Cancer incidence: life table risk versus cumulative risk

Leo J Schouten, Huub Straatman, Lambertus A L M Kiemeney, André L M Verbeek

Abstract
Study objective – To study differences between various methods of estimating cancer risk in individuals.

Design – Information was obtained from a cancer registry. Lifetime risk was calculated using the life table method, taking competing causes of death into account and using well recognised methods such as the cumulative rate and risk.

Setting – Regional cancer registry of the Comprehensive Cancer Centre IKL, located in Maastricht, The Netherlands.

Patients – Patients with a first primary malignancy in 1989–90 in the catchment area of the cancer registry.

Main results – The lifetime risk of developing cancer was estimated to be 34.8% for males and 30.4% for females. In males, especially, the risk estimates were considerably higher when the cumulative risk was used.

Conclusions – The life table method is convenient for estimating the probability that a person will develop cancer during a defined period or during his or her lifetime. The other estimates overestimate the risk of the disease under investigation, especially in the very elderly.

(J Epidemiol Community Health 1994;48:596–600)

When the first annual report of The Netherlands Cancer Registry1 was published in 1992, media attention was focussed on the estimates of the risk of developing cancer. Many newspapers stated that one in every three men and one in every four women would have cancer before the age of 75 years. These estimates were based on calculations of the cumulative risk.2 In essence, cumulative risk is a conditional probability. Because it is conditional on survival until a certain given age, it will overestimate the risk of developing cancer at an advanced age, when there is a greater risk of dying from a competing cause. It is possible to adjust for the effect of competing causes of death, however, by using life expectancy tables. This method was first described in the 1950s,3,4 but has received widespread attention only recently.5

We applied the life table method and the traditional cumulative risk method to the cancer incidence data from a regional cancer registry in The Netherlands to study the differences in risk outcomes.

Methods
CUMULATIVE RISK AND CUMULATIVE RATE VERSUS LIFE TABLE RISK
In the reports of cancer registries, first occurrences of cancer in a given geographical area and in a defined period of time are usually organised into gender and five year age groups. The age groups are 0–4, 5–9, ..., 80–84, 85+.

The total number of incident cancers in age group x is denoted by \( c_x \), \( x = 1, 2, ..., 18 \). Let \( n_x \) denote the number of person-years of age group x. The cumulative risk \( P_{nx} \) until the age of 85 years is defined as:

\[
P_{nx} = 1 - \exp \left( - \sum_{x=1}^{17} \frac{c_x}{n_x} \right)
\]

(1.1)

The cumulative rate \( P_{nx} \) until the age of 85 years is defined as:

\[
P_{nx} = \sum_{x=1}^{17} \frac{c_x}{n_x}
\]

(1.2)

The cumulative rate approximates the cumulative risk for low risk values less than approximately 10%. Both expressions are conditional upon survival. Neither the cumulative rate nor the cumulative risk take competing causes of death in the population into account. A more realistic estimate of the cancer risk can be made if life expectancy tables are available for the study population. In this case a hypothetical cohort of 100 000 individuals is defined at age 0. Subsequently, the number of person years, denoted by \( y_x \), can be calculated in each five year age group of the hypothetical cohort. The total number of individuals who will develop cancer in this hypothetical cohort, \( G_x \), in a particular life period is given by:

\[
G_x = \sum_{i=x}^{85-x} \frac{y_x}{n_x}
\]

(1.3)

The lifetime risk of developing cancer is then given by: \( G_x/100 000 \). The risk \( P_{nx(i)} \) of developing cancer from birth until the end of age group i is:

\[
P_{nx(i)} = \left( \sum_{x=1}^{i} \frac{y_x}{n_x} \right)/100 000
\]

(1.4)

In this study we refer to this method in the expression of life table cancer risk. The risk of developing cancer before the age of 80 years for an individual who is alive and free from cancer at age 40 years is:

\[
P_{nx(40-x)} = \left( \sum_{x=1}^{40-x} \frac{y_x}{n_x} \right)/n_0
\]

(1.5)

where \( n_0 \) is the total number of cancer free individuals at the age of 40 years. This number
Table 1  Summary of the life expectancy table for the IKL area in the period 1989–90

<table>
<thead>
<tr>
<th>At age (y)</th>
<th>Males</th>
<th>Females</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Life expectancy (y)</td>
<td>Persons alive</td>
</tr>
<tr>
<td>0</td>
<td>72.77</td>
<td>100000</td>
</tr>
<tr>
<td>25</td>
<td>49.04</td>
<td>98054</td>
</tr>
<tr>
<td>50</td>
<td>25.43</td>
<td>94081</td>
</tr>
<tr>
<td>75</td>
<td>7.38</td>
<td>52156</td>
</tr>
</tbody>
</table>

Based on specific population and mortality data obtained from the Central Bureau of Statistics

of cancer free individuals at age $x$ can be estimated if the number of cancer deaths in all the age groups of the population under study are available. For details, see Zdeb.  

CALCULATION OF VARIANCE OF THE CANCER RISK

The variance in the total number of individuals who develop cancer from birth until the end of age category $i$ is given by:

$$\text{VAR} \left[ C_{0 \rightarrow i} \right] = \text{VAR} \left[ \sum_{x=1}^{i} \frac{y_{x}c_{x}}{n_{x}} \right]$$

(2.1)

Assuming cancer incidence has a Poisson distribution, the variance of $c_{x}/n_{x}$ is given by:

$$\text{VAR} \left[ \frac{c_{x}}{n_{x}} \right] = \frac{c_{x}}{n_{x}^{2}}$$

(2.2)

Therefore:

$$\text{VAR} \left[ C_{0 \rightarrow i} \right] = \sum_{x=1}^{i} \frac{y_{x}^{2}c_{x}}{n_{x}^{2}}$$

(2.3)

The 95% confidence interval for $C_{0 \rightarrow i}$ is given by:

$$C_{0 \rightarrow i} \pm 1.96 \sqrt{\sum_{x=1}^{i} \frac{y_{x}^{2}c_{x}}{n_{x}^{2}}}$$

(2.4)

Dividing the lower and upper limits of the confidence interval for $C_{0 \rightarrow i}$ by the total number of individuals at birth (100 000), gives the confidence interval for the risk $P_{0 \rightarrow i}$ of developing cancer from birth to the end of age category $i$.

The variance in the cumulative rate based on the Poisson distribution is given by:

$$\text{VAR} \left[ P_{0 \rightarrow i} \right] = \sum_{x=1}^{i} \frac{c_{x}}{(n_{x})^{2}}$$

(2.5)

The 95% confidence interval for $P_{0 \rightarrow i}$ is given by:

$$P_{0 \rightarrow i} \pm 1.96 \sqrt{\sum_{x=1}^{i} \frac{c_{x}}{(n_{x})^{2}}}$$

(2.6)

and for $P_{0 \rightarrow i}$ the 95% confidence interval is given by:

$$1 - \exp \left( -\sum_{x=1}^{i} \frac{c_{x}}{(n_{x})^{2}} \pm 1.96 \sqrt{\sum_{x=1}^{i} \frac{c_{x}}{(n_{x})^{2}}} \right)$$

(2.7)

Interested readers can obtain a computer programme for SAS from the first author to calculate the life table risk.

INCIDENCE AND MORTALITY DATA

Incidence rates were obtained from the regional cancer registry IKL (= Integraal Kankercentrum Limburg (Comprehensive Cancer Centre, Limburg)), located in Maastricht. All malignancies, including non-infiltrating malignancies, were recorded, with the exception of basal cell carcinoma of the skin and non-infiltrating cervical cancer. Completeness of case ascertainment was studied by comparison with an independent registry by general practitioners. From this study the completeness of cancer registration was estimated to be at least 96%.

Incidence rates in the years 1989 and 1990 were used in this study. For total cancer incidence, only the first recorded primary invasive malignancy was considered in the case of multiple primary cancers. For the incidence at specific sites, only the first recorded invasive malignancy at a specific site (according to the 9th revision of the International Classification of Diseases (ICD-9)) was used for calculating the incidence rates. For example, after a woman has developed cancer of the breast, the diagnosis of cancer in the contralateral breast was disregarded, but she remained eligible to develop cancer at a different site.

Figure 1  Age specific cancer incidence rates according to gender. Only first primary malignancies are included, whereas basal cell carcinomas of the skin are excluded (IKL cancer registry, 1989–90)
Table 2  Cumulative risk (%), life table cancer risk (%) and 95% confidence intervals (CI) for cancer all sites, lung cancer, prostate cancer, and breast cancer (IKL, 1989–90)

<table>
<thead>
<tr>
<th>Gender</th>
<th>ICD-9 (site) and method</th>
<th>Age (y)</th>
<th>0–40</th>
<th>0–50</th>
<th>0–60</th>
<th>0–70</th>
<th>0–75</th>
<th>0–80</th>
<th>0–85</th>
<th>Life time</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Males:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>All sites*</td>
<td>Cumulative risk (95% CI)</td>
<td>1.0</td>
<td>2.6</td>
<td>7.1</td>
<td>18.5</td>
<td>27.3</td>
<td>36.9</td>
<td>45.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(95% CI)</td>
<td>(0.8, 1.3)</td>
<td>(2.2, 3.0)</td>
<td>(6.4, 7.8)</td>
<td>(17.4, 19.6)</td>
<td>(25.9, 28.7)</td>
<td>(35.2, 38.6)</td>
<td>(43.0, 47.1)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Life table risk</td>
<td></td>
<td>1.0</td>
<td>2.6</td>
<td>7.1</td>
<td>18.5</td>
<td>27.3</td>
<td>36.9</td>
<td>45.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(95% CI)</td>
<td>(0.8, 1.3)</td>
<td>(2.2, 3.0)</td>
<td>(6.4, 7.8)</td>
<td>(17.4, 19.6)</td>
<td>(25.9, 28.7)</td>
<td>(35.2, 38.6)</td>
<td>(43.0, 47.1)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>162 lung and bronchus</td>
<td>Cumulative risk (95% CI)</td>
<td>0.0</td>
<td>0.3</td>
<td>1.5</td>
<td>4.9</td>
<td>7.3</td>
<td>9.2</td>
<td>10.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(95% CI)</td>
<td>(0.0, 0.0)</td>
<td>(0.2, 0.5)</td>
<td>(1.3, 1.9)</td>
<td>(5.1, 6.5)</td>
<td>(8.4, 10.4)</td>
<td>(12.1, 14.8)</td>
<td>(14.4, 17.7)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Life table risk</td>
<td></td>
<td>0.0</td>
<td>0.3</td>
<td>1.5</td>
<td>4.9</td>
<td>7.3</td>
<td>9.2</td>
<td>10.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(95% CI)</td>
<td>(0.0, 0.0)</td>
<td>(0.2, 0.5)</td>
<td>(1.3, 1.9)</td>
<td>(5.1, 6.5)</td>
<td>(8.4, 10.4)</td>
<td>(12.1, 14.8)</td>
<td>(14.4, 17.7)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>185 prostate</td>
<td>Cumulative risk (95% CI)</td>
<td>0.0</td>
<td>0.0</td>
<td>0.3</td>
<td>2.0</td>
<td>4.1</td>
<td>7.0</td>
<td>10.4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(95% CI)</td>
<td>(0.0, 0.0)</td>
<td>(0.0, 0.0)</td>
<td>(0.1, 0.4)</td>
<td>(1.6, 2.4)</td>
<td>(3.3, 4.8)</td>
<td>(5.9, 8.0)</td>
<td>(8.8, 12.0)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Life table risk</td>
<td></td>
<td>0.0</td>
<td>0.0</td>
<td>0.3</td>
<td>2.0</td>
<td>4.1</td>
<td>7.0</td>
<td>10.4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(95% CI)</td>
<td>(0.0, 0.0)</td>
<td>(0.0, 0.0)</td>
<td>(0.1, 0.4)</td>
<td>(1.6, 2.4)</td>
<td>(3.3, 4.8)</td>
<td>(5.9, 8.0)</td>
<td>(8.8, 12.0)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Females:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>All sites*</td>
<td>Cumulative risk (95% CI)</td>
<td>1.4</td>
<td>4.1</td>
<td>8.4</td>
<td>15.1</td>
<td>19.6</td>
<td>24.0</td>
<td>27.8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(95% CI)</td>
<td>(1.1, 1.7)</td>
<td>(3.6, 4.6)</td>
<td>(7.7, 9.2)</td>
<td>(14.1, 16.1)</td>
<td>(18.5, 20.7)</td>
<td>(23.5, 26.1)</td>
<td>(28.8, 31.9)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Life table risk</td>
<td></td>
<td>1.4</td>
<td>4.1</td>
<td>8.4</td>
<td>15.1</td>
<td>19.6</td>
<td>24.0</td>
<td>27.8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(95% CI)</td>
<td>(1.1, 1.7)</td>
<td>(3.6, 4.6)</td>
<td>(7.7, 9.2)</td>
<td>(14.1, 16.1)</td>
<td>(18.5, 20.7)</td>
<td>(23.5, 26.1)</td>
<td>(28.8, 31.9)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>162 lung and bronchus</td>
<td>Cumulative risk (95% CI)</td>
<td>0.0</td>
<td>0.2</td>
<td>0.6</td>
<td>1.0</td>
<td>1.5</td>
<td>1.7</td>
<td>1.9</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(95% CI)</td>
<td>(0.0, 0.1)</td>
<td>(0.1, 0.3)</td>
<td>(0.3, 0.8)</td>
<td>(0.7, 1.3)</td>
<td>(1.1, 1.8)</td>
<td>(1.3, 2.1)</td>
<td>(1.4, 2.3)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Life table risk</td>
<td></td>
<td>0.0</td>
<td>0.2</td>
<td>0.5</td>
<td>0.9</td>
<td>1.3</td>
<td>1.5</td>
<td>1.6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(95% CI)</td>
<td>(0.0, 0.1)</td>
<td>(0.1, 0.3)</td>
<td>(0.3, 0.7)</td>
<td>(0.7, 1.2)</td>
<td>(1.0, 1.6)</td>
<td>(1.1, 1.8)</td>
<td>(1.2, 1.9)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>174 breast</td>
<td>Cumulative risk (95% CI)</td>
<td>0.4</td>
<td>1.7</td>
<td>3.3</td>
<td>5.1</td>
<td>6.6</td>
<td>7.9</td>
<td>9.4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(95% CI)</td>
<td>(0.2, 0.5)</td>
<td>(1.3, 2.0)</td>
<td>(2.8, 3.6)</td>
<td>(4.7, 6.0)</td>
<td>(5.9, 7.4)</td>
<td>(7.1, 8.7)</td>
<td>(8.4, 10.4)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Life table risk</td>
<td></td>
<td>0.4</td>
<td>1.7</td>
<td>3.3</td>
<td>5.1</td>
<td>6.6</td>
<td>7.9</td>
<td>9.4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(95% CI)</td>
<td>(0.2, 0.5)</td>
<td>(1.3, 2.0)</td>
<td>(2.8, 3.6)</td>
<td>(4.7, 6.0)</td>
<td>(5.9, 7.4)</td>
<td>(7.1, 8.7)</td>
<td>(8.4, 10.4)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* Excluding basal cell carcinoma of the skin.

Population and mortality data (both in one year age groups) and cancer mortality data (in five year age groups) from 1989–90 for the catchment area of the IKL cancer registry were obtained from the Dutch Central Bureau of Statistics. A life expectancy table was constructed using the population and mortality data (summarised in table 1).

**Results**

For males, the life expectancy at birth in the period 1989–90 in the area of the IKL cancer registry was 72.8 years. The life expectancy for females was 79.5 years. (see also table 1).

The annual age adjusted incidence rate per 100 000 person years (according to the European Standard Population15) was 392.6 for males and 273.3 for females (first primary malignancies only). In figure 1 age specific incidence rates according to gender are presented.

Table 2 shows the cumulative cancer risk and the life table cancer risk for all malignancies and for a number of common cancer sites. Using the cumulative risk, the risk of males developing cancer up to the age of 85 years was estimated to be 45.1%. The cumulative risk of females was estimated to be 30.4%. The life table cancer risk up to this age was estimated to be 33.3% for males and 27.8% for females. The cumulative risk estimate was much higher than this.
Life table risk estimates do not differ significantly until the age of 85 years.

In figure 3, the probability of eventually developing cancer from a given age onwards (for people free of cancer at the beginning of the given interval) is shown for males and females. In males, after a slight increase, the probability remains stable until the age of 50 years. Beyond this age the probability of developing cancer declines rapidly. In females, the probability is constant until the age of 40 years. Thereafter the decline in probability is less rapid than for males.

Figure 4 shows the probability of developing prostatic cancer in males and breast cancer in females. The difference is striking. The probability of developing prostatic cancer rises from 5-5% at birth to 6-1% at the age of 60 years, but declines sharply thereafter. In contrast, the risk of developing breast cancer starts to decline from the age of 30 years onwards. The risk of developing breast cancer from birth is estimated to be 8-4%.

### Discussion

More realistic cancer risk estimates were obtained using the life table method. There were large discrepancies with the other available methods – that is, cumulative rate and cumulative risk for the elderly – especially in males. The life table method produced a lower estimate because it is not conditional upon survival to an advanced age.

There was less difference between the various estimates for women because at all ages the mortality rates for women are lower than those for men. As a consequence, the male:female ratio based on the life table cancer risk is higher than the ratios based on the other measures commonly used in descriptive epidemiology, for example, the age adjusted incidence rates or the cumulative risk. For example, the male:female ratio for the age adjusted incidence rate according to the European Standard Population is 1:43, while the ratio for the life table cancer risk is estimated to be 1:14.

When calculating the individual risk, we considered only the first primary invasive malignancy to be eligible. In 1989 and 1990 together, 7% of all the invasive malignancies recorded by the IKL cancer registry were second or subsequent cancers. When these second primary malignancies were included in the estimates, the lifetime risk increased from 34-8% to 38-0% for males and from 30-4% to 32-7% for females.

All life table risk estimates are based on incidence rates. Theoretically, incidence rates are calculated as the ratio between the number of new cancer cases and the number of person-years free from cancer. However, a population derived from a population register will include prevalent cancer cases. Therefore, the true incidence of cancer will be underestimated.

This phenomenon affects the cumulative rate and risk. Thus, although these methods over-estimate the risk (because they are conditional on survival), the estimates would be even higher if they were based on prevalence adjusted in-
cision rates. Feuer et al recently suggested a method for correcting these prevalent cases when calculating the life table cancer risk. They calculated the number of expected cancer cases in the cohort as being the product of the age specific incidence rates and the number of persons in the cohort alive and free from cancer. The correction for prevalent cancer cases was performed by multiplication by the ratio R of the number alive at age x and the number alive and free from a given type of cancer at age x. A reanalysis of our data using this correction for prevalent cases did not, however, change our results significantly. This can be explained by an inherent correction for prevalent cases with the life table method. The number of expected cancer cases is calculated as the product of the (underestimated) incidence rates and the number of person years. In these person years, the prevalent cancer cases are also included.

It is possible to calculate the lifetime risk of developing cancer using the life table method because the number of person years and the cancer incidence rates in the very elderly are known. The lifetime risk represents the average risk at birth that a member of the cohort will develop cancer during his or her lifetime.

It seems remarkable that the risk of developing cancer during an individual's remaining lifetime rises until a certain age (see figs 3 and 4). This can be explained by the fact that the number of cancer cases (the numerator) decreases only slowly with increasing age early in life (most of the cancer cases occur at advanced ages), while the number of people alive and free from cancer (the denominator) decreases rapidly with increasing age. This phenomenon is especially notable for prostatic cancer, where the risk of developing cancer during the remaining lifetime rises until the age of 60 years.

Although the life table method has clear advantages over the cumulative risk method, it cannot replace the other two epidemiological measures. Firstly, life expectancy tables are not available for all populations. Secondly, age adjusted incidence rates, the cumulative risk and the cumulative rate, are better suited for comparing the cancer incidence rates of different populations in different time intervals and at different locations. The fact that these measures are not adjusted for life expectancy means that they are more suitable for making direct comparisons of cancer occurrence between different populations. For example, an increasing life expectancy (fewer deaths from competing causes) with stable age specific incidence rates will result in a higher estimate of the lifetime risk calculated with the life table, but will not influence the other risk estimates.

In conclusion, the life table method is a convenient method for estimating the probability that a person will develop cancer during a defined period or during his or her lifetime. The life table method is based on the assumption that the current death and cancer incidence rates will be maintained in the future, thus it shares this basic assumption with the other risk estimates. Because the life table methods include life expectancy, it is a good method of estimating the 'cancer burden' in a population. Other risk estimates overestimate the risk of developing any given disease, especially at advanced ages. The life table method cannot replace age standardised incidence rates.

We are grateful to the Central Bureau of Statistics and the Comprehensive Cancer Centre IKL, (Integraal Kankercentrum Limburg) for providing the data. The Comprehensive Cancer Centres IKL and IKO (Integraal Kankercentrum Oost) are acknowledged for providing financial support. We thank J. Altena-Hill for linguistic comments.

Addendum
List of notations:
\[ C_x = \text{number of incident cancer cases in age group } x, \]
\[ C_n = \text{number of persons who develop cancer in the hypothetical cohort in a defined life period,} \]
\[ I_x = \text{life table population at age } x \text{ alive and free from cancer.} \]
\[ n_x = \text{number of inhabitants in the area of cancer registry in age group } x, \text{ obtained from the Bureau of Statistics.} \]
\[ P_x = \text{cumulative rate.} \]
\[ P_x = \text{cumulative risk.} \]
\[ P_x = \text{life table risk of developing cancer in a defined period.} \]
\[ y_x = \text{total person years in the hypothetical cohort in age group } x. \]