AMERICAN JOURNAL OF EPIDEMIOLOGY
Vol. 106, No. 1

Printed in U.S.A.

# THE PROBABILITY OF DEVELOPING CANCER

MICHAEL S. ZDEB[1]

In 1956, Goldberg (1) described the probability of developing cancer during two time periods—1940 and 1950. To compute probabilities, data on cancer morbidity and mortality from 1942–1944 and 1949–1951 were combined, respectively, with life-table data for the three year-periods 1939–1941 and 1949–1951. The analysis was based on the population of upstate New York (New York State exclusive of New York City). Ferber (2) updated these probabilities by using similar data centered around 1959. The present analysis uses 1969–1971 cancer morbidity and 1970 life-table data to estimate the current probability of developing cancer, and to compare current probabilities with those presented by Goldberg and Ferber.

## MATERIALS

The cancer incidence and mortality data used for calculation of probabilities are for upstate New York for the period 1969–1971 (Figure 1). The data are for sites 140–209 (8th Revision ICDA). Rates were calculated using the average number of cases and deaths for the three years, and the 1970 census population. The 1970 census data were also used to construct life tables needed for probability calculations. All data are for residents of upstate New York.

Data on cancer incidence were obtained from the New York State Cancer Registry, which collects data provided by physicians, hospitals, and laboratories where surgical specimens are examined. Death certificates also provide data on incidence, i.e., any diagnosis of neoplastic disease listed on the death certificate of a resident of upstate New York is checked to see whether the case had previously been reported. If not, the death becomes a recorded case with the site as listed on the death certificate.

The percent-completeness of reporting of cases to the Registry is estimated to be between 85 and 90 per cent. A study conducted by the US Public Health Service (3) found that the completeness of reporting in 1945 was between 84 and 96 per cent for major site groups, with the average for all sites being 88.7 per cent. Both Goldberg and Ferber cited these figures as evidence of the completeness of reporting in their study periods.

Since 1966, data regarding all multiple primary neoplasms derived from live-report cases have been placed in the Registry. For the period 1969–1971, multiple primaries comprised less than 2 per cent of the cases used to compute incidence rates. Due to this low degree of multiple cases, the probabilities presented later may be interpreted as being for a population free of cancer.

Another feature of the data is the absence of skin as an individual site. In the papers by Goldberg (1) and Ferber (2), skin was reported as a high probability site, but, since 1969, non-melanotic skin cancer cases have not been added to the Registry, and melanotic skin cancer is now included in the category "other and unspecified". Therefore, in this paper when the Goldberg and Ferber data are compared to the current data, the probability of developing

---

[1] Office of Biostatistics, New York State Department of Health, Tower Building, Empire State Plaza, Albany, NY 12237

The author acknowledges the assistance of Dr. William Burnett, Assistant Director, Bureau of Cancer Control, New York State Department of Health, in providing the data from the New York State Cancer Registry, and for providing helpful suggestions. The editorial assistance of A. Donald Emerich and graphics of Charlotte Ronish are also acknowledged.

This study was supported, in part, by Grant No CA 12707 from the National Cancer Institute.

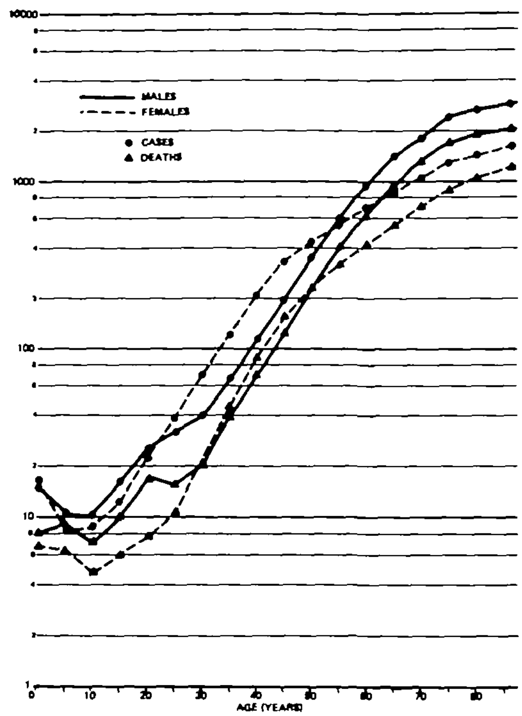


FIGURE 1 Cancer incidence and mortality, average rate per 100,000 population, 1969–1971 (upstate New York)

skin cancer has been removed from the total probability.

The 1970 upstate New York population upon which life-table construction was based was 10,346,021. Approximately 95 per cent of this population was white, with 75 per cent of the total population living in areas that may be described as urban. The abridged life tables constructed for males and females based on this population are shown in Appendix table A1. Chiang's method (4) was used for life-table construction.

## METHOD

The following symbols are used in discussing the method (with $n = 5$ years): $q_x$ = probability of dying in the age interval $x$ to $x + n$; $l_x$ = number of persons surviving to exact age $x$; $d_x$ = number of deaths in age interval $x$ to $x + n$; $L_x$ = life table population in age interval $x$ to $x + n$; $I_x$ = cancer incidence rate in age interval $x$ to $x + n$; $C_x$ = number of cancer cases occurring in age interval $x$ to $x + n$; $M_x$ = death rate in population upon which life table is based in age interval $x$ to $x + n$; $K_x$ = number of non-cancer-caused deaths occurring in the life-table population with cancer; and $t$ = number of cancer cases in life-table population in any chosen age interval.

When differentiation between groups is needed, superscripts are used: no superscript, total; prime, cancer-free; $c$, with cancer.

The number of cancer cases expected to occur in the life-table population was calculated by multiplying the life-table population by the appropriate age-specific cancer-incidence rate,

$$C_x = I_x L_x.$$

This was done for each five-year age group. The number of cancer cases expected to occur among the life-table population from birth through the last age interval (i.e., 85+ years) is the sum of all the age-specific $C_x$. Between any two specific ages, the total number of cancer cases is the sum of all $C_x$ in the chosen interval. These numbers of cases provide the numerators for calculation of probabilities. The denominators are the number of survivors from the original cohort ($l_0$ = 100,000) without cancer at each specific age interval. Thus, two populations were carried through the probability calculations, assuming a cancer-free population at birth ($l_0{}^c = 0$),

$$l_x = l_x{}' + l_x{}^c$$
(total = without + with cancer).

It was assumed that life-table deaths and the probability of dying in a specific age interval could be divided into cancer-related and non-cancer-related deaths and probabilities by using the ratio of the cancer death rate to the death rate from all causes prevailing in the population upon which the life table was based,

$$q_x{}^c = d_x{}^c/l_x,\ d_x{}^c = d_x\,(M_x{}^c/M_x)$$
$$q_x{}' = d_x{}'/l_x,\ d_x{}' = d_x - d_x{}^c.$$

However, as noted by Goldberg (1) $d_x'$ includes not only deaths among persons free of cancer, but also deaths among the cancer population from non-cancer causes. To estimate the number of these non-cancer-related deaths, the following assumptions were made: 1) the cancer-free population has the same risk of dying from causes other than cancer as does the population with cancer, and 2) cancer cases and deaths occur uniformly during the life table interval $x$ to $x + n$.

New cases entering the cancer cohort were considered subject to the risk of dying from other causes for one-half the interval $x$ to $x + n$. Also, those dying from cancer were subject to this risk for half the period. The cancer population exposed to $q_x'$ for any interval was

$$l_x^c + {}^1/_2C_x - {}^1/_2d_x^c,$$

while the number of deaths was

$$K_x = q_x' \, (l_x^c + {}^1/_2C_x - {}^1/_2d_x^c).$$

Once $K_x$ was determined, $d_x$ and $d_x^c$ were corrected as follows:

$$d_x' = d_x' - K_x, \; d_x^c = d_x^c + K_x.$$

This correction properly allocated the non-cancer deaths among the cancer population, with

$$l'_{x+n} = l_x' - C_x - (d_x' - K_x).$$

The denominator appropriate for calculation of the probability over any interval of interest is the $l_x'$ – the cancer free survivors at the start of the interval. For example, probabilities calculated from birth will use $l_0 = 100{,}000$, since $l_0^c = 0$. At age 20 to any endpoint, $l'_{20}$ will be used, etc.

The same method was used to calculate site-specific probabilities. Appendix table A2 contains the populations and cases used for computation of probabilities for the ten leading sites for males and females.

## Results

*Comparisons over time (1950–1960–1970)*

As Goldberg noted, the probability of developing cancer is a function not only of cancer incidence, but also of the risk of dying from non-cancer causes. This risk affects comparisons made over time, since the life tables used for calculating probabilities will change as the mortality experience of the population changes. Thus, if one were to generate probabilities from 1969–1971 cancer data and a 1970 life table, comparisons with probabilities derived from 1949–1951 cancer data and a 1950 life table would be confounded by changes in the life expectancy of the population between 1950 and 1970.

The magnitude of this effect was investigated by applying the 1970 incidence rates to the population represented in 1950 and 1960 life tables. The probabilities (× 100) calculated with and without this adjustment are given in table 1. For males, if the 1950 life table population prevailed in 1970, the probability of developing cancer from birth to age 85+ years for all forms of cancer combined would be 25.35 – 1.78 less than the value of 27.13 calculated using the 1970 population. Using the 1960 population, the adjusted probability is 26.37 – 0.76 less than the 1970 values. Examination of the changes in the life-table populations used in these analyses (Appendix table A1) shows the reason for the adjustment effects. The population surviving into later age intervals has increased substantially. These are the age groups where the incidence of cancer is highest for most cancer sites. With more males at risk at the high-risk ages, part of the rise in the probability of developing cancer from birth to age 85+ must be attributed to changes in life expectancy of the population.

Using the adjusted probabilities, one can determine the per cent increase in probabilities over time attributable to changes in the life tables. Between 1950 and 1970, the overall male probability rose from 18.58 to 27.13. If the 1970 adjusted to 1950 probability is used, the rise was from 18.58 to 25.35. This shows that approximately 20 per cent of the male increase over the 20-year period is attributable to

TABLE 1

*Probability (× 100) of developing cancer, comparisons over twenty-year period*

| Site | 1970 | 1960 | 1950 | 1970 adjusted to | |
|---|---|---|---|---|---|
| | | | | 1960 | 1950 |
| | | | *Males* | | |
| TOTAL | 27.13 | 20.95 | 18 58 | 26.37 | 25.35 |
| Lung | 5.83 | 3.31 | 2.11 | 5.70 | 5.48 |
| Prostate | 3.78 | 2.72 | 2.61 | 3.62 | 3.46 |
| Large intestine | 2.86 | 2.20 | 1.85 | 2.76 | 2.65 |
| Bladder | 1 88 | 1.49 | 1.14 | 1.82 | 1.75 |
| Rectum | 1.54 | 1.30 | 1.22 | 1.49 | 1.43 |
| Stomach | 1.23 | 1.59 | 1.99 | 1 19 | 1.14 |
| Pancreas | 1.05 | 0.79 | 0.62 | 1.02 | 0.98 |
| Kidney | 0.67 | 0.44 | 0.32 | 0.65 | 0.63 |
| Larynx | 0.61 | 0 46 | 0.40 | 0.60 | 0.58 |
| Lymphosarcoma | 0.58 | 0.25 | 0.22 | 0.57 | 0.55 |
| | | | *Females* | | |
| TOTAL | 27.85 | 23.82 | 22.48 | 26.19 | 24.41 |
| Breast | 7.17 | 5.93 | 5 54 | 6.81 | 6.40 |
| Large intestine | 3.79 | 3.04 | 2.65 | 3.49 | 3.20 |
| Corpus uteri | 1.58 | 1.48 | 1 60 | 1.53 | 1.44 |
| Lung | 1.42 | 0.52 | 0.37 | 1.36 | 1.28 |
| Ovary | 1.39 | 1.12 | 0.99 | 1.33 | 1.26 |
| Rectum | 1.37 | 1 02 | 1.02 | 1.26 | 1.16 |
| Cervix uteri | 1.16 | 2 16 | 2.32 | 1.13 | 1.07 |
| Pancreas | 0.92 | 0.70 | 0.53 | 0.84 | 0.77 |
| Stomach | 0.90 | 1.18 | 1 35 | 0.81 | 0.73 |
| Bladder | 0.75 | 0.66 | 0 53 | 0.66 | 0.62 |

changes in the life-table population. A similar comparison between 1960 and 1970 results in 12 per cent of the increase being due to life-table changes.

The adjustment effects are similar for the female population, but of a different magnitude. The 1950 to 1970 rise for females was from 22.48 to 27.85, with an adjusted increase from 22.48 to 24.41. Between 1960 and 1970, the unadjusted increase was from 23.82 to 27.85, and the adjusted to 26.19. Thus approximately 62 per cent of the 1950–1970 increase, and 41 per cent of the 1960–1970 increase, are associated with changes in the life-table population.

Cancer morbidity and life-table effects can also be compared for individual cancer sites. For example, probability of male lung cancer has risen from 2.11 in 1950 to 5.83 in 1970. The adjusted rise is to 5.48 in 1970, showing 90 per cent of the increase to be attributable to increased cancer morbidity. The same proportions also hold for the 1950–1970 rise in female lung cancer. For an individual site among those given for males and females in table 1, female breast cancer is the one most affected by changes in the life-table population. For 1950–1970 and 1960–1970, the per cent increases attributable to life-table changes were 47 per cent and 29 per cent, respectively.

Discussion of the effects of adjustment so far has dealt with probabilities at birth through 85+ years of age. Very little difference is caused by adjustment through the lower age groups. Adjustment begins to effect the probabilities at approximately age 60, with the effects from then on being greater for females than for males. Again, inspection of the life-table populations (Appendix table A1) demonstrates the reason for this effect. Changes in these populations have been minimal at lower ages. Also, at later ages, where most change has occurred, the female increases have been of greater magnitude than those of males.

*Current probabilities at birth*

Table 2 shows the current probabilities (× 100) of developing cancer at birth for various sites. One should remember that comparisons of these figures, and all further 1970 data, to other time periods will be confounded by previously mentioned changes in non-cancer mortality and by differences in the accuracy of diagnoses occurring between time periods. Also, the pattern of age-specific probabilities will be affected by earlier diagnoses of certain cancers. Within the current data, however, these factors present no problem.

The probabilities of table 2 are cumulative from birth, assuming a life span of 85+ years. The total male probability–27.13–is approximately equal to the total female probability of 27.85. Earlier data consistently resulted in a higher probability for females. For males, the leading three sites (lung, prostate, large intestine) account for 46 per cent of the total probability. If the leading ten sites are grouped, they account for about three-quarters of the eventual 27 cases per 100 population.

Many of the leading male sites are also high-probability female sites–i.e., large intestine, lung, rectum, pancreas, stomach and bladder. Breast cancer, with a probability of 7.17, accounts for more than one-quarter of the total probability. The three leading sites (breast, large intestine, corpus uteri) comprise 45 per cent of the total probability. Lung cancer, a lower ranking site in both 1950 and 1960, now is ranked fourth among the leading female sites. The only sites among the top ten in 1970 to have declined over the period 1950–1960–1970 are stomach for both sexes, and cervix uteri for females.

If the cutoff point for determining probabilities at birth is age 65, the overall probability of developing cancer decreases dramatically (table 3). For males, the drop is from 27.13 to 9.96 cases per 100, while females decrease from 27.85 to 11.55. The largest change in probability for an individual site between the 85+ and 65 year endpoints is prostate cancer. To age 65, the probability of developing prostate cancer is only 13 per cent of that through age 85+, a drop from 3.78 to 0.48 cases per 100. Among the leading female sites, none decrease as much as prostate for males. Of the leading sites at birth through age 85+, the probabilities of female stomach and gall bladder cancer decline most, both decreasing to about 19 per cent of their through 85+ value.

The rank order of both male and female probabilities also changes if 65 is taken as the end of the interval from birth. Lung cancer remains the leading male site (one-quarter of the total). Prostate drops from second to fifth, with the order of the leading ten sites remaining almost the same as through 85+. Cancer of the esophagus drops from eleventh to fourteenth, its position being taken by brain cancer.

The order of the leading three female sites remains the same, with breast cancer now accounting for almost one-third of the total probability. Lung cancer drops from a rank of fourth to sixth. The largest drop is for stomach cancer, from ninth to thirteenth, although the difference between the probabilities for sites ranked tenth through fourteenth before age 65 (i.e., skin, thyroid, brain, bladder, and stomach) is not very large.

*Current probabilities at various ages*

The previous section illustrated the effect of changing the endpoint of the interval from birth on the probability of developing cancer. The intervals can also be changed by moving the starting point, and the effect of this modification can be investigated. Table 3 and figure 2 show these changes.

TABLE 2

*Current probability (× 100) of developing cancer from birth to age 85+*

| Males | | | | Females | | | |
|---|---|---|---|---|---|---|---|
| Rank, site & ICDA code (8th revision) | | | Probability | Rank, site & ICDA code (8th revision) | | | Probability |
| | TOTAL | | 27.13 | | TOTAL | | 27.85 |
| 1 | Lung | 162 | 5.83 | 1 | Breast | 174 | 7.17 |
| 2 | Prostate | 185 | 3.78 | 2 | Large intestine | 153 | 3.79 |
| 3 | Large intestine | 153 | 2.86 | 3 | Corpus uteri | 182 0 | 1.58 |
| 4 | Bladder | 188 | 1.88 | 4 | Lung | 162 | 1.42 |
| 5 | Rectum | 154 | 1.54 | 5 | Ovary | 183 | 1.39 |
| 6 | Stomach | 151 | 1.23 | 6 | Rectum | 154 | 1.37 |
| 7 | Pancreas | 157 | 1 05 | 7 | Cervix uteri | 180 | 1.16 |
| 8 | Kidney | 189 | 0 67 | 8 | Pancreas | 157 | 0.92 |
| 9 | Larynx | 161 | 0.61 | 9 | Stomach | 151 | 0.90 |
| 10 | Lymphosarcoma | 200 | 0.58 | 10 | Bladder | 188 | 0.75 |
| 11 | Esophagus | 150 | 0.49 | 11 | Lymphosarcoma | 200 | 0.51 |
| 12 | Lymphatic leukemia | 204 | 0 39 | 12 | Kidney | 189 | 0.41 |
| 13 | Myeloid leukemia | 205 | 0.35 | 13 | Gallbladder | 156 | 0.39 |
| 14 | Brain | 191 | 0.34 | 14 | Genital organs (other) | 184 | 0.34 |
| 15 | Hodgkin's disease | 201 | 0 31 | 15 | Myeloid leukemia | 205 | 0.34 |
| 16 | Skin | 172 | 0.30 | 16 | Skin | 172 | 0.32 |
| 17 | Multiple myeloma | 203 | 0.29 | 17 | Lymphatic leukemia | 204 | 0.32 |
| 18 | Tongue | 141 | 0.24 | 18 | Thyroid | 193 | 0.29 |
| 19 | Liver | 155 | 0.21 | 19 | Multiple myeloma | 203 | 0.28 |
| 20 | Gallbladder | 156 | 0 20 | 20 | Brain | 191 | 0.27 |
| 21 | Lip | 140 | 0.20 | 21 | Hodgkin's disease | 201 | 0.21 |
| 22 | Lymphoid tissue (other) | 202 | 0.20 | 22 | Lymphoid tissue (other) | 202 | 0.19 |
| 23 | Testis | 186 | 0.16 | 23 | Esophagus | 150 | 0.19 |
| 24 | Leukemia (other) | 207 | 0.16 | 24 | Connective tissue | 171 | 0.14 |
| 25 | Connective tissue | 171 | 0.15 | 25 | Liver | 155 | 0.13 |
| 26 | Mouth (other) | 145 | 0.15 | 26 | Leukemia (other) | 207 | 0.13 |
| 27 | Oropharynx | 146 | 0.13 | 27 | Larynx | 161 | 0.12 |
| 28 | Floor of mouth | 144 | 0.13 | 28 | Tongue | 141 | 0.11 |
| 29 | Bone | 170 | 0.10 | 29 | Mouth (other) | 145 | 0.08 |
| 30 | Thyroid | 193 | 0.10 | 30 | Nervous system (other) | 192 | 0.07 |
| 31 | Nervous system (other) | 192 | 0.08 | 31 | Bone | 170 | 0.07 |
| 32 | Salivary glands | 142 | 0.08 | 32 | Salivary gland | 142 | 0 06 |
| 33 | Hypopharynx | 148 | 0.08 | 33 | Small intestine | 152 | 0.06 |
| 34 | Nose, naval cavity | 160 | 0.07 | 34 | Polycythemia vera | 208 | 0.05 |
| 35 | Small intestine | 152 | 0.07 | 35 | Floor of mouth | 144 | 0.05 |
| 36 | Polycythemia vera | 208 | 0.06 | 36 | Myelofibrosis | 209 | 0.05 |
| 37 | Penis | 187.0 | 0.06 | 37 | Nose, nasal cavity | 160 | 0.05 |
| 38 | Breast | 174 | 0.06 | 38 | Oropharynx | 146 | 0.05 |
| 39 | Myelofibrosis | 209 | 0.05 | 39 | Eye | 190 | 0 04 |
| 40 | Gum | 143 | 0.05 | 40 | Gum | 143 | 0.04 |
| 41 | Nasopharynx | 147 | 0.05 | 41 | Monocytic leukemia | 206 | 0.03 |
| 42 | Eye | 190 | 0.04 | 42 | Endocrine glands (other) | 194 | 0.02 |
| 43 | Endocrine glands (other) | 194 | 0.04 | 43 | Nasopharynx | 147 | 0.02 |
| 44 | Monocytic leukemia | 206 | 0.02 | 44 | Lip | 140 | 0.02 |
| | Other and unspecified | | 1.62 | 45 | Hypopharynx | 148 | 0.02 |
| | | | | 46 | Chorionepithelioma | 181 | 0.01 |
| | | | | | Other and unspecified | | 1 90 |

TABLE 3

*Current probability (× 100) of developing cancer, various age intervals*

| Site | At birth | | At 20 years | | At 45 years | | At 65 years |
|---|---|---|---|---|---|---|---|
| | To 85+ | To 65 | To 85+ | To 65 | To 85+ | To 65 | To 85+ |
| *Males* | | | | | | | |
| TOTAL | 27.13 | 9.96 | 27 80 | 10.03 | 28.16 | 9.23 | 26 81 |
| Lung | 5.83 | 2.40 | 6 02 | 2 47 | 6.24 | 2.46 | 5.36 |
| Prostate | 3.78 | 0.48 | 3.91 | 0 50 | 4.16 | 0.52 | 5.16 |
| Large intestine | 2.86 | 0.86 | 2.95 | 0 89 | 3 05 | 0.85 | 3.12 |
| Bladder | 1 88 | 0.58 | 1.95 | 0.60 | 2.01 | 0.58 | 2.03 |
| Rectum | 1.54 | 0.53 | 1 59 | 0 55 | 1.64 | 0.53 | 1.57 |
| Stomach | 1.23 | 0.36 | 1 27 | 0 37 | 1 32 | 0 66 | 1.36 |
| Pancreas | 1.05 | 0.35 | 1 09 | 0 36 | 1.13 | 0.35 | 1.10 |
| Kidney | 0.67 | 0.31 | 0.68 | 0.31 | 0.68 | 0.28 | 0.56 |
| Larynx | 0.61 | 0.32 | 0 58 | 0.29 | 0.55 | 0.24 | 0.44 |
| Lymphosarcoma | 0 58 | 0.30 | 0 58 | 0 29 | 0.55 | 0 24 | 0.44 |
| Esophagus | 0 49 | 0.19 | 0.51 | 0.19 | 0.53 | 0.20 | 0.48 |
| Lymphatic leukemia | 0.39 | 0.16 | 0.35 | 0 11 | 0 35 | 0.10 | 0.37 |
| Myeloid leukemia | 0 35 | 0.16 | 0 35 | 0.15 | 0 30 | 0 18 | 0.14 |
| Brain | 0 34 | 0 25 | 0 32 | 0.22 | 0.28 | 0.18 | 0 14 |
| Hodgkin's disease | 0.31 | 0.24 | 0 31 | 0.23 | 0.17 | 0.10 | 0.11 |
| *Females* | | | | | | | |
| TOTAL | 27 85 | 11.55 | 28.30 | 11.60 | 27.23 | 9.73 | 21.56 |
| Breast | 7.17 | 3.64 | 7 34 | 3.72 | 6 83 | 3.04 | 4 67 |
| Large intestine | 3 79 | 0 94 | 3 88 | 0.96 | 3.98 | 0.92 | 3 77 |
| Corpus uteri | 1 58 | 0.83 | 1.62 | 0.85 | 1.61 | 0 80 | 1.00 |
| Lung | 1.42 | 0 71 | 1 45 | 0.73 | 1.44 | 0.68 | 0 94 |
| Ovary | 1 39 | 0 74 | 1 41 | 0.75 | 1.31 | 0.62 | 0.86 |
| Rectum | 1.37 | 0.42 | 1.40 | 0.43 | 1.42 | 0 40 | 1.26 |
| Cervix uteri | 1.16 | 0 75 | 1 19 | 0.77 | 0 95 | 0 50 | 0.55 |
| Pancreas | 0 92 | 0 21 | 0 94 | 0.22 | 0.96 | 0 21 | 0.93 |
| Stomach | 0.90 | 0 17 | 0.92 | 0.17 | 0 94 | 0.16 | 0 96 |
| Bladder | 0.75 | 0.17 | 0.77 | 0 17 | 0 78 | 0.16 | 0.77 |
| Lymphosarcoma | 0.51 | 0.22 | 0 52 | 0.22 | 0.50 | 0.19 | 0 38 |
| Kidney | 0.41 | 0.15 | 0.40 | 0 14 | 0.40 | 0.13 | 0.34 |
| Gallbladder | 0.39 | 0.07 | 0.40 | 0 07 | 0 41 | 0 07 | 0.43 |
| Genital organs (other) | 0.34 | 0.10 | 0.35 | 0 10 | 0.35 | 0 09 | 0.32 |
| Myeloid leukemia | 0.34 | 0.13 | 0.34 | 0 13 | 0 31 | 0 09 | 0.27 |

In table 3, the probability of developing any form of cancer for a male through age 85+ is highest at age 45, and for a female at age 20. This rise in probability as age increases is due to the higher incidence of cancer at greater ages and the decrease in size of the non-cancer population. Figure 2 shows the changes in probabilities that occur when the beginning and endpoints of the age intervals are varied. Each line running from left to right represents a constant age at the end of an interval. Points on the horizontal axis represent the ages at the start of the intervals. Vertical lines drawn from a starting age to an age at end of interval can be used to determine the probability of developing cancer for any combination of starting and ending ages. Thus, if a male is age 45, his probability of developing cancer by age 65 is read as 9.2 on the vertical axis, while the female probability for the interval 45 to 60 is read as 6.2 (figure 2A).

Figure 2A can be used to examine patterns in change of probability with age for both males and females. The probabilities

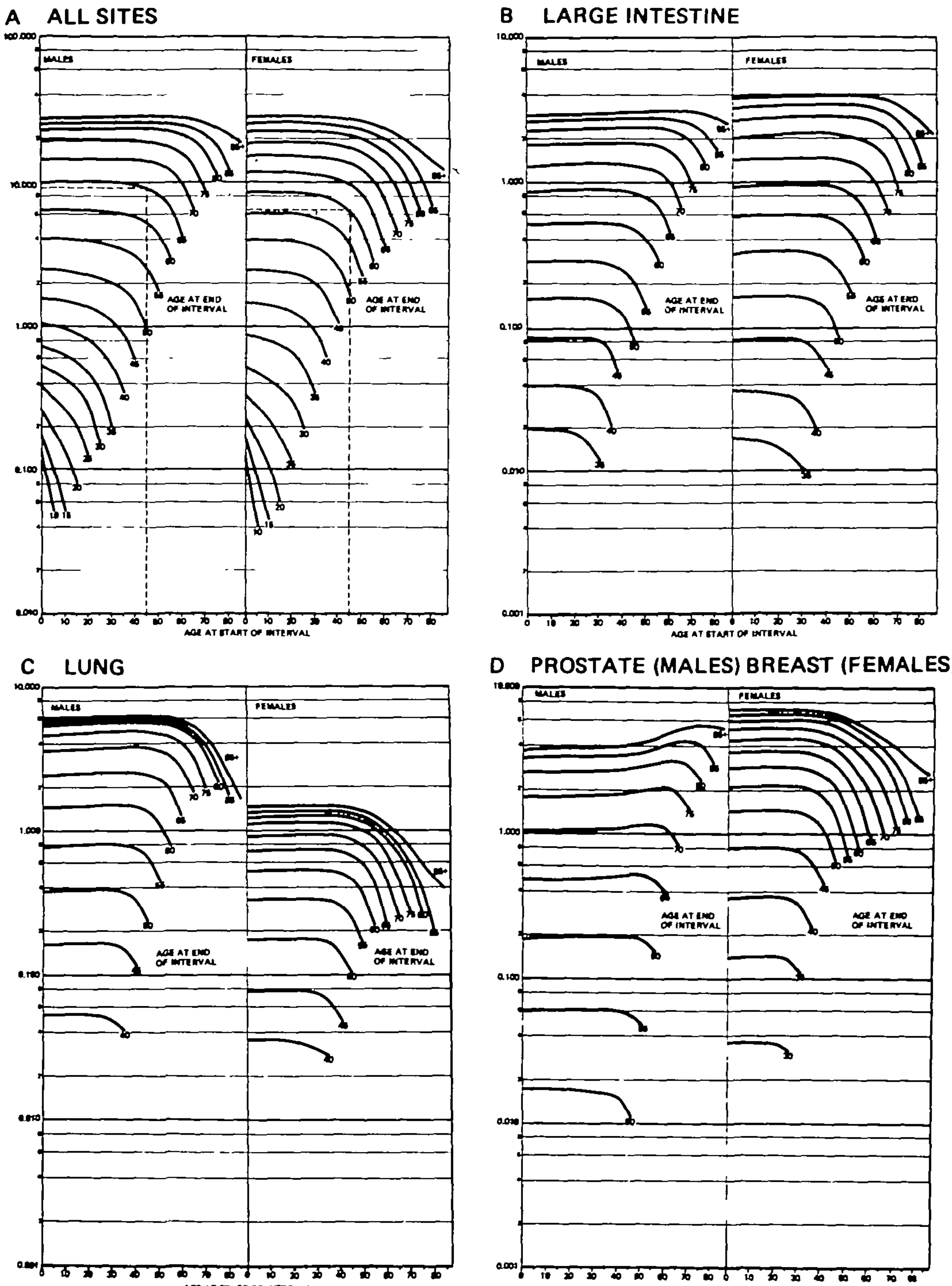


FIGURE 2. Probability × 100 of developing cancer.

plotted represent all possible five-year age-intervals calculated using cases and populations in Appendix table A1. At any specific starting point, for either sex, the probability of developing cancer increases as length of the age interval increases. If the opposite is done—i.e., holding the age at the end of the interval constant and varying the starting point—the pattern of change in probability depends on the endpoint chosen. For example, a male's probability of developing cancer by the time he reaches 65 stays constant at approximately 10.0 until he reaches 35. From then until age 60, the probability continues to drop, to a low point of 4.2 for the interval 60 to 65. This contrasts with selecting an endpoint of age 85, where the probability holds at about 25.0 until age 60, where it begins to decline to a low of 12.0 for the interval 80 to 85.

Until an endpoint of age 30, male and female patterns differ very little. From an endpoint age 35 to one of age 65, female probabilities at birth are consistently greater than those of males, while from ages 70 to 85+ they are approximately equal. At the last three age endpoints, the pattern of change in probability differs between sexes. Female probabilities begin to decline at a starting age 40, while male probabilities do not begin to decline until age 55. The female probabilities also decline at a faster rate. The trends shown in figure 2A are mirrored in the age-specific incidence rates of figure 1.

Several site-specific probabilities are shown in figures 2B through 2D. These sites were chosen since they demonstrate different patterns of probability change with age. The large intestine (figure 2B) demonstrates a site where neither the level nor shape of probability patterns differ greatly between sexes, although female probabilities are greater starting with the endpoint age 60. Lung cancer (figure 2C) differs between sexes in both pattern and level. Figure 2D shows male prostate and female breast cancer. The probability of the former increases with age beginning with an endpoint of 60 (the exception being the last age group on each plotted line). The probability of female breast cancer is shown to be greatest at early ages, declining at intervals that begin approximately at age 30.

REFERENCES

1. Goldberg ID, Levin ML, Gerhardt PR, et al: The probability of developing cancer. J Natl Cancer Inst 17:155–173, 1956
2. Ferber B, Handy VH, Gerhardt PR, et al: Cancer in New York State, Exclusive of New York City, 1941-1960. Albany, NY, Bureau of Cancer Control, New York State Health Department, 1962
3. West MD, Perkins G: A study of the completeness of cancer reporting in New York State. Unpublished study
4. Chiang CL: Introduction to Stochastic Processes in Biostatistics. New York, Wiley, 1968

## APPENDIX

TABLE A1

*1970 abridged life tables for upstate New York (New York State exclusive of New York City)**

| Age (years) X to X + N | Q | SM L | D | LG L | T | E | % change in LG L | |
|---|---|---|---|---|---|---|---|---|
| | | | | | | | 1950 to 1970 | 1960 to 1970 |
| | | | | *Males* | | | | |
| 0–5 | 0 02275 | 100000 | 2275 | 492605 | 6824626 | 68.2 | 0 8 | 0.4 |
| 5–10 | 0.00204 | 97725 | 200 | 488084 | 6332021 | 64.8 | 1 2 | 0.7 |
| 10–15 | 0.00199 | 97525 | 194 | 487178 | 5843937 | 59 9 | 1.4 | 0 8 |
| 15–20 | 0.00595 | 97331 | 579 | 485409 | 5356759 | 55.0 | 1.4 | 0 8 |
| 20–25 | 0 00989 | 96752 | 956 | 481321 | 4871350 | 50.3 | 1 2 | 0.7 |
| 25–30 | 0.00756 | 95796 | 724 | 477167 | 4390029 | 45.8 | 1.1 | 0.6 |
| 30–35 | 0 00759 | 95071 | 722 | 473624 | 3912862 | 41.2 | 1.1 | 0 6 |
| 35–40 | 0.01239 | 94349 | 1169 | 469059 | 3439238 | 36 5 | 1.3 | 0 5 |
| 40–45 | 0.01923 | 93181 | 1792 | 461783 | 2970179 | 31.9 | 1 5 | 0.3 |
| 45–50 | 0.03056 | 91389 | 2793 | 450520 | 2508397 | 27.4 | 2 1 | 0.5 |
| 50–55 | 0.05194 | 88596 | 4601 | 432165 | 2057876 | 23 2 | 3 0 | 0.9 |
| 55–60 | 0.08289 | 83994 | 6962 | 403263 | 1625711 | 19 4 | 4.0 | 1 3 |
| 60–65 | 0.12907 | 77032 | 9943 | 361298 | 1222448 | 15.9 | 5.0 | 1 5 |
| 65–70 | 0.19048 | 67089 | 12780 | 304776 | 861150 | 12 8 | 5.8 | 1.7 |
| 70–75 | 0.26626 | 54310 | 14461 | 236120 | 556374 | 10.2 | 7.2 | 2 2 |
| 75–80 | 0.37295 | 39849 | 14862 | 162834 | 320253 | 8.0 | 8 7 | 2.4 |
| 80–85 | 0 47937 | 24987 | 11978 | 93793 | 157419 | 6 3 | 11.6 | 4.8 |
| 85+ | 1 00000 | 13009 | 13009 | 63626 | 63626 | 4.9 | 20.8 | 17.3 |
| | | | | *Females* | | | | |
| 0–5 | 0.01719 | 100000 | 1719 | 494412 | 7513019 | 75 1 | 0.6 | 0 3 |
| 5–10 | 0.00144 | 98281 | 142 | 491020 | 7018607 | 71.4 | 1.0 | 0 5 |
| 10–15 | 0.00148 | 98139 | 145 | 490361 | 6527587 | 66 5 | 1 0 | 0.5 |
| 15–20 | 0 00274 | 97994 | 268 | 489393 | 6037226 | 61.6 | 1 1 | 0.5 |
| 20–25 | 0.00313 | 97726 | 306 | 487847 | 5547834 | 56.8 | 1 1 | 0.5 |
| 25–30 | 0.00382 | 97420 | 372 | 486167 | 5059986 | 51.9 | 1 2 | 0.4 |
| 30–35 | 0.00468 | 97047 | 454 | 484147 | 4573819 | 47.1 | 1 3 | 0.4 |
| 35–40 | 0.00793 | 96593 | 766 | 481205 | 4089672 | 42.3 | 1.5 | 0.4 |
| 40–45 | 0.01212 | 95827 | 1162 | 476464 | 3608467 | 37 7 | 1.7 | 0.3 |
| 45–50 | 0.01856 | 94665 | 1757 | 469285 | 3132003 | 33 1 | 2.1 | 0.2 |
| 50–55 | 0.02825 | 92908 | 2625 | 458372 | 2662718 | 28 7 | 2.8 | 0.2 |
| 55–60 | 0.04266 | 90283 | 3851 | 442174 | 2204346 | 24 4 | 3 8 | 0.3 |
| 60–65 | 0 06524 | 86432 | 5639 | 418627 | 1762172 | 20.4 | 5.3 | 0.7 |
| 65–70 | 0.10147 | 80793 | 8198 | 384289 | 1343545 | 16.6 | 7.8 | 1.6 |
| 70–75 | 0.15934 | 72595 | 11567 | 334635 | 959256 | 13.2 | 12.2 | 3 4 |
| 75–80 | 0.25080 | 61028 | 15305 | 267640 | 624621 | 10 2 | 19.8 | 7.1 |
| 80–85 | 0.37154 | 45722 | 16987 | 184444 | 356981 | 7.8 | 31.7 | 14.1 |
| 85+ | 1.00000 | 28735 | 28735 | 172537 | 172537 | 6 0 | 67.8 | 41.8 |

* Where $Q$ = proportion dying in interval $X$ to $X + N$; $SM\ L$ = number living at age $X$; $D$ = number dying in interval $X$ to $X + N$; $LG\ L$ = number of years lived in interval $X$ to $X + N$; $T$ = total number of years lived beyond age $X$; and $E$ = observed expectation of life at age $X$.

TABLE A2

*Populations and numbers of cases for calculation of probabilities of developing cancer*

*Males*

| Age (years) X to X + N | Non-cancer population | Total | Lung (162) | Prostate (185) | Large intestine (153) | Bladder (188) | Rectum (154) | Stomach (151) | Pancreas (157) | Kidney (189) | Larynx (161) | Lymphosarcoma (200) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 0–5 | 100000 | 77.8 | 2.2 | 0.4 | 1.1 | 1.1 | 0.4 | 0.0 | 0.4 | 5.9 | 0.0 | 2.9 |
| 5–10 | 97687 | 51.3 | 0.3 | 0.3 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 1.5 | 0.0 | 4.9 |
| 10–15 | 97479 | 51.0 | 0.0 | 0.0 | 0.3 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 4.1 |
| 15–20 | 97268 | 79.5 | 0.7 | 0.3 | 0.0 | 1.3 | 0.7 | 0.7 | 0.3 | 0.3 | 0.0 | 6.4 |
| 20–25 | 96658 | 124.1 | 2.4 | 0.0 | 3.3 | 2.4 | 0.9 | 0.5 | 0.9 | 1.9 | 0.0 | 6.6 |
| 25–30 | 95659 | 152.5 | 2.1 | 1.1 | 3.2 | 3.7 | 2.6 | 0.0 | 1.6 | 3.7 | 1.1 | 5.3 |
| 30–35 | 94856 | 190.3 | 8.2 | 0.6 | 11.7 | 4.1 | 5.3 | 4.1 | 2.9 | 3.5 | 3.5 | 11.7 |
| 35–40 | 94042 | 316.0 | 38.1 | 0.0 | 19.3 | 12.7 | 11.6 | 11.1 | 8.8 | 11.6 | 2.8 | 14.9 |
| 40–45 | 92745 | 536.9 | 108.7 | 4.8 | 46.4 | 30.4 | 26.1 | 14.0 | 11.1 | 23.7 | 19.3 | 24.2 |
| 45–50 | 90744 | 899.6 | 216.3 | 9.9 | 71.9 | 54.1 | 46.6 | 32.9 | 35.3 | 37.1 | 27.7 | 32.4 |
| 50–55 | 87622 | 1536.3 | 394.8 | 42.3 | 121.3 | 87.5 | 90.0 | 64.1 | 60.2 | 53.2 | 55.2 | 50.7 |
| 55–60 | 82529 | 2461.0 | 653.4 | 126.8 | 239.1 | 150.5 | 137.6 | 103.2 | 99.9 | 76.3 | 84.9 | 67.7 |
| 60–65 | 74893 | 3480.3 | 967.8 | 294.3 | 339.1 | 235.4 | 209.8 | 128.8 | 125.9 | 90.3 | 127.6 | 69.3 |
| 65–70 | 64061 | 4331.6 | 1159.4 | 543.9 | 435.6 | 282.9 | 255.7 | 191.2 | 167.3 | 114.2 | 90.3 | 72.4 |
| 70–75 | 50432 | 4445.3 | 1004.0 | 747.2 | 480.7 | 345.4 | 258.9 | 213.6 | 191.6 | 87.9 | 103.0 | 86.5 |
| 75–80 | 35562 | 3955.3 | 752.9 | 830.9 | 481.5 | 304.8 | 235.9 | 203.2 | 160.1 | 87.0 | 46.6 | 59.8 |
| 80–85 | 21012 | 2567.7 | 349.5 | 656.0 | 343.0 | 199.4 | 150.8 | 144.4 | 110.1 | 47.2 | 33.6 | 38.6 |
| 85+ | 10048 | 1877.7 | 164.7 | 524.9 | 257.9 | 167.4 | 102.3 | 116.7 | 77.8 | 23.5 | 18.1 | 23.5 |

*Females*

| Age (years) X to X + N | Non-cancer population | Total | Breast (174) | Large intestine (153) | Corpus uteri (182.0) | Lung (162) | Ovary (183) | Rectum (154) | Cervix uteri (180) | Pancreas (157) | Stomach (151) | Bladder (188) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 0–5 | 100000 | 79.3 | 1.2 | 1.2 | 0.0 | 0.8 | 1.5 | 0.4 | 0.8 | 0.4 | 0.4 | 0.4 |
| 5–10 | 98236 | 40.5 | 0.0 | 0.3 | 0.3 | 0.0 | 0.3 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| 10–15 | 98085 | 43.7 | 0.0 | 0.0 | 0.0 | 0.3 | 1.8 | 0.0 | 0.3 | 0.0 | 0.0 | 0.0 |
| 15–20 | 97919 | 60.9 | 0.0 | 1.4 | 0.7 | 0.3 | 4.1 | 0.0 | 0.3 | 0.0 | 0.7 | 0.3 |
| 20–25 | 97620 | 114.9 | 5.0 | 1.7 | 2.9 | 2.1 | 8.7 | 0.8 | 7.9 | 0.4 | 1.2 | 2.1 |
| 25–30 | 97236 | 192.8 | 31.1 | 2.5 | 2.0 | 1.5 | 14.8 | 1.5 | 36.6 | 1.5 | 1.0 | 1.5 |
| 30–35 | 96724 | 345.3 | 106.8 | 10.0 | 8.3 | 3.3 | 18.4 | 4.4 | 63.4 | 1.1 | 3.3 | 5.6 |
| 35–40 | 96031 | 589.4 | 223.4 | 19.2 | 21.9 | 26.7 | 43.3 | 6.9 | 81.8 | 4.8 | 4.3 | 3.7 |
| 40–45 | 94894 | 1018.7 | 434.5 | 48.6 | 47.7 | 42.5 | 74.1 | 26.4 | 88.3 | 12.8 | 13.2 | 9.4 |
| 45–50 | 93151 | 1541.4 | 638.6 | 79.7 | 95.5 | 94.1 | 110.8 | 40.8 | 114.9 | 14.8 | 15.3 | 15.3 |
| 50–55 | 90593 | 2039.1 | 673.2 | 161.3 | 186.7 | 154.4 | 135.9 | 74.7 | 118.0 | 39.8 | 19.4 | 27.4 |
| 55–60 | 87047 | 2480.1 | 711.3 | 264.4 | 227.3 | 191.9 | 152.1 | 109.0 | 124.4 | 44.2 | 48.1 | 44.8 |
| 60–65 | 82248 | 3005.8 | 810.8 | 351.8 | 233.7 | 193.6 | 174.7 | 152.2 | 112.0 | 95.6 | 62.1 | 58.4 |
| 65–70 | 75600 | 3442.1 | 827.3 | 501.8 | 233.9 | 192.6 | 188.5 | 191.9 | 116.0 | 124.8 | 97.6 | 82.1 |
| 70–75 | 66561 | 3602.3 | 769.8 | 618.0 | 196.4 | 183.3 | 145.6 | 192.3 | 103.7 | 150.4 | 144.2 | 121.5 |
| 75–80 | 54602 | 3589.1 | 758.6 | 657.8 | 164.5 | 149.6 | 142.2 | 205.2 | 90.4 | 171.1 | 164.5 | 113.3 |
| 80–85 | 39614 | 2745.7 | 544.1 | 549.0 | 91.6 | 82.6 | 90.0 | 158.7 | 54.0 | 122.7 | 141.5 | 110.5 |
| 85+ | 24009 | 2918.0 | 631.8 | 520.8 | 69.7 | 99.2 | 82.7 | 200.8 | 49.6 | 132.3 | 179.5 | 151.2 |