The Statistical Approach to the Cancer Problem in Massachusetts*

EVELYN A. POTTER AND MILDRED R. TULLY

Biometricians, State Department of Public Health, Boston, Mass.

STATISTICAL studies inspired the Massachusetts Cancer Program, determined its scope, evaluated its activities, changed its policies, and elicited new ideas in cancer control.

The results of many of these studies have been reported in the literature, some have been utilized in official documents, while the findings of still others have never been published. The data used in these studies fall into five general classifications: the death record, the clinic record, the hospital record, the house-to-house survey, and the selected sample survey. The statistical procedures utilized have for the most part been simple and only the occasional study has demanded the more advanced statistical treatment.

The objective of this paper is a demonstration (through exemplification) of the integration of statistics in the Massachusetts Cancer Program. Most of the material selected is from studies that have not been incorporated into the literature, either because these studies are confirmation of previous ones or because their brevity does not warrant an independent report.

DEATH RECORDS

A type of data readily available and one from which much information can be procured is the death record. A study of the accuracy of the cancer death certificates in Massachusetts in 1932 was made by Macdonald.[1] The results were obtained by comparing the death certificate data with the composite information received by interviews with the family and all physicians who attended the patient, together with a study of the case history from the hospital records. The analysis indicated that the identification of cancer deaths was sufficiently accurate to warrant statistical compilations of data on the disease as a whole.

A comparable study was conducted using a sample of the 1939 cancer deaths. The results indicated that there had been considerable improvement in cancer consciousness. This sample showed that the accuracy of the cancer death record had improved 2 per cent regarding the presence of cancer and nearly 7 per cent regarding primary location of the disease. In Table 1 the percentage agreement of

TABLE 1

Percentage Agreement of Verified Diagnoses with Death Certificates

	Number of Death Certificates		*Percentage Agreement*	
	1932	*1939*	*1932*	*1939*
Buccal cavity	79	74	83.5	85.1
Digestive tract	1,028	972	75.5	84.3
Respiratory tract	64	102	68.8	70.6
Uterus	213	140	80.8	82.9
Breast	225	210	98.7	98.1
Other female genital organs	43	40	76.7	85.0
Male genitourinary organs	169	155	83.4	88.4
Skin	28	25	57.2	76.0
Other and unspecified organs	180	239	66.7	70.7
All locations	2,029	1,957	78.2	83.5

* Presented before the Cancer Symposium, in association with the Seventy-third Annual Meeting of the American Public Health Association, New York, N. Y., October 2, 1944.

verified diagnoses with death certificate diagnoses, subdivided by location of cancer, showed an improvement in 1939 over 1932 which was significant by the chi-square test.

On the basis of these findings the recorded deaths from cancer are sufficiently accurate to warrant their use for statistical procedures, but equal reliance cannot be placed on studies based on the data of individual sites of cancer.

The death records may be used to confirm or disprove tentative hypotheses which have been advanced by other research workers. In 1940 Ciocco,[2] using Washington County, Maryland, data, reported an apparent tendency for both husbands and wives to die of cancer. He compared his observed finding of 29 such couples with an expected number of 19.9 and a control number of 19.0. The chi-square test on his data was 4.98 with a probability of 0.0257. Considerable interest was aroused in the medical profession following the publication of this paper and it was felt that if the findings represented general truths the public might postulate contagion.

In order to confirm or refute this study a comparable one was done using Massachusetts figures. Two types of data were used: one comprising material comparable to Ciocco's in the number of married couples but covering twice as long a period of time; the other comparable to his in length of time but having one-half as many married couples. Duplicating the method which Ciocco used in obtaining his expected values, the number of husbands and wives dying of cancer in both Massachusetts series showed no association between marital status and cancer.

It was thought that, due to the change in the cancer death rate over an extended period of time, a better expected value could be computed. This was accomplished by applying to each surviving spouse of the husband or wife who died of cancer the probability of his or her death being due to cancer at the age and the time at which each died. The summation of these comprised the expected number of deaths from cancer among spouses. The expected values for both Massachusetts sets of data were used to compute chi-square tests. The results showed no significance. Ciocco obtained the number of control non-marital couples who had cancer and found a difference between these and the marital couples. In Massachusetts no significant difference was found. Ciocco states in his report: "In the case of widows and widowers who died of cancer, 13.3 per cent of their spouses had died from this cause and only 9.4 per cent of the spouses of all the widowed. The difference is probably significant from a statistical standpoint since it is over twice its standard error." Similar compilations from Massachusetts data did not show significant differences.

Inasmuch as the Massachusetts figures did not substantiate those of Maryland, it is believed that the Maryland figures represented the unusual occurrence and that the cancer incidence in husbands and wives is not related. Further investigation may be desirable.

The death certificate furnishes the best data for measuring the efficiency of the cancer control program. The decrease in the cancer death rate after adjustments have been made for changes in the population structure, and after allowances have been made for chance variations, furnishes a clear-cut picture of the cancer situation. Due to the life span of cancer patients, the effect of control measures is not reflected in the death rate immediately, but a lag must occur between improvement in the cancer situation and evidence of this from the death certificate. The death rate, adjusted for age and sex, has been

TABLE 2

Comparison of Maryland and Massachusetts Data in Respect to Deaths from Cancer Among Husbands and Wives

	Maryland *	Massachusetts	
	1898–1938	1841–1932	1893–1932
Both husband and wife died of cancer	29	18	12
Expected number of husbands and wives to die of cancer. Method A †	19.9	19.0	15.4
Expected number of husbands and wives to die of cancer. Method B ‡	....	17.8	13.5
Number of control non-marital couples both of whom died of cancer	19	18	11
Chi-square test based on expected values computed by Method A †	4.98	1.23	2.42
Among widowed persons who died of cancer, percentage of spouses who also died of cancer	13.2	7.8±1.8	8.3±2.3
Among all widowed persons, percentage of spouses who died of cancer	9.4±0.6	8.4±0.5	10.9±0.9
Total married couples	2,571	2,600	1,210
Number of deaths from cancer among husbands	187	194	115
Number of deaths from cancer among wives	274	255	162
Number of deaths from cancer among widowed persons	219	231	145
Number of deaths from cancer among spouses of widowed persons	242	218	132

* This part of table constructed from data in "On the Mortality of Husbands and Wives." *Human Biol.*, Dec., 1940.

† Method A for computing expected values. Apply the product of the probability of husbands dying of cancer and that of wives dying of cancer to the total married couples.

‡ Method B for computing expected values. Apply time-age-sex rate to surviving spouses of husbands or wives with cancer.

periodically examined to determine changes in the trends.[3] During the early years of the Massachusetts program the female cancer rate adjusted for age was almost trendless. From 1932 to 1943 inclusive, it dropped from 127.0 to 117.5 per 100,000 population. The male rate adjusted for age continued as an upward trend until 1941 and since then has fallen slightly, possibly indicating the beginning of a downward trend. While the drop in the cancer death rate is not great it has been consistent. It is a tangible evidence of accomplishment and warrants an optimistic point of view.

CANCER CLINIC RECORDS

The records of the Massachusetts cancer clinics are an important source of data. The original record listed ten items of information to be obtained from the clinic patients. At the suggestion of Professor E. B. Wilson, Chairman of the Statistical Advisory Committee for the Cancer Program, the record was limited to a minimum number of items to avoid disturbing the patients unduly. This principle has been continued, with only a slight increase in the original number. These items included identifying data, reason for coming to the clinic, duration of symptoms before first visit to a doctor and before first visit to a clinic. These were arranged in the form of a code so that underlining alone would be necessary in filling out the record. Other items to facilitate short studies have been incorporated into the card, used for a few years, and then discarded. On the original card four items dealing with diagnosis and recommendations for treatment were filled out by either the physician or the follow-up worker. In the present revision this has increased to eight items. Progress notes are incorporated with the original record until the report is closed. All cancer cases, with the exception of approximately 2 per cent lost, are followed until death. The clinic cards have been used for four general types of study: the evaluation of the program, etiological studies, material for educational purposes, and for the measurement of routine activities.

The various periods of delay between first suspicious symptom and treatment have been recorded on the clinic cards.[4]

In the first year of the Massachusetts Cancer Program the delay between the first symptom and consulting the first physician was 6.5 months. For the next 8 years this period of delay remained practically constant (trend coefficient $= -.06 \pm .05$). Since 1935 there has been a statistically significant downward trend with a new low of 3.3 months reached in 1943 (trend coefficient $= -.32 \pm .03$). Perhaps a more informative appraisal is the percentage of individuals visiting their physicians within 1 month of noticing suspicious symptoms. In the 4 year period 1930–1933, 12.7 per cent of individuals with cancer who attended the clinics consulted their physicians within 1 month of recognition of suspicious symptoms. In a corresponding period 10 years later the rate had increased to 22.4 per cent. The greatest improvement was noted in the cases of skin cancer. Improvement in rapidity of seeking diagnosis is an excellent method of evaluating a cancer program.

Since an earlier study from data collected by the Visiting Nurse Associations had suggested a possible relationship between tobacco and cancer, the chairman of one of the clinic committees requested that this be substantiated by data collected from clinic patients.[5] Information on the use of tobacco was obtained from 2,927 male clinic patients over the age of 40. They were divided into those who used no tobacco, those who reported slight use, moderate use, and excessive use. Attack rates were computed for cancer of the buccal cavity, digestive tract, respiratory tract, skin, and for cancer of all other sites. There was a definite association between cancer of the buccal cavity and the use of tobacco. There also appeared to be some association between the use of tobacco and cancer of the respiratory tract. These findings confirm the opinion held by many clinicians.

An important part of the Massachusetts Cancer Program consists of furnishing information to the public regarding prevention and the importance of early recognition and treatment. One of the difficulties encountered is the skepticism on the part of many regarding the curability of the disease. When actual figures demonstrating cures can be produced, a large part of the skepticism disappears. The clinic cards have furnished such information.[4] Individuals with skin cancer treated within the first month showed 80 per cent living at the end of 10 years while those who delayed more than 1 month had a survival rate of only about half this figure (43.9 per cent). Similarly, cancer of sites other than skin showed a 10 year survival rate of about 38 per cent among those treated within 1 month contrasted with 20 per cent among those who were not.

The data from the clinic cards have been used as a guide to administrative procedure, such as the measurement of

TABLE 3

Tobacco and Cancer—3 Year Period

Rate per 100, by Site

Extent of Use of Tobacco	*Males Over Age of 40 in Massachusetts Cancer Clinics*	*Cancer of the Buccal Cavity*	*Cancer of the Digestive Tract*	*Cancer of the Respiratory Tract*	*Cancer of the Skin*	*Cancer of All Other Sites*
None	655	3.7	5.6	0.5	11.3	5.5
Slight	357	8.1	9.5	1.1	7.8	7.8
Moderate	1,155	11.5	8.3	2.0	12.5	5.2
Excessive	760	17.9	5.8	1.7	13.0	4.6

the effectiveness of the follow-up service in the clinics.[4] Perhaps one of the most informative figures of this type is that which showed that in 1943 81.1 per cent of recommendations made by clinic physicians had been put into execution within 1 month. Another interesting compilation is the survival rate of cancer patients which has been computed from these cards for the past 17 years.

HOSPITAL DATA

Reporting of cancer is not attempted in Massachusetts but the periodic collection of hospital data furnishes a somewhat similar type of information. Hospital morbidity is not total morbidity, but recent studies indicate it is approaching that optimum. In 1932 approximately 70 per cent of cancer cases were hospitalized. This has increased to about 85 per cent at the present time.[4]

HOUSE-TO-HOUSE SURVEYS

Data collected by means of house-to-house surveys have been voluminous. The book *Cancer and Other Chronic Diseases in Massachusetts*[6] was based on such surveys. The most recent use of this method has been several public opinion polls to determine the knowledge of the people regarding cancer, and their sources of obtaining this knowledge.[7] Partial correlations have been computed to determine the real value of the various educational media. About one-quarter of the individuals surveyed displayed some knowledge of cancer by answering correctly three pertinent questions relative to the disease. The principal source of their information was reading material.

SELECTED SAMPLE SURVEY

Another form of survey is that utilizing the selective sample rather than the random one. An example is the true-false test conducted in the Lynn High Schools as a measurement of the value of the experimental program of cancer education.[8] Four groups of students received different types of instruction in the hope of ultimately determining the best approach to this problem. The methods employed were the general assembly, detailed instruction in the biology class and in the public health class, and bulletin board exhibits. A fifth group of students did not receive instruction. As a measurement of this experiment, true-false tests were given. In order to minimize the possibility of a student obtaining a high mark by guesswork, twelve thoughts were presented in this questionnaire, each expressed four times in a different manner, and a score was given only when all four parts of the same thought were answered correctly.

The results of this test indicated a definite upward trend in cancer knowledge as the students progressed from the lower to the higher grades, and suggested that the most satisfactory approach is detailed instruction in the biology class. The experiment is to be repeated and should substantiate or disprove this opinion.

SUMMARY

This paper is a report of some of the statistical work that is being done as a part of the Massachusetts Cancer Program. The results from several studies furnish examples of the application of various types of data.

SOURCE OF DATA AND RESULTS

Death Records

The death records can be freely used in studies of cancer as a whole but are not sufficiently accurate to warrant too much credence in cancer of individual locations.

Massachusetts figures do not indicate any association between cancer and marital status.

One measurement of the efficacy of the Massachusetts Cancer Program has been the drop in the adjusted death rate.

Clinic Records

The decrease in delay between first symptoms recognized by the patient and first visit to a physician is tangible evidence that the public has become more cancer conscious.

The use of tobacco is strongly associated with cancer of the mouth and to a lesser degree with cancer of the respiratory tract.

The clinic follow-up service of cancer patients has shown that among those who sought treatment within the first month, twice as many were cured as among those who did not seek treatment within that period.

The fact that over 80 per cent of the recommendations for clinic patients have been carried out within 1 month is a criterion of the excellence of the follow-up service.

Hospital Data

There has been a steady increase in hospitalization of cancer patients in Massachusetts.

House-to-House Surveys

The knowledge of the public regarding cancer and the sources of obtaining this knowledge has been ascertained by house-to-house surveys. About one-quarter of the individuals surveyed displayed some knowledge of cancer by answering correctly three pertinent questions relative to the disease. The principal source of their information was reading material.

Selected Sample Surveys

The true-false test incorporating a multiple presentation of similar thoughts to limit good scores by guesswork has demonstrated improvement in cancer knowledge of high school students following an experimental course of instruction on the disease.

CONCLUSION

A complete list of the problems in cancer which have been studied statistically in Massachusetts would be far outside the limits of this report. These few examples have been chosen to show different types of studies and some of the different approaches. The findings in some of the studies influenced the program as a whole, while those of others only a part. The decrease in the death rate and the duration of delay before seeking diagnosis were measures of the whole program and aided administrative evaluation, while the value of the cancer poll was limited to the educational groups.

The statistical approach is fundamental in attacking the cancer problem from the public health standpoint. A program based on this approach provides a graphic picture of the extent of the problem and of past accomplishments, and it suggests new measures of control. Appraisals of activities elicit information as to the relative merits of each endeavor, and are a necessary adjunct to any control program.

Comprehensive data aid not only in the solution of the immediate, local problem, but they may be of use to those who are delving into the many ramifications of the larger problem.

REFERENCES

1. Macdonald, Eleanor J. Accuracy of the Cancer Death Records. *A.J.P.H.*, 28, 7:818 (July), 1938.

2. Ciocco, Antonio. Studies on the Biological Factors in Public Health. III. On the Mortality in Husbands and Wives. *Human Biol.*, 12, 4 (Dec.), 1940.

3. Lombard, Herbert L., and Macdonald, Frances A. Results of Fifteen Years of the Cancer-Control Program in Massachusetts. *New England J. Med.*, 226, 3 (Jan. 15), 1942.

4. Massachusetts Department of Public Health Annual Report, 1943 (not published).

5. Lombard, Herbert L., and Doering, Carl R. Cancer Studies in Massachusetts: 2. Habits, Characteristics and Environment of Individuals with and without Cancer. *New England J. Med.*, 198, 10 (Apr. 26), 1928.

6. Bigelow, George H., and Lombard, Herbert L. *Cancer and Other Chronic Diseases in Massachusetts.* Houghton Mifflin, 1933.

7. Lombard, Herbert L., Tully, Mildred R., and Potter, Evelyn A. Evaluation of the Cancer Educational Program in Massachusetts. *Human Biol.*, 16, 2 (May), 1944.

8. The Committee on Publication. *Cancer Control, The What-Whither-How.* 2nd ed. Rumford Press, 1944.