Wilhelm Weinberg, German physician, founder, and president of the Stuttgart Society for Racial Hygiene, published in 1913 the results of a large, retrospective cohort study entitled *Die Kinder der Tuberkulosen* (Children of the Tuberculous). The exposed cohort comprised 18,212 children whose 3,246 fathers and 2,022 mothers died of tuberculosis between 1873 and 1902. The unexposed cohort comprised 7,574 children whose 1,830 parents died of causes other than tuberculosis in 1876, 1879, or 1886. He found that children of tuberculous parents had higher mortality rates and lower fertility than did children of nontuberculous parents. Because of its size, rigorous design, and meticulous analysis, Weinberg’s cohort study stands as one of the major epidemiologic works carried out before 1945.

cohort studies; epidemiologic studies; history of medicine; research design; retrospective studies; tuberculosis
social than to hereditary criteria in the determinants of tuberculosis. Weinberg died poor in 1937, 4 years after Hitler came to power, and does not seem to have adhered to the Nazi ideology (3).

Even though the eugenic context is puzzling and important, the main objective of this paper is to highlight and discuss the place of Weinberg’s work in the history of epidemiologic methods, in particular, the history of cohort studies. This is the largest known retrospective cohort study conducted before the Second World War. This book has been previously cited in the epidemiology literature (4–7), but we have good reason to believe that it has been studied haphazardly and that the significance and innovative aspects of this work have not been appropriately valued. Die Kinder der Tuberkuloesen is difficult to read and understand, and apparently few people before us felt stimulated to take on the challenge. We provide therefore detailed information on the study design and results. We also have prepared an English summary of the book, chapter by chapter, to facilitate access to it by other scholars. The summary can be downloaded from www.epidemiology.ch/index3.htm.

**MATERIALS AND METHODS**

**Weinberg’s study**

The study was planned to determine whether 1) children of tuberculous parents had a higher mortality from all causes and from tuberculosis; 2) the age of the parents at death, number of siblings, age of the children at the death of their parents, or order of sibship influenced the children's mortality; and 3) tuberculous persons had a higher fertility.

The study identified two cohorts of children born several decades before Weinberg started collecting the data. Their parents had died of either tuberculosis (we will refer to them as the “exposed cohort”) or causes other than tuberculosis (i.e., the “unexposed cohort”). Weinberg compared their risks of death up to age 20 years. The exposed cohort comprised 18,212 children whose 3,246 fathers and 2,022 mothers died of tuberculosis between 1873 and 1902, representing 201,475.5 person-years of follow-up. The unexposed cohort comprised 7,574 children whose 1,830 parents died of causes other than tuberculosis in 1876, 1879, or 1886. Weinberg considered the deaths during these 3 years as “representative” of all deaths between 1873 and 1889. Weinberg tracked all selected children until 1909. Thus, all children whose parents had died before 1890 were followed up at least until their 20th birthday. Children whose parents died after 1889 could have been observed for a shorter time. The outcomes were age at death and cause of death. Weinberg first tabulated the number of child deaths by year and age at death of the parent and size of sibship (4, pp. 12–14).

To manage the large amount of information gathered on about 7,098 parents and 25,786 children without automated technology, Weinberg produced a family card for each person who died of tuberculosis including name, dates of birth and death, and data on the marriages and children (4, p. 13). Children’s data comprised the date of birth, emigration, or death and the cause of death. These cards were summarized in 60 basic tables (“Grundtabellen”), one for each year of death and sex of parents, including all relevant information to perform his analyses: year of birth of all children, age of children upon the death of parents, deaths and emigrations of all children with corresponding ages, and survival time of children. In modern terminology, this is a retrospective (or historical) cohort study (7).

Computations based on the basic tables yielded 26 source tables (“Urtabellen”), which mainly contained information on deaths, emigrations, and population sizes at different points of time. Weinberg created specific tables for looking into the effect on children’s mortality of the age of tuberculous parents, number of siblings, and so on. For creation of the source tables, he sometimes had to conduct one or even more intermediate steps, presented as “support tables” (“Hilfstabellen,” an example of which is shown in Web figure 1). This example, referred to as “Web figure 1,” is posted on the Journal’s website (http://aje.oxfordjournals.org/).

**Life table analysis**

Weinberg understood and applied life-table techniques to estimate the risk of death from birth to age 20 years. He tabulated the different components of the life-table calculation separately for children of mothers and fathers of different tuberculosis status. Children whose parents both had tuberculosis appeared twice.

Weinberg first tabulated the number of child deaths by year and age at death of the parent and size of sibship (4, pp. 40, 41, 44). In another set of source tables, he summarized the deaths, losses to follow-up, and person-years. Finally, using these numbers, he computed logarithms of the probabilities of surviving (4, pp. 74–75), which he summed to get the logarithm of the cumulated survival up to age 20 years. The cumulative mortality risk was 1 minus the cumulative survival up to age 20 years.

In order to reanalyze the data using a survival analysis/life table approach, we transformed the categorical data into
a new data set comprising one record per observation and per time interval and then attributed the outcomes (dead, alive, or lost to follow-up) proportionally to these records. Table 1 gives the number of children we worked with (4, pp. 53, 59, 62, 68–73). For the children of nontuberculous parents, we used Weinberg’s table 5 (4, p. 62). In Weinberg’s analysis, deaths were counted as if they had all occurred at the end of the 1-year interval, whereas withdrawals were counted as if they had occurred in the middle of the time interval. All time intervals had a fixed, 1-year length. We only knew if a person died within the interval rather than the exact time of death. We therefore gave the time 0.1 to the perinatal deaths, 0.5 to the deaths during the first year of life, 1.5 to the second year of life, and so on. In doing so, we treated the deaths as if they had occurred in the middle of the time interval. We counted children lost to follow-up as 0.5 person-year in the survival rates; for example, a child who emigrated in its fifth year of life got 4.5 observed life-years and “not dead” as an outcome.

As Weinberg, we used an actuarial life-table approach and treated the children as if they were all born in the same year. Therefore, the x axis on the life table is not the birth year but the age of children.

RESULTS

Total mortality risk and rates

Table 2 shows that the 20-year risk of death and average yearly mortality rates were larger for children of tuberculous parents in general and slightly larger for children of tuberculous mothers. Weinberg did not compute risk differences or risk ratios.

As indicated in table 3, the child mortality risk was lower for parents who died at an older age in both the exposed and the unexposed cohorts. Weinberg noted, therefore, that the younger age at death of tuberculous parents compared with nontuberculous parents explained part of the excess mortality associated with tuberculosis (4, p. 67). Still, as table 3 shows, the mortality of children of tuberculous parents was, for almost all ages at the parent’s death, higher compared with the mortality of children of nontuberculous parents.

Table 4 shows the lower fertility of tuberculous parents, which persisted when the analyses were stratified by the age of death of the parents. However, because most of the differences disappeared after age standardization, the fertility deficit appeared to be due to shorter survival and less opportunity to reproduce by the tuberculous parents.

<table>
<thead>
<tr>
<th>TABLE 1. Numbers of children, deaths, and losses by characteristics of the parent, Baden-Württemberg, Germany, 1873–1909*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exposure</td>
</tr>
<tr>
<td>-----------</td>
</tr>
<tr>
<td>Nontuberculous</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Tuberculous</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>All</td>
</tr>
</tbody>
</table>

* Source: Die Kinder der Tuberkuloßen (Leipzig, Germany: S. Hirzel, 1913) (4).

<table>
<thead>
<tr>
<th>TABLE 2. Mortality risks and rates and life expectancy, by gender and cause of death, Baden-Württemberg, Germany, 1873–1909*,†</th>
</tr>
</thead>
<tbody>
<tr>
<td>Parent</td>
</tr>
<tr>
<td>-----------------</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Father</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Mother</td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>

* Source: Die Kinder der Tuberkuloßen (Leipzig, Germany: S. Hirzel, 1913) (4).
† Exposed cohort: children of tuberculous parents who died between 1873 and 1889; unexposed cohort: children of parents who died of other causes in 1876, 1879, or 1886.
‡ Average life expectancy up to age 20 years + ((1 – risk of dying during the first 20 years of life) × 40 years), assuming a further life duration of 40 years for those aged 20 years.
Additional findings

In chapter 14, Weinberg speculated that there should be no differences in mortality risk between two siblings if the “Elbzat” or genetic constitution explained the excess mortality of children of tuberculous parents. He found, however, as shown in table 5 for the first and second halves of sibship order, that first-born children had a lower mortality than did their siblings. Actually, the higher (i.e., first, second, and so on) in the sibling order of birth, the higher the mortality. Pearson’s correlation coefficient between birth order and survival was stronger for tuberculous than for nontuberculous parents (4, p. 116).

Weinberg also found that the children of tuberculous parents had a higher mortality if they were younger at the time of death of their parents (4, p. 80) and if they belonged to a larger sibship (4, p. 97).

In chapter 15, Weinberg considered the effect of social class defined by profession, origin, and income. Figure 1 shows a clear inverse trend of social class and mortality before age 20 years. Weinberg found a higher mortality risk among innkeepers and wine growers, both professions related to alcohol, than among workers such as masons and stonemasons (4, pp. 117–118).

Standardized mortality ratios

In addition to the life-table analysis, Weinberg compared the observed and expected child deaths of tuberculous parents, which he referred to as the “percentage of the expected” and which we would call today “standardized mortality ratio.” The expected child deaths were computed by use of information on age-specific death rates from the total population (from the annual abstract of the Statistical State Agency of Württemberg) and the total number of exposed children (4, pp. 122–123 and 129).

These standardized mortality ratios (observed/expected child deaths) standardized on age were, for the overall
mortality, 118 for children of tuberculous fathers and 133 for children of tuberculous mothers (excluding stillbirths: 122 and 138 for children of, respectively, tuberculous fathers and mothers); the corresponding standardized mortality ratios for mortality from tuberculosis were 198 when fathers were tuberculous and 246 when it afflicted the mothers.

Reanalysis of Weinberg’s data

Figure 2 depicts the survival curve of our reanalysis of Weinberg’s data. The y axis is truncated to magnify the observed differences. About 30 percent of the children of all types of parents died within the first year of life. After that, children of tuberculous parents had a weaker survival. In addition, the older the parents were at death, the lower the mortality of their children (figure 3).

The cumulative mortality (1 – survival) up to age 20 years was 48.1 percent, 46.7 percent, 41.5 percent, and 40.0 percent for children of, respectively, tuberculous mothers, tuberculous fathers, nontuberculous mothers, and nontuberculous fathers. These mortality risks were virtually identical to those shown in table 2. More generally, our reanalysis confirmed all mortality risks shown in table 3.

DISCUSSION

Conclusions drawn by Weinberg

Weinberg drew the following conclusions. 1) During the first year of life, children of tuberculous parents had a higher overall mortality compared with children of nontuberculous parents, but afterwards their excess mortality was attributable mainly to tuberculosis. 2) Children of tuberculous parents had shorter lives if they were younger siblings, they had many siblings, or their parents died at a younger age or were from a lower social class. 3) Tuberculous parents were less fertile than nontuberculous people. Part of the difference was due to tuberculous parents’ deaths at a younger age and therefore having less time to reproduce.

He therefore inferred that the fragility of tuberculous persons was more environmental than genetic in nature and that tuberculous persons did not represent a threat for the human race:

This study shows the strong social dependence of tuberculosis, its association with living with a tuberculous relative, with more clarity than any of my previous work in the field of tuberculosis statistics . . . . The question of whether there is also a constitutional factor beyond the influence of the environment on the mortality from children of the tuberculous is open. Its existence cannot be completely ruled out . . . . From a racial hygienic perspective of the tuberculosis problem it is reassuring to note that the fertility of the tuberculous is not excessive and should remain far below that of the non tuberculous [4, pp. 156–157].

Among the study limitations, Weinberg indicated that it did not include illegitimate children and did not separate children by gender. He also felt that follow-up could not
have been extended beyond the 20th birthday, because children who left home to marry or do military service were more difficult to trace.

**Methodological contribution of Die Kinder der Tuberkulosen**

Weinberg explicitly conceived *Die Kinder der Tuberkulosen* as a retrospective cohort study in order to avoid the survival biases that would have plagued a prevalence survey of the Stuttgart population. From this point of view, the study is qualitatively superior to all 19th century epidemiologic research known to us. The design of 19th century studies is, as a rule, nondirectional. The data collected by John Snow (8), Pierre Louis (9), or Ignaz Semmelweis (10) can be analyzed equivalently as exposed versus unexposed or as diseased versus nondiseased comparisons. In contrast, Weinberg was preoccupied to assemble an exhaustive cohort with an exhaustive follow-up. He spent much energy in trying to track the children that had left Stuttgart and demonstrated that the complete losses to follow-up were not numerous and could not have threatened the validity of the conclusions. Weinberg carefully defined the exposed cohort as being children of parents who had died of tuberculosis during a specified time period and just as carefully selected an unexposed but comparable cohort.

Weinberg is also to be commended for his ability to manipulate the enormous amount of data he had collected and to share with his readers the techniques he used to organize the data and facilitate the analyses. His work really qualifies as herculean, as he apparently did this alone in his spare time.

On the other hand, *Die Kinder der Tuberkulosen* belongs to the 19th century in the sense that it is essentially based on linkage of records collected for vital statistic purposes. Weinberg was not in direct contact with his participants, did not send out questionnaires, and so on. Weinberg distinguished risks and rates, but the distinction was well established by then (11). Moreover, with the exception of the standardized mortality ratios, which he called appropriately “percentages of the expected,” there is no measure of relative risk or risk difference.

**Historical place of Die Kinder der Tuberkulosen**

Weinberg’s pioneering contribution to the history of cohort studies has been essentially missed, probably because few people had access to the original documents. Weinberg’s mathematics has been described as abstruse (12), and *Die Kinder der Tuberkulosen* confirms this judgment. Weinberg’s book was widely known and cited, but people failed to realize its historical importance. Comstock (6) being probably the exception. American and British epidemiologists had only a superficial knowledge of the study and were not aware that its retrospective cohort design had preceded that of the study conducted by Wade Hampton Frost (5) in the Black population of Kingsport, Tennessee, the preliminary results of which had been reported in 1933. This resulted in the widespread notion that the use of historical records to reconstruct exposed and unexposed cohorts by Frost was novel (6, 7, 13).

Frost’s work is probably the third report of an intentionally designed retrospective cohort study. Winkelstein (14) has rightly given credit for the first (known so far) publication to Janet Lane-Claypon, who reported in 1912 that she had assembled a cohort of newborns on the basis of historical information obtained from children attending an infant consultation in Berlin. The aim was to determine whether being fed boiled cow’s milk as opposed to breastfed influenced child weight gain. The Lane-Claypon study comprised about 500 infants and was much smaller than Weinberg’s, which was very large for 1900 standards. Who would have expected that such a large study could have been conducted before 1945? In his history of cohort studies, Doll wrote as follows:

Weinberg’s study [4] of the mortality of children born to tuberculous parents, which was cited by Frost [5] and which I have not been able to trace, may have been a prospective cohort study by the strictest modern definition. It was, however, almost certainly small and the large studies that we now usually envisage under this head were not developed until after the second world war, when they were initiated independently and more or less contemporaneously in the UK [United Kingdom] and the USA [United States of America] [7, p. 251].

*Die Kinder der Tuberkulosen* also contradicts Liddell’s credit to Frost for the first description of life-table analysis applied to an epidemiologic study (13, p. 1,224). In 1913, Weinberg mastered actuarial survival analysis. We don’t know if he invented it for his purpose or whether he was familiar with its use by Lawrason Brown and E. G. Pope in 1904, mentioned by Philip Sartwell (15), or its description in 1910 by William Palin Elderton and Sidney James Perry (16).

**Conclusion**

This rediscovery of *Die Kinder der Tuberkulosen* confirms that the first cohort studies have been of the retrospective type. Janet Lane-Claypon (1912, United Kingdom), Wilhelm Weinberg (1913, Germany), and Wade Hampton Frost (1933, United States) apparently discovered the design independently of one another. Contrary to belief, Weinberg’s study was large and used sophisticated analytical methods. *Die Kinder der Tuberkulosen* is probably the most important early epidemiologic work before the work in South Carolina by Goldberger et al. in 1920 (17).

**ACKNOWLEDGMENTS**

Conflict of interest: none declared.

**REFERENCES**