FRAMINGHAM HEART STUDY

CANCER DATA SET PROTOCOL

The methodology used to generate
the Framingham Cancer data set
is described in the methods section of
the following paper

The Cancer Experience in the Framingham Heart Study Cohort

Bernard E. Kreger, MD, MPH,*† Greta L. Splansky, MS,*†
and Arthur Schatzkin, MD, DrPH‡

The almost 40-year records of The Framingham Heart Study (FHS) cohort were reviewed to establish the cancer experience of this noninstitutionalized group of white subjects. Diagnoses were confirmed from pathology and laboratory reports and clinical notes. Age-specific incidence rates were compared with Connecticut Surveillance, Epidemiology, and End Results (SEER) data. Among the 5209 subjects, 1201 malignancies were confirmed. Median age at diagnosis was 69 for men and 65 for women. Lung, prostate, skin, and colon accounted for more than half of men's cancers; breast, colon, and skin made up half of the women's. FHS and Connecticut SEER rates matched closely, with the same primary tumor sites appearing commonly in both groups. Thus, the FHS cohort should provide a fair database for analysis of risk factors in cancer incidence, as it has done in cardiovascular diseases. *Cancer* 67:1–6, 1991.

THE ONGOING STUDY of the natural history of cardiovascular diseases among adults in Framingham, Massachusetts, has provided valuable clues concerning significant risk factors over the last 40 years.[1] During the course of this study, diagnostic end points other than cardiovascular conditions have been recorded, and we have had the opportunity to review and document all suspected malignancies. The unique design of the Framingham Study allows us to describe here what amounts to the lifetime cancer experience of a cohort of 5209 men and women followed biennially for almost 4 decades so far. Initially aged 30 to 62 years, the survivors are now 70 or older. This article presents prevalence and incidence information and compares it with available regional survey data from a nearby, free-standing population.

Methods

As part of a previous analysis of cancer in the Framingham Study (FHS) cohort, the records of the entire cohort were reviewed during biennial Examination 12, approximately 1970 to 1971.[2] At the end of that review, although most cases had been confirmed as malignant or benign, many remained with incomplete information. We began by examining all the suspect cases from this review, confirming topography (the localization of a tumor within an organ), morphologic characteristics, and date of first tissue diagnosis for all malignancies, for which those data were available.

Cases of tissue-diagnosed nonmalignancy were recorded, as well as patients in whom there appeared to be no reason, from the record, to suspect malignancy at all. Few of these initial suspects remained without a definite classification because of lack/unavailability of data. From this base, we brought the case list up to date as of December 31, 1986. To accomplish this task, we reviewed the following FHS cohort records:

- Subjects whose deaths, confirmed by the FHS "Death Review Committee," had been attributed to cancer,
- Subjects dead of causes other than cancer but with mention of cancer as a secondary diagnosis,
- Subjects with "cancer" (unavailable as a specified diagnosis before Examination 14) circled as a diagnosis at FHS Examinations 14 to 18,
- Subjects not examined at FHS subsequent to Examination 12 but not known to be dead (searching for outside information, *e.g.*, hospital discharge summaries, mentioning malignancy occurring after the comprehensive record review carried out at Examination 12), and

From the *Sections of General Internal Medicine and of Preventive Medicine and Epidemiology, The Evans Memorial Department of Clinical Research, Boston University Medical Center, Boston, Massachusetts; †The Framingham Heart Study, Framingham, Massachusetts; and ‡The National Cancer Institute, Cancer Prevention Studies Branch, Division of Cancer Prevention and Control, Bethesda, Maryland.

Supported in part by grant #5 RO1 CA39766-02 from the National Cancer Institute, Bethesda, Maryland, and grant #N01-HC-38038 from the National Heart, Lung, and Blood Institute, Bethesda, Maryland.

Address for reprints: Bernard E. Kreger, MD, University Hospital, 88 East Newton Street, Boston, MA 02118.

Accepted for publication July 2, 1990.

- A 10% random sample of subjects never suspected to have malignancy (*i.e.*, not in any of the categories mentioned above).

For these record reviews, the entire file of each subject was available for examination. The biennial letters to physicians and death certificates usually served to focus attention on the episode to be considered. All hospitalization records were reviewed, even if the hospitalization was for reasons other than cancer. When necessary because of insufficient information from these sources, we could search through all of the remaining FHS forms, including those regarding medical history, physical examination, and test results.

Whenever the diagnosis of definite malignancy or nonmalignancy could not be confirmed from the FHS file or when it appeared that the first cancer diagnosis antedated what was found in the file, documentation was sought from subjects' physicians forms or hospital records. Most often, we already had the major part of a hospitalization record and needed only the pathology report and/or operative note.

Eighty cases of uncertain diagnosis, confusing descriptions, or questionable date of diagnosis were adjudicated by a surgical oncologist.

All malignancies were coded according to International Classification of Diseases for Oncology (ICD-O) format for topography and morphologic characteristics, with non-Hodgkin's lymphomas scored with a more recently published formulation.[3,4] Coding was done independently by two reviewers (B.E.K. and G.L.S.). Coding questions and disagreements were settled in consultation with Constance L. Percy at the National Cancer Institute.

Age-specific and age-adjusted incidence rates were calculated for each sex for overall cancer and for particular malignancies. Chi-square analysis was used, with suitable adjustment for small numbers.

The comparison population for these FHS data was the white population included in the Connecticut SEER (Surveillance, Epidemiology, and End Results) program; data on black people were not considered because the FHS cohort includes almost no black patients.[5] We calculated ratios of observed (FHS) to expected (SEER) cases for all cancer sites combined and for specific cancer sites. The expected number of cases for a given site was determined by partitioning person years at risk within the cohort into age, sex, and race categories and then multiplying by the appropriate site-specific comparison rates from the Connecticut Tumor Registry.

Results

Among the 2873 Framingham women, 572 (20%) had a total of 627 malignancies in their lifetimes. Diagnoses had been made before entry in the study in only 23 women. Five hundred five (22%) of the 2336 men had 574 malignancies; prestudy diagnoses had been made in only 11 men. In the youngest third of the cohort—patients younger than 40 years of age at entry—the percentages for those ever diagnosed as having cancer were only 16% and 17% in men and women, respectively, whereas for those older than 49 at entry, the oldest third, the percentage with malignancies was 32% for men and 25% for women.

For all cancer sites, median age at diagnosis was 69 for men and 65 for women.

Table 1 displays the primary sites for cancer in men and in women (prestudy and poststudy entry). Among the men, 14 sites accounted for more than 90% of the cancers; and lung, prostate, skin, and colon cancers constituted 54% themselves. Breast, colon, and skin constituted half of the cancers in women, with the 14 most common malignancies again adding up to 90% of the total. Both lung and bladder cancer were found much more frequently in men than women ($P < 0.001$ by chi-square analysis); hematologic and upper airway malignancies appeared to be more common in men and brain cancers in women, but these differences did not achieve statistical significance at the $P < 0.05$ level. For the rest of the sites available to both men and women, there were no differences by gender.

Multiple cancers were not rare: 60 of the 505 men with cancer (12%) and 48 of the 572 women (8.4%) were in this category, six of each sex having three malignancies apiece, the rest only two each. Only skin cancer was overrepresented in people with multiple cancers, accounting for 28% of the men's cancers (*versus* 7.2% of cancers in those with only one malignancy) and 18% of women's (*versus* 8.8%) ($P < 0.001$ and <0.01, respectively). It should be noted that multiple skin tumors of the same histologic type were counted as a single malignancy, no matter where or how many years apart they appeared. Diagnoses of multiple cancers were documented generally within a few years of each other, but many were found at the same time, several were confirmed more than 20 years apart, and one woman had a brain malignancy in 1933 and breast cancer in 1983—a 50-year gap.

Age at first diagnosis of the most commonly found cancers (prestudy plus poststudy entry) is shown in Table 2. Clearly, malignancy accompanies advancing age in both sexes, with median ages mainly in the sixties. However, cases of most of these primary malignancies begin to appear in subjects in their thirties and forties, exceptions being prostate and esophageal cancer in men (no women in this cohort had esophageal malignancy) and cancer of unknown primary in both sexes. Only 4.5% of men's cancers and 9.9% of women's cancers were diagnosed before the patients reached the age of 50. Among the men, ten of the 88 colorectal cancers were in this category. Among

TABLE 1. Distribution of Primary Cancer Sites—The Framingham Study: 38-Year Follow-Up

Name of site	ICD-O	Men		Women	
		n	%*	n	%*
Tongue	141	3	0.5	1	0.2
Major salivary glands	142	1	0.2	1	0.2
Floor of mouth	144	2	0.3	1	0.2
Other and unspecified parts of mouth	145	3	0.5	2	0.3
Oropharynx	146	2	0.3	2	0.3
Nasopharynx	147	2	0.3	0	—
Hypopharynx	148	2	0.3	1	0.2
Pharynx and ill-defined sites in lip, oral cavity, and pharynx	149	0	—	1	0.2
Esophagus	150	11	1.9	0	—
Stomach	151	24	4.2	18	2.9
Small intestine	152	1	0.2	7	1.1
Colon	153	66	11.5	83	13.2
Rectum, rectosigmoid, anal canal	154	22	3.8	23	3.7
Liver and intrahepatic bile ducts	155	3	0.5	3	0.5
Gallbladder and extrahepatic bile ducts	156	2	0.3	6	1.0
Pancreas	157	19	3.3	16	2.6
Peritoneum and retroperitoneum	158	4	0.7	2	0.3
Nasal cavities, sinuses, middle and inner ear	160	2	0.3	0	—
Larynx	161	14	2.4	0	—
Trachea, bronchus, lung	162	97	16.9	32	5.1
Pleura	163	1	0.2	0	—
Hematopoietic and reticuloendothelial	169	30	5.2	21	3.3
Bones, joints, and articular cartilage	170	2	0.3	2	0.3
Connective, subcutaneous, and other soft tissue	171	5	0.9	2	0.3
Skin	173	68	11.8	65	10.4
Female breast	174	—	—	165	26.3
Male breast	175	4	0.7	—	—
Uterus, not otherwise specified	179	—	—	2	0.3
Cervix uteri	180	—	—	27	4.3
Corpus uteri	182	—	—	35	5.6
Ovary, fallopian tube, broad ligament	183	—	—	31	4.9
Other and unspecified female genital organs	184	—	—	4	0.6
Prostate	185	81	14.1	—	—
Testis	186	1	0.2	—	—
Penis and other male genital organs	187	2	0.3	—	—
Urinary bladder	188	41	7.1	17	2.7
Kidney and other urinary organs	189	20	3.5	8	1.3
Eye and lacrimal gland	190	3	0.5	0	—
Brain	191	4	0.7	11	1.8
Other and unspecified parts of nervous system	192	1	0.2	1	0.2
Thyroid	193	1	0.2	6	1.0
Other endocrine glands	194	0	—	1	0.2
Lymph nodes	196	15	2.6	13	2.1
Unknown	199	15	2.6	17	2.7

ICD-O: International Classification of Diseases for Oncology.
* To calculate percent distribution omitting nonmelanoma skin cancer, multiply by 1.11 for men, 1.10 for women.

the women, the major cancers in this young group were breast (18 of 165 cases), cervical (13 of 27), and ovarian (seven of 31) cancer.

Details of Specific Cancer Sites

Female breast These 165 cases constituted one-fourth of all malignancies in Framingham Study women and thus constituted the most common cancer in either sex. Only seven cases had appeared by the time the patients reached the age of 40, and by age 55 only 37 (22% of breast cancers) of the cancers had developed; median patient age at diagnosis was 63. Breast cancer represented 29% of malignancies confirmed in women younger than 50, 29% of those found in women between the ages of 50 and 59, and 25% of those in patients 60 and older.

Tumor location could not be determined from available records in 49 instances. Of the remaining 116 cases, the most common site was upper outer quadrants (42 of 116 = 36%), followed by two adjacent quadrants (36 of 116 = 31%); and when tumor occupied two quadrants, the upper outer quadrant was included in 25 of these 36 instances, so that most breast cancers (67 of 116 known locations) were found in that area.

Ductal carcinoma accounted for 40% of the diagnoses, adenocarcinoma 37% (34 of 61 of them scirrhous). Over the years of the study, an increasing proportion of breast

TABLE 2. Age at First Diagnosis: Most Common Cancers—The Framingham Study: 38-Year Follow-Up

	Men		Women	
Name of primary site	Median	Range	Median	Range
Bladder	67	36–83	68	56–81
Brain	54	50–66	58	24–78
Breast, female	—		63	34–89
Cervix uteri	—		50	38–85
Colon/rectum	70	36–85	68	42–87
Endometrium/corpus uteri	—		61	30–82
Esophagus	73	57–87	—	
Hematopoietic	68	49–87	66	33–91
Kidney	67	48–89	66	56–81
Larynx	64	49–76	—	
Lung	68	49–85	67	49–80
Lymphatic	70	50–80	65	41–89
Ovary	—		57	34–82
Pancreas	68	49–85	70	49–81
Prostate	72	53–86	—	
Skin	67	45–87	68	40–89
Stomach	69	41–85	62	43–93
Unknown	71	54–85	70	51–80

cancers were read by pathologists as ductal carcinoma, whereas the proportion of adenocarcinoma and scirrhous adenocarcinoma decreased; just the former trend was statistically significant at the $P < 0.05$ level. Only six cancers were noninfiltrating: two intraductal carcinoma, three intraductal papillary adenocarcinoma, and one medullary carcinoma. No relation appeared between tumor morphologic characteristics and tumor location within the breast.

Colon and rectum Median patient age for colorectal cancer diagnosis was 70 for the 88 cases in men and 68 for the 106 cases in women. The earliest colon cancer appeared in Framingham men at age 44 and women at 42. A third were found by the time the patients were 64, and median age at diagnosis was 70, in both sexes. The youngest man to have rectal cancer diagnosed was 36, and the youngest woman was 48. Median patient age at diagnosis was 64 in women and 71 in men.

Rectal cancer constituted 25% of all men's colorectal cancers and 22% of women's. Rectosigmoid tumors accounted for almost half of these in men (11 of 23) but only 27% (six of 22) in women. In the colon, the sigmoid colon was the most common tumor location: one-third (29 of 88) of men's colorectal cancers were found there and one-fourth (26 of 106) of women's. Almost another third of colorectal malignancies in both sexes occurred in the cecum and ascending colon. All other parts of the colon were affected infrequently.

The pathologists called all but a few colorectal cancers adenocarcinomas. Without qualifier, this term described these tumors in 56 of the 88 men (64%) and 65 of the 106 women (61%). Mucinous and mucin-producing adenocarcinomas were the diagnoses in another 25% of men (22 of 88) and 21% of women (22 of 106). Adenocarcinomas were found in villous adenomas and in adenomatous polyps in several patients: four of 88 men (4.6%) and nine of 106 women (8.5%). No association was seen between histologic characteristics and colorectal tumor site.

Hematopoietic/lymphatic When all hematologic malignancies were grouped together, median age at diagnosis was 68 years for men and 66 for women. Lymphomas were diagnosed only half as frequently as the rest in men, two-thirds as often in women. No one entity accounted for more than a few cases, except for chronic lymphocytic leukemia (11 of 29 nonlymphomas in men, five of 21 in women).

Lung Ninety-seven men and 32 women had lung cancer develop, at median ages of 68 and 67 years, respectively. The earliest tumors were seen at age 49 in both men and women. This was the most common malignancy in Framingham men.

In 14 men and eight women, the precise location of the primary tumor in the lung could not be determined. Among the rest, most men's cancers (48 of 83, 58%) and almost half of women's (11 of 24, 46%) originated in an upper lobe; the lower lobe accounted for most of the rest (in men, 20 of 83 or 24%; in women, eight of 24 or 33%).

Histologically, squamous cell carcinoma (not further subclassified) was seen most often in both men (31 of 97, 32%) and women (nine of 32, 28%). In men, three others constituted at least 10% of the cases: carcinoma, not otherwise specified by the pathologist (11 of 97, 11%); oat cell carcinoma (ten of 97, 10%); and adenocarcinoma (ten of 97, 10%). In women, adenocarcinoma (five of 32, 16%) and bronchoalveolar carcinoma (four of 32, 13%) were the other relatively prominent types. When all subcategories of squamous cancers and adenocarcinomas were combined, 36% of men and 34% of women had squamous cancer and 20% of men and 38% of women had adenocarcinoma ($P < 0.05$).

There was no association among age at diagnosis, tumor location, and histopathologic findings.

Prostate This tumor was detected in 81 men and was the second most common malignancy among men. It was found as early as age 53, with a median patient age of 72 at diagnosis. In all but two cases, morphologic type was adenocarcinoma.

Skin Skin cancers were identified in 68 Framingham men and 65 Framingham women. The earliest confirmed diagnosis was at age 45 in men and age 40 in women; median patient ages were 67 and 68, respectively.

In men, 48 of the 68 skin cancers (71%) occurred at sites from the neck up, 37 of them on the face and ears.

The distribution in women was similar: 42 of 65 tumors (65%) were on the head and neck and 36 on the face. No cancers appeared on women's ears.

Melanoma accounted for a small proportion of skin cancers: only eight of 68 cases (12%) in men, six of 65 cases (9%) in women. Other serious skin malignancies were seen in three women: one each with hemangiosarcoma, well-differentiated lymphocytic lymphoma, and reticulosarcoma. The rest had an assortment of morphologic types, most of them basal cell carcinoma (68% of men's skin cancers, 60% of women's) and squamous cell carcinoma (15% of men's skin cancers and 20% of women's). Basal cell carcinomas were diagnosed simultaneously at multiple sites in two women and two men; another woman had squamous cell carcinoma in multiple sites.

Comparison With Connecticut Data

Table 3 displays the ratios between FHS age-specific incidence rates and those reported from Connecticut SEER, along with 95% confidence limits for these ratios, for all malignancies except nonmelanoma skin cancer and for the most common individual primary sites for which data exist from Connecticut. Only two disparities appear between the two studies: rectal cancer is less common in both sexes at Framingham than would have been expected from Connecticut figures, and so is cancer of the endometrium.

TABLE 3. Ratio of Framingham Study Cancers to Those Predicted by Connecticut SEER Rates

Primary site	Men	Women
All sites	0.95 (0.87–1.04)*	0.96 (0.88–1.04)*
Bladder	1.04 (0.76–1.40)	0.90 (0.50–1.48)
Brain	0.62 (0.20–1.44)	1.46 (0.73–2.62)
Female breast	—	0.95 (0.81–1.12)
Cervix uteri	—	0.92 (0.57–1.39)
Colon	1.00 (0.77–1.27)	0.96 (0.76–1.19)
Corpus uteri	—	0.70 (0.48–0.98)
Esophagus	0.80 (0.38–1.47)	0†
Kidney	1.14 (0.76–1.80)	0.81 (0.35–1.59)
Lung	0.89 (0.73–1.08)	1.00 (0.70–1.39)
Ovary	—	0.95 (0.63–1.37)
Pancreas	0.85 (0.48–1.37)	0.90 (0.52–1.47)
Prostate	0.97 (0.77–1.20)	—
Rectum	0.57 (0.35–0.87)	0.63 (0.40–0.96)
Stomach	0.86 (0.55–1.28)	0.86 (0.49–1.39)

* Ratio (95% confidence limits).
† No cases.

Discussion

These data demonstrate the relative importance of cancer morbidity, about half that of cardiovascular disease in a free-living (*i.e.*, noninstitutionalized) population. Over 36 to 38 years of follow-up, 1077 of these men and women had a total of 1201 malignancies, whereas in the same cohort 2347 had some manifestation of cardiovascular disease develop (coronary heart disease, congestive heart failure, stroke/transient ischemic attack, intermittent claudication) during this same period.

It is reassuring that the cancer experience of the Framingham cohort, population-based and representative by documentation, so closely approximates that of the SEER population. These results indicate that case ascertainment has been reasonable. Connecticut's most common malignancies are the same as Framingham's: female breast, colon and rectum, lung and bronchus, prostate.[5] A few more cases among the FHS cohort will be found eventually and verified, raising some rates a bit; but it is unlikely that such additions will disturb the comparability of the results. The discrepancies in incidence of rectal cancer, Framingham's being less than two-thirds the expected rate for both women and men, and of endometrial cancer have no obvious explanation. It is possible that these two differences in rates result purely from chance, there being more than 20 comparisons. Armed with details of diet, other habits, and Metropolitan Relative Weight among the two populations, one could attempt an explanation, but that analysis lies beyond the scope of this presentation.[6–8]

As for risk variables within the study, cigarette smoking and alcohol consumption among Framingham women have been modest or even low from the beginning, likely influencing both the gender disparity for cancers of the lung, upper airway, esophagus, and urinary tract and lower overall cancer incidence.[9,10]

Finally, the reliability of the skin cancer data needs to be addressed. Undoubtedly, all forms of cutaneous malignancy must be underdiagnosed, many people ignoring lesions that a physician might consider suspicious. Only in intensive screening programs or clinical studies might one expect to get accurate and complete information; and even then the value of the results would depend on mandatory, not voluntary, examination and on the representative nature of the subjects examined.[11] The high prevalence of skin cancer in those with multiple malignancies may reflect only greater sensitivity to other possible malignancies once one has been discovered.

At least for those sites for which this population has yielded a fair number of cases, this data set is already providing a valuable resource for analysis of potential causative factors, with breast and colon cancer having been our initial focus.[12–14]

REFERENCES

1. Shurtleff D. Some characteristics related to the incidence of cardiovascular disease and death, Framingham Heart Study: 18-year follow-

up. U.S. Department of Health, Education and Welfare publication #74-599; section 30. Bethesda: National Institutes of Health, 1974.

2. Williams RR, Sorlie PD, Feinleib M, McNamara PM, Kannel WB, Dawber TR. Cancer incidence by levels of cholesterol. *JAMA* 1981; 245: 247–252.

3. World Health Organization. International Classification of Diseases for Oncology. Geneva: World Health Organization, 1976.

4. Percy CL, O'Conor GT, Ries LG, Jaffe ES. Non-Hodgkin's lymphomas. Application of the International Classification of Diseases for Oncology (ICD-O) to the Working Formulation. *Cancer* 1984; 54:1435–1438.

5. Young JL Jr, Percy CL, Asire EJ, eds. Surveillance, epidemiology, and end results: Incidence and mortality data, 1973–77. *National Cancer Institute Monograph* 57, 1981; 66–67.

6. Klatsky AL, Armstrong MA, Friedman GD, Hiatt RA. Relations of alcoholic beverage use to colon and rectal cancer. *Am J Epidemiol* 1988; 128:1007–1015.

7. Wynder EL, Escher GC, Mantel N. An epidemiological investigation of cancer of the endometrium. *Cancer* 1966; 19:489–520.

8. Jensen H. Relationship of premorbid state of nutrition to endometrial carcinoma. *Acta Obstet Gynecol Scand* 1986; 63:301–306.

9. Gordon T, Kannel WB, Dawber TR, McGee D. Changes associated with quitting smoking: The Framingham Study. *Am Heart J* 1975; 90: 322–328.

10. Gordon T, Kannel WB. Drinking habits and cardiovascular disease: The Framingham Study. *Am Heart J* 1983; 105:667–673.

11. Scotto J, Fears TR, Fraumeni JF Jr. Incidence of Nonmelanoma Skin Cancer in the United States. U.S. Department of Health and Human Services, NIH publication #83-2433. Bethesda: National Institutes of Health 1983; 1–13.

12. Schatzkin A, Carter CL, Green SB *et al.* Is alcohol consumption related to breast cancer? Results from the Framingham Heart Study. *J Natl Cancer Inst* 1989; 81:31–35.

13. Ballard-Barbash R, Schatzkin A, Carter CL *et al.* Body fat distribution and breast cancer in the Framingham Study. *J Natl Cancer Inst* 1990; 82:286–290.

14. Ballard-Barbash R, Schatzkin A, Albanes D *et al.* Physical activity and risk of large bowel cancer in the Framingham Study. *Cancer Res* 1990; 50:3610–3613.