George W. Comstock

School of Hygiene and Public Health,
Johns Hopkins University, Hagerstown, MD

Cohort analysis: W.H. Frost's contributions to the epidemiology of tuberculosis and chronic disease

Summary

Although Wade Hampton Frost was not the first to develop cohort analysis, it was the post-humous publication of his study of age and time trends of tuberculosis mortality that directed attention to this method of analysis. Frost's developing interest in and contributions to the epidemiology of chronic disease are reviewed in connection with a summary of his professional career.

Although Wade Hampton Frost's bibliography contains 57 scientific publications, only four of them deal with tuberculosis [1]. This small proportion hardly suggests a major interest in or influence on the field of tuberculosis. However, a review of his biography shows that these four papers represent the culmination of a teaching career aimed at developing theoretical bases for epidemiology and expanding its limits.

Frost was born on the third of March, 1880, in rural Virginia, the son of a country doctor [1]. His pre-college education was obtained at home under the tutelage of his mother except for a final two years in boarding schools. He entered the University of Virginia, and was granted his medical degree in 1903. Following the lead of many graduates of that institution, he sought a post in one of the Uniformed Services. In 1905, he was appointed Assistant Surgeon in the Public Health and Marine Hospital Service (now the Public Health Service). He was first assigned to one of the medical care facilities of the Service, as was customary for newly commissioned officers. With the exception of a temporary assignment to help investigate a yellow fever epidemic in New Orleans, and another to examine immigrants at Ellis Island, he spent several years in Baltimore, first at the U.S. Marine Hospital and then at the Training School for the U.S. Revenue Cutter Service (now the U.S. Coast Guard). Two summers were spent on extensive cruises with the cadets, first to the New England area and then to Europe and North Africa.

In 1908, the arcane workings of the Service fortuitously resulted in an assignment that proved to be extremely fortunate for Frost, for the Service, and for public health. His new post was the Hygienic Laboratory (the forerunner of the National Institutes of Health) where he associated with some of the best investigators of the time. Maxcy, Frost's successor at Johns Hopkins, says Frost emerged from four

Soz.- Präventivmed. 46 (2001) 007–012
0303-8408/01/010007-06 $ 1.50 + 0.20/0
© Birkhäuser Verlag, Basel, 2001

years' service in this environment "a trained and highly competent investigator in epidemiology"[1]. He was then assigned to posts that gave him opportunities to apply his new skills to studies of the health problems related to stream pollution, to poliomyelitis, and to influenza. This last assignment brought him into a life-long collaboration with Edgar Sydenstricker, the first national public health statistician[2]. Together they struggled with the problems of analysing morbidity surveys and family studies. Their solutions laid the groundwork for Frost's later work with tuberculosis.

In 1919, there came another major turning point in Frost's career. William H. Welch, the first director of the newly established School of Hygiene and Public Health of the Johns Hopkins University was able to obtain Frost's assignment to the School as Resident Lecturer in the Department of Epidemiology and Public Health Administration. In 1922, he was appointed Professor and Chairman of the Department. Three years later, in 1925, Epidemiology and Public Health Administration became separate departments, Frost remaining with Epidemiology as Professor and Chairman. During his first decade at the School, Frost was largely occupied in defining the field of epidemiology[3,4], developing appropriate courses of instruction, and working with his students on problems related to diphtheria, the common cold, and other acute infectious diseases.

During the 10 years from 1928 to 1938, the last decade of his life, he began to turn his attention to diseases of longer duration. What led him to expand his interests cannot be ascertained, but it seems reasonable to assume that Frost's inquisitive and logical mind would eventually lead to speculation about applying epidemiologic methods to chronic diseases. That such a transition did occur can be documented from his published and unpublished writings[5]. Elizabeth Fee, in her history of The Johns Hopkins University School of Hygiene and Public Health, has three illustrative citations[5]. In 1919, Frost's definition of epidemiology was "the natural history of the infectious diseases, with special reference to the circumstances and conditions which determine their occurrence in nature"[5]. This was broadened but still restrictive in 1927: "It is (…) good usage to speak of the epidemiology of tuberculosis; (…) and also to apply the term to the mass-phenomena of such noninfectious diseases such as scurvy, but not to those of the so-called constitutional diseases, such as arteriosclerosis and nephritis"[5]. By 1937, he referred to "epidemiology as comprising the whole of the unremitting effort being made to clarify the relation between the disease and disabilities which men suffer and their way of life"[5].

Frost also left no indication of why he selected tuberculosis as the principal springboard for his venture into the field of chronic diseases. Maxcy surmises that there were several reasons[1]. First, tuberculosis was then a major cause of death and disability. It was the sixth leading cause of death in 1930[6], accounting for 6.3 percent of all deaths in the United States[7]. Frost also had a personal reason for appreciating its importance. In 1917, he had been diagnosed as having incipient pulmonary tuberculosis and had spent several months in a sanatorium[1]. Furthermore, in making a break with traditional epidemiology, it is not unreasonable to concentrate on a single disease, especially one like tuberculosis that was both infectious and chronic. Ruth Puffer, one of Frost's major collaborators in his tuberculosis studies, suggests that he initially selected tuberculosis and rheumatic fever in order to compare the familial characteristics of these two chronic and infectious diseases[8]. Her suggestion is supported by the doctoral theses and publications of his students and colleagues[1]. Two dealt with families of rheumatic fever patients and far more, 18 in number, had tuberculosis as their subject (1, and theses on file in Department of Epidemiology, Johns Hopkins School of Hygiene and Public Health).

It is likely that the emphasis on tuberculosis was influenced by the accessibility of new and challenging data for analysis. Dr. E.L. Bishop, Commissioner of Health for the State of Tennessee, was concerned about the high tuberculosis mortality rates in his state, and puzzled by the unusual concentration of tuberculosis deaths among elderly persons and rural residents[9]. With the help of the Rosenwald Fund, he began to study tuberculosis in the little town of Trenton. When the type of analyses employed in investigations of acute infectious diseases did not seem to make sense, Bishop sought the help of his former teacher, Frost. After reviewing Bishop's findings, Frost identified some deficiencies and set down a detailed plan for a subsequent investigation[9]. Requisites of the new study included its being based on an unselected series of tuberculosis cases reported to the state from a single rural county of about 25000 persons. Detailed histories were to be taken of the cases and their families, and the subsequent incidence of tuberculosis. Comparisons were to be made with families without tuberculosis. In addition, Frost recommended tuberculin skin tests of the families, and for comparison, similar testing of children in the community schools.

Before embarking on such an ambitious investigation, Frost recommended a pilot survey in Kingsport. This was supported by the Rockefeller Foundation, and included virtually all the black families in the town[10]. Rather than setting out to follow these families for the many years required by the long incubation periods commonly seen in tuberculosis, Frost suggested a retrospective approach by obtaining "simple

Soz.- Präventivmed. 46 (2001) 007–012
© Birkhäuser Verlag, Basel, 2001

facts (i.e. deaths and tuberculosis cases) as lie within the knowledge and memory of the average householder" [10]. To analyze the data, Frost made adaptations to the life table methods used by Elderton and Perry in their prospective studies of the fate of persons discharged from tuberculosis sanatoria [11] and by Weinberg in his study of the fate of children with tuberculous parents [12]. Surprisingly, Frost does not cite the earlier work by Lawrason Brown, a clinician at the Adirondack Cottage Sanitarium in Saranac Lake, New York, and E.G. Pope, an actuary and patient at the Sanitarium [13]. Frost's addition to these early applications of survival analysis was to recognise that the techniques could be applied to historical data and to show how to do this [10]. His approach was subsequently used in studies of the risk of tuberculosis among families of tuberculosis patients in Philadelphia by Persis Putnam, a biostatistics graduate of the Johns Hopkins School of Hygiene and Public Health [14]; in Cattaraugus County, New York by Jean Downes, a colleague of Edgar Sydenstricker [15]; and in Williamson County, Tennessee by Ruth Puffer, one of Frost's students [8,9].

The Williamson County Study was the major field study of tuberculosis in the United States during the 1930s and 1940s [9]. Its design was based on the prior experience in Trenton and Kingsport, and was laid out in considerable detail in a 1931 memorandum from Frost [9]. In addition to the use of historical survival analysis, a major feature was an adaptation of Chapin's secondary attack rate [16]. In a chronic infectious disease like tuberculosis, there could be multiple cases within a single family. In such instances, it was often impossible to differentiate "primary" and "secondary" cases. To avoid this uncertainty, Frost suggested that the first case to be identified be called the "index case". Like primary cases in acute infections, index cases were to be excluded from calculations of risk associated with living in a tuberculosis family. Although not a perfect substitute for the primary case, the index cases could be clearly defined, and the attack rate among the other members of the household could be treated like a secondary attack rate.

Frost's first paper dealing with tuberculosis was basically an exposition of how to apply survival analysis to historical data, using the findings from the Kingsport survey mainly for illustrative purposes [10]. The risk of dying from tuberculosis was found to be twice as high among families exposed to a tuberculous member than among families not so exposed, a finding very similar to those of Weinberg in an earlier prospective study in Germany [12]. The Kingsport data also showed that the highest case rates occurred among young children, adolescents and young adults, and the elderly. This pilot study served its purpose. It allowed Frost and his students to develop methods for handling historical data and suggested some interesting findings to be investigated further in the Williamson County Study. Unfortunately, definitive results from the Williamson County Study would not be available until after Frost's death in 1938.

Perhaps the most widely cited of Frost's four papers on tuberculosis was the one on cohort analysis [17]. This was found in his desk drawer and published after his death. The material had been presented to the Southern Branch of the American Public Health Association and apparently laid aside to await additional data and analyses [1]. To illustrate cohort analysis, Frost first arranged tuberculosis mortality rates from Massachusetts, supplied by his friend Sydenstricker, in a table with age on one axis and year of death on the other (Tab. 1). Arranged in this way, one could quickly see the age-specific mortality for each of the available years on one axis, and the time trend for each age group on the other. What proved to be most interesting in this instance

Age	1880	1890	1900	1910	1920	1930
Males						
0–4	1760	578	309	209	108	41
5–9	143	49	31	21	24	11
10–19	126	115	90	63	49	21
20–29	444	361	288	207	149	81
30–39	378	368	296	253	164	115
40–49	364	336	253	253	175	118
50–59	366	325	267	252	171	127
60–69	475	340	304	246	172	95
70+	672	396	343	163	127	95
Females						
0–4	1658	595	354	162	101	27
5–9	171	82	49	45	24	13
10–19	265	213	145	92	78	37
20–29	537	393	290	207	167	92
30–39	422	372	260	189	135	73
40–49	307	307	211	153	108	53
50–59	334	234	173	130	83	47
60–69	434	295	172	118	83	56
70+	584	375	296	126	68	40

* They were obtained as follows: For the years 1910, 1920 and 1930 – based on U.S. Mortality Statistics – deaths from tuberculosis, all forms. For the years 1880, 1890 and 1900, the rates used are calculated from data compiled by the late Dr. Edgar Sydenstricker from the state records. Because of differences of classification in deaths, it has been necessary to base the rates on the deaths recorded as "tuberculosis of the lungs" to get comparable data for these years. The rate calculated from the state records for "tuberculosis of the lungs" has been multiplied by a factor based on the proportion such deaths bore to those from tuberculosis, all forms. This factor varied with the year and age considered.
(Permission has been granted to reproduce Table 1. The original appeared in 17.

Table 1 Deaths rates* per 100,000 from tuberculosis, all forms, for Massachusetts, 1880 to 1930, by age and sex,with rates for cohort of 1880 indicated

© Birkhäuser Verlag, Basel, 2001

were the rates in the cells of the table that lay on the diagonals, starting with the youngest ages and earliest years. These "diagonal rates" were analogous to tuberculosis mortality rates experienced by a group of persons born in a specified time period, the "cohort". They represented the tuberculosis death rates that were those experienced by each cohort of persons as they simultaneously aged and passed through time. One had to assume, of course, that immigrants and emigrants were generally similar in their risks of dying from tuberculosis.

As he studied the diagonal "cohort" curves, Frost first noted that in every cohort, the highest rates occurred among infants and very young children, and also among young adults. After this second peak, the tuberculosis death rates tended to decrease with age. The latter finding was in marked contrast to the death rates in any specified year in which the highest rates occurred among older persons.

Frost also noted that the pattern of high rates in infancy, lower rates among children, high rates again in young adult life, and then falling with increasing age was similar for each cohort, although rates at every age became consistently lower in more recently born cohorts. Frost concluded that the present high rates in old age must be "residuals of higher rates in earlier life". If the cause of the decreasing rates was that the frequency of exposure to tuberculosis had become progressively less, the similar age patterns for each cohort gave no indication that postponement of infection to later years caused more serious disease as Frost had once feared, based on his experience with acute infectious diseases[18]. Thus, a theoretical objection to postponing infections with *Mycobacterium tuberculosis* was rebutted by this finding.

The findings in Frost's cohort analysis of tuberculosis death rates now appear to be the resultant of two different trends: a decreasing risk of becoming infected with time and a consistent pattern of change with age. His findings were in marked contrast to those of Andvord who had published cohort analyses of tuberculosis death rates in Norway six years prior to Frost[19]. Andvord's cohort curves were virtually identical with the age specific curves for each year during his observation period[20], presumably because tuberculosis infection rates were not changing appreciably during the early years of this century in Norway. In retrospect, it is difficult to see why Andvord felt that rates for cohorts offered any advantage over age-specific death rates.

Although Frost referred to Andvord's paper, one wonders if Frost did not develop the idea on his own, only later finding Andvord's work. This speculation is supported by the similarity of Andvord's age-specific and cohort curves. If Frost's introduction to cohort analysis had come from Andvord's work, it is hard to see why he would consider cohort data an improvement over the more readily available age-specific curve for a recent year. His correspondence with Sydenstricker further strengthens the assumption that he arrived at cohort analysis independently[1]. In it, he gives no hint of the idea having been suggested to him by anyone else. That Andvord's report might have served to raise doubts about the generalisability of cohort analysis is indicated by a comment to Sydenstricker that in spite of having obtained similar results with data from England and Wales and also the United States, Frost still wanted to "get together material for a somewhat more orderly study later"[1].

Frost's other two papers on tuberculosis had essentially the same theme, namely what would be needed to control the disease and the probability of eventual eradication[18–21]. He emphasised that the principal goal was the avoidance of infection. His contemporaries must have considered it nearly impossible to avoid tuberculous infection at a time when almost all adults in much of the country reacted to the strong dose of tuberculin, a reaction then considered indicative of tuberculosis infection[22]. Now, reactions to a strong dose of tuberculin are considered to be almost entirely the result of nontuberculous mycobacterial infections[23]. But even though Frost could not have known this, he did know that close and prolonged contacts with infectious cases were more likely to result in tuberculous disease than casual contacts. He therefore emphasised the desirability of reducing the dose of infection by isolating as many infectious cases as possible in sanatoria. Another feature of tuberculosis that he felt favored its human hosts was the fact that most infections with tubercle bacilli were rather quickly walled off. These organisms could only become infectious if they were located in the lung where some, as a result of some uncommon circumstance, were able to erode through the encapsulating tissue and escape into an airway. Finally, he noted that the steady fall in tuberculosis death rates in Western countries was evidence that each case of tuberculosis was, on average, giving rise to less than one new case. Thus, "the biological balance is against the survival of the tubercle bacillus", and "the eventual eradication of tuberculosis requires only that the present balance against it be maintained"[18]. Looking forward to a future when tuberculosis might become uncommon, Frost anticipated the modern disease control technique of concentrating more and more on infectious cases, advising "that the protection thrown around these infectious cases and their immediate contacts be not relaxed, but steadily and progressively increased"[18].

To Frost's credit, it should be pointed out that he realised that isolation of cases and surveillance of contacts could impose major hardships on patients and their families. "If we are to require the isolation of open tuberculosis as a matter

© Birkhäuser Verlag, Basel, 2001

of public protection, it becomes a public responsibility to bear not only the cost of medical care, but the whole cost to the patient's family, or as large a share as may be required. Moreover, it should be recognised that what is needed is not bare maintenance on a minimum or average 'relief' standard, that it is not sufficient merely to prevent their dropping lower in the economic scale; it may often be necessary to raise them to a higher level"[18]. Even with less frequent and shorter durations of hospital care at the present time, Frost's words still need to be heard and heeded.

It is hard to assess Frost's influence on the field of tuberculosis. A search through a nonrandom selection of books on tuberculosis reveals only scattered brief references to the age selection of tuberculosis and to his statement about the biologic balance being against the tubercle bacillus. Many epidemiologic textbooks, if they mention Frost at all, do so only in relation to cohort analysis. In my opinion, Frost's published work has had very little effect on phthisiologists and only a slight effect on epidemiologists. His influence appears to have been on his students, a high proportion of whom rose to high positions in public health. Through them, he had some influence on tuberculosis but very much more on epidemiologic thinking in all fields of public health. Frost was interested in tuberculosis as a means to an end, an area in which to develop methods for understanding and eventually controlling all the ills of mankind. He and his students gave us all a good start.

Zusammenfassung

Kohortenanalyse: W.H. Frosts Beiträge zur Epidemiologie von Tuberkulose und anderer chronischer Krankheiten

Obwohl Wade Hampton Frost nicht der erste war, der die Kohortenanalyse entwickelte, war es die posthume Veröffentlichung seiner Studie zu Alter und Zeittrends von Tuberkulose-Mortalität, die die Aufmerksamkeit für diese Analysemethode auf sich zog. Frosts fortschreitendes Interesse an und seine Beiträge zur Epidemiologie von chronischen Krankheiten werden in diesem Artikel zusammen mit einem Überblick über seine berufliche Karriere nachgezeichnet.

Résumé

Analyse de cohorte: la contribution de W.H. Frost a l'épidémiologie de la tuberculose et des maladies chroniques

Bien que Wade Hampton Frost ne fut pas l'inventeur de l'analyse de cohorte, c'est la publication posthume de son étude sur les tendances temporelle et selon l'âge des décès par tuberculose qui a attiré l'attention sur ce type d'analyse. Cet article montre le lien entre la carrière professionnelle de Frost, son intérêt croissant et sa contribution a l'épidémiologie des maladies chroniques.

Soz.- Präventivmed. 46 (2001) 007–012
© Birkhäuser Verlag, Basel, 2001

References

1 *Maxcy KF,* ed. Papers of Wade Hampton Frost, M.D. A contribution to epidemiological method. New York: The Commonwealth fund, 1941. (This is the major source of material for this article).

2 *Kasius RV*, ed. The challenge of facts. Selected public health papers of Edgar Sydenstricker. New York: Prodist, 1974.

3 *Frost WH*. The importance of epidemiology as a function of health departments. Medical Officer (London), 1923; *23*: 113–4.

4 *Frost WH*. Epidemiology. Nelson Loose-Leaf System, Public Health – Preventive Medicine. Volume 2, Chapter 7. New York: Thomas Nelson & Sons, 1927. [Reprinted in *(1).*]

5 *Fee E*. Disease and discovery. A history of the Johns Hopkins School of Hygiene and Public Health, 1916–1939. Baltimore: The Johns Hopkins Press, 1987.

6 *Linder FE, Grove RD*. Vital statistics rates in the United States, 1900–1940. Washington: U.S. Government Printing Office, 1947.

7 U.S. Public Health Service, Division of Public Health Methods and Tuberculosis Control Division. Tuberculosis in the United States, graphic presentation. Vol 2: Proportionate mortality statistics for states and geographic divisions by age, sex and race. Washington: Medical Research Committee, National Tuberculosis Association, 1944.

8 *Puffer RR*. Familial susceptibility to tuberculosis. Cambridge, MA: Harvard University Press, 1946.

9 *Zeidberg LD, Gass RG, Dillon A, Hutcheson RH*. The Williamson County Study. A twenty-four-year epidemiologic study. Am Rev Respir Dis 1963; *87*(No. 3, Part 2): 1–88.

10 *Frost WH*. Risk of persons in familial contact with pulmonary tuberculosis. Am J Public Health 1933; *23:* 426–32.

11 *Elderton WP, Perry SJ*. Drapers' Company Research Memoirs. Studies in National Deterioration. VI: A third study of the statistics of pulmonary tuberculosis. The mortality of the tuberculous and sanatorium treatment. London: Cambridge University Press, 1910.

12 *Weinberg W*. Die Kinder der Tuberkulosen. Leipzig: Hirzel, 1913.

13 *Brown L, Pope EG*. The postdischarge mortality among the patients of the Adirondack Cottage Sanitarium. Am Med 1904; *8:* 879–82.

14 *Putnam P*. Tuberculosis incidence among white persons and negroes following exposure to the disease. Am J Hyg 1936; *24:* 536–51.

15 *Downes J*. A study of the risk of attack among contacts in tuberculous families in a rural area. Am J Hyg 1935; *22:* 731–42.

16 *Cassedy JH*. Charles V. Chapin and the Public Health Movement. Cambridge, MA: Harvard University Press, 1962.

17 *Frost WH*. The age selection of mortality from tuberculosis in successive decades. Am J Hyg 1939; *30:* 91–6. (Reprinted in Am J Epidemiol 1995; *141:* 4–9).

18 *Frost WH*. How much control of tuberculosis? Am J Public Health 1937; *27:* 759–66.

19 *Andvord KF*. Hvad kan vi laere ved a folge tuberkulosens gang fra generasjon til generasjon? Norsk Magasin Laegevidenskapen 1930; *91:* 642–60. (Engl. transl. by Gerard Wijsmuller).

20 *Comstock GW*. Early studies of tuberculosis. National Cancer Institute Monograph 67. Selection, follow-up, and analysis in prospective studies: a workshop. Washington: U.S. Government Printing Office, 1985. (NIH Publication No. 85-2713).

21 *Frost WH*. The outlook for the eradication of tuberculosis. Am Rev Tuberc 1935; *32:* 644–50.

22 National Tuberculosis Association, Committee on Diagnostic Standards. Diagnostic standards and classification of tuberculosis. 1940 edition. New York: National Tuberculosis Association, 1940.

23 American Thoracic Society, Committee on Revision of Diagnostic Standards. Diagnostic standards and classification of tuberculosis. New York: National Tuberculosis and Respiratory Disease Association, 1969.

Address for correspondence

George W. Comstock, MD, DrPH, FFPHM (Hon)
The Johns Hopkins University
School of Hygiene and Public Health
Washington County Health Department
Box 2967
Hagerstown, MD 21742-2067
USA

Tel.: 001-301-791-3230
Fax: 001-301-797-3669

Soz.- Präventivmed. 46 (2001) 007–012
© Birkhäuser Verlag, Basel, 2001