Int. J. Epidemiol. Advance Access published February 8, 2010

Published by Oxford University Press on behalf of the International Epidemiological Association

International Journal of Epidemiology 2010;1–13
doi:10.1093/ije/dyp392

Measuring cancer survival in populations: relative survival vs cancer-specific survival

Diana Sarfati,[1]* Tony Blakely[1] and Neil Pearce[2]

[1]Department of Public Health, University of Otago Wellington, Wellington, New Zealand and [2]Centre for Public Health Research, Massey University, Wellington, New Zealand

*Corresponding author. Department of Public Health, University of Otago, PO Box 7343, Wellington 6242, New Zealand. Email: diana.sarfati@otago.ac.nz

Accepted 14 December 2009

Background Two main methods of quantifying cancer patient survival are generally used: cancer-specific survival and relative survival. Both techniques are used to estimate survival in a single population, or to estimate differences in survival between populations. Arguments have been made that the relative survival approach is the only valid choice for population-based cancer survival studies because cancer-specific survival estimates may be invalid if there is misclassification of the cause of death. However, there has been little discussion, or evidence, as to how strong such biases may be, or of the potential biases that may result using relative survival techniques, particularly bias arising from the requirement for an external comparison group.

Methods In this article we investigate the assumptions underlying both methods of survival analysis. We provide simulations relating to the impact of misclassification of death and non-comparability of expected survival for cause-specific and relative survival approaches, respectively.

Results For cause-specific analyses, bias through misclassification of cause of death resulted in error in descriptive analyses particularly of cancers with moderate or poor survival, but had smaller impact in analyses involving group comparisons. Relative survival ratio (RSR) estimations were robust in relation to non-comparability of comparison populations for single RSR but were less so in group comparisons where there was large variation in survival.

Conclusions Both cause-specific survival and relative survival are potentially valid epidemiological methods in population-based cancer survival studies, and the choice of method is dependent on the likely magnitude and direction of the biases in the specific analyses to be conducted.

Keywords Cancer, survival, survival analysis, relative survival, excess mortality

Introduction

Quantifying cancer patient survival is an important but challenging task. Two main methods of survival analysis are generally used: cancer-specific survival and relative survival. The former uses cancer-specific deaths as the end-point of interest, and patients who die from other causes are considered to be 'censored'. The latter uses death from any cause as the end-point

of interest, and compares the observed survival with that which would have been expected if the cancer patients had had the same mortality rates as an external comparison population. A number of authors have argued that relative survival is the most, and possibly the only, appropriate measure to use in population-based cancer survival studies.[1–4] The basis for this argument is that there is likely to be misclassification of cancer-specific deaths, resulting in biased estimates for cancer-specific survival. However, there has been little discussion, or evidence, as to how strong such biases may be, or of the possible other biases that may result from the alternative approach of using relative survival techniques.

In this article we provide an overview of the two methods and their assumptions, with a particular focus on the types of systematic error that may affect each method.

At the outset, it should be noted that cancer-specific survival and relative survival approaches are often used in different situations and contexts. The methodological issues involved will be discussed in more detail in the rest of the article, but the general issues should be briefly considered and contrasted here.

First, relative survival analyses are most commonly used in 'descriptive', often population-based studies, and involve estimating risks (proportions). On the other hand, cancer-specific analyses are most often used in etiologic, often clinical studies, and involve using multiple regression to estimate ratios of mortality rates (hazard ratios). The aims and approaches in these two contexts are completely different, and direct comparisons of the two methods will suggest greater differences than there actually are. In fact, both methods can be effectively used in both descriptive and etiologic studies.

Secondly, although in principle both methods are aiming to measure the same underlying construct, i.e. net cancer survival, they are, in fact, measuring slightly different things. Cancer-specific survival measures deaths that are identified as due to a specific type of cancer. Therefore, an explicit decision has to be made for each death whether it was fully attributable to the cancer or not. Some deaths are clearly related to the cancer in question, and assuming correct diagnosis and coding, will be identified as such. These should include deaths directly related to treatment for the cancer. We have referred to these as cancer-specific deaths (Figure 1). For other deaths, the cancer will be a contributing factor, but may not be wholly responsible for the death. It is therefore problematic to identify and assign these deaths to being (entirely) cancer specific or not. The line between cancer-specific and cancer-consequent deaths is a matter of judgement and interpretation.

Relative survival, in contrast, aims to measure deaths that are in excess of what would be expected for the study population if it did not have cancer. There is no need for differentiation between cancer-specific and cancer-consequent deaths (Figure 1). This is the key strength of the relative survival approach. However, although this approach obviates the need to categorize individual deaths, it introduces the important requirement for an external comparison group, and the resulting assumption that the comparisons being made are valid. These distinctions are important for understanding the assumptions underlying each method, and the potential biases that may result as a consequence of them.

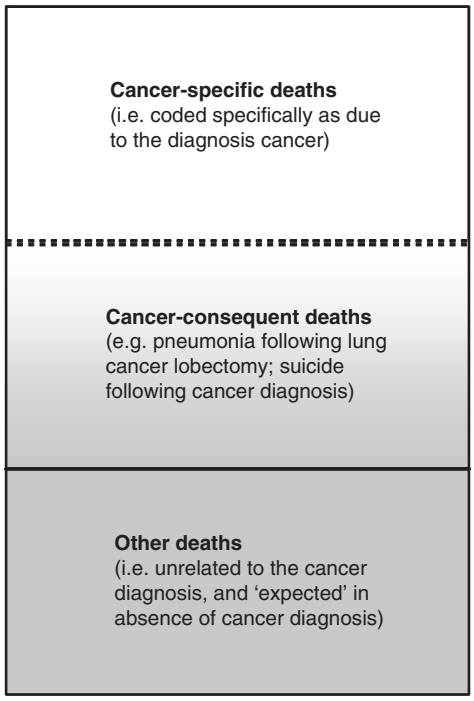

Figure 1 Categories of death relevant to cause-specific and relative survival/excess mortality estimation of survival

Cancer-specific survival (and mortality)

Cancer-specific survival analysis involves using only deaths identified as being due to a specified cancer as the outcome of interest. Follow-up typically starts on the date of cancer diagnosis, and continues until death, loss-to-follow-up or the end of the study period. Patients who die of a cause other than the cancer under study or who are lost to follow-up are 'censored' at the last date on which they were known to be alive. The survival for a given time period can be calculated directly using standard methods such as Kaplan Meier.[5,6]

Systematic error in single population cancer-specific analyses

Table 1 summarizes the main possible sources of systematic error (bias) in cancer-specific analyses and relative survival analyses in single populations and in comparisons of populations. Some sources of

Table 1 Epidemiological characteristics and systematic sources of error of relative survival and cancer-specific survival analyses

	Cancer-specific survival/mortality		Relative survival/excess mortality	
	Single population **Cancer-specific survival estimates**	**Two or more sub-population comparison** **Cox proportional hazards modelling (usually)**	**Single population** **Relative survival estimates**	**Two or more sub-population comparison** **Excess mortality modelling (or direct comparison of RSRs)**
Basic epidemiological characteristics				
Comparison groups	NA (internal comparison)	Two or more sub-populations with cancer	External comparison with general population	Two or more sub-populations with cancer, (each of which is compared with the external population comparison group/s)
Outcome	Death due to specified cancer (with other deaths censored)	Mortality (hazard) due to specified cancer (with other deaths censored)	Probability of survival of each group	Excess mortality (or RSRs) of each subpopulation
Measure (of association)	Survival proportion (or mortality rate/ hazard)	Mortality hazard ratio	RSR (or excess mortality rate)	Excess mortality rate ratios
Systematic error[a]				
Selection bias[b]	Competing risks	Sub-population differences in bias due to competing risks	NA	NA
Information bias	Misclassification of cancer-specific cause of death	Non-differential and differential (between subpopulations) misclassification of cancer-specific cause of death	NA	NA
Non-comparability of external comparison group[b]	NA	NA	The external population may be non-comparable (non-exchangeable) with the cancer group. This may occur due to: (i) inherent non-comparability of background mortality risk; (ii) misclassification of exposure (e.g. ethnicity) in the external population mortality rates	The bias from non-comparability of the external population(s) may be of different strengths for the sub-populations being compared

RSR, relative survival rate; NA, not applicable.
[a]The table only lists the potential sources of bias that are relatively unique to the study design under consideration. It does not list other potential sources of bias (e.g. general selection bias, uncontrolled confounding, misclassification of 'exposure' (e.g. ethnicity) on the cancer registry, etc.) that can occur in all epidemiological studies.
[b]May also be thought of as a type of confounding.

error will result in bias regardless of the measure of survival that is used (cancer-specific or relative survival) and will not be discussed further here (e.g. misclassification of whether death occurred, or misclassification of the main exposure variable). In this section we will focus on certain key biases relevant to estimating cancer-specific survival, particularly competing risks and misclassification of cause of death.

Competing risks

In addition to the usual possible selection biases inherent in any observational epidemiological study, the issue of competing risks is important for cancer-specific survival estimation. A competing risk is one whose occurrence either precludes or fundamentally alters the probability of another event under study.[7] In the case of cancer-specific survival, death from other (non-cancer) causes precludes the

possibility of death from cancer. For this reason, those who experience non-cancer deaths in a cancer-specific analysis are censored at the time of death. This will not necessarily cause bias in itself, as those who died from other causes may have had the same cancer-specific survival experience as those who did not, if they had not been censored. However, bias will occur if those who censored because they died from other causes would have had different cancer-specific survival than those who were not censored in this way.[8] For substantial bias in cause-specific estimates of survival to arise from competing risks, the competing risks have to be common compared with the outcome of interest (e.g. elderly men with prostate cancer have a high mortality from causes other than prostate cancer); and those censored due to competing risks must have a different counterfactual risk of the outcome of interest than the non-censored population (e.g. men who died from other causes would have a poorer survival from prostate cancer than those who did not die from other causes). Obviously, whether these two conditions are met is highly context-specific, and further methodological research and empirical examples of the likely magnitude of such biases due to competing risks is required.[7–10]

Misclassification of cause of death

The most important potential source of bias in relation to estimating cancer-specific survival is misclassification of the underlying cause of death.[1,11–14] The underlying cause of death is defined as 'the disease or injury which initiated the train of morbid events leading directly to death'.[15] These, then, should include both deaths from cancer, and from cancer treatment. Misclassification of deaths can be usefully considered as two issues. The first is 'genuine' misclassification and relates either to deaths that are clearly due to the specific cancer or cancer treatment but are misclassified as deaths due to other causes (lack of sensitivity), or deaths from other causes that are misclassified as being due to the specific cancer under study (lack of specificity). This is a data quality issue that can be minimized with high-quality data.

The second issue relates to deaths that may be consequent on the cancer diagnosis, at least in part, but are not correctly classified as cancer-specific deaths. For example, a distressed patient who commits suicide after being informed of a diagnosis of cancer may be appropriately classified as a non-cancer death despite the fact that the diagnosis of cancer may have played a role in the death.

Between these two relatively clear extremes is a spectrum of deaths that may be more or less correctly classified as cancer specific. Examples might include a woman on hormone treatment for breast cancer who dies of a pulmonary embolism or a man who has a successful curative lobectomy for lung cancer but who dies 2 years later from pneumonia. In these intermediate cases the cancer is likely to have contributed to the death to some extent, but it is impossible to accurately ascribe such individual deaths as being wholly cancer specific or not. This is not a data quality issue, but a conceptual issue, and is represented by the dashed line separating cancer-specific and cancer-consequent deaths in Figure 1. Evidence relating to this issue is reflected in studies which show that even when all available data are known to both clinicians and pathologists, there can be disagreements between them as to the underlying cause of death.[16]

Several studies have attempted to estimate the extent to which deaths directly attributable to cancer therapy have been misclassified as non-cancer related. For example, Welch and Black[17] argued that up to 41% of cancer deaths that occurred within the immediate post-treatment period were not counted as such. Nevertheless, the impact of this issue on 5-year survival estimates may not be that great, with studies suggesting a 2–7% underestimate of cancer mortality as a result.[17,18]

Potential biases introduced by misclassification of cause of death will be affected both by data quality, and by the conceptual issue of 'cancer-consequent' deaths. However, the empirical evidence relates to the former, and so in the rest of this section, we will leave aside the latter issue of 'cancer-consequent' deaths.

It is also noteworthy that some cancer registries may not have access to data on specific cause of death, and therefore cancer-specific survival cannot be calculated. Nevertheless even registries that do have such data are dependent on the quality of the death data available to them. A few studies have attempted to determine the sensitivity and specificity of routine death data. However, this is complicated by the fact that accuracy of death data varies depending on a range of factors such as the level of diagnostic detail, the 'gold standard' against which death certificate data are measured and hence cancer-specific death defined (e.g. autopsy or medical notes review), and the level of clinical certainty underlying specific diagnoses. The quality of data is also likely to vary over time, place and possibly social grouping.[14,16] Several authors have shown that death certificate classification of cause of death may be more accurate for cancer than for some other causes of death, at least at a major diagnostic grouping level.[16,19,20] For example, Goldacre[20] found that cancer fell into a category of conditions that when present at death were usually entered as the underlying cause of death.

Although the quality of mortality data is invariably cited as the main reason for the preference of relative survival over cancer-specific survival in population level studies, the impact of the quality of these data on cancer-specific survival is rarely (if ever) assessed.

Table 2 Estimated 5-year cancer-specific survival rates for varying levels of misclassification of cancer, and non-cancer deaths for cancers with good, moderate and poor survival, assuming a fixed mortality rate (2.3% per year) from non-cancer causes over the follow-up period

	Sensitivity (%)	Specificity (%)		
		100	90	80
5-year cancer-specific survival				
Good survival	100	**0.79**	0.78	0.77
	90	0.81	0.80	0.79
	80	0.83	0.82	0.81
Moderate survival	100	**0.48**	0.48	0.47
	90	0.52	0.51	0.51
	80	0.56	0.55	0.55
Poor survival	100	**0.16**	0.15	0.15
	90	0.19	0.19	0.19
	80	0.23	0.23	0.22

The bold estimates are the true 5-year cancer-specific survival rate.

Table 2 shows estimated 5-year cancer-specific survival rates for rates of misclassification of cause of death of up to 20% for cancers with good (79%), moderate (48%) and poor (16%) 5-year survival. We have assumed that all deaths are reported, but that there is a variable level of misclassification of cause of deaths, i.e. that misclassification occurs for cause of death, but not for the occurrence of death itself. Sensitivity and specificity refer to the proportion of cancer and non-cancer deaths, respectively, that are recorded as such. Thus, for example, a sensitivity of 90% means that 10% of cancer deaths are counted as deaths from other causes, whereas a specificity of 90% means that 10% of deaths from other causes are misclassified as cancer deaths. We have also assumed the majority of cancer deaths occur in the first 2 years of follow-up and that there is a constant rate of death from other causes (2.3% per year) over the follow-up period. The 'true' estimate is shown in bold.

Misclassification of cause of death had little impact on estimates of survival for a cancer with good survival with estimates ranging from 0.77 to 0.83 for a cancer with a true 5-year survival rate of 0.79. The effect of misclassification was stronger for cancers with moderate or poor survival. In the former situation, the range of estimates was 0.47–0.56 for a cancer with a survival rate of 0.48, and in the latter the estimates ranged from 0.15 to 0.23 with the true survival rate being 0.16.

The biased estimates in Table 2 were 1.3–19% different from the true estimates, depending most importantly on the underlying survival rate of the cancer. For a cancer type with good 5-year survival, even with 20% misclassification of cause of death, the biased estimate was <4% different than the true estimate. Currently, there is very little information available on the extent of misclassification of cause of death. In countries with good cancer registration and death registration systems, it might be reasonable to expect that the sensitivity and specificity of cancer-specific cause of death on death certificates to be >90%—especially for aggregated groupings (e.g. studying colorectal cancer rather than studying colon and rectal cancer separately).[14] In this situation, the bias from misclassification of cause of death would be less than the estimates shown in Table 2.

Systematic error in comparing sub-population cancer-specific mortality rates

Cancer-specific mortality rates are usually compared using maximum likelihood techniques that are used to estimate the hazard ratio (i.e. the mortality rate ratio) for an 'exposed' group (e.g. patients with late-stage disease, or patients in a particular ethnic group) compared with an 'un-exposed' group (e.g. patients with early-stage disease, or patients in other ethnic groups) while adjusting for potential confounders (e.g. age, gender, socio-economic status, smoking). In this context, the most commonly used method is the Cox proportional hazards model, which is directly related to other maximum likelihood methods such as Poisson regression, and will yield the same findings when applied in the same manner to the same data set.[21] Note that when patient survival is modelled using multiple regression it is usually mortality (hazard) rates that are modelled—not survival itself.

Selection bias

In the previous section we discussed the issue of competing risks for cancer-specific survival analyses. We now consider the same issue for comparisons of cancer-specific mortality between two sub-populations. If sub-populations are defined by a factor, for example, ethnicity, that is related to death from other causes, and if those who die from other causes would have had a different risk of cancer-specific death than those who did not die from other causes, then the censoring on deaths from other causes introduces bias. This is illustrated in Figure 2, a directed acyclic graph (DAG),[22] showing an example of competing risks whereby some third factor 'U' is a cause of both other mortality (D_o) and cancer-specific mortality (D_c). The subgroups being compared are ethnic groups (E), and ethnicity has no direct causal association with cancer-specific death but is a cause of deaths from other causes (other mortality). These deaths from other causes are censored, i.e. the analysis is censored on D_o; however, D_o is a collider with both ethnicity and the unknown third factor (U) as parents. Conditioning on D_o will therefore cause a spurious association of E and D_c. In other words, if ethnicity is a cause of deaths from other causes, and people who die of other causes would have had a different cancer-specific mortality

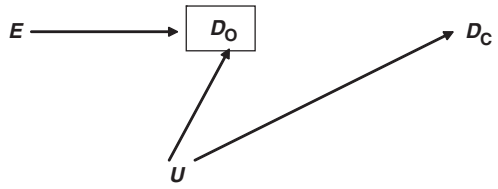

Figure 2 Directed acyclic graph showing potential bias due competing mortality risks when comparing cancer-specific mortality rates between ethnic sub-populations. *E*, ethnicity; D_C, cancer-specific mortality; D_O, other mortality; *U*, unknown common causes of D_O and D_C; conditioning on censoring (because of other mortality) introduces a spurious association between ethnicity and cancer-specific mortality, because those who are censored have different cancer-specific mortality compared with those who are not, and ethnicity is a risk factor for mortality from other causes, which results in censoring

rate, then the censoring on deaths from other causes introduces bias due to competing risks. Further methodological research is required to assess the likely magnitude of such bias under plausible scenarios.

Information bias

When cancer-specific mortality rates are compared using multiple regression techniques such as Cox proportional hazards modelling, the bias due to misclassification of cause of death may not be particularly strong provided that misclassification is non-differential. Furthermore, when exposure is dichotomous, the bias in the estimated hazard ratio will always be towards the null value of 1.0.

Table 3 presents simulations comparing the hazard ratios of the poor, moderate and good survival groups, where the mortality hazard of each of these three separate groups has been back-calculated to be consistent with 5-year survival reported in Table 2 (using the exponential formula linking rates and survival proportions).[22] The simulated hazard ratios are presented for non-differential misclassification only (i.e. with the same sensitivity and specificity between poor, moderate and good survival groups). Under the scenarios shown, the hazard ratio was, at most, biased by 11% to the null when comparing moderate with good survival with 80% sensitivity and specificity (i.e. a shift from the true moderate:good hazard ratio of 3.12 to 2.88).

The greater concern in cancer-specific analyses is bias resulting from differential misclassification of cause of death. This situation is illustrated in Figure 3, where ethnicity is a cause of misclassification of cause of death; thus, in this hypothetical example, ethnicity is not associated with true cancer-specific mortality (provided that the analysis is adjusted for confounders such as age, gender and stage at diagnosis), but it is associated with 'measured' cancer-specific mortality.

Table 3 Hazard ratios assuming non-differential misclassification bias of cause of death

		Specificity (%)		
	Sensitivity (%)	100	90	80
Moderate[a] vs good[a]	100	**3.12**	3.02	2.92
survival	90	3.11	3.00	2.90
	80	3.11	2.99	2.88
Poor vs moderate	100	**2.55**	2.53	2.51
survival	90	2.54	2.52	2.49
	80	2.53	2.50	2.48
Poor[a] vs good	100	**7.96**	7.63	7.33
survival	90	7.92	7.56	7.23
	80	7.88	7.48	7.13

[a]Five-year survival proportions of 0.79, 0.48 and 0.16 for good, moderate and poor survival, respectively, as shown in Table 2, but with back-estimated mortality hazard rates for the simulation (see text).
The bold estimates are the true hazard ratios.

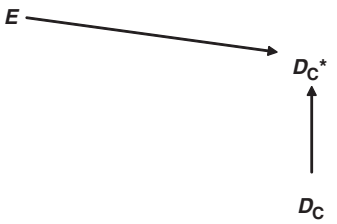

Figure 3 Directed acyclic graph showing potential bias due to differential misclassification of cancer-specific mortality when comparing between ethnic sub-populations. *E*, ethnicity; D_C, cancer-specific mortality; D_C*, measured cancer-specific mortality. There is a spurious association between ethnicity and measured cancer-specific mortality, because measurement error of cancer-specific mortality is associated with ethnicity, and affects the measured cancer-specific mortality

There is some evidence that this may occur, for example, it has been suggested that quality of death certificate data may vary systematically between socio-economic groups.[16] In this situation, the bias can be either towards or away from the null value. Table 4 shows simulated hazard ratios for the most extreme combinations of 80 or 100% sensitivity and specificity, and using the moderate and good survival groups as comparison groups. There is now a considerably larger range of observed hazard ratios from 2.26 to 4.02 given the true hazard ratio of 3.12, and these can be biased either towards or away from the null value.

In summary, provided that it can be reasonably assumed that misclassification of cause of death is non-differential, such misclassification is a relatively minor concern in cancer-specific analyses comparing subgroups, when compared with all of the other sources of bias such as uncontrolled confounding and misclassification of exposure (which apply to

Table 4 Hazard ratios assuming differential misclassification bias of cause of death

Se/Sp of moderate[a] survival group (%)	Se/Sp of good[a] survival groups (%)			
	100/100	100/80	80/100	80/80
100/100	**3.12**	2.83	3.89	3.12
100/80	3.22	*2.92*	4.02	3.22
80/100	2.49	2.26	*3.11*	2.49
80/80	2.59	2.35	3.24	*2.59*

[a]Five-year survival proportions of 0.79 and 0.48 for good and moderate survival, respectively, as shown in Table 2, but with back-estimated mortality hazard rates for the simulation (see text). The bold estimate is the true hazard ratio, and the estimate in italics shows the hazard ratios under non-differential misclassification bias as shown in Table 3.

both sets of techniques). However, if there is reason to suspect that there is considerable differential misclassification of cause of death, the estimated hazard ratios should be interpreted with caution, and ideally the likely extent and direction of information bias should be estimated.

Relative survival (excess mortality)

In relative survival analysis the survival of a cohort with cancer is compared with that of an external comparison group without cancer (usually the general population).[23] The difference between the two is assumed to be due to cancer-related deaths in the first group, that is the sum of the 'cancer-specific' and 'cancer-consequent' deaths in Figure 1.

Survival probabilities (risks) in the cancer and comparison groups are compared for regular time intervals from the point of origin (the date of diagnosis in the cancer group). Cumulative expected survival is calculated using population life tables usually matched to the cancer cohort by age, sex and calendar year, to estimate the probability of survival for the comparison group. There are several approaches to estimating expected survival, the most common of which are the Ederer I and II[23,24] and the Hakulinen methods.[25] The difference in estimations of cumulative expected survival using these approaches is minimal in many situations, and will not be discussed further here. For each time interval an RSR (strictly speaking, a ratio of two proportions) is calculated. The probability of surviving until the end of that interval in the cancer group is divided by the probability of surviving until the end of that interval for the external comparison group. The (usually) lower survival in the cancer group is assumed to reflect cancer-specific deaths in this group. Alternatively, rather than taking the ratio of the two survival proportions, the risk of death (cumulative mortality) in the cancer group can be compared with that in the general population group, and the excess mortality in the cancer group can be estimated.

Systematic error in single population relative survival estimates

The main sources of systematic error in the calculation of relative survival and excess mortality are summarized in Table 1. For relative survival analyses we do not have to differentiate between cancer-specific mortality and cancer-consequent mortality. The most important source of bias specific to relative survival analysis is the potential for lack of comparability between the cancer group, for whom observed survival is calculated, and the external comparison group, which is used to calculate expected survival.

Non-comparability between observed and expected mortality

The assumption of comparability will be invalid if a factor that influences mortality from other causes is distributed differently between the cancer and the external comparison group, i.e. if the external comparison group does not provide a valid estimate of the expected mortality (in the absence of cancer) in the cancer group. For example, Baade *et al*. found that patients with melanoma had a reduced risk of non-cancer death. They postulated that this might be because melanoma is more common among more-affluent people, who tend to have better health.[26] If this is correct, then using general population background mortality rates to estimate relative survival among melanoma patients will underestimate their expected survival and therefore overestimate their relative survival.

Patients with smoking-related cancers such as lung cancer will have a notably higher tobacco exposure compared with the general population, so their risk of death from other tobacco-related conditions will be considerably greater.[1,27] There are a number of other factors that are associated with cancer and also increase the risk of other diseases including obesity, physical activity and diet,[28–32] and might potentially cause bias in the RSR calculation in the absence of comparable life tables.

There are few studies that have investigated the issue of non-cancer deaths among cancer patients. Brown *et al*.[18] found that cancer patients had considerably higher mortality rates from non-cancer causes, but were unable to establish whether this was due to miscoding of deaths or to a higher risk of other deaths among cancer patients. More recently, Baade *et al*.[26] found that cancer patients in Australia had 50% higher background mortality than the general population, but that this varied considerably by cancer type. So, for example, melanoma patients actually had significantly reduced background mortality, whereas lung cancer patients had 400% higher background mortality than the general population. Of note is that in order to make these estimations, the authors were dependent on accurate cause of death data. It is likely that there was some misclassification of cancer-related deaths as non-cancer

Table 5 Estimated 5-year RSRs and excess mortality rate ratios for varying ratios of background mortality rates for 'cancer' compared with 'non-cancer' cohort for cancers with good, moderate and poor survival, assuming cohorts with identical age structure and a fixed mortality rate from non-cancer causes over the follow-up period

	Background mortality rate in 'non-cancer' and 'cancer' cohorts per year	Ratio of background mortality in 'cancer' cf 'non-cancer' cohorts	Relative survival	Excess mortality rate ratio compared with 'good survival'	Excess mortality rate ratio compared with 'moderate survival'
Good survival	0.023; 0.023	1.0	**0.79**	**–**	**–**
	0.023; 0.025	1.1	0.78	–	–
	0.023; 0.027	1.2	0.77	–	–
	0.023; 0.034	1.5	0.75	–	–
	0.023; 0. 046	2.0	0.70	–	–
Moderate survival	0.023; 0.023	1.0	**0.48**	**3.12**	**–**
	0.023; 0.025	1.1	0.48	3.02	–
	0.023; 0.027	1.2	0.47	2.90	–
	0.023; 0.034	1.5	0.45	2.69	–
	0.023; 0.046	2.0	0.43	2.42	–
Poor survival	0.023; 0.023	1.0	**0.16**	**7.98**	**2.56**
	0.023; 0.025	1.1	0.15	7.62	2.53
	0.023; 0.027	1.2	0.15	7.30	2.52
	0.023; 0.034	1.5	0.15	6.57	2.44
	0.023; 0.046	2.0	0.14	5.61	2.32

The bold estimates are the true RSRs and excess mortality rate ratios.

deaths, so the latter may have been inflated. However, the authors proposed that only a small proportion of the excess mortality was likely to be related to treatment, and the remainder was likely to be due to cancer patients having different demographic characteristics (such as socio-economic and ethnic differences) and a higher prevalence of risk factors (such as smoking) associated with both cancer and other causes of death than the general population.

Conversely, in the calculation of RSRs, the external comparison population is assumed to be free from cancer, allowing the assumption that any excess mortality among the cancer patients is, in fact, due to cancer. In practice, the comparison group is usually a population that includes some patients with the cancer in question. However, because cancer is a rare outcome in the general population, this effect is very small, and of little practical importance.[8]

There is relatively little quantitative information on the actual size of this non-comparability bias in specific studies. We have calculated 5-year relative survival estimates for hypothetical cancer cohorts using the same scenarios as for the cancer-specific survival simulations (Table 5). We have assumed the cancer and comparison groups are identical in all respects apart from the excess mortality due to cancer in the former group. Non-cancer mortality is, as previously, constant throughout the follow-up period for the initial calculation of RSRs, with the timing of the deaths being identical to the previous example. In our simulations we have then assumed that the non-cancer mortality rate among those with cancer is, in fact, higher than those without by specified ratios (1.1, 1.2, 1.5, 2.0). We have then recalculated the RSR using the observed survival, including the higher mortality rate from other causes among the cancer group, and the original expected survival from the comparison group, as would occur in real-life situations. Table 5 shows the results. When the background mortality for the cancer and comparison groups is identical, we obtain precisely the same findings as for the unbiased cancer-specific survival estimates (Table 2). As the background mortality increases among those with cancer, the relative survival estimates are progressively underestimated (Table 5). If we assume that the background mortality for cancer patients is 50% higher than that of the general population,[26] the relative survival estimates are biased downwards by ~5%. However, given that Baade *et al.*[26] found that there is considerable variability in background mortality rates among patients with different cancer types, this bias may vary in importance with cancer type.

Systematic error in comparing sub-population relative survival

Relative survival estimates are often used to monitor population survival rates, and so direct comparisons between relative survival in different populations may be made. Although such comparisons are also theoretically possible with cancer-specific survival

rates, it is less commonly done. This is partly because, particularly in cross country comparisons, it is likely that cause-of-death data will vary systematically between regions and countries introducing differential misclassification bias. For this reason we will first discuss computational issues relating to direct comparisons of RSRs and to excess mortality rate modelling, before returning to the consideration of potential sources of bias relevant to both approaches.

Computational issues: direct standardization and excess cancer mortality analyses

When comparing relative survival in two populations, or sub-populations, two potential methods are available to control for potential confounders: direct standardization and excess mortality modelling.

When direct comparisons are made between RSRs, age (and possibly sex) standardization is likely to be required because survival rates may differ by such factors, and they may also vary between the comparison populations. Standardization enables a comparison of RSRs, which is unbiased with regards to the factors that have been standardized on. There are various approaches used to age standardize RSRs and these are summarized elsewhere.[33,34]

The techniques used in standardization of RSRs are analogous to those used for direct standardization of cancer incidence rates, and therefore carry the same advantages and disadvantages. In particular, directly standardized incidence or survival rates have the advantage that they can be compared across populations provided that the same standard population distribution has been used in each instance. However, directly standardized incidence or survival rates have the disadvantage that they are highly unstable when there are small numbers in particular strata (e.g. an age–gender–stage strata).[22] This problem is overcome by using excess mortality modelling with maximum likelihood methods of estimation.

Excess mortality is the difference between the observed all-cause mortality of the cancer patients and the expected mortality estimated from the comparison group. Excess mortality models subtract the latter estimates from the former to estimate the excess mortality rate. Modelling excess mortality allows assessment of the effect of multiple variables on excess mortality (and therefore survival). Cox regression used in cause-specific survival and Poisson regression used in excess mortality modelling are very similar, with the only difference being that follow-up time is treated as a continuous variable in the former, in the latter it is categorized into discreet time units (e.g. person-years or person-months).[3,35] Both methods are statistically robust and not subject to the same problems of instability as standardized RSRs.

Non-comparability between observed and expected mortality

Whichever computational methods are used, bias may occur in comparisons of relative survival in sub-populations if the observed and expected mortality for each sub-population are not comparable. This situation is shown in Figure 4. This differs from Figures 2 and 3 in which the outcome under study was cancer-specific mortality. In Figure 4 the outcome is excess total mortality (i.e. the observed mortality in the cancer patient group minus the mortality expected on the basis of the mortality rates in the external comparison group). In this hypothetical example, ethnicity is not a cause of excess mortality from any cause (cancer-specific, cancer-consequence or other mortality), but is associated with measured excess other mortality. In this situation, exposure (e.g. ethnicity) will show a spurious association with measured excess total mortality. This problem of measurement error for expected mortality can occur for two main reasons: (i) the external comparison group is non-comparable (non-exchangeable) and is therefore not appropriate for estimating expected mortality in the cancer group; or (ii) exposure (e.g. ethnicity) is not classified in the same way in cancer registrations and national mortality rates, and the external comparison mortality rates are therefore invalid. We will consider each of these two issues in turn.

Non-comparability of the cancer group and the external comparison group

As noted above, for comparisons of RSRs in sub-populations, it is important to ensure that the

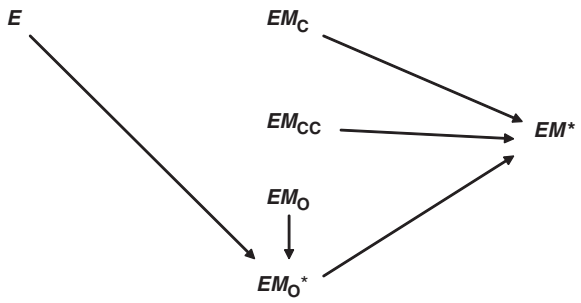

Figure 4 Directed acyclic graph showing potential bias due to measurement error of expected mortality after comparing ethnic sub-populations. *E*, ethnicity; *EM**, measured excess total mortality; EM_C, excess cancer-specific mortality; EM_{CC}, excess mortality that is a consequence of cancer, but is not cancer specific; EM_O, excess other mortality; $EM_O{}^*$, measured excess other mortality. Ethnicity is associated with measurement error for excess other mortality (which occurs because the national death rates are not appropriate for estimating excess mortality in the cancer patients, or because ethnicity is not classified in the same way in cancer registrations and national mortality rates); this biases the association of ethnicity with measured excess mortality.

appropriate life tables have been used, otherwise the RSRs (the outcomes being compared) are mismeasured—and often differentially so. For example, Māori New Zealanders have a lower life expectancy than non-Māori. If general population life tables are used to calculate expected survival of Māori patients with cancer, their expected survival is over estimated, leading to an under-estimation of net cancer survival and exaggerating disparities in cancer survival between Māori and non-Māori. The most common way of dealing with this issue is to use sub-population specific, in this case, ethnic-specific life tables. Examples of this approach include a study by Dickman *et al*.,[2] who stratified both cancer and comparison populations to calculate RSRs by social class, and Jeffreys *et al*., who used ethnic-specific life tables to calculate RSRs for different ethnic groups in New Zealand.[36] Table 6 shows an example using New Zealand data for colon cancer. The 5-year RSR was 0.42 for Māori when using general population life tables, but it increased to 0.45 when using a Māori-specific life table. Of course, this approach may be limited by the availability of appropriate life tables, and may become problematic if multiple strata are required.

Even having used the correct socio-demographic life tables, as described above, people who develop cancer may also have higher (or lower) background mortality rates than the same demographic group from the general population. Table 5 illustrates this bias for RSRs. The last two columns of Table 5 show the excess mortality rate ratios for the moderate and poor survival groups compared with the good and moderate survival groups, using the underlying excess mortality rates for each given scenario that are consistent with the RSRs. The simulated excess mortality rate ratios are for scenarios where the ratios of background mortality in the cancer and comparison groups are the same for the compared good, moderate and poor survival groups (i.e. non-differential misclassification bias). This might be the case if, for example, we were calculating mortality ratios for regional (moderate survival) and distant (poor survival) stage of disease at diagnosis compared with localized stage (good survival). In this scenario, there is notable underestimation of the excess mortality rate ratios. For example, if even after using the correct subpopulation population mortality rates the 'other' mortality among the cancer patients was actually 2-fold greater than that in the equivalent subgroup in the general population, the excess mortality rate ratio was underestimated by 33% for the moderate compared with good survival groups (i.e. a shift in excess rate ratio from 3.12 to 2.42).

In excess mortality models, variables that are associated with excess mortality only through their association with the observed survival among the cancer patients, particularly cancer-related variables such as grade of cancer, can be adjusted for fully

Table 6 Five-year survival data by age for Māori and non-Māori patients with colon cancer diagnosed between 1994 and 2002

	Māori						Non-Māori						All
Age group (years)	No. of cases	Cumulative observed survival (So)	Cumulative expected survival (Se)[a]	Cumulative expected survival (Se)[b]	Cumulative So/Se[a] (RSR)	Cumulative So/Se[b] (RSR)	No. of cases	Cumulative observed survival (So)	Cumulative expected survival (Se)[a]	Cumulative expected survival (Se)[b]	Cumulative So/Se[a] (RSR)	Cumulative So/Se[b] (RSR)	No. of cases
15–44	51	0.58	0.99	0.99	0.58	0.58	375	0.63	0.99	0.99	0.64	0.64	426
45–54	60	0.38	0.98	0.95	0.39	0.40	1045	0.59	0.98	0.98	0.60	0.60	1105
55–64	128	0.45	0.94	0.87	0.48	0.52	2799	0.58	0.94	0.95	0.62	0.61	2927
65–74	98	0.24	0.84	0.74	0.28	0.32	4586	0.52	0.85	0.86	0.61	0.60	4684
75–99	54	0.18	0.64	0.56	0.28	0.32	5353	0.37	0.60	0.61	0.61	0.60	5407
All ages	391	0.37	0.89	0.82	0.42	0.45	14 158	0.48	0.79	0.80	0.61	0.61	14 549

[a]Using general population life tables.
[b]Using ethnic-specific life tables.

(assuming no measurement error) by including them directly in the model, just as in proportional hazards modelling. However, factors that are associated with both survival of the cancer patients and background estimated survival are more challenging.

For example, suppose that an analysis involves comparing excess mortality in patients from two ethnic groups. Suppose also that national mortality rates are available by age, gender and ethnicity, but not by smoking. An excess mortality rate model would estimate the excess mortality separately for each age–gender–ethnicity group, and then additionally adjust for smoking in a multiple regression model. For example, we may wish to investigate survival disparities between Māori and non-Māori patients with lung cancer. Because smoking rates are higher among Māori, and cancer survival may be lower among smokers, we would include smoking as a covariate in an excess mortality model comparing excess mortality between Māori and non-Māori lung cancer patients. In doing so we would aim to adjust for any difference in prevalence of smoking among Māori and non-Māori lung cancer patients, and the impact of that on excess mortality. However, given that a higher proportion of Māori smoke in the general population, and that smokers have a lower life expectancy from causes other than lung cancer, the smoking variable will capture both the impact of smoking on survival from the given cancer and the misclassification bias in background estimated mortality arising from not have smoking specific life tables. Thus, although factors such as smoking can be included directly in the excess mortality rate model, full adjustment and correct interpretation of the variable's coefficient will only be possible if stratum-specific life tables are also used to generate stratum-specific excess mortality. Once again, the likely magnitude of such potential biases is context specific, and there are relatively few quantitative estimates; we will be exploring this further in subsequent publications using linked census-mortality data, which includes information on smoking.

Misclassification of exposure

Two lesser, but potentially non-trivial, other misclassification biases include numerator–denominator bias in the subpopulation life tables and misclassification of exposure (subpopulation membership) in the external population mortality rates. Regarding the former, national (or regional) life tables used in the calculation of expected mortality (or survival) may suffer from numerator–denominator bias because the classification of factors such as ethnicity and socio-economic status may differ between the numerator data (usually obtained from death certificates), and the denominator data (usually obtained from the census). For example, in the New Zealand Census-Mortality Study, census data were linked to mortality data allowing the extent of the undercounting of Māori deaths to be assessed. In that setting, during the 1980s and early 1990s there was undercounting of Māori deaths by about a quarter.[37] This means that over that time, if appropriate adjustments were not made, expected mortality would be underestimated and excess mortality overestimated for Māori. This bias can be eliminated by using mortality rates from population-based cohort studies in which the numerator data are the same as the denominator data, but this approach depends on the availability of such linked cohort data, and there are currently few examples internationally of such data being routinely available.

Secondly, even when numerator–denominator bias has been eliminated, bias may occur because the classification of the 'exposure' of interest (e.g. ethnic subpopulation group membership) is different in the cancer group and the comparison group. This is likely to be of much smaller magnitude than numerator–denominator bias, but nevertheless is a potential bias that should be considered. It may occur, for example, if the quality of data on factors such as ethnicity or socio-economic status differs between the cancer registry and the source of the comparison mortality rates, or simply if the questions on ethnicity are asked in different ways in the two data sets. For example, if socio-economic status can be measured more accurately using information (e.g. occupation or address) from the Cancer registry than when using information from the Census, then the comparison population mortality rates will show smaller socio-economic differences in mortality than would have been obtained if more accurate data had been available. As a result, the estimates of excess mortality for a particular socio-economic group, and the comparisons of excess mortality (or survival) between socio-economic groups, may be biased.

Both of these potential biases resulting from misclassification will only occur in excess mortality (or relative survival comparisons), and cannot occur in cancer-specific survival analyses because the latter involve a single data set with each person having a single classification for factors such as ethnicity.

Summary and conclusions

In summary, the main concern in cancer-specific analyses, in addition to the usual concerns that apply to all observational studies, is the potential for bias due to misclassification of cancer-specific cause of death on death certificates. The likelihood that this will occur will be specific to particular countries, populations, comparisons (e.g. rates of misclassification may differ by ethnicity but not by socio-economic status, or vice versa), and the cancer site under study. Thus, it is not possible to draw general conclusions about how strong such biases may be, or even in what direction they may be. However, assuming a range of plausible values for sensitivity and specificity

of the cause of death indicates that although this bias might be of some concern in descriptive studies for cancers with poor or moderate survival, its impact is likely to be relatively small in etiological studies where hazard ratios are calculated, unless there is considerable differential misclassification of death between comparison groups. Beyond issues of valid classification of cancer-specific deaths, however, cause-specific methods are likely to have poor validity for analyses of cancer-consequent mortality where it may simply be impossible to assign each death as 'cancer-consequent' or not.

For relative survival analyses, the main potential for bias is in lack of comparability in 'background' mortality between the cancer group and the external general population comparison group. Again, it is difficult to draw general conclusions about the strength and direction of these biases, but they are more likely to be problematic for cancers for which there are risk factors strongly associated with other causes of death (e.g. smoking-related cancers). Our simulations suggest that, although relative survival estimates are reasonably robust in relation to this bias, lack of comparability in background mortality can result in quite considerable bias in the RSRs when two groups with very different survival (e.g. localized compared with regional extent of disease) are compared.

The choice of method therefore depends on the study objectives, the type of data available and the appropriateness of the assumptions underlying the two methods, particularly the availability of accurate cause of death data for cancer-specific analyses and the availability of an appropriate comparison group for relative survival analyses. In some situations, it may be possible and desirable to apply both methods, compare results and attempt to explain any differences in study findings. Of note is when such comparisons are made, the differences in findings between different analytic approaches are not large, but tend to vary by cancer site, stage and underlying mortality.[2,8,38,39]

We therefore conclude that both cause-specific survival and relative survival are potentially valid epidemiological methods in population-based cancer survival studies, and that the choice of method is specific to the population and cancer type under study, and whether the proposed analyses are descriptive or analytic. A comprehensive understanding of the likely biases arising from each of the two methods is necessary for appropriate study design and interpretation of study findings.

Funding

The Health Inequalities Research Programme is supported by funding from the Health Research Council of New Zealand, Cancer Society and Ministry of Health. The Centre for Public Health Research is supported by a Programme Grant from the Health Research Council of New Zealand.

Acknowledgements

The authors would like to thank Sarah Hill, Mona Jeffreys, Kristie Carter, Matt Soeberg and Sander Greenland for helpful comments on earlier drafts of this article.

Conflict of interest: None declared.

KEY MESSAGES

- There are two main methods of quantifying cancer patient survival: cancer-specific survival and relative survival.
- The main source of bias for cancer-specific survival approaches relates to potential misclassification of cause of death.
- Our simulations for cancer-specific survival suggest that misclassification bias is likely to be particularly of concern in descriptive studies of cancers with poor or moderate survival, but has relatively small impact in etiological studies comparing 2 or more groups if biases are assumed to be of similar magnitude across groups.
- The main source of bias for relative survival and excess mortality approaches relates to the potential for lack of comparability in background mortality between the cancer group and the external comparison group/s.
- Our simulations suggest that this non-comparability bias has a relatively small impact on descriptive RSR estimations for single populations unless there are risk factors strongly associated with other causes of death.

References

[1] Dickman PW, Adami HO. Interpreting trends in cancer patient survival. *J Intern Med* 2006;**260:**103–17.

[2] Dickman PW, Auvinen A, Voutilainen ET, Hakulinen T. Measuring social class differences in cancer patient survival: is it necessary to control for social class differences in general population mortality? A Finnish population-based study. *J Epidemiol Community Health* 1998;**52:**727–34.

[3] Dickman PW, Sloggett A, Hills M, Hakulinen T. Regression models for relative survival. *Stat Med* 2004;**23:**51–64.

[4] Rachet B, Woods LM, Mitry E *et al*. Cancer survival in England and Wales at the end of the 20th century. *Br J Cancer* 2008;**99:**S2–S10.

[5] Kalbfleisch JD, Prentice RL. *The Statistical Analysis of Failure Time Data*. New York: John Wiley, 1980.

[6] Kaplan EL, Meier P. Nonparametric estimation from incomplete observations. *J Am Stat Assoc* 1958;**53:**457–81.

[7] Gooley TA, Leisenring W, Crowley J, Storer BE. Estimation of failure probabilities in the presence of competing risks: new representations of old estimators. *Stat Med* 1999;**18:**695–706.

[8] Gamel JW, Vogel RL. Non-parametric comparison of relative versus cause-specific survival in Surveillance, Epidemiology and End Results (SEER) programme breast cancer patients. *Stat Methods Med Research* 2001;**10:**339–52.

[9] Satagopan JM, Ben-Porat L, Berwick M, Robson M, Kutler D, Auerbach AD. A note on competing risks in survival data analysis. *Br J Cancer* 2004;**91:**1229–35.

[10] Prentice RL, Kalbfleisch JD, Peterson AV Jr, Flournoy N, Farewell VT, Breslow NE. The analysis of failure times in the presence of competing risks. *Biometrics* 1978;**34:** 541–54.

[11] Platell CF, Semmens JB. Review of survival curves for colorectal cancer. *Dis Colon Rectum* 2004;**47:**2070–75.

[12] James DS, Bull AD. Information on death certificates: cause for concern? *J Clin Pathol* 1996;**49:**213–16.

[13] Maudsley G, Williams EM. ''Inaccuracy'' in death certification—where are we now? *J Public Health Med* 1996;**18:**59–66.

[14] Percy C, Stanek E 3rd, Gloeckler L. Accuracy of cancer death certificates and its effect on cancer mortality statistics. *Am J Public Health* 1981;**71:**242–50.

[15] World Health Organization. *Manual of the International Statistical Classification of Diseases, Injuries, and Causes of Death, based on the recommendations of the ninth revision conference, 1975*. Geneva: WHO, 1977.

[16] Samphier ML, Robertson C, Bloor MJ. A possible artefactual component in specific cause mortality gradients. Social class variations in the clinical accuracy of death certificates. *J Epidemiol Community Health* 1988;**42:**138–43.

[17] Welch HG, Black WC. Are deaths within 1 month of cancer-directed surgery attributed to cancer? *J Natl Cancer Inst* 2002;**94:**1066–70.

[18] Brown BW, Brauner C, Minnotte MC. Noncancer deaths in white adult cancer patients. *J Natl Cancer Inst* 1993;**85:** 979–87.

[19] Albertsen PC, Walters S, Hanley JA. A comparison of cause of death determination in men previously diagnosed with prostate cancer who died in 1985 or 1995. *J Urology* 2000;**163:**519–23.

[20] Goldacre MJ. Cause-specific mortality: understanding uncertain tips of the disease iceberg. *J Epidemiol Community Health* 1993;**47:**491–96.

[21] Pearce N, Checkoway H, Dement J. Exponential models for analyses of time-related factors, illustrated with asbestos textile worker mortality data. *J Occup Med* 1988;**30:**517–22.

[22] Rothman K, Greenland S, Lash TL (eds). *Modern Epidemiology*, 3rd edn. Philadelphia: Lippincott Williams and Wilkins, 2008.

[23] Ederer F, Axtell LM, Cutler SJ. The relative survival: a statistical methodology. *Natl Cancer Instit Monograph* 1961;**6:**101–21.

[24] Ederer F, Heise H. *Instructions to IBM 650 Programmers in Processing Survival Computations. Methodolgical Note No. 10*. Bethesda, Maryland: End Results Evaluation Section, National Cancer Institute, 1959.

[25] Hakulinen T. Cancer survival corrected for heterogeneity in patient withdrawal. *Biometrics* 1982;**38:**933–42.

[26] Baade PD, Fritschi L, Eakin EG. Non-cancer mortality among people diagnosed with cancer (Australia). *Cancer Causes Control* 2006;**17:**287–97.

[27] Ezzati M, Lopez AD. Estimates of global mortality attributable to smoking in 2000. *Lancet* 2003;**362:**847–52.

[28] Warburton DER, Nicol CW, Bredin SSD. Health benefits of physical activity: the evidence. *CMAJ* 2006;**174:**801–9.

[29] Kushi L, Giovannucci E. Dietary fat and cancer. *Am J Med* 2002;**113(Suppl 9B):**63S–70S.

[30] Key TJ, Allen NE, Spencer EA, Travis RC. The effect of diet on risk of cancer. *Lancet* 2002;**360:**861–68.

[31] Hutter RV. Cancer prevention and detection. Status report and future prospects. *Cancer* 1988;**61:**2372–78.

[32] Blakely T, Sarfati D, Shaw C. What proportion of cancer is due to obesity? *NZ Med J* 2009;**122:**3478.

[33] Pokhrel A, Hakulinen T. Age-standardisation of relative survival ratios of cancer patients in a comparison between countries, genders and time periods. *Eur J Cancer* 2009;**45:** 642–47.

[34] Pokhrel A, Hakulinen T. How to interpret the relative survival ratios of cancer patients. *Eur J Cancer* 2008;**44:** 2661–67.

[35] Giorgi R, Hedelin G, Schaffer P. Relative survival: comparison of regressive models and advice for the user. *J Epidemiol Biostat* 2001;**6:**455–62.

[36] Jeffreys M, Stevanovic S, Tobias M *et al*. Ethnic differences in cancer survival in New Zealand: linkage study. *Am J Public Health* 2004;**95:**834–37.

[37] Ajwani S, Blakely T, Robson B, Atkinson J, Kiro C. Unlocking the numerator-denominator bias III: adjustment ratios by ethnicity for 1981-1999 mortality data. The New Zealand Census-Mortality Study. *N Z Med J* 2003;**116:**U456.

[38] Boer R, Ries L, van Ballegooijen M, Feuer E, Legler J, Habbema D. *Ambiguities in Calculating Cancer Patient Survival: The SEER Experience for Colorectal and Prostate Cancer, Technical Report #2003–05*. Statistical Research and Applications Branch: National Cancer Institute, 2003.

[39] Kravdal O. A cancer survival model that takes sociodemographic variations in ''normal'' mortality into account: comparison with other models. *J Epidemiol Community Health* 2002;**56:**309–18.