If the Framingham Heart Study Did Not Invent the Risk Factor, Who Did?

David Shumway Jones* and Gerald M. Oppenheimer†

ABSTRACT Most historians, epidemiologists, and physicians credit the Framingham Heart Study for introducing the term "risk factor" to public health and medicine. Many add that the term came from life insurance companies. This familiar history is incorrect. Taking advantage of the expanding availability of digitized and full-text searchable journals, textbooks, newspapers, and other sources, we have uncovered a deeper and broader history. Antecedent concepts (such as risk, factor, predisposition) have ancient roots. "Risk factor" began to appear in the late 19th and early 20th centuries in many industries, not just in insurance but also in finance, agriculture, and manufacturing. The term appeared in the occupational health literature in 1922. It reappeared in the 1950s in many different areas of medicine including psychiatry, surgery, cardiology, epidemiology, and aerospace medicine. Furthermore, despite the influential appearance of "risk factor" in a 1961 Framingham Heart Study publication, the term did not gain momentum in medicine and public health until the mid-1970s. While our analysis is not exhaustive, our findings are extensive enough to require a substantial revision to the history of the risk factor.

*Department of the History of Science, Harvard University; and the Department of Global Health and Social Medicine, Harvard Medical School.

†School of Public Health and Health Policy, City University of New York

Correspondence: David Jones, Harvard University, Department of the History of Science, Science Center 371, 1 Oxford Street, Cambridge, MA 02138.

E-mail: dsjones@harvard.edu.

Perspectives in Biology and Medicine, volume 60, number 2 (spring 2017): 131–150.

MEDICAL THEORY AND PRACTICE in the second half of the 20th century were transformed by the idea of risk, and, in particular, by the concept of the "risk factor." Many historians have described how the concept of the risk factor emerged in the actuarial science of the life insurance industry in the early 20th century and entered medicine through the Framingham Heart Study (FHS), specifically with its July 1961 article, "Factors of Risk in the Development of Coronary Heart Disease," in the *Annals of Internal Medicine* (Kannel et al. 1961). This story, while plausible, is incorrect.

New techniques of digital history enable a new kind of research: full-text searching of publications in medicine, public health, and other fields for specific terms and their antecedents. This reveals that "risk factor" appeared in medicine 40 years earlier, in 1922. While it might have come to medicine through insurance, this is not clear. The phrase circulated in many domains before 1961. Moreover, the concept of factors that influence risk was widespread in medicine and other industries long before the phrase itself. The long history of "risk factor" and the longer prehistory of relevant concepts create new puzzles and opportunities for historical research.

In this essay we excavate the history and prehistory of "risk factor" in medicine, public health, and other discourses. Our analysis is not comprehensive: digital databases remain incomplete, and earlier instances of "risk factor" will emerge as more material is searched and digitized. That said, the available record is sufficient to support a substantial rewriting of the history of the risk factor. The concept emerged independently in many domains before 1961. While it is possible to speculate about links between these threads, none of them are clear-cut. Moreover, the uptake of the term after 1961 was slow. This suggests that the risk factor did not revolutionize health care in the decades after World War II. Instead, the risk factor concept underwent a slow transformation that drew on many leads and accrued many meanings before gaining momentum in medicine in the 1970s. These developments reflected broader changes in how risk was recognized and conceptualized throughout society.

THE CONVENTIONAL NARRATIVE

There is no doubt that medical theory and practice have been transformed by notions of risk. William Rothstein described the transition well in 2003:

> At the beginning of the twentieth century, most people believed that their health was a matter of concern only when they were sick. By the end of the century, most people accepted the statistical evidence that specific behaviors and characteristics of healthy persons, called "risk factors," can increase the probability of developing disease, especially chronic disease. (xi)

People are no longer well, but rather in a state of "susceptibility," steadily accumulating preconditions for future diseases. Doctors take histories, obtain laboratory tests to detect pre-disease states such as insulin resistance, diagnose risk, and prescribe drugs to reduce the risk of moving from susceptibility to pathology. Jeremy Greene (2007) has described this as "prescribing by numbers." Robert Aronowitz (2015) has explored the deep consequences of the risk revolution: "Reducing risk is no longer a means to health but often its very definition" (202). Public health took up risk factors with equal enthusiasm, from studies of the probability of morbidity and mortality in prospective cohorts, to analyses of the ways in which income inequality and relative position in status hierarchies are risk factors for disease and premature mortality. Risk, its distribution, and its management are now a central preoccupation of life and governance at the individual and social levels. As sociologists have described, we now live in "risk societies" (Aronowitz 2015, 23; see also Beck 1992).

Even though risk discourse reflects a social evolution with deep roots, many historians have argued that the critical concept, the "risk factor," crystalized in the work of the Framingham Heart Study, initiated in 1947. As Aronowitz noted in 2015, "many observers simply equate the two as cause and effect" (71).[1] The standard narrative begins in the late 19th century, as life insurers explored new ways to offer policies to a broader swath of the population, including people who were at high risk. Rothstein (2003) has described how the Actuarial Society of America developed new ways to estimate and price risk in 1903 "by constructing life tables that quantified the risks of factors such as build, occupation, medical history, and residence" (63). By 1919, actuaries had developed a rating system in which the "state of each risk factor for an applicant was assigned a number, the numbers for all factors were summed, and the total score was compared to a scale that assigned applicants to specific risk categories" (64). Rothstein asserts that the term "risk factor" appeared in this context (282). However, he later admitted that he did not actually see that phrase in the early insurance literature.[2]

Insurance companies could easily have introduced the risk factor concept to medicine: they hired many physicians to examine and assess their applicants. As discussed in detail below, recognizable notions of risk and factors percolated through medicine from the 1920s through the 1950s. Élodie Giroux (2011) has described how cardiologists recognized a "cardiac" or "coronary profile" (316). For instance, Paul Dudley White (1957) asserted that clinicians "know from our own experience that basic factors behind coronary heart disease are of very great importance," including heredity, sex, race, age, stress, strain, tobacco and alco-

[1]Aronowitz discounts the importance of the FHS but does not trace alternative origin stories.
[2]Rothstein to Jones, email, March 18, 2017. Rothstein explained that "I don't think the life insurance industry used the term in the early years. Rather, I think they used the phrase 'factors of risk,' which I vaguely recall seeing."

hol use, rest, relaxation, and diet (2). Mikko Jauho (2012) has described parallel developments in Finland, where life insurance underwriting linked insurers, physicians, and applicants into a mutually shaped discourse on risk, rationality, and, eventually, risk factors.

These developments converged on the FHS. Gerald Oppenheimer (2006) has shown that the "language of risk, specifically, the probability of developing CHD, given the presence and magnitude of certain personal attributes, permeated early Framingham publications" (725). The researchers developed 28 initial hypotheses, including, for instance, "'degenerative cardiovascular disease appears earlier and progresses more rapidly in persons who habitually use tobacco'" (725). The FHS 1949 *Manual of Operation* discussed both "constitutional factors" and "conditioning factors" (Giroux 2013, 102). In 1951, the FHS issued a statement about the multifactorial etiology of heart disease (Rothstein 2003). The researchers' thinking and language evolved over the decade that followed. Giroux (2013) has described how "'predisposition' and 'susceptibility' or even 'constitutional factors' were progressively replaced first by 'risk characteristics or attributes', and then by 'risk factors'" (108). When researchers published their first results in 1957, they used *risk* repeatedly to indicate an increased probability of future disease (Oppenheimer 2006). However, at that early point (six-year follow-up) there had been too few events for them to identify specific risks conclusively (Aronowitz 2015). In 1958, Stamler wrote about "actuarial predictions of risk" (qtd. in Oppenheimer 2006, 721). Meanwhile, they used *factor* to indicate a characteristic or clinical attribute (Giroux 2013). In 1959, the FHS described "factors believed to be important in the development of coronary heart disease" and discussed "disease-prone individuals" and "environmental and host factors associated with the development of disease" (qtd. in Rothstein 2003, 282–83). They focused on probabilistic disease prediction and not on etiology: "The FHS investigators clearly wanted to develop instruments which would allow the pre-diagnosis of cardiac cases before it was too late to intervene" (Giroux 2013, 108).

By 1961, the FHS researchers had growing confidence in the robustness of their findings. Giroux (2013) argues that as soon as they had enough data to calculate the relative risk of specific factors, they began to discuss risk factors. They used both "factors of risk" and "risk factor" in the 1961 *Annals* article, in the title and then the text, respectively (Kannel et al. 1961). Giroux (2013) argued that this was a new usage, borrowed from the insurance industry: she claimed that "Before the 1960s, the notions of 'risk' and 'risk factor' were not used by medical professionals but by life insurance industries, especially when talking about obesity and hypertension" (108). It is clear that the researchers did not coin a new usage deliberately. Aronowitz (2015) notes that the "term was used without definition or special emphasis, suggesting that there was little consciousness that the term was new or suggested any new conception of the disease" (85). Oppenheimer (2006) has similarly described how "risk factor" was used "in a rather

casual manner. Kannel, the lead author, claims to have soon forgotten that they had coined the term, or at least first applied it to epidemiology, until reminded some years later" (725).

Historians have emphasized how quickly the concept caught on. Aronowitz (2015) writes that the "expression quickly took off. By 1965, the term was used in most heart disease epidemiology publications and in most lay magazine and newspaper coverage of research into heart disease epidemiology" (85). Oppenheimer (2006) describes the phrase's broad appeal:

> In 1961, "risk factor" seemed to capture a moment. It was a phrase that signaled a new approach to disease. It was more felicitous than terms like "actuarial predictions of risk," a short-hand formula that could be applied aptly to each dimension of CHD epidemiological thought, a term that joined them together. (725)

Aronowitz (2015) adds that "risk factor" offered "a new *style* of explaining cause and responsibility, one that used probabilistic language to link quantifiable and elementary properties of individual physiology, behavior, and social and familial background to specific untoward outcomes" (69–70). It served many functions: "It also provides a consoling framework with which to minimize the frightening randomness of disease. Finally, it provides a sanctioned vehicle, in a secular era, with which to say certain individuals and groups are doing the right thing and others the wrong thing (e.g., overeating, smoking, not taking medicines, never exercising)" (92). This became the rationale for public health and education campaigns.

This is a plausible and appealing narrative. It makes sense that insurers would think this way. It makes sense that the logic would be applied to heart disease, then the leading cause of death, whose clinical appearance was often a fatal or disabling event and for which there was little effective treatment. There were specific links between the insurance industry and FHS. The history also fits comfortably into broader narratives of heart disease triumphalism: in the decades that followed World War II, the science of heart disease improved, and the disease has been substantially vanquished (Jones and Greene 2013).

The Problem, Our Method, and Its Limitations

There are, however, several problems with this narrative. While conducting other research, we stumbled across many occurrences of "risk factor" in the medical literature before July 1961. The phrase also circulated beyond medicine. The *Oxford English Dictionary* reports that the earliest occurrence of "risk factor" was in 1907, but we have found it used in 1895. It occurred in many domains, of which insurance was just one. This raises important questions. First, what is the fuller history of the phrase, beyond that provided in the familiar life insurance and FHS

narrative? Second, what is the prehistory of relevant concepts before the specific phrase "risk factor" appears? Third, how did the phrase's meanings and uses develop over time? Fourth, how do the concepts and phrase manifest in languages beyond English? In the past, such questions would have been prohibitively difficult, requiring extensive reading of the literature of medicine and other areas in search of the needle in a haystack. However, the advent of full-text searchable digital sources makes this kind of research newly possible. One can search swaths of the medical literature for *risk, factor,* and "risk factor," as well as for "factors of risk," "relative risk," "causative factors," or anything else. It is also possible to search broader databases for occurrences in other literatures.

Despite these new capacities, the process is still not straightforward. The ideal features of a digital resource are clear: (1) a repository of sources that are full-text searchable (for example, scans that have been accurately processed with optical character recognition, or OCR, software); (2) a search engine that allows searching by word or exact phrases; and (3) search results that can be limited in time and sorted in useful ways. Few collections meet all of these criteria. The current search engine for the *New England Journal of Medicine* comes close.[3] The search engine for *JAMA* offers this functionality but does not actually provide it.[4] PubMed must be searched carefully (a search for "risk factor" misses occurrences of "risk factors"), and its results are not wholly reliable.[5] Searches are often compromised by inaccurate metadata. Google Books (especially its Ngram Viewer) offers an excellent starting point, as it provides an enormous digitized corpus going back to 1800.[6] However, while it yields many valuable hits, as described below, it also lists sources with dates before 1961 that actually date from well after 1961.[7] Each online repository must be examined closely to determine its capabilities and limitations.

Because of current limitations with existing digital repositories and search engines, we cannot offer a conclusive history of "risk factor." Instead, we pursued a multi-pronged approach to find many early occurrences of the term: (1) PubMed; (2) prominent American and British medical and public health journals that are full-text searchable from their first volumes, specifically *New England Journal of Medicine*, *Lancet*, *BMJ*, *JAMA*, *Circulation,* and *American Journal of Public Health*; (3)

[3]*New England Journal of Medicine* (http://www.nejm.org/medical-search). While the full text has been run through OCR and is searchable, many of the actual PDFs are not searchable: you may know that the term exists in the article, but you have to read the article to find it.
[4]*JAMA* (http://jamanetwork.com/advanced-search). A search for "risk factor" instead returns hits to *risk* and *factor* co-occurring in the same article.
[5]PubMed.gov (https://www.ncbi.nlm.nih.gov/pubmed). PubMed reports "risk factors" in four articles between 1952 and 1960 that do not actually contain the phrase.
[6]Google Books Ngram Viewer (https://books.google.com/ngrams).
[7]For example, a 1949 hit about the risk factor of coliform bacteria in water (Jerry D. Stoner's *Water-Quality Indices for Specific Water Uses*, Geological Survey Circular 770, U.S. Department of the Interior, 1949) is actually from 1970 (as printed on the volume's title page).

Oxford English Dictionary; (4) Google Books; (5) Proquest Historical Newspapers; (6) financial magazines, specifically *Barron's* and *Forbes*; and (7) marketing reports (available through www.marketresearch.amdigital.co.uk). In this essay, we discuss important instances of "risk factor" before July 1961 that we found through these sources. We did not attempt a systematic analysis of occurrences after July 1961, but we offer an initial sketch in our epilogue. We did not look beyond English language sources.

We also looked for earlier occurrences of the risk factor concept appearing under different names, since scholars have done this productively with other concepts. For instance, while the terms *gene* and *genetics* date to the early 20th century, Staffan Müller-Wille and Hans Jörg Rheinberger (2007) have excavated earlier histories of genetics by examining *heredity, ancestry, inheritance,* and other antecedent concepts. This endeavor, however, is inevitably open-ended, requiring subjective judgments about whether an early usage of "factors of risk" or "causative factors" is meaningfully related to "risk factor" as now used.

We offer a provisional analysis here. While our method is not exhaustive, it does generate an initial genealogy of ideas about risk and risk factors that justifies a substantial revision of the existing history of "risk factor."

THE EVOLUTION OF *RISK* AND *FACTORS*

The basic notion that specific characteristics or exposures were associated with current or future disease is an ancient one in medicine, evident in Hippocratic texts. The old ideas have been reformulated repeatedly in medical theory and practice. According to the *OED, risk*, for instance, entered English in the 16th century from the French *risqué*, "danger or inconvenience, predictable or otherwise," usually in discussions of commerce. By the 17th century, risk could be found in discussions of politics, crime, reputation, or financial investments. By 1867, it had assumed the meaning seen in insurance and medicine: "A person or thing regarded as likely to produce a good or bad outcome in a particular respect." *Factor*, which the *OED* reports first appeared in English in the 14th century, took on many meanings. It was used to denote an "agent who buys and sells, or transacts other business, on behalf of another person or company," or, more generically, a person or thing "that makes or does something," for instance a "component affecting the outcome, nature, or perception of something."

Both words have long been used in medicine. There is no easy way to search the early modern literature systematically, but relevant examples are easy to find.[8] A 1690 treatise on the treatment of "French-pox," for instance, warns patients to choose their doctors carefully before they "run a Risk" and accept mercury,

[8]Early English Books Online (https://eebo.chadwyck.com/home) produces no hits for "risk factor," but 1,307 hits in 843 sources for *risk*. However, there is no obvious way to limit this result to medical texts.

fumigation, or other potentially dangerous remedies (*A New Method* 1690, 28). As modern medical journals emerged in the 19th century, the concept of risk was found in the earliest issues. The *New England Journal of Medicine and Surgery* published its first issue in January 1812; *risk* appeared in April. Many early occurrences involve the risks of therapeutic intervention, though discussions of the risk of acquiring a disease also appear. *Lancet* first published in 1823; *risk* appeared in its initial issues, in discussions of the dangers of graverobbing, pompous doctors, horseback riding while intoxicated, or surgery. *Factor* occurred with various meanings (merchant factors, mathematical factors) in the *Boston Medical and Surgical Journal* in the 1830s and 1840s. It first appeared in a discussion of disease causation in an 1861 lecture on therapeutics (Ware 1861).

The terms began to converge in the late 19th century, often in discussions of surgery. An 1892 article in *JAMA*, for instance, discussed factors that contributed to the idiosyncrasy of anesthetic risk. A 1909 article discussed the various risks of surgery and the factors that influenced recovery, especially the surgeon's knowledge and skill (Simpson 1909). A 1924 article advised surgeons about how to estimate factors (such as the patient's age, obesity, and blood pressure, the lesion type, and the surgeon's skill) that influenced operative risk (Lemon and Moersch 1924).[9] These discussions closely resemble the modern concept of risk factor. While surgeons might have acquired this approach from insurers (many surgeons presumably worked as insurance examiners), this surgical literature might also represent a distinct origin story. Surgeons and life insurers share a similar goal: the desire to predict mortality in order to protect someone or something (patient, surgeon, insurer) from a bad outcome. Thomas Schlich's (2013) work on surgeons, railway accidents, and risk management lends credence to the suggestion that surgeons were thinking along these lines by the 1880s, but more extensive research would be needed to ferret out the relationship between surgery and insurance in the late 19th and early 20th centuries.

Similar suggestive language appears in the public health literature. An article in *Public Health Reports* in 1896 described the danger of plague infection in Cuba, where "personal risk is a differential of personal hygiene." A communication in the *American Journal of Public Health* in 1913 warned that the new Panama Canal would likely increase trade to Guayaquil, Ecuador and, consequently, "the risk of spreading its diseases." In 1926, *Public Health Reports* described how workers who painted motorcars were "exposed to risk of lead poisoning" (Badham 1926). *Fac-*

[9]The hope that a single factor—vital capacity—could estimate operative risk was dashed: "It was found that the surgeon considered the operative risk higher in seventy-two instances than the vital capacity determinations indicated, whereas the reverse was true in thirty-two instances. Several factors account for this variation. The operative risk calculated from the vital capacity was based simply on the patient's physical fitness. On the other hand, the surgeon had the benefit of his own knowledge of the type of disease and of his skill in performing the particular operation under consideration, and also of the internist's opinion regarding the patient's general condition. None of these factors was included in the calculation from the vital capacity readings" (Lemon and Moersch 1924, 129).

tor, roughly meaning a fact that produces an effect, also appeared often. In 1922, C. E. A. Winslow discussed in the *American Journal of Hygiene* the possible factors that helped explain the explosiveness and timing of the influenza epidemic within a given community. Researchers studying disease in populations often sought to compare and contrast the degree of risk represented by a factor. A 1930 analysis in *Public Health Reports* discussed the "relative risk" of different types of laborers to industrial accidents (Britten 1930). By comparing the risks of different types or degrees of exposure, relative risk provided evidence of association, one of the key features of "risk factor" in postwar epidemiology.

A related concept, "factors of risk," also circulated. An 1897 article in the *British Medical Journal* defended the safety of the typhoid vaccine, citing the successful experience of Waldemar Haffkine in India to argue that "there is no unknown factor of risk in the case of the injection of bacterial toxines [sic]" (Wright and Semple 1897, 258). A 1927 comparison of two surgical procedures for removal of kidney stones warned that one procedure, nephrotomy, had "added factors of risk," including hemorrhage, scarring, and infection (Hamer 1927). A 1937 analysis of medical insurance in Nazi Germany praised the way in which the capitated fee schedule for doctors who treated workers considered "the factor of risk of the industry concerned" (Haedenkamp 1938, 161). What is the substantive relationship between "factors of risk" and "risk factors"? No single answer can be given, since different authors meant different things by each phrase. The 1961 *Annals* article is revealing: while the article used "factors of risk" only in its title, it used "risk factors" (four times) in the text, suggesting that the FHS authors saw them as equivalent (Kannel et al. 1961).

By the late 1950s, many usages close to the modern "risk factor" can be found. A 1958 article in *California Medicine* described the "multiple factors" that cause coronary artery disease, especially cigarette smoking and aging (Breslow and Buechley 1958). Another article that year in *JAMA* discussed the factors behind the risk of myocardial infarction (Pell and D'Alonzo 1958). In June 1961, Ernest Wynder argued that cigarette smoking was a "causative factor" for disease, "a factor that increases the risk that a given disease will develop and in the absence of which the disease would occur less frequently" (1236, 1238; for a discussion, see Parascandola 2011/12). Each of these included analysis of specific factors, quantification, and a notion of probability though not mechanistic cause.

The Emergence of "Risk Factors"

The specific phrase "risk factor" appeared by the 1890s. An 1895 article in the *New York Times*, reporting on the annual meeting of the International Underwriters Association, noted that "the important matter considered was the bicycle as a risk factor in accident insurance": "the opinion prevailed that wheelmen are subject to greater risks of injury than are pedestrians, and presumably future

accident policies will specifically classify wheelmen." This could be interpreted two ways. One reading is that bicycle riding increased the potential for accidents and injuries, a probabilistic account that fits current epidemiological usage of the term. Another interpretation is that the factor of interest is risk itself: underwriters price risk, and if bicyclists have a higher risk of injury, then their higher-risk policies would need to be priced accordingly. Other writers about insurance understood "risk factor" in the second sense: a potential hazard the consequences of which should be avoided, transferred, or priced. The *Manchester Guardian* reported a 1933 talk in which Sir Josiah Stamp stressed how entrepreneurs could gain an economic advantage if they could "isolate the risk factor": "the essence of insurance is transfer of a factor of risk from the field of interest to that of insurance." The *Guardian* later wrote that where such insurance did not exist, as when football (soccer) teams paid a large transfer fee to obtain a player, the "risk factor, and it is a considerable factor touching on unpredictable misfit as well as on injury, must remain as the chill wind that blows in these affairs" (*Manchester Guardian* 1938). While most of these uses of "risk factor" focus on risk, some did emphasize the various factors. In a 1953 article in the *Hartford Courant* titled "Other Driver Is Unknown Risk Factor," the paper warned about the uncertain sobriety or attentiveness of other drivers: "you know nothing about him, but he figures in accidents" (Marez 1953). In this case, these factors raised the risk of accidents (and insurance claims).

Risk factors also appeared outside of insurance. The first instance offered by the *Oxford English Dictionary* is a 1907 textbook about the economics of interest rates. In a single occurrence in this 465-page textbook, the author described how to assess risk factors to set interest rates: "the element of risk introduces disturbances into those determining conditions which were expressed in previous chapters as explaining the rate of interest. To summarize these disturbances, we may apply the risk factor to each of the six conditions which were originally stated as determining interest" (Fisher 1907, 217). In 1921, the *Manchester Guardian*, under the headline "Risk Factor in Interest Rate, Moneylender's Claim Fails," reported the outcome of a legal case that adjudicated a reasonable rate of interest given the risk the lender ran. In 1942, the *New York Times* noted that "the war risk factor" had reduced shipments of petroleum from Mexico (Cianfarra 1942). In 1947, an article in *Forbes*, "Where Risk Factors Appear Greater," explored the hazards of investing in industries facing a post-war recession, including automotive, tire and rubber, electrical equipment, and railroads (Krauss 1947). Just five years later, the economic tide had turned: stocks surged to record highs after President Eisenhower's election. A *Forbes* analysis offered "a reappraisal of the risk factor" (Biel 1952). From April 1950 through January 1951, *The Banker* ran a six-part series on "Risk Factors in 'Foreign Banking,'" with articles on export trade risks, import trade risks, exchange, travel services, overseas customers, foreign securities, and dividends. In these and other cases risk was the factor of interest, as it represented a hazard that decision-makers had to recognize and manage.

Google Books allows the broadest search, returning hits on hog pricing, industrial process control, international currency exchange, utility regulation, and petroleum exploration. In 1935, for instance, the Bureau of Agricultural Economics published a guide on *The Direct Marketing of Hogs,* in which the uncertainties involved in pork processing and storage were the "risk factor" that hog producers had to manage. A 1952 analysis of assembly line manufacturing described "risk factor" as the probability of producing products outside acceptable standards of quality (Heide 1952). A 1953 report from the International Monetary Fund used "risk factor" twice in a discussion of how to set exchange rates in order to stabilize post-World War II (and Korean War) currencies. A 1955 textbook on utility regulation discussed different factors that influence risk, "risk factors," as part of an effort to determine what rate of profit utilities could demand given the financial risk they took on (Nichols 1955). As before, the object of concern in these examples was risk itself and how to protect oneself from its hazards. Such a focus made sense in industries that sought to price risk in order to profit from it.

New technologies opened new venues for risk factors. In July 1961—the same month that the FHS article appeared in *Annals*—two newspaper articles addressed the risk inherent in human space flight. On July 16, the *Boston Globe* reported that NASA "makes no effort to hide the fact that there is a high risk factor in these experimental flights," in that case, a 15-minute trajectory into space (Menzies 1961). The next week, the *Christian Science Monitor* reported the successful flight by Gus Grissom to the outer edges of space, noting that "the safety factor in these shots is now 90 per cent, with a ten per cent risk factor" (Cowen 1961). Given that the risk in these cases was almost certainly a fiery death, such risks were substantial.

Risk Factors in Medicine

Given the long interest in risk and factors in medicine, and the broad circulation of the phrase "risk factor" since the 1890s, it should come as no surprise that the phrase appeared in medicine before July 1961.

Our first medical instance of "risk factor" occurred in the *Lancet* in April 1922, in a discussion of health conditions in the Royal Air Force. The article examined disease and injury rates within the RAF in 1920, comparing troops stationed at home or abroad, and troops involved in different types of work (trades, unskilled work, and so forth). Units actively involved in flight, such as air squadrons and flying schools, had the highest accident rates, but the accidents were not from flight alone: "there are many other risk factors to be considered—e.g., propeller accidents, starting car accidents, and accidents among skilled trades while at work, all of which apply to these units in particular" (656). The largest share came from athletics, especially football (soccer). The emphasis here was on factors that increase risk, as in the prior surgical literature about factors associated with operative risk.

The term appeared again in August 1922 in the *British Medical Journal*, in a discussion of industrial accidents. The article addressed accident prevention, worker's compensation, and the costs of treatment, but not predisposing factors: "I must leave out of consideration meantime such factors contributing to accidents as a monotony of employment; overtime; the shift factor; piece work; the illumination factor; and the risk factor involved in the particular occupation itself" (Kerr 1922, 378). While writing about specific factors, here the risk factor was risk itself. Further research is needed to understand why these first two occurrences came in England in 1922, in discussions of occupational health. While the *BMJ* article discusses workmen's compensation (an insurance issue), the *Lancet* article does not.

The term disappeared from medical publications that we searched for nearly 30 years before resurfacing in the American medical literature in the 1950s. A New York psychiatrist published a guide to suicide risk assessment in the *New England Journal of Medicine* in 1951. After reviewing a range of "suicide factors," including depression, anxiety, sexual panic (homosexuality and incest), alcohol, and "suicidiferous over-reactions to despair," he advised other clinicians that "a detailed evaluation of the risk factor may not be necessary, since the need for psychiatric hospitalization is usually clear in these cases" (Oliven 1951, 492). Careful risk assessment and "prompt recognition of a latent suicidal risk may actually save the patient's life" (494). In this case, risk factors were clinical characteristics associated with an imminent likelihood of suicide.

Five years later, *JAMA* published an extensive discussion of the various factors that contributed to surgical risk. Such risk assessment remained an imperfect science, "a sketchy, intuitive evaluation of the probability of dying during an operation and convalescence. The word probability bears statistical implications and infers the formulation of a judicious estimate of the magnitudes of the factors jeopardizing the person's life" (Moyer and Key 1956, 853). Since surgical techniques, especially skill with antibiotics, anesthetics, fluids, and blood transfusions, had improved over the 20th century, only three principal risks remained: the patient's age, the patient's debility, and the skill of the surgical team. "Risk factor" was invoked once, in reference to the surgeon: "the surgeon should look upon himself as a far more important operative risk factor than anesthesia and the anesthetist and recognize that for some surgically remediable diseases the surgeon, or he who calls himself one, may be the major factor in the assessment of the operative risk today" (855). This use involves the idea of a specific factor (agent), associated in a probabilistic way with an undesirable outcome. The long-standing discourse on operative risk had evolved into a recognizably modern usage.

In 1958, *BMJ* published a discussion of which patients would benefit from anticoagulant therapy after myocardial infarction. Anticoagulants had clear value in patients with "bad-risk factors," including "previous history of infarction, intractable pain, cardiac failure, arrhythmia, significant cardiac enlargement, diabetic

acidosis, and severe shock and hypotension" (Toohey 1958, 254). The author favored their use in "good-risk" cases as well: "the most important factor is an experienced clinician" (253). This is a distinct meaning of "risk factor": specific clinical traits associated with bad therapeutic outcomes.

In 1959, two researchers at the Douglas Aircraft Company published "Calculations of the Radio-Biologic Risk Factors in Nuclear Powered Space Vehicles." At this early moment in space travel, spacecraft designers worked to assess the risks faced by humans in space: "The problems and risks associated with man in space include the obvious variations in acceleration, weightlessness, changes in the closed ecological system, possible decompression, meteoroidal collisions, extra-terrestrial radiation (cosmic and Van Allen's radiation belt), temperature variation and psychologic factors" (Konecci and Trapp 1959, 487). The researchers worried most about radiation: "Of all the risk factors involved in space flight, exposure to ionizing radiations has perhaps received the most attention. The use of a nuclear reactor for propulsion will increase the radiation hazard to which the spacemen will be exposed" (487). By identifying factors associated with increased risk, spacecraft designers would be able to "avoid pitfalls and unjustifiable compromises at the expense of the human occupants in the initial design of an optimal nuclear space system" (488).[10]

In 1960, Lester Breslow, then a rising young epidemiologist, and Philip Buell wrote of the association of heart disease mortality and death from other causes, referring to a "general mortality risk factor" as a commonly shared prior variable (434). They then analyzed the relationship between the general mortality risk factor and both social class and physical activity. In a follow-up article later that year, they posited "the existence of some risk factor, or factors" that had "an obscuring effect on the association." They also discussed a "risk-associated factor" (Buell and Breslow 1960, 621, 624). For Buell and Breslow, risk factors mediate between variables and have a direct effect on dependent variables, such as the clinical outcome. This is quite close to current meanings.

In April 1961, "risk factor" appeared in the *Milbank Memorial Fund Quarterly*. As part of a discussion of the "proximal factors" that precipitate mental illness, psychiatrist Paul Lemkau spoke of a "genetically distributed risk factor" that was a possible cause of schizophrenia (Reid 1961, 248).

[10]A 1949 article had discussed a new disease introduced by the spread of commercial aviation, "aero-otitis media." The authors discussed risks and factors, but not risk factors: "The increasing volume of passenger traffic has resulted in a higher incidence of this condition. The precautions taken by military aviation personnel are not observed in commercial aviation. The risk of flying with infections of the upper respiratory tract is not realized by the average civilian passenger. In the seasoned flight passenger the press of business often occasions a disregard of this risk. This combination of factors is producing the higher incidence of acute aero-otitis media in commercial aviation" (Trowbridge 1949, 255).

Newspaper coverage of medicine and health also began to use the phrase in the late 1950s. A 1958 *Chicago Tribune* article, about an atoms-for-peace conference held in Geneva, noted that nuclear reactors in the Soviet Union had a 5/8 mile "safety zone" around them, in which no building or farming was permitted. Delegates hailed "proposals for licensing all the world's reactors to evaluate their 'risk factor'" (10). This usage is reminiscent of efforts in insurance and finance to manage risk. In 1960, the United Nations hosted an international meeting in Washington, D.C., of physicians and scientists to discuss the growing problem of burns. Howard Rusk, famous for developing rehabilitative medicine as a specialty, covered the meeting for the *New York Times*. He ascribed the rising incidence of serious burns to technological change: "this trend will be further accentuated as we come to rely even more on the products of nuclear fission, high-combustion fuels and generalized mechanization through technological progress. Unfortunately, as the risk factor of exposure to the hazards of burns has increased, our medical knowledge of prevention, treatment and effect of burns has not kept pace" (Rusk 1960, 78). This use of "risk factor" is ambiguous, referring either to risk per se, or expressing an association between new sources of heat exposure and the incidence of severe burns. A second article in 1960 used the phrase differently, as possibly causally associated. The *Washington Post* described a study of Seventh Day Adventist men, whose religion forbade cigarette smoking and whose rate of lung cancer was 90 times lower than other males. Because the study was conducted in Los Angeles (already infamous for its smog), "it thereby discounts air pollution as the great risk factor" for cancer (Haseltine 1960).

Overall, we have identified nine instances of "risk factor" in the medical literature and three in medical reporting in newspapers that predate the Framingham Heart Study.[11] We do not know if the FHS researchers were familiar with any them. Regardless, it is clear that the concept and the specific phrase, with varying meanings, were available and meaningful for medical researchers before July 1961.

Risk Factors After 1961

Despite prior claims by historians about the rapid uptake of the "risk factor" concept after 1961, the idea was actually slow to catch on, possibly because the concept was already familiar to some readers and had been slowing gathering steam over decades. This can be shown in several ways.

First, the term was slow to reappear in the journals where it had initially appeared. After its first appearance in the *Lancet* in 1922, it did not appear there

[11]In 1959, British social epidemiologist Alice Stewart presented a talk at the Society for Social Medicine entitled "Physique as a Risk Factor in Pregnancy," which was announced in the *British Journal of Preventive Social Medicine* 13 (1959): 162. We have not been able to locate a copy of her talk itself.

again until 1964, in a discussion of risk factors for heart disease. After its appearance in the *New England Journal of Medicine* in the 1951, it did not reappear until 1968 (twice) and 1969 (twice), and only one of those involved heart disease (in an article on atherosclerosis). After appearing in *BMJ* in 1922 and 1958, the term only recurred in 1964 (pregnancy and birth factors that are a risk for abnormal child development) and 1965 (seven times, including two articles about heart disease). Writing in *JAMA* in 1963 about "factors that contribute to disease," Ernest Wynder made extensive use of risk and factor, but not risk factors. *JAMA's* first clear mention of the risk factor concept came in a 1966 article about heart disease and the "risk-factor" status of patients.

Second, as the term "risk factor" began to appear in the public health literature, its use seemed unrelated to the FHS or heart disease. A 1962 article in the *American Journal of Public Health* reported that although the 1957 influenza pandemic doubled the risk of fetal malformations in Dublin, the "same risk factor" did not operate in New York (Hewitt 1962, 1688). A 1963 description of suicide rates in *Public Health Reports* suggested that "Another possible risk factor is the method used in the attempt," whether drowning, firearms, or jumping (Tuckman and Youngman 1963, 764). A discussion of tuberculosis eradication that year explained that "Infection with tuberculosis provides both a degree of risk and a degree of resistance, and when there is little infectious tuberculosis in a community, the risk factor becomes more important than the resistance factor" (CDC 1963, 508). As in the medical literature, some occurrences emphasize factors (such as suicide method), while other emphasize risk (as of tuberculosis).

Third, the incidence of the phrase can be traced through the PubMed database. Care must be taken here to avoid the flaws that undermine most analyses of PubMed time series. Since the PubMed corpus has grown over time (from 151,635 citations in 1965 to 869,999 citations in 2016), a count of the occurrences of any randomly distributed term will increase nearly six-fold over these decades. A subtler issue is that while early citations in PubMed might only contain the title, subsequent citations increasingly include the abstract and even the full text: this also produces counts that increase over time simply because a larger share of each article is being searched. The most conservative approach searches only the titles and then divides the count by the total number of PubMed citations in each year to report the incidence rate. The term remained rare until the mid-1970s (see Figure 1).[12]

The eventual popularity of risk factor may not have been driven entirely, perhaps even primarily, by an internal logic of medical and public health theory and practice. Instead, risk thinking increased broadly across American society in

[12]It is revealing, and relevant for our work, that "risk factor" does not appear in Raymond Williams's *Keywords: A Vocabulary of Culture and Society* (1983). Had the concept achieved the importance in the 1970s that is now has, it is likely that Williams would have examined it.

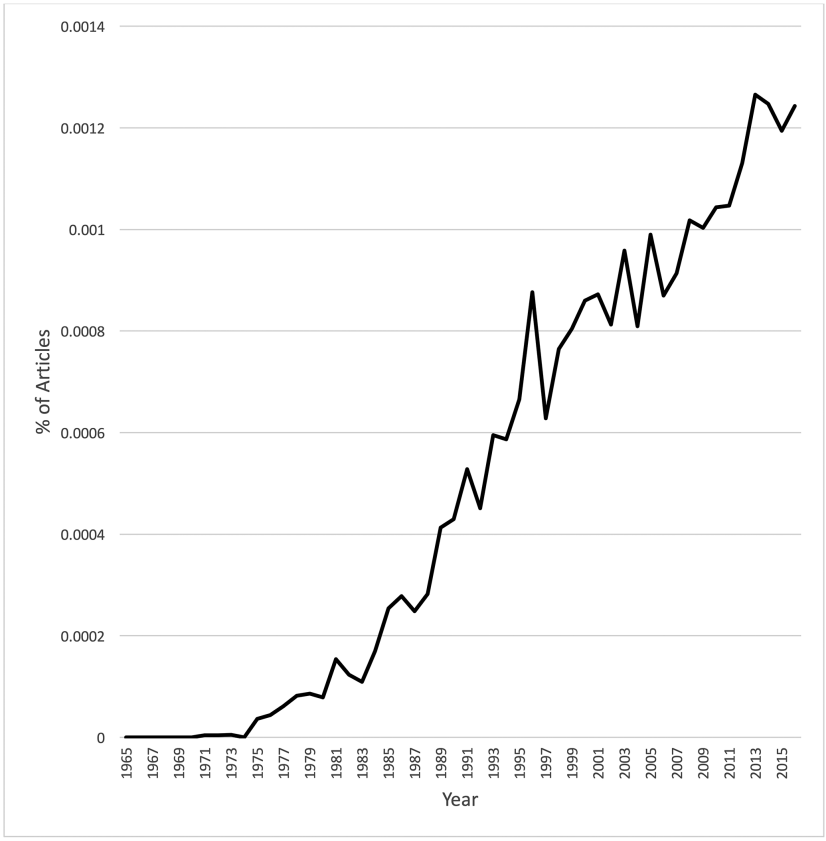

Figure 1

Percentage of articles in the PubMed corpus with "risk factor" in the title

the second half of the 20th century. The prevalence of *risk* in the Google Books corpus increases steeply and steadily after the early 1940s (perhaps triggered by Pearl Harbor and World War II), in parallel with a rise in *nuclear*. While more research would be required to ferret this out, it is plausible that Cold War American culture grew increasingly preoccupied with risk. The slow rise of "risk factor" in medicine after the 1960s may have had as much to do with broader cultural developments than with medical theory and practice, of which the FHS was but one, albeit significant, example.

Conclusion

The emerging tools of digital humanities allow a new kind of historical research. It is possible to scan quickly vast troves of primary sources. However, full-text searches of digital repositories cannot be fully automated: because of inaccurate

OCR or corrupt metadata, researchers need deliberate strategies for finding and assessing leads. Nonetheless, it is possible to search broadly in many areas, including medicine, public health, and further afield, for specific phrases and their possible antecedent concepts. We found nine instances of "risk factor" in medical and public health journals between 1922 and April 1961, along with a handful of occurrences of the phrase in medical journalism, a few occurrences of "factors of risk," and diverse occurrences of "risk factor" outside of medicine and public health. The much-acclaimed occurrence of "risk factors" in the FHS *Annals* article may have been a non-event in the history of the phrase. While we cannot yet present a decisive history, several important revisions to the conventional wisdom are clear, as are several avenues for future research.

First, antecedent concepts, whether of predisposition, constitution, environmental factors, or risk, have ancient roots. Extensive research would be required to map this genealogy. Such research, however, would require subjective assessments about whether any given author's language and conceptual apparatus is substantively (and usefully) similar to the modern concept. Second, the phrase "risk factor" circulated in many fields by the early 20th century, including insurance, finance, manufacturing, and agriculture. Third, it is unclear how the concept entered medicine, where it appeared in writing about occupational health in the 1920s and then in diverse settings in the 1950s. Fourth, important developments took place in surgery, where discussions of operative risk invoked recognizably modern concepts since the 1890s. More extensive research in the surgical literature could trace these developments. Fifth, new technology, whether space flight or nuclear power, triggered discussions of risk factors. Sixth, different authors emphasized different aspects of "risk factor," sometimes the risk itself (for example, of economic loss or surgical mortality), and other times the factor (for example, occupation, blood pressure). While FHS researchers emphasized factors, risk-based usages continued in the 1960s. Seventh, the phrase itself only slowly caught on in the 1960s. Extensive research would be required to map the networks of influences (citations) and meanings as "risk factor" became more widespread.

Finally, it would be valuable to explore how the concept and language evolved in languages other than English. For instance, did the risk discourse emerge differently in communist countries, in which industries of risk management (insurance, finance) played a less central role than they did in England and the United States? Each language will presumably have its own phrases that circulated distinctly or in tandem across health care, insurance, and other industries. We hope that our initial work on this topic will stimulate ongoing scholarship.

References

American Journal of Public Health. 1913. "Compulsory Sanitation." 3: 195.

Aronowitz, R. 2015. *Risky Medicine: Our Quest to Cure Fear and Uncertainty*. Chicago: University of Chicago Press.

Badham, C. 1926. "An Investigation Concerning the Incidence of Lead Poisoning in Motor-Car Painters." Abstracted by E. L. Collis in *Public Health Rep* 41: 2656.

The Banker. 1950. "Risk Factors in 'Foreign' Banking. I. Export Trade Risks." April 1.

Beck, U. 1992. *Risk Society: Towards a New Modernity*, trans. M. Ritter. London: Sage.

Biel, H. H. 1952. "Stock Analysis: Stock Values Reappraised." *Forbes*, Dec. 15.

Breslow, L., and R. Buechley. 1958. "Factors in Coronary Artery Disease: Cigarette Smoking and Exercise." *California Med* 89: 175–78.

Breslow, L., and P. Buell. 1960. "Mortality from Coronary Heart Disease and Physical Activity of Work in California." *J Chronic Dis* 11: 421–44.

Britten, R. H. 1930. "Occupational Mortality as Indicated in Life-Insurance Records for the Years 1915–1926." *Public Health Rep* 45:1250–59.

Buell, P., and L. Breslow. 1960. "Mortality from Coronary Heart Disease in California Men Who Work Long Hours." *J Chronic Dis* 11: 615–25.

Bureau of Agricultural Economics. 1935. *The Direct Marketing of Hogs*. Misc. Publn. 222. Washington, DC: U.S. Dept. of Agriculture.

Centers for Disease Control and Prevention (CDC). 1963. "Prospect of Tuberculosis Eradication." *Public Health Rep* 78: 507–9.

Chicago Tribune. 1958. "Red Nuclear Plants Have Safety Zones." Sept. 9.

Cianfarra, C. M. 1942. "Mexico Increases Strategic Output." *NY Times*, Sept. 24.

Cowen, R. C. 1961. "Grissom Ride Smooth, Safety Training Pays Off." *Christian Sci Monitor*, July 21.

Elliot, G. F. 1948. "Elliot Urges Rearmament of France." *Hartford Courant*, June 12.

Fisher, I. 1907. *The Rate of Interest: Its Nature, Determination, and Relation to Economic Phenomena*. New York: Macmillan.

Giroux, É. 2011. "The Origins of the Prospective Cohort Study: American Cardiovascular Epidemiology and the Framingham Heart Study." *Rev Hist Sci* 64: 297–318.

Google Books Ngram Viewer. 2017. https://books.google.com/ngrams.

Greene, J. A. 2007. *Prescribing by Numbers: Drugs and the Definition of Disease*. Baltimore: Johns Hopkins University Press.

Haedenkamp, K. 1938. "Medical Practice in Germany." *BMJ* 2: S161–S162.

Hamer, H.G. 1927. "Enlarged Pyelotomy for Stone." *Boston Med Surg J* 197: 819–23.

Haseltine, N. 1960. "Lung Cancer 90 Times Less Likely for Religious Non-Smoker, Is Report." *Washington Post*, June 6.

Heide, J. D. 1952. *Industrial Process Control by Statistical Methods*. New York: McGraw-Hill.

Hewitt, D. 1962. "A Study of Temporal Variations in the Risk of Fetal Malformation and Death." *Am J Public Health* 52: 1676–88.

International Monetary Fund (IMF). 1953. *Annual Report of the Executive Directors for the Fiscal Year Ended April 30, 1953*. Washington, DC: IMF. https://www.imf.org/external/pubs/ft/ar/archive/pdf/ar1953.pdf.

Journal of the American Medical Association (JAMA). 1892. "Idiosyncrasy." 18: 357.

Jauho, M. 2012. "The Mutual Shaping of Life Insurance and Medicine in Finland." *Soc Stud Sci* 45: 501–24.

Jones, D. S., and J. A. Greene. 2013. "The Decline and Rise of Coronary Heart Disease: Understanding Public Health Catastrophism." *Am J Public Health* 103: 1207–18.

Kannel, W. B., et al. 1961. "Factors of Risk in the Development of Coronary Heart Disease: Six-Year Follow-Up Experience." *Ann Intern Med* 55: 33–50.

Kerr, J. 1922. "The Treatment of Industrial Accidents." *BMJ* 2: 377–80.

Konecci, E. B., and R. Trapp. 1959. "Calculations of the Radio-biologic Risk Factors in Nuclear Powered Space Vehicles." *Aerospace Med* 30: 487–506.

Krauss, E. A. 1947. "Where Risk Factors Appear Greater." *Forbes,* Jan. 15.

Lancet. 1922. "Health Report of the Royal Air Force for 1920. II. Disease and Injury in Different Types of Units and Trades in the Royal Air Force During 1920." 199: 655–57.

Lemon, W. S., and H. J. Moersch. 1924. "Vital Capacity in Relation to Operative Risk." *Arch Intern Med* 33: 128–29.

Manchester Guardian. 1921. "Risk Factor in Interest Rate, Moneylender's Claim Fails." March 4.

Manchester Guardian. 1933. "Money & Stocks: Cheques and Trade—Insurance and Interest." Jan. 3.

Manchester Guardian. 1938. "Association Football: Transfer Fees—The Outlay and the Rewards." Aug. 9.

Marez, J. F. 1953. "Other Driver is Unknown Risk Factor." *Hartford Courant*, June 17.

Menzies, I. 1961. "Here's How 2nd Astronaut Shot Will Differ from 1st: Tuesday Set to Blast Off 'Liberty Bell.'" *Boston Globe,* July 16.

Moyer, C. A., and J. A. Key. 1956. "Estimation of Operative Risk in 1955." *JAMA* 160: 853–55.

Müller-Wille, S., and H.-J. Rheinberger. 2007. *Heredity Produced: At the Crossroads of Biology, Politics, and Culture, 1500–1870*. Cambridge: MIT Press.

A New Method of Curing the French-Pox Written by an Eminent French Author; Together with the Practice and Method of Monsieur Blanchard; as also Dr. Sydenham's Judgment on the Same; to which Is Added Annotations and Observations by William Salmon. 1690. Amsterdam: John Taylor and Thomas Newborough.

New York Herald Tribune. 1930. "A Transatlantic Service." Nov. 15.

New York Times. 1895. "Among the Thousand Islands." June 30.

Nichols, E. 1955. *Ruling Principles of Utility Regulations*. Washington, DC: Public Utilities Reports.

Oliven, J. F. 1951. "The Suicidal Risk: Its Diagnosis and Evaluation." *N Engl J Med* 245: 488–94.

Oppenheimer, G. M. 2006. "Profiling Risk: The Emergence of Coronary Heart Disease Epidemiology in the United States (1947–70)." *Int J Epidemiol* 35: 720–30.

Parascandola, M. 2011/12. "The Epidemiologic Transition and the Changing Concepts of Causation and Causal Inference." *Rev Hist Sci* 64: 243–62.

Pell, S., and C.A. D'Alonzo. 1958. "Myocardial Infarction in a One-Year Industrial Study." *JAMA* 166: 332–27.

Public Health Reports. 1896. "Spread of Plague in China." 11: 498.

Reid, D. D. 1961. "Precipitating Proximal Factors in the Occurrence of Mental Disorders: Epidemiological Evidence." *Milbank Q* 39: 229–58.

Rothstein, W. G. 2003. *Public Health and the Risk Factor: A History of an Uneven Medical Revolution*. Rochester: University of Rochester Press.

Rusk, H. A. 1960. "Burns: Medical Puzzle: Scientists and Clinicians of 17 Nations Meeting in Washington to Share Findings." *NY Times,* Sept. 25.

Schlich, T. 2013. "Railways, Industry, and Surgery: The Introduction of Risk Management," *N Engl J Med* 369: 1978–79.

Simpson, F. F. 1909. "Factors Which Contribute to a Reduction in Mortality in Abdominal Surgery." *JAMA* 53: 1173–79.

Stoner, J. D. 1949. *Water-Quality Indices for Specific Water Uses*. Geological Survey Circ. 770. Washington, DC: U.S. Department of the Interior.

Toohey, M. 1958. "Anticoagulants in Myocardial Infarction." *BMJ* 1: 252–55.

Trowbridge, B. C. 1949. "A New Treatment of Acute Aero-Otitis Media." *Arch Otolaryngol* 50: 255–63.

Tuckman, J., and W. F. Youngman. 1963. "Identifying Suicide Risk Groups among Attempted Suicides." *Public Health Rep* 78: 763–66.

Ware, J. 1861. "Lectures on General Therapeutics." *Boston Med Surg J* 64: 9–17.

White, P. D. 1957. "The Cardiologist Enlists the Epidemiologist." *Am J Public Health* 47 (suppl. 4): 1–3.

White, W. R. 1928. "The Investor: Economist Issues Tests of Foreign Government Securities." *Los Angeles Times,* Aug. 16.

Williams, R. 1983. *Keywords: A Vocabulary of Culture and Society*. Rev. ed. New York: Oxford University Press.

Winslow, C.-E. A., and C. C. Grove. 1922. "Note on Certain Correlation Factors of the 1918 Influenza Epidemic." *Am J Hygiene* 2: 240–45.

Wright, A. E., and D. Semple. 1897. "Remarks on Vaccination Against Typhoid Fever." *BMJ* 1: 256–59.

Wynder, E. L. 1963. The Physician, the Patient, and Statistics." *JAMA* 186: 1150.