Chapter 2

Statistical methods for the analysis of cancer survival data

R.J. Black¹ and R. Swaminathan²

¹International Agency for Research on Cancer
Lyon,
France

²Cancer Institute (WIA)
Chennai (Madras) 600036
India

Introduction

In the context of population-based cancer registry data, the aim of survival analysis is to estimate the probability of survival, expressed as time elapsed since diagnosis, for individuals within groups defined by, say, type of cancer diagnosed, sex, age and place of residence. Even though it is common practice to use the term ‘survival rate’ to describe this quantity, it is important to realize that we seek to estimate an individual probability rather than a ‘rate’. Despite this, the terms ‘survival rate’, ‘survival probability’ and simply ‘survival’ are used interchangeably in this publication, since these will be commonly understood as having the same meaning by many readers. This chapter sets out some basic concepts in survival analysis, and describes how these have been used to estimate the survival of subjects from the developing-country populations included in the study. The methods of analysis used are essentially the same as those of the EUROCARE study (Berrino et al., 1995), the only other recent systematic comparative analysis of survival in a number of cancer registry populations. A comprehensive description of survival analysis methods for cancer registry data is given by Estève et al. (1994).

Follow-up of subjects in a survival study

If we were able to obtain complete follow-up information for each individual in a group under study, then the probability of survival during a time period \( t \) could be estimated simply from the proportion of survivors at the end of the period among all subjects alive at the beginning of \( t \). It would be sufficient to know each individual’s survival status at the beginning of the period, \( t_i \), and at its end, \( t_i + t \). With cancer registry data, we are usually concerned with periods of elapsed time between the date of incidence and some fixed point of follow-up time, such as five years after the date of incidence (‘five-year survival’). In practice, follow-up of persons registered with cancer is not complete, either because subjects become ‘lost to follow-up’ during the period \( t \) (say by moving out of the area of surveillance of the cancer registry), or because the end of the period of possible follow-up by the registry occurs before the end of \( t \). This is illustrated in Fig. 1, which shows follow-up of three subjects A, B and C.

Fig. 1 is divided into two parts. The first shows follow-up of the three subjects in terms of calendar time, while the second shows the same information in terms of the duration of follow-up. It will be noted that the period of registration is less than the follow-up period. This is typical of the data analysed in this publication and other cancer registry datasets. Subject A is diagnosed with cancer during the first year of the period of registration, with the date of incidence shown as \( i \), and dies between \( y_j \) and \( y_k \), shown as \( d \), which is within the period of possible follow-up by the registry. In the second part of the figure, this is shown in terms of the duration of follow-up, as three units of time \( t \) between incidence and death. Subject B is diagnosed at the beginning of \( y_1 \) but is lost to follow-up (LF) between \( y_2 \) and \( y_3 \), after a duration of follow-up 1.5 units of \( t \). Finally, subject C is diagnosed between \( y_1 \) and \( y_2 \) and is still alive at the end of the follow-up period of the study \( y_4 \); subject C is thus said to be ‘withdrawn alive’ after 2.5 units of follow-up time \( t \). It will be seen that the characters \( d \), LF and \( w \) have been replaced with the values 1, 0 and 0, respectively, in the second part of the figure. This reminds us that, when we come to enumerate the number of deaths during follow-up, only subject A’s death is known to us. Subjects B and
C have incomplete follow-up and will be censored from the analysis at the point in follow-up time at which they were either lost to follow-up or withdrawn. We are not aware of the deaths of subjects B and C, but we can use the information that they did not die during the period in which they were being followed up in estimating the probability of survival for the study group as a whole. This, indeed, is the key to formal survival analysis.

In practical terms, to prepare data for survival analysis, we require the time elapsed between the date of incidence and the date of death or date of loss to follow-up or date of withdrawal for each individual in the group under study, whichever occurs first. The accuracy of these survival times calculated from cancer registry data depends on the method of follow-up used by the registry. Some registries employ passive follow-up, which relies on notifications of deaths of cancer patients to the registry by national statistical organizations. Other registries use active follow-up, in which information on the survival status of patients is sought by the registry at fixed points in time after the date of incidence, usually on the anniversaries of this date. Active follow-up by cancer registries can be achieved by using clinical follow-up systems, by contacting patients' physicians, or by contacting the patients or their families directly by means of postal enquiries or even home visits. In the present study, most registries used a mixture of active and passive methods. Typically, registries undertook special follow-up exercises specifically for this study, in order to augment incomplete passive notification systems. The precise methods used by each registry are detailed in their respective chapters. The mixture of active and passive follow-up methods means that the data for analysis were composed of exact survival times for some subjects and less precise data representing cases censored as a result of follow-up enquiries.

<table>
<thead>
<tr>
<th>Interval (years)</th>
<th>Alive at beginning of interval</th>
<th>Last known alive during interval (censored)</th>
<th>No. of deaths during interval</th>
<th>Effective no. at risk</th>
<th>Conditional probability of death</th>
<th>Conditional probability of survival</th>
<th>Cumulative prob. of survival (to end of year)</th>
</tr>
</thead>
<tbody>
<tr>
<td>t-t_{0}</td>
<td>n_i</td>
<td>w_i</td>
<td>d_i</td>
<td>N_i</td>
<td>q_i</td>
<td>s_i</td>
<td>S_{si}</td>
</tr>
<tr>
<td>0-1</td>
<td>3289</td>
<td>166</td>
<td>365</td>
<td>3206.0</td>
<td>0.114</td>
<td>0.886</td>
<td>0.886</td>
</tr>
<tr>
<td>1-2</td>
<td>2758</td>
<td>275</td>
<td>301</td>
<td>2620.5</td>
<td>0.115</td>
<td>0.885</td>
<td>0.784</td>
</tr>
<tr>
<td>2-3</td>
<td>2182</td>
<td>37</td>
<td>278</td>
<td>2163.5</td>
<td>0.128</td>
<td>0.872</td>
<td>0.683</td>
</tr>
<tr>
<td>3-4</td>
<td>1867</td>
<td>30</td>
<td>191</td>
<td>1852.0</td>
<td>0.103</td>
<td>0.897</td>
<td>0.613</td>
</tr>
<tr>
<td>4-5</td>
<td>1646</td>
<td>20</td>
<td>106</td>
<td>1636.0</td>
<td>0.065</td>
<td>0.935</td>
<td>0.573</td>
</tr>
</tbody>
</table>
Estimation of survival

There are two related approaches to the estimation of survival: the Kaplan-Meier and the actuarial, or life table, methods. The former is particularly useful when exact survival times are available, since smooth estimates of survival as a function of time since diagnosis can be obtained. Given the data available, and in order to achieve consistency with other studies, the actuarial method was used in the present study of survival in developing countries. The actuarial method involves the construction of a life table, which permits the calculation of the cumulative probability of survival at time \( t_{n+1} \) from the conditional probabilities of survival during consecutive intervals of follow-up time up to and including \( t_{n} \). The layout and method of calculation of the elements of a life table are shown in Table 1.

For each time period \( t_i \) to \( t_{i+1} \), \( n_i \) is the number of subjects at risk of death at the beginning of the interval. The number of cases censored during the interval, because they were lost to follow-up or withdrawn alive at the end of the follow-up period, is shown as \( w_i \). The symbol \( d_i \) denotes subjects who died during each interval. Values for \( n_i \), \( w_i \) and \( d_i \) can be obtained directly from the survival times and outcomes (1=death, 0=censored) from the data set out in Fig. 1. The effective number of subjects at risk during each interval is calculated as

\[
N_i = n_i - \frac{w_i}{2}
\]

In this way, subjects who were alive and at risk of death during the interval \( t_i \) to \( t_{i+1} \), but who were censored at some point during the interval, are assumed to have been followed up for, on average, half of the interval. Having estimated the effective number of subjects at risk, it is possible to calculate the probability of death during the interval from

\[
q_i = \frac{d_i}{N_i}
\]

The probability of survival during the interval beginning \( t_i \) is then calculated as

\[
s_i = 1 - q_i
\]

from which the cumulative probability of survival up to time \( t_{i+1} \) is derived from the product of the \( s_i \)

\[
S_{i+1} = \prod_{j=0}^{i} s_j
\]

The final quantity estimated, \( S_{i+1} \), is often multiplied by 100 to give the ‘percentage survival’ at time \( t_{i+1} \). In the context of the present study, use of the actuarial method has the advantage that information from all cases is used in the estimation of survival, including cases which were lost to follow-up at the point of time at which they were last ‘known to be alive’ by registries using active follow-up methods. However, the estimation of the effective number of subjects at risk in the actuarial method assumes that censored cases are actually followed up for, on average, half the length of a given interval and that such cases are subject to the same probability of death as the cases with complete follow-up during the same interval. In the analysis of registry data, these assumptions may be invalid. For example, at the beginning of follow-up, when cases are under short-term clinical surveillance but are then lost to follow-up by the registry, the average survival times of the censored cases may be less than half of the length of the first interval, as assumed. Furthermore, the true probability of death of the censored cases may be greater than assumed if cases of poor prognosis are more likely to be lost to follow-up.

Relative survival

The above method of calculating observed survival relates to deaths from all causes among the group of cancer patients under follow-up. However, cancer patients are at risk of death both from the cancer with which they have been diagnosed and from other causes of death. The observed survival is therefore influenced by mortality both from the cancer of interest and from other causes. Indeed, the presence of a tumour may increase an individual’s risk of death from other causes. If we wish to compare survival in groups which are heterogeneous in terms of their risk of death from causes other than a particular cancer of interest, then observed differences between the groups concerned may be due in part to variations in the risk of death from these other causes, rather than the risk from the cancer under study. Estève et al. (1990) describe ‘net survival’, which is the survival which would pertain if deaths from other causes did not occur. In other words, net survival is the inverse of cause-specific mortality. One way of estimating the net survival is to censor cases at the point of death from causes other than the cancer of interest. This is called the ‘corrected survival’. In the present study, and indeed in many other contexts, information about the cause of death of cancer patients is not available. This may be because of incomplete follow-up of all subjects, or because
the death certification system is not sufficiently accurate to discriminate between deaths due to the cancer under study and deaths due to other causes. Recognition of this problem has led to the development of 'relative survival' methodology. The relative survival at the end of an interval beginning at \( t_i \) is defined as

\[
R_{i+1} = \frac{S_{i+1}}{S_{i+1}^*}
\]

where \( S_{i+1} \) is the observed survival for subjects with a particular cancer and \( S_{i+1}^* \) is the expected survival of a group of individuals with the same demographic characteristics who are at risk of death only from causes of death other than the cancer under study (Ederer et al., 1961). As long as the cancer of concern does not make a large contribution to overall mortality, we can estimate the probability of survival, in the absence of this cancer, of an individual of a given age from \( t_i \) to \( t_{i+1} \) from life tables based on general population mortality. Life tables present age-specific probabilities of surviving intervals of age, usually of one year, given an initial age \( x \). Separate tables for males and females are generally available. The probability of survival at \( t_{i+1} \) of each subject \( h \) alive at \( t_i \) is calculated as

\[
e_h(t_{i+1}) = I_{x+0.5} / I_{x+0.5}
\]

where \( I_{x+0.5} \) is the average of the values of the survival function at ages \( x \) and \( x+1 \) from the life table and \( t \) represents the units of time elapsed between \( t_i \) to \( t_{i+1} \). Age \( x+0.5 \) is used to find the baseline value for the survival function, since the average of cancer patients who have attained age \( x \) will be closer to \( x+0.5 \) than exactly \( x \). Similar adjustments can be made when using abridged life tables in which values of the survivor function are given at intervals of five years of age. The overall expected survival of the group of subjects under study at \( t_{i+1} \) is obtained from expected number of survivors at \( t_{i+1} \) divided by the number of subjects alive at \( t_0 \).

\[
S_{i+1}^* = \frac{\sum_{h=1}^{n_0} e_h(t_{i+1})}{n_0}
\]

In the present study, detailed general population mortality life tables were not generally available. Instead, we relied on published life tables which were representative of the period of registration of the subjects under study, or generic life tables for developing countries published by the United Nations (UN, 1982). The life tables used for each population are detailed in the relevant chapters.

The above method of estimating expected survival has some important limitations, particularly in analysing long-term survival. There are three possible end-points for each subject: death from the cancer under study, death from another cause, and withdrawal from the study due to, for example, emigration from the cancer registry area. Normally we are obliged to estimate survival in quite large groups, which may be heterogeneous with respect to the probability of these events occurring. This means that the composition of a group changes over time (since subjects with relatively high probabilities of death or withdrawal would tend to be removed from the study before others), and the expected survival of all subjects alive at the beginning of follow-up becomes less representative of the surviving subjects. In the context of cancer registry data, age is commonly associated with risk of death from the cancer of concern (young patients may be more likely to withstand treatment), risk of death from other causes (young patients may be less likely to die from an unrelated condition such as heart disease), and the probability of becoming a censored case (young patients may be more likely to emigrate from the cancer registry area).

Despite our concern in the present study with survival up to only five years, we have used a computer package which takes account of heterogeneity in expected survival and withdrawal of subjects (Hakulinen et al., 1994). Specifically, the method used was that of Hakulinen (1982), both for overall and for age-specific and sex-specific estimates of relative survival.

**Age-standardized survival**

It is important to realize that, when comparing survival in different groups, the method of relative survival takes account of variations in the age structure of the groups only to the extent that age is correlated with risk of death from causes other than the cancer under study. For many types of cancer, the risk of dying as a result of the cancer itself is clearly associated with the subject’s age at diagnosis. In the present study we were interested in comparing survival in developing and developed country populations, in which the age structures of cancer patient groups are grossly different. For this reason, we used direct standardization of age-specific relative survival estimates to derive an overall summary statistic, *age-standardized relative survival (ASRS)*

\[
ASRS_i = \frac{\sum_x r_{x, w_x}}{\sum_x w_x} - 1
\]
where $r_{ix}$ are age-specific relative survival estimates at the end of the follow-up period $t_i$ and $w_i$ are age-specific proportions from the World Standard Cancer Patient Population for the appropriate site of cancer (see Chapter 3).

References


UN (1982) Model Life Tables For Developing Countries (Population Studies No. 77). New York, United Nations Department of International Economic and Social Affairs
Chapter 3

World standard cancer patient populations: A resource for comparative analysis of survival data

R.J. Black and S.A. Bashir

International Agency for Research on Cancer
Lyon,
France

Introduction
As noted in Chapter 2 ‘Statistical methods for the analysis of cancer survival data’, cancer patient survival is influenced by age in two ways: the risk of dying as a result of the cancer with which a patient has been diagnosed tends to be greater for elderly persons, and elderly subjects tend to be at greater risk of death from other causes. In comparing two groups of subjects, one of which has a larger proportion of elderly patients than the other, the relatively greater risk of death from causes other than the cancer under study will be reflected in a lower value for expected survival in the older group. However, relative survival is determined by deaths from the cancer under study as well as other causes of death. In these circumstances, comparisons of relative survival within age bands can be recommended. However, it is common for investigators to seek to summarize differences between groups using overall estimates of survival. For this purpose, direct standardization of relative survival estimates has been advocated (Parkin & Hakulinen, 1991). In the EUROCare study, for example, the data from all registries were combined, within categories of tumour site, in order to establish the standard populations of persons registered with cancer to which individual registry survival estimates were standardized (Berrino et al., 1995). This approach was not possible in the present study, since we wished to make comparisons between the developing country populations and published data for European and US populations. For this reason, we constructed a set of abstract World Standard Cancer Patient Populations for use in the present study and, we hope, other comparative studies of cancer survival.

Data and methods
We obtained global estimates of incidence rates of major cancers in 1985 from Parkin et al. (1993). The data were in the form of rates for the age groups 15–44, 45–54, 55–64 and 65+. In order to provide standard populations for the more detailed five-year age groups (0–4, 5–9, ..., 85+), we used polynomial regression models to estimate incidence rates for intermediate points in the age range 15–64. Worldwide incidence rates for childhood cancer were obtained from Parkin et al. (1988). For age groups from 65–69 to 85+, incidence rates were estimated by linear projection of the trend in incidence between the point estimates of rates at ages 55–59 and 60–64. These incidence rates were then applied to United Nations estimates of the total world population in 1985 (UN, 1991) to obtain estimates of annual numbers of new cases. Finally, for each cancer site, percentages of cases in each age group were calculated.

Results
The standard populations (in percentage terms) are presented in Table 1.

Discussion
As described in Chapter 2, the standard populations are required for direct standardization (i.e. by summing age-specific relative survival estimates weighted by the standard percentages). The standard populations presented are approximations to the true age distributions of new cases of cancer globally. The accuracy of the estimates is not of great concern, since the intention is that they should be used in the intermediate calculations necessary to calculate age-standardized relative survival. Further refinements would not have any material bearing on such results. The regression method used to obtain age-specific estimates of global incidence rates imposes a degree of smoothing on the proportions. Therefore their use is unlikely to produce distorted standardized survival
## Table 1. World standard cancer patient populations (percentages): males and females combined

<table>
<thead>
<tr>
<th>Age</th>
<th>140–208</th>
<th>140–149</th>
<th>150</th>
<th>151</th>
<th>153–154</th>
<th>155</th>
<th>157</th>
<th>161</th>
<th>162</th>
<th>172</th>
</tr>
</thead>
<tbody>
<tr>
<td>0–4</td>
<td>0.5</td>
<td>0.1</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.1</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>5–9</td>
<td>0.6</td>
<td>0.2</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.3</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.1</td>
</tr>
<tr>
<td>10–14</td>
<td>0.7</td>
<td>0.4</td>
<td>0.0</td>
<td>0.0</td>
<td>0.1</td>
<td>0.4</td>
<td>0.0</td>
<td>0.1</td>
<td>0.0</td>
<td>0.4</td>
</tr>
<tr>
<td>15–19</td>
<td>0.9</td>
<td>0.7</td>
<td>0.1</td>
<td>0.1</td>
<td>0.2</td>
<td>0.8</td>
<td>0.0</td>
<td>0.2</td>
<td>0.1</td>
<td>0.9</td>
</tr>
<tr>
<td>20–24</td>
<td>1.3</td>
<td>1.2</td>
<td>0.2</td>
<td>0.2</td>
<td>0.4</td>
<td>1.3</td>
<td>0.1</td>
<td>0.4</td>
<td>0.2</td>
<td>1.8</td>
</tr>
<tr>
<td>25–29</td>
<td>1.8</td>
<td>1.9</td>
<td>0.5</td>
<td>0.7</td>
<td>0.8</td>
<td>2.1</td>
<td>0.4</td>
<td>0.9</td>
<td>0.5</td>
<td>3.3</td>
</tr>
<tr>
<td>30–34</td>
<td>2.5</td>
<td>3.1</td>
<td>1.4</td>
<td>1.7</td>
<td>1.5</td>
<td>3.2</td>
<td>1.3</td>
<td>1.7</td>
<td>1.1</td>
<td>4.9</td>
</tr>
<tr>
<td>35–39</td>
<td>3.4</td>
<td>4.6</td>
<td>3.1</td>
<td>3.6</td>
<td>2.3</td>
<td>4.7</td>
<td>3.2</td>
<td>3.0</td>
<td>2.1</td>
<td>6.6</td>
</tr>
<tr>
<td>40–44</td>
<td>4.7</td>
<td>6.5</td>
<td>5.3</td>
<td>5.7</td>
<td>3.4</td>
<td>6.5</td>
<td>5.3</td>
<td>5.1</td>
<td>3.6</td>
<td>7.8</td>
</tr>
<tr>
<td>45–49</td>
<td>6.3</td>
<td>8.6</td>
<td>7.5</td>
<td>7.3</td>
<td>4.7</td>
<td>8.3</td>
<td>6.6</td>
<td>7.7</td>
<td>5.6</td>
<td>8.7</td>
</tr>
<tr>
<td>50–54</td>
<td>8.0</td>
<td>10.5</td>
<td>9.1</td>
<td>8.2</td>
<td>6.1</td>
<td>10.0</td>
<td>7.1</td>
<td>10.7</td>
<td>8.0</td>
<td>9.0</td>
</tr>
<tr>
<td>55–59</td>
<td>9.7</td>
<td>11.9</td>
<td>10.3</td>
<td>8.6</td>
<td>7.5</td>
<td>11.1</td>
<td>7.5</td>
<td>13.2</td>
<td>10.4</td>
<td>8.9</td>
</tr>
<tr>
<td>60–64</td>
<td>10.5</td>
<td>11.2</td>
<td>12.4</td>
<td>10.8</td>
<td>9.5</td>
<td>11.6</td>
<td>9.9</td>
<td>12.9</td>
<td>12.6</td>
<td>9.1</td>
</tr>
<tr>
<td>65–69</td>
<td>10.5</td>
<td>10.0</td>
<td>11.8</td>
<td>11.0</td>
<td>10.6</td>
<td>10.2</td>
<td>10.6</td>
<td>11.5</td>
<td>12.3</td>
<td>8.7</td>
</tr>
<tr>
<td>70–74</td>
<td>10.3</td>
<td>8.9</td>
<td>11.0</td>
<td>10.9</td>
<td>11.7</td>
<td>9.0</td>
<td>11.2</td>
<td>10.1</td>
<td>11.9</td>
<td>8.3</td>
</tr>
<tr>
<td>75–79</td>
<td>9.9</td>
<td>7.8</td>
<td>10.1</td>
<td>10.7</td>
<td>12.8</td>
<td>7.8</td>
<td>11.8</td>
<td>8.7</td>
<td>11.3</td>
<td>7.8</td>
</tr>
<tr>
<td>80–84</td>
<td>9.5</td>
<td>6.8</td>
<td>9.1</td>
<td>10.5</td>
<td>13.8</td>
<td>6.8</td>
<td>12.2</td>
<td>7.4</td>
<td>10.5</td>
<td>7.2</td>
</tr>
<tr>
<td>85+</td>
<td>8.9</td>
<td>5.6</td>
<td>8.2</td>
<td>10.0</td>
<td>14.6</td>
<td>5.8</td>
<td>12.8</td>
<td>6.4</td>
<td>9.8</td>
<td>6.5</td>
</tr>
<tr>
<td>Total</td>
<td>100.0</td>
<td>100.0</td>
<td>100.0</td>
<td>100.0</td>
<td>100.0</td>
<td>100.0</td>
<td>100.0</td>
<td>100.0</td>
<td>100.0</td>
<td>100.0</td>
</tr>
</tbody>
</table>

## Table 1. World standard cancer patient populations (percentages): males and females combined

<table>
<thead>
<tr>
<th>Age</th>
<th>174</th>
<th>180</th>
<th>182</th>
<th>183</th>
<th>185</th>
<th>186</th>
<th>188</th>
<th>189</th>
<th>200-203</th>
<th>204-208</th>
</tr>
</thead>
<tbody>
<tr>
<td>0–4</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>4.5</td>
<td>2.8</td>
<td>8.1</td>
<td></td>
</tr>
<tr>
<td>5–9</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>2.1</td>
<td>2.7</td>
</tr>
<tr>
<td>10–14</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>0.7</td>
<td>0.0</td>
<td>0.1</td>
<td>0.0</td>
<td>0.0</td>
<td>1.3</td>
<td>2.7</td>
</tr>
<tr>
<td>15–19</td>
<td>0.3</td>
<td>0.3</td>
<td>0.2</td>
<td>1.1</td>
<td>0.0</td>
<td>3.8</td>
<td>0.2</td>
<td>1.0</td>
<td>2.8</td>
<td>4.6</td>
</tr>
<tr>
<td>20–24</td>
<td>0.9</td>
<td>1.1</td>
<td>0.4</td>
<td>1.9</td>
<td>0.0</td>
<td>12.5</td>
<td>0.4</td>
<td>1.0</td>
<td>3.1</td>
<td>4.3</td>
</tr>
<tr>
<td>25–29</td>
<td>2.1</td>
<td>2.7</td>
<td>0.8</td>
<td>2.9</td>
<td>0.0</td>
<td>18.6</td>
<td>0.6</td>
<td>1.3</td>
<td>3.5</td>
<td>4.1</td>
</tr>
<tr>
<td>30–34</td>
<td>4.0</td>
<td>5.2</td>
<td>1.6</td>
<td>4.2</td>
<td>0.1</td>
<td>19.0</td>
<td>1.1</td>
<td>1.6</td>
<td>4.0</td>
<td>4.2</td>
</tr>
<tr>
<td>35–39</td>
<td>6.4</td>
<td>8.2</td>
<td>3.0</td>
<td>5.8</td>
<td>0.2</td>
<td>16.1</td>
<td>1.8</td>
<td>2.3</td>
<td>4.6</td>
<td>4.3</td>
</tr>
<tr>
<td>40–44</td>
<td>8.5</td>
<td>10.8</td>
<td>5.1</td>
<td>7.5</td>
<td>0.3</td>
<td>9.8</td>
<td>2.8</td>
<td>3.5</td>
<td>5.4</td>
<td>4.5</td>
</tr>
<tr>
<td>45–49</td>
<td>10.0</td>
<td>12.1</td>
<td>7.9</td>
<td>9.0</td>
<td>0.8</td>
<td>6.3</td>
<td>4.2</td>
<td>5.3</td>
<td>6.2</td>
<td>4.8</td>
</tr>
<tr>
<td>50–54</td>
<td>10.5</td>
<td>12.0</td>
<td>11.0</td>
<td>10.2</td>
<td>1.6</td>
<td>3.9</td>
<td>6.0</td>
<td>7.7</td>
<td>7.0</td>
<td>5.2</td>
</tr>
<tr>
<td>55–59</td>
<td>10.2</td>
<td>10.9</td>
<td>13.6</td>
<td>10.9</td>
<td>3.3</td>
<td>2.9</td>
<td>7.9</td>
<td>10.6</td>
<td>7.7</td>
<td>5.6</td>
</tr>
<tr>
<td>60–64</td>
<td>9.7</td>
<td>9.8</td>
<td>13.6</td>
<td>10.4</td>
<td>6.2</td>
<td>1.8</td>
<td>10.0</td>
<td>10.9</td>
<td>8.5</td>
<td>6.2</td>
</tr>
<tr>
<td>65–69</td>
<td>9.0</td>
<td>8.0</td>
<td>11.7</td>
<td>9.2</td>
<td>8.7</td>
<td>1.3</td>
<td>11.1</td>
<td>10.6</td>
<td>8.4</td>
<td>6.4</td>
</tr>
<tr>
<td>70–74</td>
<td>8.2</td>
<td>6.5</td>
<td>9.9</td>
<td>8.0</td>
<td>12.1</td>
<td>1.1</td>
<td>12.2</td>
<td>10.1</td>
<td>8.2</td>
<td>6.6</td>
</tr>
<tr>
<td>75–79</td>
<td>7.5</td>
<td>5.2</td>
<td>8.3</td>
<td>6.9</td>
<td>16.4</td>
<td>0.9</td>
<td>13.1</td>
<td>9.4</td>
<td>7.9</td>
<td>6.6</td>
</tr>
<tr>
<td>80–84</td>
<td>6.7</td>
<td>4.1</td>
<td>6.9</td>
<td>5.9</td>
<td>21.8</td>
<td>0.7</td>
<td>13.9</td>
<td>8.8</td>
<td>7.5</td>
<td>6.6</td>
</tr>
<tr>
<td>85+</td>
<td>5.9</td>
<td>3.0</td>
<td>5.9</td>
<td>4.8</td>
<td>28.5</td>
<td>0.5</td>
<td>14.6</td>
<td>8.0</td>
<td>7.0</td>
<td>6.3</td>
</tr>
<tr>
<td>Total</td>
<td>100.0</td>
<td>100.0</td>
<td>100.0</td>
<td>100.0</td>
<td>100.0</td>
<td>100.0</td>
<td>100.0</td>
<td>100.0</td>
<td>100.0</td>
<td>100.0</td>
</tr>
</tbody>
</table>
values, as long as there is reasonable precision in the original estimates of age-specific relative survival. If the original survival data are very sparse, it may be wise to combine them in larger age groups, for which the standard proportions can be combined additively. A further recommendation would be to show a truncated standardized relative survival, say to age 74, as we have done in the present report. If an investigator wishes to standardize data for a cancer which has not been included in the tables, then it would be a reasonable approach to choose a set of proportions for a similar type of cancer, or for all cancer combined. It should be noted that the site-specific cancer patient populations are not suitable for comparisons of survival of patients in a single population with different types of cancer. If this is an aim of a particular study, then, again, the standard populations given in Table 1 for all cancers combined should be used.

The effect of standardizing using a worldwide standard is that greater weight is given to younger patients than would be the case if an age distribution based on developed countries only had been used. For developed countries in which elderly cancer patients predominate, use of the World Standard Cancer Patient Populations tends to raise the age-standardized relative survival above the unstandardized value. The interpretation of results of this kind is that they indicate the overall relative survival which would pertain if the developed country's age-specific survival values applied in a 'worldwide average' group of patients. This may seem artificial but, in the spirit of the World Standard Population (Segi, 1960), which has now gained universal acceptance, use of the World Standard Cancer Patient Populations will enhance the comparability of survival results published by individual cancer registries.

References


Segi, M. (1960) Cancer Mortality for Selected Sites in 24 Countries (1950-1957). Sendai, Japan, Tohoku University School of Medicine, Department of Public Health

Chapter 4

Interpretation of population-based cancer survival data

R. J. Black, R. Sankaranarayanan and D.M. Parkin
International Agency for Research on Cancer
Lyon,
France

Introduction

In order to describe completely the experience of cancer in a population, it is necessary to know not only its incidence and mortality, but also the survival of cancer patients. There are three main sources of information about survival: the randomized controlled clinical trial, which represents the 'gold standard' for the evaluation of forms of treatment; the hospital-based study, which aims to provide information about the outcome of treatment in particular settings; and population-based survival from cancer registries, which reflects a broader range of cancer control activities, including screening and the organization of treatment services. Each of these has its limitations: survival information from trials and published hospital series is often biased by patient selection, whereas population-based survival data may lack the details of stage and treatment which are of particular interest to the clinician.

The rationale of a randomized clinical trial is to eliminate the confounding effects of factors such as age and comorbidity in order to isolate the effects of treatment. This is achieved in two ways: by adopting selection criteria which exclude some subjects (such as those with comorbid conditions) and random allocation of the remainder into groups in which the only systematic differences are the treatments to be received. This approach is essential in order to determine the efficacy of particular treatments. However, the effectiveness of cancer services in general depends not only on the efficacy of particular treatments but also on the context in which they are applied.

Evaluating effectiveness requires estimation of survival in unselected groups of cancer patients, which is a key aim of most cancer registries. Estimates of survival in such groups may be influenced by a range of prognostic and other factors (see Table 1). In accounting for these, the methodology of clinical trials cannot be deployed, since selection would invalidate the generality of results and it is impossible, of course, to randomize cancer patients to different health care systems. Therefore another approach is required when evaluating survival at the population level and making comparisons between population groups which are heterogeneous in respect of prognostic factors. Some factors such as age and sex can be accounted for in the statistical methodology used to estimate survival. Information on other factors such as comorbidity may not even be present in cancer registry data. For these, the best we can do is to be aware of their possible influences or, as has been attempted in the EUROCare study (Berrino et al., 1995), augment the basic cancer registry information with 'high resolution' data. The remainder of this chapter provides a review of comparability issues: that is, data quality factors (e.g. methods of ascertainment and follow-up), host factors (e.g. age, sex, risk of death from other diseases), tumour-related factors (e.g. extent of disease) and health care factors (e.g. availability and quality of diagnosis and treatment services) which influence population-based estimates of survival.

Data quality factors

Inevitably, the quality of cancer registration data will vary according to the availability of source data, the experience of registry staff and other factors. This variation complicates the interpretation of survival data based on routine cancer registry data (Hanai & Fujimoto, 1985). Of particular concern is the completeness of ascertainment. If cases of cancer which are not registered represent a random sample from the total, then there should be no systematic bias in survival results. However, this is unlikely to be the case, since the probability of being registered tends to be correlated with prognosis: for example, elderly patients not seen in hospital are less likely to be registered than younger patients, for whom curative treatment may have been attempted. Estimates of survival may therefore be artificially
Table 1. Factors influencing population-based survival data

**Data quality factors**
- Completeness of ascertainment
- Accuracy of registration
- Completeness of follow-up
- 'Death certificate only' (DCO) registrations

**Host factors**
- Age
- Sex
- Race/Ethnicity
- Comorbidity
- Socioeconomic status
- Behaviour (including awareness of cancer symptoms and compliance with treatment)

**Tumour-related factors**
- Extent of disease
- Site (including subsite) of tumour
- Morphology of tumour
- Tumour biology

**Health care-related factors**
- Screening
- Diagnostic facilities
- Treatment facilities
- Quality of treatment
- Follow-up care

raised for a particular registry area if ascertainment is not complete. Similarly, the accuracy of diagnostic information for cancer patients tends to be correlated with prognosis. For example, a registry relying exclusively on a particular pathology laboratory for diagnostic information might tend to classify cases from other sources in nonspecific categories for primary site. Such cases would be excluded from tumour-site-specific survival analyses, whereas data for another registry might include cases with clinical diagnoses. Again, the effect of this aspect of data quality would be to increase the survival estimate for a registry which allocated a large proportion of cases to nonspecific diagnostic categories.

In cancer registry data, there are usually some subjects for whom the registration of cancer was based on information from the death certificate only (DCO). By convention, such cases are excluded from survival analyses since — by definition — their survival time is zero. This convention was adopted in the present study. If the proportion of DCO cases is relatively low, say less than 10%, then excluding them from the analysis does not greatly influence survival estimates. However, larger proportions of DCO cases are problematic, since they may mean that cases of poor prognosis (which would have been registered by other means if the registry had had better ascertainment procedures) are excluded, thus artificially increasing estimates of survival. Some of the registries in the present study have quite large proportions of DCO cases.

Under either active or passive follow-up systems, individuals can be lost due to migration, breaking off contacts with local authorities or other changes in living conditions. Normally, registries assume that a subject is alive until a notification of death is received, or active follow-up results in a confirmation of death. Many of the individuals who are lost to follow-up will, in fact, have died, so that a registry with a large proportion of individuals with whom they have lost contact will report artificially high survival. However, the direction of the bias is unpredictable, and will depend on local circumstances. For example, loss to follow-up may occur when subjects with a relatively good prognosis are obliged to move away from their original cancer registry area to receive treatment.

A common feature of most of the aspects of data quality discussed above is that poor data quality tends to increase estimates of survival. A key aim of the international network of cancer registries is to standardize data collection methods and indicators of data quality. We believe that the differences in survival between the registries reported in the present study are mainly due to factors other than data quality (see the discussion in Chapter 16). However, it is important to be aware that apparently high survival rates for some registries may have been influenced by data quality factors. Detailed information on data quality, including data quality indicators such as the proportion of DCO registrations and the proportion of cases with histological verification, can be seen in the individual registry chapters.

**Host factors**
Age at diagnosis is an independent prognostic factor for many types of cancer. This operates in two ways: age may be correlated both with the risk of dying from a particular type of cancer and with the risk of dying from some other cause. In the present study, we adjusted for age using age-standardized relative survival (to take account of variations in age-specific
background mortality and differences in the age distributions of the populations being compared).

Sex is less commonly associated with variations in survival and, for this reason, many registries combined data for males and females in the interests of increasing the precision of survival estimates. However, survival from some cancers, such as malignant melanoma, has been seen to be greater for women in some developed countries, which is probably due to a greater recognition of early symptoms and a willingness to seek medical attention.

Comorbid conditions experienced by cancer patients may vary substantially between registry populations. Comorbidity affects survival by presenting an additional source of risk of death, making it less likely that a patient will be offered curative treatment and, if it is offered, less likely that the patient will be able to withstand the effects of the treatment itself.

Socioeconomic differences in survival have been reported for many sites of cancer within populations in Europe (Kogevinas, 1991) and the USA (Berg et al., 1977). Socioeconomic status tends to be correlated with strong prognostic factors such as extent of disease at diagnosis, but it has been shown to have a residual effect which may be due to inequalities in access to medical facilities, compliance with treatment regimens, coping strategies or social support. Within developed countries, race is also associated with survival, although the extent to which this operates independently of socioeconomic status is unclear (Howard et al., 1991). Berg et al. (1977) propose a host vulnerability hypothesis in which the poor nutritional status, general health and immunological status (related to alcoholism) of some social and racial groups leads to lower survival from cancer. Clearly, socioeconomic conditions in developed and developing countries are grossly different, to the extent that inequalities in access to medical care are likely to be of particular significance in the present study.

**Tumour-related factors**

By convention, cancer registry data are aggregated within categories defined by the anatomical site of the tumour. When comparing international survival data, caution must be exercised when the distributions of tumour subsites vary. The same point applies to variations in the frequency of morphological types of tumour within categories of site, and variations in tumour biology, as expressed by differences in natural history and aggressiveness of clinical course (e.g. breast cancers with variations in the frequencies of tumour markers such as hormonal receptor status; variations in the grade of non-Hodgkin lymphomas).

The stage of disease at diagnosis is generally the most important factor determining the survival of cancer patients. This is because certain treatments may be available only for early-stage tumours, and any treatment is more likely to be successful if initiated before metastasis has occurred. Therefore, variations in the stage distributions of tumours in populations being compared are of particular concern. Some of the registries involved in the present study were able to supply data on extent of disease, which we have used in interpreting results for these regions. However, even when such data are available, variations in diagnostic technology such as those between developed and developing countries are likely to lead to measurement error. Stage of disease at diagnosis is influenced both by the general level of health awareness in the population, and by the presence or absence of programmes of early detection for cancer. The effect of the former is seen clearly with respect to cervical cancer statistics recorded in patients admitted to the Radiumhemmet Hospital, Stockholm, Sweden since 1920. Even before the introduction of screening programmes, there was a dramatic change in the proportion of cases diagnosed at early stages (I and II), from less than 20% in 1920 to some 80% in 1965. (Pontén et al., 1995).

**Health care-related factors**

Factors relating to the health care of cancer patients in developed and developing countries are of particular concern in the present study. There are a number of ways in which the availability of, and access to, screening services and diagnostic and treatment facilities can influence survival. Screening programmes aim to detect early-stage cancers or premalignant tumours so that the disease can be treated at an early stage, which is generally more effective. However, interpreting survival statistics in terms of the benefit to patients resulting from screening is problematic, since one consequence of early detection is to bring forward the date of diagnosis of a condition, whether or not this has the desired effect of reducing risk of death from the disease. This is called 'lead-time bias'. In addition, screening programmes may result in the detection of disease that would not otherwise have been diagnosed at all during the life of the patient — so-
called 'overdiagnosis bias' (Morrison, 1985). This latter will necessarily result in a marked improvement in survival, and one that is independent of any ‘downstaging’ effect of screening which, theoretically at least, could be monitored in cancer registry data. Overdiagnosis almost certainly accounts for the huge increases recently observed in the reported incidence of prostate carcinoma in the USA, and the corresponding changes in survival (Kosary et al., 1995). As far as the results in this volume are concerned, however, they are likely to be little influenced by screening programmes. With certain exceptions (described in the relevant chapters), screening programmes are not extensive or systematically implemented in developing countries and, where information on extent of disease is available, it will be observed that, for many cancers, more patients are presenting late in the course of their disease than would be expected from experience in Europe or North America.

Diagnostic facilities may also influence survival by ensuring that a specific and correct diagnosis can be made. Improvements in the sensitivity of diagnosis may have the effect of inducing ‘stage migration’ in which, for example, tumours of limited metastatic activity which at one time would have been inaccurately described as simply invasive, may be reallocated to the metastatic category, thus increasing estimates of survival of individuals in both metastatic and the nonmetastatic groups (Feinstein et al., 1985). This phenomenon can operate on a geographical as well as a temporal basis (Farrow et al., 1995). Therefore, comparisons of stage-specific survival data from settings with very different diagnostic facilities cannot be made with confidence. It should be noted that comparisons of survival of groups comprising patients with tumours of all stages combined are not subject to this problem, as long as there is no selection bias due to greater diagnostic specificity in one population compared with another.

The availability of treatment facilities for cancer patients affects the survival of those for whom curative treatment would have the potential to succeed. Therefore the issue of availability of treatment facilities is bound up with the availability of other facilities, such as screening programmes and diagnostic facilities. Survival data from cancer registries cannot be used to make direct comparisons of populations in terms of the quality of care available, although some studies have shown that the survival of some cancer patients is prolonged after treatment at specialized cancer centres (Stiller, 1994). However, results of this kind are difficult to interpret because of selection criteria for specialized care, which may determine the apparently better results rather than the quality of care received per se.

Conclusions
The previous discussion indicates the difficulty of making meaningful comparisons of survival among groups of cancer patients with varying demographic and socioeconomic characteristics, and served by very different health care infrastructures, using retrospectively collected data from cancer registries. It is certainly important to realize that variations are not simply due to the availability and quality of medical services. However, as will be seen from the discussion in Chapter 16 ‘An overview of cancer survival in developing countries’, a comparison of the magnitude of differences between countries provides at least an indirect indication of the relative importance of early detection and treatment for certain major cancers.

References

