



---

**COMMENTARY**

---

**Vignettes of the History of Epidemiology: Three Firsts by Janet Elizabeth Lane-Clayton**

**Warren Winkelstein, Jr.**

From the Division of Epidemiology, School of Public Health, University of California, Berkeley, Berkeley, CA.

*Received for publication April 22, 2004; accepted for publication April 29, 2004.*

---

In 1912, Janet Elizabeth Lane-Clayton, a British medical scientist 35 years of age who had already contributed substantial research findings in the fields of reproductive physiology and the bacteriology and biochemistry of milk, reported the results of a retrospective cohort study of weight gain during the first year of life among 204 infants fed boiled cows' milk compared with 300 infants fed human breast milk. The results of her investigation revealed that, up to the age of 208 days, breastfed infants gained more weight than infants fed boiled cows' milk. After that time period, weight gain was equal in the two groups. Lane-Clayton described, discussed, and analyzed her data for the possibility that her findings were due to sampling variation or confounding, and she used Student's *t* test to evaluate observed differences in weight gain in small subsets of the study population. As far as is known, this was the first use of the retrospective (historical) cohort design and the *t* test in an epidemiologic study.

confounding factors (epidemiology); history; retrospective studies

---

Interest in the history of epidemiology has been sporadic and fragmentary. Many years ago, the Delta Omega Society (a public health honorary society) initiated a project to reprint facsimile copies of four classic epidemiologic studies (1–4). In 1941, the Commonwealth Fund published selected papers of Wade Hampton Frost, the first American professor of epidemiology, with commentary by Kenneth Maxcy (5). Charles-Edward A. Winslow provided a thoughtful essay on the philosophical basis of epidemiology in 1943 (6). Several textbooks have included more or less lengthy chapters on the subject (7–10). However, as yet, no text on the history of epidemiology exists (11).

This situation was somewhat ameliorated by the publication of a series of historical papers prepared as the result of a symposium, "Measuring our Scourges," held in Annecy,

France, in 1996 and organized by Swiss epidemiologist Alfredo Morabia. All of the articles from the symposium were published in the journal *Social and Preventive Medicine* (volumes 46 and 47) and can be accessed online (<http://www.epidemiology.ch/history>), along with additional papers and editorial comments on the subject. In his editorial introducing the symposium papers, Morabia described the workshop (*sic*) as "focused on the historical emergence of the corpus of epidemiologic methods used today" (12, pp. 3–4).

The purpose of this commentary is to provide examples of what may have been the first epidemiologic implementation of a retrospective (historical) cohort study, the first modern description of confounding with an accompanying analysis, and the first use of Student's *t* test to assess the difference of

means in small samples. All three examples can be found in Janet Elizabeth Lane-Claypon's 1912 classic, albeit forgotten, publication, *Report to the Local Government Board upon the Available Data in Regard to the Value of Boiled Milk as a Food for Infants and Young Animals* (13).

Lane-Claypon's 60-page report is divided into five parts. The first three introduce and review the then-existing evidence to support the assertion that "experimental evidence confirms the conclusion derived from clinical experience as to the superior results obtained by feeding infants or young animals with the breast milk of an animal of the same species ... and emphasizes the opinion that infants should be fed on the breast unless there is urgent reason to the contrary" (13, p. 1). Nevertheless, she also pointed out that "there remains ... a small minority for whom artificial feeding is necessary ...; and for them ... the relative nutritive value of raw and boiled milk is of great importance" (13, p. 1). After concluding that the evidence indicated that raw and boiled cows' milk were equal in nutritive value, she addressed the question of the effects of a diet of boiled cows' milk. Part IV describes the design and results of the retrospective cohort study that she carried out to evaluate the relative nutritive value of boiled cows' milk and human breast milk. The index of nutritive value used in the study was body weight at 8-day intervals from birth to 368 days of age. Part V provides a summary and conclusions.

### THE THREE FIRSTS

#### Lane-Claypon designs and implements the first retrospective (historical) cohort epidemiologic study

In the introduction (part I) to the report, Lane-Claypon spelled out the requirements for implementing the study successfully: 1) a large number of healthy babies under medical supervision who had been fed boiled cows' milk during the first year of life, 2) a similar number of healthy babies subjected to the same supervision and fed an alternative diet (breast milk), and 3) the babies in each group being "as far as possible" (13, p. 3) from the same social environment. To obtain such a study population, Lane-Claypon required access to a large clinical facility (she called them "Infant Consultations") that served healthy children by providing regular periodic examinations. Unable to find such a facility in England, she took advantage of contacts made in Berlin, Germany, while traveling in Europe as a Jenner Research Fellow (1909–1911). At the time of the study, there were seven Infant Consultations in Berlin, all of which served infants "exclusively of the working classes" (13, p. 28). The Consultation chosen for the study (Naunyn Strasse) served a clientele of about 100 babies every day. Sick children were routinely referred to hospitals or private practitioners. Extensive data were obtained on each baby at the time of its first visit to the Consultation and on each subsequent visit. The critical data obtained at the first visit included dates of birth and first visit, type of feeding (breast or artificial), weight, and wages of the father. At subsequent visits, the data obtained included date of visit, weight, and type of feeding.

Study subjects were selected seriatim from attendees in 1907–1908 and 1908–1909, with the following exclusions: infants over the age of 4 months at the first visit, infants who did not attend regularly over a period of 4 months, infants who suffered from "constitutional diseases" (13, p. 31), and infants who died during their attendance at the Consultation. The resultant study population consisted of 204 infants whose primary milk source was boiled cows' milk and 300 infants whose primary milk source was the breast. (Lane-Claypon considered the breastfed series as controls.) Table I (reproduced here as figure 1) and table II (identical except for the numbers in each cell; not reproduced here) of her report describe the two series with regard to age at entry and exit. The boiled-milk-fed series yielded 5,444 weight measures and the breastfed series 6,297. Because each series included a small proportion of infants whose diets were mixed, Lane-Claypon carried out various analyses based on selected portions of the cohorts (not considered herein).

The basic analyses consisted of calculating the mean weights of the babies in the two series for each 8-day period of follow-up for a year. Results were graphed as shown in diagram I of the report (reproduced here as figure 2). Lane-Claypon described the findings as follows: "Diagram I. shows at once that a considerable divergence between the two curves starts in the early days of life, and continues well-marked up to about the 208th day, after which it disappears fairly rapidly. The question suggested by these curves is, – Is the difference between the average weight of breast-fed and of babies ... fed upon boiled cow's milk due to the method of feeding?" (13, pp. 38–39). Lane-Claypon then considered the possibility that the differences might be due to "Error of Sampling" (13, p. 39). She then tested the probability that the means of three consecutive 8-day periods, from the 137th to the 160th day after birth, in each series could be due to sampling. She did so by calculating the ratio of the observed difference of means for the babies in each series and the probable error (analogous to the standard error) of the difference. The "critical" ratios for the three differences were similar, 7.6–8.4, an observation indicating that the differences were unlikely to be due to sampling error. (A "critical" ratio of 1.96 based on the standard error of the difference as the denominator is equivalent to a 0.05 probability that the observed differences were due to sampling variability.)

#### Lane-Claypon describes confounding and analyzes her data to investigate the possibility that it explained the findings

Having satisfied herself that the differences in weight gain in the two series were not due to sampling, Lane-Claypon considered the possibility that the differences were due to confounding. She described it as follows: "It does not, however, necessarily follow that the difference of food has been the causative factor, and it becomes necessary to ask whether there can be any other factor at work which is producing the difference found. ... The social class of the children seemed a possible factor, and ... it was considered advisable to investigate the possible significance of any difference which existed between the social conditions of the homes" (13, p. 40). To accomplish this, Lane-Claypon

TABLE I.

*Showing the age of first attendance and of leaving the Consultation of the Babies of the Control or Breast-fed Series.*

I. Age in weeks.	II. No. brought at each day.	III. Age on leaving (in months).							IV. Total in each week.
		4-5.	5-6.	6-7.	7-8.	8-9.	9-10.	10-12.	
1	3. 2. 0. 3. 0. 4. 4.	2	1	1	0	4	0	8	16
2	5. 7. 8. 6. 6. 7. 10.	9	4	5	3	1	3	24	49
3	8, 11, 9, 15, 12, 5, 7.	8	11	5	6	8	9	20	67
4	9. 4. 3. 8. 0. 3. 4.	2	3	7	5	1	2	11	31
5	1. 2. 4. 5. 2. 3. 2.	—	1	3	1	2	1	11	19
6	5. 2. 3. 5. 2. 2. 3.	—	6	3	1	3	0	9	22
7	3. 1. 2. 3. 4. 2. 2.	—	4	1	5	1	2	4	17
8	1. 2. 1. 0. 3. 3. 0.	—	3	2	2	1	0	2	10
9	2. 2. 3. 4. 3. 2. 2.	—	3	6	2	2	1	4	18
10	1. 1. 1. 1. 2. 0. 1.	—	—	3	1	1	0	2	7
11	2. 4. 3. 3. 0. 0. 1.	—	—	4	1	1	1	6	13
12	3. 0. 0. 0. 3. 5. 1.	—	—	—	4	1	2	5	12
13	2. 1. 0. 1. 2. 0. 2.	—	—	—	1	0	2	5	8
14	1. 1. 0. 0. 0. 0. 2.	—	—	—	1	0	1	2	4
15-18	—	—	—	—	—	3	1	3	7
		21	36	40	33	29	25	116	300

FIGURE 1. Lane-Claypon's tabulation of ages at entry and at study termination of infants in the breastfed series (13, p. 33).

compared the weekly wages of the fathers of the infants in each series and found that, although wages were low overall and reflected the artisan class from which the study subjects were drawn, their distributions were essentially the same. She also noted that "no baby is allowed to have a deficient food-supply, since if it is artificially-fed[,] the milk is supplied free if the family cannot afford to pay for it, and the nursing mothers receive a nursing bonus" (13, p. 41).

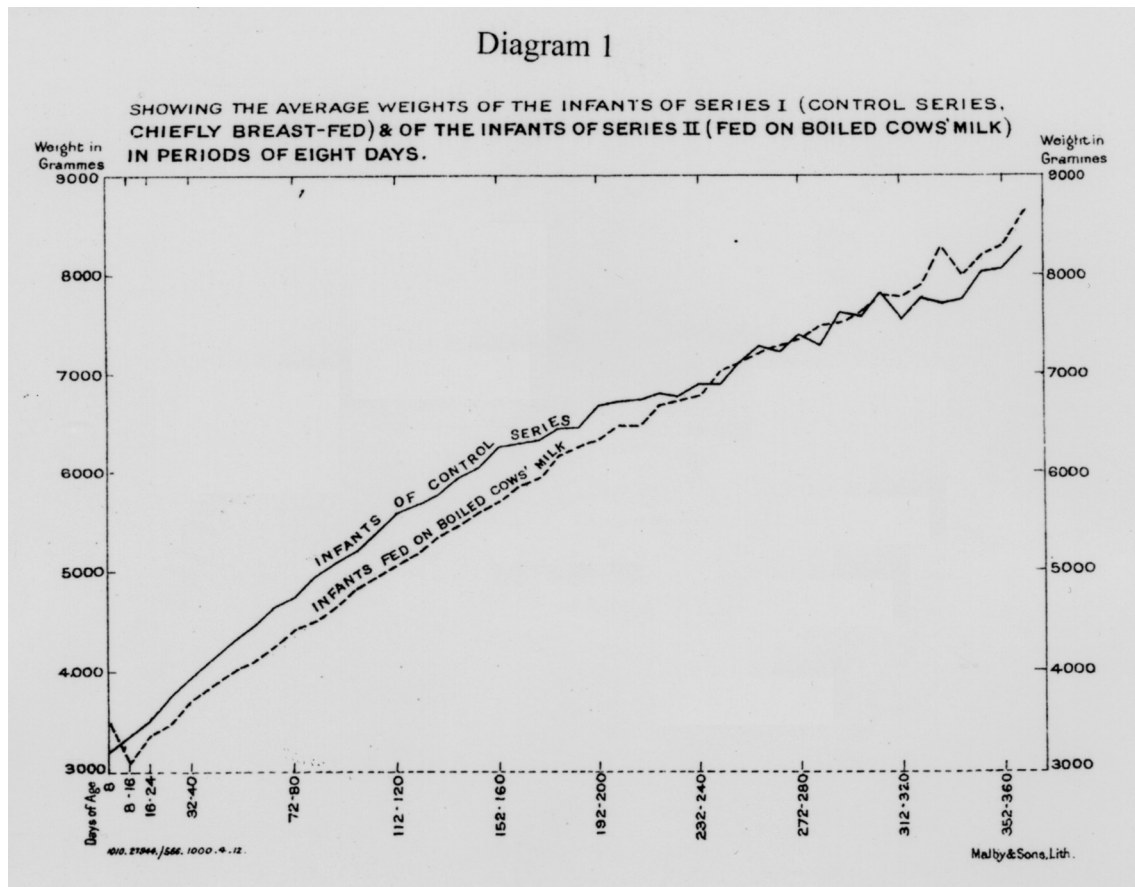
#### Lane-Claypon uses Student's *t* test to evaluate the observed differences in weight during the first 8 days of life of the boiled-milk and breastfed infants

Toward the end of the report, Lane-Claypon took note of the crossover of the weights for the two series of babies during the first 8 days of life shown in figure 2 (diagram I). She wrote, "The average weight of the babies fed upon boiled cows' milk is higher for the first eight-day period than that of the breast-fed babies. The former value is based on 10 observations, and the latter upon 24; it becomes a question whether any importance can be attributed to this difference ... or whether it may not be due to an error introduced by the

extremely small number of observations available for the cows' milk series" (13, p. 45). To deal with this possibility, Lane-Claypon took advantage of a statistical procedure that had been reported 4 years before publication of her report. She described it as follows, "The method introduced by 'Student' is applicable for small numbers of observations" (13, p. 45). She was, of course, referring to the *t* test developed by W. S. Gosset in his classic paper, "On the Probable Error of a Mean," which provided a technique for analysis of data from small samples (14). Application of Student's *t* test to the analysis of the difference of means in the data for the first 8 days of life indicated that the probability of a sampling-error explanation was relatively high, about 14 percent, compared with a sampling-error probability of about 2 percent given by the standard method of testing.

#### DISCUSSION

Lane-Claypon concluded her report with the statement, "The evidence dealt with throughout this report emphasizes very forcibly the importance of breast-feeding for the young of all species and shows the special importance of breast-



**FIGURE 2.** Lane-Clayton's graphical presentation of the weights of infants in the two study series at 8-day intervals during 12 months of follow-up (13, p. 39).

feeding during the early weeks of life" (13, p. 56). Nevertheless, she also wrote, "The Berlin figures dealing with infants fed on boiled cow's milk give extremely favorable results" (13, p. 56).

Lane-Clayton's report would appear to predate by almost three decades other retrospective cohort studies. In his review of retrospective cohort studies, Doll (15) cites Frost's 1933 study of familial acquisition of tuberculosis in the Black population of Kingsport, Tennessee, as an early example (16). However, cohort studies did not become a prominent part of the epidemiologic armamentarium until after the Second World War. Pioneering studies of occupational hazards in England (17, 18); a study of atomic bomb explosion survivors in Hiroshima and Nagasaki, Japan (19); and a long-term follow-up study of cardiovascular risk factors in a general population sample in Framingham, Massachusetts (20), opened the door to an abundance of this type of study that is ongoing today.

In his insightful essay on the history of confounding, Vandembroucke wrote, "Confounding is not a statistical or analytic concept. It is a concept that has to do with the logic of scientific reasoning. In particular the logic of inferring causality from observations" (21, p. 216). Lane-Clayton's skepticism with regard to her statistically significant obser-

vations of the differences in weight gain of the infants fed on different regimens led her to a clear description of the dilemma posed by possible confounding variables in the interpretation of an observed association as causal. Although Vandembroucke cites Claude Bernard as having described, during the 19th century, the complexity of causal inference in observational studies, Bernard's exposition, in his 1866 classic, *Introduction to the Study of Experimental Medicine*, was not based on real data as was Lane-Clayton's 1912 consideration of the issue.

Lane-Clayton's use of "Student's *t* test" was undoubtedly suggested by Major Greenwood, who had been on the staff of the Lister Institute of Preventive Medicine in the United Kingdom during Lane-Clayton's tenure there (1907–1912) and who was subsequently professor of epidemiology and biostatistics at the London School of Hygiene and Tropical Medicine. Lane-Clayton credited Greenwood in a footnote of her report, as follows: "For instruction in the statistical methods employed and for supervision of the results obtained, I am deeply indebted to Dr. Major Greenwood ... of the Lister Institute" (13, p. 29). It should be noted that Gosset's seminal article was highly technical and understandable to only the most sophisticated statisticians (14).

This report, the 11th publication attributed to Janet Elizabeth Lane-Claypon, was her first epidemiologic paper. Her previous publications reported the results of laboratory studies of reproductive physiology or the bacteriology and chemistry of milk. These latter studies and the report discussed here led to her selection by the Medical Research Committee (forerunner of the Medical Research Council) to prepare a comprehensive review of the hygienic aspects of milk. The resulting book, *Milk and its Hygienic Relations*, was published simultaneously in New York and London in 1916 (22).

## REFERENCES

1. Budd W. Typhoid fever: its nature, mode of spreading, and prevention. New York, NY: American Public Health Association, 1931.
2. Snow J. Snow on cholera: being a reprint of two papers. New York, NY: The Commonwealth Fund, 1936.
3. Panum PL. Observations made during the epidemic of measles on the Faroe Islands in the year 1846. Delta Omega Society. (<http://www.deltaomega.org/PanumFaroeIslands.pdf>).
4. Baker GW. An essay concerning the cause of the endemial colic of Devonshire. Delta Omega Society. (<http://www.deltaomega.org/devon.pdf>).
5. Maxcy KF, ed. Papers of Wade Hampton Frost. New York, NY: The Commonwealth Fund, 1941.
6. Winslow C-EA. The conquest of epidemic disease: a chapter in the history of ideas. Princeton, NJ: Princeton University Press, 1943.
7. Greenwood M. Epidemics and crowd-diseases: an introduction to the study of epidemiology. London, United Kingdom: Williams and Norgate Ltd, 1935.
8. Lilienfeld AM. Foundations of epidemiology. New York, NY: Oxford University Press, 1976.
9. Rothman KJ, Greenland S, eds. Modern epidemiology. 2nd ed. Philadelphia, PA: Lippincott-Raven, 1998.
10. Stolley PD, Lasky T. Investigating disease patterns: the science of epidemiology. New York, NY: W H Freeman, 1995.
11. Winkelstein W Jr. From the editor. (Editorial). *Am J Epidemiol* 2002;155:776.
12. Morabia A. History of epidemiologic methods. (Editorial). *Soz Praventivmed* 2001;46:1–6.
13. Lane-Claypon JE. Report to the local government board upon the available data in regard to the value of boiled milk as a food for infants and young animals. No. 63. London, United Kingdom: His Majesty's Stationary Office, 1912:1–60.
14. Selvin S. Statistical analysis of epidemiologic data. 2nd ed. New York, NY: Oxford University Press, 1996.
15. Doll R. Cohort studies: history of the method. II. Retrospective cohort studies. *Soz Praventivmed* 2001;46:152–60.
16. Frost WH. Risk of persons in familial contact with pulmonary tuberculosis. *Am J Public Health* 1933;23:426–32.
17. Hill AB. Principles of medical statistics. 8th ed. London, United Kingdom: The Lancet Ltd, 1966.
18. Doll R. The causes of death among gas-workers with special reference to cancers of the lung. *Br J Ind Med* 1952;9:180.
19. Folley JH, Borges W, Yamasaki T. Incidence of leukemia in survivors of the atomic bomb in Hiroshima and Nagasaki. *Am J Med* 1952;13:311–21.
20. Dawber TR, Meadors GF, Moore FE Jr. Epidemiological approaches to heart disease: the Framingham study. *Am J Public Health* 1951;41:279–86.
21. Vandembroucke JP. The history of confounding. *Soz Praventivmed* 2002;47:216–24.
22. Lane-Claypon JE. *Milk and its hygienic relations*. New York, NY, and London, United Kingdom: Longmans Green & Co, 1916.