterns of mortality. "Yes" or "No" in a column (and color) indicates whether the theory in the column predicts the observation, or the opposite. This table does not compare predictions versus evidence.

idea despite substantial evidence to the contrary"). The fact, reflected in row (5a), is that all theories predicted water as a causal factor, so that evidence did not distinguish between them. Simon (1856) clearly states the case, first providing strong evidence for water as causal ("The above conclusions [that fecalized water causes cholera mortality] rest on so large a basis of facts, that I venture to believe they will be accepted as the final solution of any existing uncertainty as to the dangerousness of putrefiable drinking-water during visitations of epidemic cholera" Simon 1856, 9), but then explicitly states that this does not rule out other causes ("The present contribution therefore aims only at giving a more exact knowledge of one cause, not at gainsaying the existence of other causes" Simon 1856, 13).

We can also use Table 4 to understand why Snow turned from demonstrating water as *a* causal factor (Snow, 1855) to trying to demonstrate water as the *only* causal factor (Snow, 1856b). Snow (1855, Table IX) compared across quasi-randomized water customers to try and demonstrate a large difference in mortality between contaminated-water customers (served by the Southwark & Vauxhall Company) versus clean-water customers (served by the Lambeth Company) – to demonstrate water as causal. Due to data limitations, Snow could only compare across a large area, which included subdistricts served by both companies (where customers were mixed and quasi-randomized) *and* subdistricts served by Southwark & Vauxhall only. Snow recognized, and Edmund Parkes emphasized, that combining all regions left the comparison open to problems of confounding: "the Lambeth Company supplies, to a considerable extent, a good neighbourhood on elevated ground (including the healthy districts of Streatham, Foresthill, and Syndenham); while the Southwark and Vauxhall Company supplies the greater part of the poorest, lowest, and marshiest district in London" (Parkes, 1855, 461).

With the publication of detailed by-supplier population statistics in Simon (1856), Snow now had access to data to re-analyze the quasi-randomized data, but he did not. Instead, in Snow (1856b), he turned to demonstrating that variation across subdistricts was due to water alone, and not other factors. Such a shift in focus, however, is understandable when we recognize that prediction (5b) – water as causal – does not separate theories, but evidence *against* predictions (5c-f) – other factors causal – would separate theories. Showing that water was the only causal factor, and other factors did not affect cholera mortality, would sharply separate Snow's theory from alternatives.

Laying out the predictions for alternate theories, as in Table 4, provides guidance for experimental design and interpretation. In the present circumstances, when all experiments were completed over 150 years ago, it provides guidance for understanding the competition between theories. In the 1850s, such a table could have been used to direct research towards particular predictions, and collection of data, that could help separate between theories.

C Summary of Snow's Evidence

Snow's analysis of the large (400,000+) population in South London (Snow, 1855) exploited two research designs which are in widespread use today – randomization as an instrumental variable and difference-in-differences (DiD). First, for randomization, Snow compared mortality for customers of two water-supply companies, one of which supplied contaminated water (the Southwark & Vauxhall

Company or S&V) and the other clean water (the Lambeth Company). Snow argued, plausibly, that the choice of water company (and thus treatment with contaminated versus clean water) was effectively random (often unknown to the residents because the landlord paid the water bill), and that customers were well-mixed on all characteristics (such as location, age, sex, income, housing). These circumstances allowed comparison of those treated with contaminated versus dirty water in a natural experiment (an *experimentum crucis* following Farr 1853, 401 and Bacon 1620) that well-approximated a randomized control trial to decide between rival hypotheses. Snow's comparison of the mixed South London population in Snow (1855, Table IX) is cited as the first use of randomization as an instrumental variable – see (Greene, 2018, 228), which also quotes (Deaton, 1997; Grootendorst, 2007); see also (Deaton, 2013, 95-96).

The second analysis, nascent DiD (Snow, 1855, 89), exploited the fact that the Lambeth company, which supplied clean water in 1854, had supplied contaminated water in the 1849 outbreak. By comparing in 1849 (when all customer were exposed to contaminated water) subdistricts (regions) supplied jointly by the two companies versus subdistricts supplied only by S&V, Snow could difference out subdistrict or regional effects. Comparing those same subdistricts in 1854 (when jointly-supplied subdistricts had some customers with clean water) allowed comparison of the water effect and controlling for subdistrict (regional) effects. The result was "a considerable diminution of mortality in the sub-districts partly supplied by the Lambeth Company" (Snow, 1855, 89). This is cited as the first use of difference-in-differences; see (Angrist and Pischke, 2014, 205), (Angrist and Pischke, 2008, 227).

In fact, both of Snow's analyses suffer from flaws. The issue with the randomization, highlighted at the time by, for example, Parkes (1855, 461), was that Snow combined disparate subdistricts (regions). Lacking detailed population data by supplier, Snow was only able to compare at the *aggregate* or regional level, and not the subdistrict level where customers were well- (quasi-randomly) mixed. For the DiD, the across-subdistrict variation in mortality rates is high enough that the observed treatment effect observed in the raw DiD is not statistically significant (see Coleman, 2020). Neither of these flaws is fatal, however. For the randomization, there was evidence in 1855 (from the DiD) that subdistricts did not differ greatly in 1849 when all customers were supplied with contaminated water, and so combining subdistricts would be unlikely to cause problems with randomization. For the DiD, Snow recognized that the treatment effect was stronger for subdistricts with more Lambeth customers[14] and this effect is statistically significant. Re-analysis using modern statistical tools (in Coleman, 2020) confirms that these issues are not significant, and that there is considerable evidence for water as a causal factor in cholera mortality.

There was other, arguably stronger, evidence provided in the 1850s which showed the causal effect of water; we highlight two instances. Sutherland (1851, 15-16) reports an outbreak in Hope Street, Manchester, in 1849. Ninety houses were on the street, 30 using water from a particular pump-well and 60 not. Twenty-five persons died, all in houses drinking from the pump and none from houses with other water sources. People were well-mixed with regard to location, and presumably with regards to other characteristics – mixing (quasi-randomization) almost as good as Snow's but here in a small and concentrated locale. In the second instance, Simon (1856) discussed at some length

[14]"In certain sub-districts, where I know that the supply of the Lambeth Water Company is more general than elsewhere, as Christchurch, London Road, Waterloo Road 1st, and Lambeth Church 1st, the decrease of mortality in 1854 as compared with 1849 is greatest" Snow (1855, 89).

the same 1854 South London outbreak and population that Snow analyzed, but with the benefit of population-by-supplier detailed by subdistricts (which was not available when Snow published in 1855). Simon concluded that putrefiable was important. Among other evidence, in Table I Simon provides a difference-in-differences analysis recognizable to a modern reader, one that arguably serves as a better case of first use than Snow's.[15]

D Sophisticated Falsification: Predictions versus Evidence

Choosing among competing theories, the *sophisticated falsification* of Lakatos, does not easily fit in a statistical testing framework, but it can be organized and structured. Table 4 (in this appendix) and Table **??** (in the body of the paper) are an example of organizing theories, predictions, and evidence to . Table 4 organizes the predictions from alternative theories. Table **??** expands on the predictions to include evidence, and specifically whether the prediction agrees with the evidence.

The content of Table **??** is rather dense – we are summarizing a wide range of predictions and evidence that varies substantially in form and quality. The first column shows "Prediction & Observation" – each row a particular prediction about cholera and observed mortality. The sixth or penultimate column contains a description of the evidence – the observation itself. First, the quality or reliability of the evidence, whether *moderate*, *good*, or *high*. Second, whether that evidence agrees with the prediction or not. Finally, a brief summary of the evidence.

The first column of Table **??**, *Prediction & Observation*, states the prediction. Consider prediction (5b), "water is causal". The prediction (column one) is that water is a causal factor – consuming contaminated water causes higher mortality. Columns two, three, and four show two pieces of information. First, whether the theory of that column agrees with (generates) the predictions shown in the first column – indicated by a "yes" or "no". Second, whether the theory's prediction is supported (shown in teal) or contradicted (shown in red) by the evidence in the sixth column. For the case of prediction (5b), the fecal-oral theory agrees with (generates) the prediction that water is causal, and it is supported by the evidence. The airborne theories also agrees with the prediction that water is causal, because after the 1849 outbreak (and in an instance of a protective auxiliary hypothesis), water was proposed as a major predisposing cause. The airborne theories were essentially protected from what would have otherwise been contrary evidence. Column five summarizes the evidence; in the case of (5b) the evidence indeed shows that water is causal, so agrees with the prediction, and so column five is labeled "Agree: Yes". The fecal-oral theory (column two) agrees with the predictions so the cell is labeled "yes". The prediction of the fecal-oral theory ("yes water is causal") matches the evidence ("evidence indicates water is causal"), so the cell is labeled "Support".

Column six is our assessment of the value of the prediction-versus-evidence comparison for discriminating between the theories. For prediction (5b) the value is low, since by 1854 all theories predicted that water would be associated with (and cause) higher cholera mortality – the fecal-oral theory as a central component of the core theory, and the airborne theories by virtue of an (ad-hoc) auxiliary hypothesis. Columns seven and eight summarize by noting whether the evidence contradicts the

[15] Snow (1856b) highlights problems with both the number of reported deaths assigned to water companies (pp 248-249) which form the numerator for mortality rates and the number of houses (pp 245-246) which form the denominator. The conclusion is that Simon's underlying data are not as reliable as Snow's.

theory (Yes-versus-No), is Mixed, or if the theory would be contradicted but is Protected by an auxiliary hypothesis.

D.1 Discussion of Detailed Predictions

Prediction (2) – Person-to-Person Contagion A central controversy and puzzle over cholera, going back to the earliest publications (Blane and Corbyn, 1821), was whether cholera was contagious (transmitted person-to-person). The airborne theories can broadly be categorized as contagious or non-contagious. Airborne contagious theory predicted that those in close proximity and sharing airspace (breathing the same air) should become sick (prediction 2a). Airborne non-contagious theory predicted that those sharing airspace would *not* become sick (at any higher rate than others). Evidence was presented on both sides (for early evidence on both sides see Blane and Corbyn 1821, 143, 148-150, 152-153). Both contagious and non-contagious theories were contradicted by evidence, but this was based on the (maintained) hypothesis that transmission was by air. In the 1830s a protective auxiliary hypothesis (*contingent contagion*) was introduced to the non-contagious airborne theory to make it consistent with contrary observations.

Snow's breakthrough in 1849 was to recognize that cholera did not *have* to be airborne, and that the extant evidence would be consistent with, indeed implied by, a disease that was intestinal and transmitted via inadvertent ingestion of contaminated fecal matter – an abductive leap (to use Peirce's meaning of the term). For airborne inhaled theories the evidence was conflicting while the fecal-oral theory, where the material causing cholera is ingested, predicted that those in close contact would transmit cholera through hand-to-mouth contact or sharing food, but *not* by sharing airspace (prediction 2b). For this set of predictions, the evidence was consistent with the prediction.

The fact that the fecal-oral theory was able to account for observations that posed a substantive puzzle for alternative theories should be viewed as the major accomplishment that it was. According to Lakatos (1980, 69) "an objective reason [to reject a research programme] is provided by a rival research programme which explains the previous success of its rival and supersedes it by a further display of heuristic power."

Summarizing person-to-person transmission we can conclude the following: 1) the fecal-oral theory is corroborated (supported) by the evidence; 2) evidence contradicts the airborne contagious theory; 3) the airborne non-contagious theory is protected by the auxiliary contingent contagious hypothesis, but the hypothesis produces no new predictions or facts – it is theoretically and empirically degenerating. The consistency of the fecal-oral theory and observed person-to-person transmission could be considered a new "fact" – the fecal-oral theory changes the interpretation and understanding of the evidence.

Prediction 4 – General Transmission by Air In the simplest case, with no auxiliary hypotheses about water as a predisposing cause or as a method of transmission, the prediction "cholera is transmitted via air" means that those sharing the same airspace should have similar mortality and illness outcomes, while those not sharing the same airspace could have different mortality outcomes. Snow recognized these predictions and provided evidence that refuted them. (The Reverend Henry

Whitehead collected evidence strongly refuting the first.) Nonetheless, auxiliary hypotheses for the airborne theories served to protect them from contradiction by the evidence.

There are three pieces of evidence from the 1854 Broad Street cholera outbreak bearing on airborne transmission. The first is mortality rates that are unusually low among particular residents near the Broad Street pump. The St. James workhouse at 50 Poland Street had 535 residents but only five died of cholera. The Lion Brewery at 50 Broad Street had more than 70 workers and none died of cholera. Snow explains that both institutions had an independent water supply, residents did not drink from the Broad Street pump, and thus were not exposed to contaminated pump water (Snow, 1855, 42). But all these people were, presumably, exposed to the same air as those nearby ("The Workhouse in Poland Street is more than three-fourths surrounded by houses in which deaths from cholera occurred" (Snow, 1855, 42)); in other words the evidence contradicts the prediction that those sharing the same air will have similar mortality.

The second piece of evidence was data collected by the Reverend Whitehead, data on the number of residents of Broad Street who drank or did not drink from the pump, and then fell ill or not (Cholera Inquiry Committee, 1855, 128 ff). Importantly Whitehead collected data both for the number who drank and those who did not drink pump water, thus avoiding the survivorship bias that could arise from only sampling those who were sick or died.[16] The data show that those who did not drink did not get sick (low mortality rate), and those who did drink got sick (high mortality rate).[17] Again, this contradicts any prediction that those breathing the same air should suffer similar mortality rates.

The third piece of evidence was the death of the widow Susannah Eley in Hampstead, far from Broad Street (Snow, 1855, 44-45). She had lived in Soho near the pump (her husband had owned the percussion-cap factory at 37 Broad Street) and her sons (the current owners) regularly sent her water from the Broad Street pump. She clearly did not share in the air of the Broad Street pump, instead breathing the (seemingly clean) air of Hampstead where few others died. She consumed water from the Broad Street pump, as did her niece, and both died. This contradicts the prediction that those breathing different air should suffer different mortality outcomes, and corroborates that those drinking the same water should suffer similar mortality.

Stating the prediction and evidence in this detail clarifies how and why this is valuable for separating the alternative theories – a value sometimes not fully recognized. Evidence contradicted the simple airborne predictions (with water not a predisposing cause). Nonetheless the refutation did little to convince those skeptical of Snow's theory. After 1849 water was incorporated as both a predisposing cause, and a potential vehicle for transmission. As a predisposing cause, it could be that contaminated air would only cause disease in those also exposed to contaminated water. And as a vehicle for transporting (possibly dissolved) airborne cholera poison, it could cause disease in someone exposed to clean air but drinking contaminated water. This is, in fact, exactly the justification used by the *Committee for Scientific Inquiries* in defending airborne theories:

> [the case of the widow] might easily be admitted, without its therefrom resulting that infection depended on the specific material alleged [fecal matter or *choleraic excrements*].

[16] As highlighted, for example, by Johnson (2007, 173, 175)

[17] See Coleman (2018, 23 ff) for a discussion of the underlying data and the presentation as a modern contingency table analysis. The probability that not drinking vs drinking and healthy vs sick are statistically independent is exceedingly small – the p-value for a Fisher exact test is far less that 0.01%

> The water was undeniably impure with organic contamination; and ... if, at the times of epidemic invasion there was operating in the air some influence which converts putrefiable impurities into a specific poison, the water of the locality ... would probably be liable to similar poisonous conversion. Thus, if the Broad Street pump did actually become a source of disease to persons dwelling at a distance ... this ... may have arisen, not in its containing choleraic excrements, but simply in the fact of its impure waters having participated in the atmospheric infection of the district.

These auxiliary hypotheses meant that breathing the same air or not no longer determined mortality outcomes – water was now a causal factor. These auxiliary hypotheses were not themselves irrational, but they did serve to protect the airborne theory from refutation by the evidence that mortality was not associated with breathing air.

Although the airborne theory is protected from refutation by these particular pieces of evidence, it does not relieve advocates of the airborne theory of responsibility for both explaining why the airborne theory is a better explanation than the fecal-oral theory, and producing *positive* evidence of airborne transmission instead of simply avoiding contradiction by contrary evidence. On both counts the airborne theory falls short. As Snow (1849a, 6) states, there is no a priori reason to conclude that airborne must be the mode of communication, and ingestion of contaminated fecal matter provides a direct explanation for the observations of *both* person-to-person transmission *and* these instances of non-airborne transmission. Further, we know of no evidence produced during the 1850s that demonstrated airborne transmission directly. In other words, even though this evidence was prevented from directly refuting the airborne theories, it does argue against airborne theories.

5b – Water is Causal Prediction (5b) is that contaminated water is associated with higher cholera mortality, and in fact that contaminated water causes cholera. The evidence supporting the (causal) association between contaminated water and cholera was very strong. Water as a causal factor was supported by multiple strands of evidence from multiple sources. Case studies of Albion Terrace and Horsleydown (Snow, 1849a, 1855) and Hope Street (Sutherland, 1851, 14). The nascent difference-in-differences (DiD) from Snow (1855, 89) and the mixed (randomized) comparison from Snow (1855, Table IX). The improved mixed comparison of Simon (1856, Table I). The analysis of drinkers versus non-drinkers for the Broad Street pump from Whitehead (Cholera Inquiry Committee, 1855, 128f). We cannot in this paper recount the details of all the evidence and testing. Some background is given in Sectiion C, and Coleman (2018) and Coleman (2020) delves into specifics regarding Snow's evidence and testing in greater detail. Generally these analyses would not satisfy modern statistical standards, but it is the repeated conclusion from multiple studies by different researchers with different methodologies that leads to our assessment that the quality of the evidence is "High."

Snow's fecal-oral theory immediately and directly predicts that water would cause cholera transmission, since water contaminated by sewage would be a mechanism for inadvertently transmitting fecal matter from an infected person to others. The situation for airborne theories is somewhat more complicated. The primary mode of transmission is airborne, so one might presume they would *not* predict water is causal. Such a presumption might have been correct prior to 1849, but by 1855 water had been, by means of auxiliary "protective" hypotheses, broadly incorporated as both a

predisposing cause and as a mode of transmission.[18] These auxiliary hypotheses were degenerating in Lakatos's terms because they protected the theories from recalcitrant observations but did not increase the empirical content of the theories by predicting *new* facts, simply explaining the existing old facts. Nonetheless, the airborne theories by 1855 *did* predict water as causal, so that in Table **??**, columns three and four of row (5b) is labeled "yes". To record that the theory is protected from contradiction by adoption of (*ad hod*) auxiliary hypotheses, however, the cells are labeled Protected.

5c-f – Other Factors Than Water as Causal The fecal-oral theory predicted that, particularly in a municipal outbreak such as South London in 1854, factors other than water would have no – or minimal – influence. Airborne theories predicted that many other predisposing conditions – overcrowding and poor ventilation, poverty and unhealthy living conditions, dampness and filth – would all contribute substantially to mortality. Failure to find influence of other factors apart from water would be a corroboration of the fecal-oral theory and contradiction of alternative airborne theories.

Snow's nascent difference-in-differences and randomized comparison (Snow, 1855, 89 and Table IX) and more definitively Simon's re-analysis with population data by supplier (Simon, 1856, 6-7) was evidence that water was causal. But none of these addressed or ruled out the possibility that other factors also contributed to cholera mortality. Comparison between clean and dirty water for a mixed (quasi-randomized) population would average out effects of, say, overcrowding, and provide estimates of a water effect purged of confounding by crowding. But this does not *rule out* crowding as a separate causal factor, it simply rules out that an association between water and crowding causes a spurious association between water and mortality.[19]

Snow may have recognized this, but in any case in Snow (1856b) he turned from trying to demonstrate water as *a* factor to water as the *primary* factor, attempting to "prove the overwhelming influence which the nature of the water supply exerted over the mortality" (Snow, 1856b, 248). He did so by a useful and ingenious idea. South London consisted of some 30 subdistricts, with varying proportions of the population supplied by the Southwark & Vauxhall Company (contaminated water) versus the Lambeth Company (clean water). He assumed mortality for S&V customers was constant across the whole region, and the same for Lambeth customers. He then predicted mortality in each subdistrict based solely on the proportion of customers by company, under the hypothesis that only water (no other factors) mattered, and compared the predicted versus actual mortality (Snow, 1856b, Table VI). At this point, however, Snow was hamstrung by the lack of statistical and quantitative tools available in 1856. He states "it will be observed that the calculated mortality bears a very close relation to the real mortality in each subdistrict ... and proves the overwhelming influence which the nature of the water supply exerted over the mortality, overbearing every other circumstance which could be expected to affect the progress of the epidemic" (Snow, 1856b, 248).

Snow's argument is correct but weak, relying as it does on comparing two sets of 30 mortality rates by eye. Today we can use regression analysis and apply formal measures for the goodness of fit and

[18] Regarding water as a cause: "a number of most severe and fatal outbursts of cholera were referable.to no other cause except the state of the water-supply" Sutherland (1851, 15). Regarding water as a mode of transmission: "Thus a second great inlet is established by which injurious substances in a liquid form as certainly find their way into the circulating blood, as noxious aerial agents do by the way of the respiratory apparatus" Grainger (1851, 91).

[19] Coleman (2020) discusses the Snow's use of the South London municipal outbreak of 1854 and issues of randomization and determining water as *a* causal factor versus the *primary* causal factor.

testing for the significance of other factors and confirm that Snow's claim is correct (see Coleman, 2020). Using modern statistical tools we can say with confidence that water is causal but other factors are not, i.e. to corroborate Snow's fecal-oral theory and contradict the alternatives. Even with a definitive result for this single prediction, however, it would be reasonable to argue that the alternatives do not need to be abandoned. The analysis here is limited to South London, and possibly there is not the necessary variation in alternative factors. For example, it would be possible that an experiment comparing a relatively low-lying area near the Thames (South London) with a higher-elevation area further from the Thames (Hampstead) would produce a different result. The fact that a single negative result rarely leads to abandonment of a theory echos Lakatos's contention that "[o]ne of the most important points one learns from studying research programmes is that relatively few experiments are really important" (Lakatos, 1980, 65). Consideration of the range of predictions and multiple strands of evidence is required for any firm conclusion in sorting between theories.

Predictions 8 & 9 – Epidemic Curve (Timing) for Neighborhood & Municipal Outbreaks
The fecal-oral theory predicts the time-pattern of outbreaks. For neighborhood outbreaks due to contamination of a point source such as the Broad Street pump, the outbreak will be explosive: "In the cases in which the cholera poison gains access to a limited supply of drinking water, such as a tank or pump-well, the outbreak it occasions is always sudden, violent, and limited" (Snow, 1856b, 250). For municipal outbreaks the outbreak will be more moderate: "when a river is the medium of the propagation of the disease, its progress is more gradual and extended" (Snow, 1856b, 250). But a municipal outbreak will also show important variation in the relative mortality of those supplied with contaminated versus clear water. Early in the outbreak, when most cases are from exposure to the contaminated municipal source, the mortality will be much higher for those drinking contaminated water. As the outbreak progresses and normal (person-to-person and within-household) propagation increases, the relative mortality will fall (but never one-to-one): "In the beginning of the epidemic the cases appear to have been almost altogether produced through the agency of the Thames water obtained amongst the sewers and the small number of cases; occurring in houses not so supplied, might be accounted for by the fact of persons not keeping always at home and taking all their meals in the houses in which they live but as the epidemic advanced it would necessarily spread amongst the customers of the Lambeth Company, as in parts of London where the water was not in fault, by all the usual means of its communication" (Snow, 1855, 82).

Airborne theories will not predict these patterns, and therefore these predictions provide a good means of separating between the alternatives.

The evidence for explosive neighborhood outbreaks is in the form of case studies (rather than statistical) but is pretty clear. Hope Street (Sutherland, 1851, 15-16), Horsleydown and Albion Terrace (Snow 1849a, 15ff; Snow 1855, 30ff), and Broad Street (Snow, 1855, 38-49) are all instances of point-source neighborhood outbreaks. The fecal-oral theory has a good explanation for the pattern of those outbreaks: sudden onset and explosive growth in mortality from an otherwise low background rate. Airborne theories do not.

Regarding the time pattern for municipal propagation, Snow argued that the influence of the water supply should be highest at the start of an outbreak – the ratio of mortality for those supplied by

contaminated versus clean water highest at the beginning. As the outbreak progressed the disparity would diminish: "as the epidemic advanced [cholera] would necessarily spread amongst the customers of the Lambeth Company [those with clean water] ... by all the usual means of its communication" Snow (1855, 82). Snow presented evidence for this (Snow, 1855, 80, 86-88) but most clearly in Snow (1856b, 243-245):

> the influence of the water supply was found to diminish in relative intensity as the epidemic progressed. In the first four weeks ... the disease was between thirteen and fourteen times as fatal to the population having the impure water [supplied by the Southwark and Vauxhall Company] as to that having the improved supply [the Lambeth Company]. ... In the next three weeks ... the mortality [rate] was ... nearly eight times as great in that supplied by the Southwark and Vauxhall Company as in that supplied by the Lambeth Company. During the last ten weeks ... the mortality [rate] was still more than five times as great amongst the population supplied by the former company as amongst that supplied by the latter.

Snow repeated these observations in Snow (1856a, 1857).[20]

Predictions about the time pattern or epidemic curve for mortality rates, for both neighborhood and municipal outbreaks, qualify as a "new facts" uncovered by the theory, in the sense that Snow looked for the evidence based on the prediction. Airborne theories had no such predictions, and thus were contradicted by these observations. The prediction and the evidence supporting the prediction are a strong corroboration of the fecal-oral theory and strong contradiction of airborne theories.

D.2 Summary of Predictions versus Evidence

To summarize, predictions (2 a&b) concerning contagion, and predictions (8 & 9) concerning the time-pattern of mortality, all strongly refuted the airborne theories. Prediction (5c-f) concerning other factors weakly refuted the airborne theories – weakly because Snow could not demonstrate (as we can today) that in fact the actual versus predicted mortality across subdistricts depended only on water and not other factors.

The two remaining predictions – (4) concerning airborne transmission and (5b) concerning water – demonstrate how auxiliary hypotheses can protect a theory from recalcitrant facts. If cholera is airborne, then those sharing the same air should suffer similar mortality, while those not sharing the same air may not. As discussed above, Snow discusses counterexamples to both.

The evidence shows that transmission was *not* airborne but waterborne. And yet this does not refute or falsify the airborne theories, because they had incorporated water as both a predisposing cause and as a mode of transmission. The airborne theories were protected from contradiction by the auxiliary hypotheses concerning water (hypotheses that were *ad hoc* according to Lakatos's criteria), and for this prediction we have entered "P" into the final column of Table **??** to denote that the airborne theories are protected from refutation by this prediction. Nonetheless, this prediction

[20] Using modern statistical tools the observations on the time pattern of mortality rates can be reproduced and strengthened – see Coleman (2020).

does have some value for discriminating between theories. First, airborne cholera requires some *positive* evidence of airborne transmission, and we know of none. Second, this case is grounds for claiming the airborne theories are degenerating – the auxiliary hypotheses indeed protects against the observations of waterborne transmission but *without* any increase in theoretical or empirical content, and with no new facts within the airborne programme.

Finally, there is the prediction of water as a causal effect for cholera mortality (5b in Tables **??** and 1), often considered the signature evidence for Snow's theory. In fact, this prediction and evidence did little to discriminate between theories, because by 1855 all theories predicted that water would be causal – at least in some if not all cases. We rate the value of this prediction as low for discriminating between theories, although as with the case of airborne transmission, the adoption of the auxiliary hypotheses within the airborne programme without any increase in content or new facts is an indication of a degenerating programme.

The fecal-oral theory's predictions and observations summarized in Table **??** fits the description of a progressive programme: new predictions (2b, 5c-f, 8, and 9) all of which were corroborated. The airborne theory fits the description of a degenerative programme: the auxiliary hypothesis of water as a predisposing cause explained given facts but did not predict new ones. Furthermore, predictions of the airborne theory were contradicted by the new facts: the fecal-oral theory exhibited higher corroborated content.

It is important to recognize that, among the wide range of predictions and evidence available in the mid-1850s bearing on alternative theories shown in Table **??**, there is no single crucial prediction or experiment that did serve, or even should have served, to convince skeptics of the primacy of the fecal-oral theory. What Table **??** does show is that, although airborne theories were not uniformly refuted by the evidence (largely due to the auxiliary hypotheses concerning water as a predisposing cause or method of transmission – see predictions 5a and 5b), the preponderance of evidence was in favor of the fecal-oral theory *and* the airborne theories were refuted in many instances; the fecal-oral theory was a theory that exhibited higher corroborated content.

D.3 Epidemiological versus Biological Evidence

The evidence presented by Snow and others is what we might call "epidemiological evidence," in contrast to individual-level "biological" evidence of the particular causative agent – in this case *vibrio cholerae* – and on the particulars of transmission – for example that choleric material existed in a contaminated water source, and that victims actually drank the contaminated water. There was scant biological evidence available in the 1850s. Filippo Pacini identified *vibrio cholerae* in 1854 but Pacini's work did not appear to have been known in London in the 1850s, and *v. cholrae* was not universally recognized until Robert Koch's re-discovery in 1883 (Carboni, 2021). The closest to biological evidence of which we are aware was Whitehead's identification of the Broad Street index case and the subsequent uncovering of the direct contamination pathway from the baby's cholera dejections to the pump. The Reverend Henry Whitehead discovered that a baby girl Frances Lewis living at 40 Broad Street – facing the pump in Broad Street – fell ill two or three days before the outbreak commenced (and died two days after). Water from rinsing diapers was emptied into the cesspool at the front of the house. The Cholera Commission had Mr. York, a local surveyor and

secretary for the Cholera Inquiry Commission, investigate the sewer lines. He found that a leaking waste pipe from 40 Broad passed within feet of the pump-well and that sewage from 40 Broad Street could leak into and contaminate the well. (See Cholera Inquiry Committee 1855, 159-160, 170 ff and Johnson 2007, 178ff.)

The addition of biological evidence to the epidemiological and statistical evidence of the 1850s would have strengthened the argument (in favor of the fecal-oral theory). Critics asked for positive evidence of water contamination, and argued that dilution in water made transmission by water improbable.[21] Lack of such biological evidence, however, does not detract from the strength of the statistical evidence. No doubt the development of biological evidence subsequent to Robert Koch's rediscovery of *v. cholerae* was the crucial factor that swung the scientific community behind the fecal-oral theory.

[21]Snow did provide reasonable counter-arguments: "The morbid matter of cholera, like the pus of small-pox and other morbid poisons, owes its properties, no doubt, to cells; and there is no more reason why diluting with water should necessarily destroy their activity, than that it should destroy the ova of fishes; and there is also no reason why every one who dips a pail into the water should draw a prize" (Snow, 1854). In any case, airborne theories were subject to the same criticisms.