Cancer incidence, mortality and survival by site for 14 regions of the world.

Colin D Mathers

Cynthia Boschi-Pinto

Alan D Lopez

Christopher JL Murray

Global Programme on Evidence for Health Policy Discussion Paper No. 13

World Health Organization
2001

1. Introduction

Cancer was estimated to account for about 7 million deaths (12% of all deaths) worldwide in 2000 *(1)*, only preceded by cardiovascular diseases (30 % of all deaths), and by infectious and parasitic diseases (19%). Cancer was also estimated to account for almost 6% of the entire global burden of disease in that same year *(1)*. More than 70% of all cancer deaths occurred in low- and middle-income countries and, although the risk of developing/dying from it is still higher in the developed regions of the world, the control of communicable diseases as well as the ageing of the population in developing countries, point to an increasing burden of cancer worldwide. In fact, Pisani et al *(2)* have projected a 30% increase in the number of cancer deaths in developed countries, and more than twice this amount (71%), in developing countries, between 1990 and 2010, due to demographic changes alone. Rising incidence will only add to this burden.

Attempts have been made to quantify the global burden of cancer, and estimate site-specific cancer mortality and morbidity (2-6). Such studies are of considerable importance in helping to better allocate resources towards the prevention and treatment of cancer. In the early 1980's, Doll & Peto *(7)* were already calling attention to the evidence about the avoidability of cancer. According to these authors, approximately 75% of the cases of cancer in most parts of the US, in 1970, could have been avoided. More recently, Parkin et al *(8)* have estimated that there would have been 22.5% fewer cases of cancers in the developing world in 1990, if infections with hepatitis B virus, hepatitis C virus, human papillomaviruses, EBV, HTLV-I, HIV, helicobacter pylori, schistossoma, and liver flukes had been prevented. Another estimate suggests that 230,000 deaths (4.4% of all cancer deaths) from liver cancer could have been avoided with only immunization against hepatitis B *(2)*. According to Murray & Lopez *(3)*, cancer of the trachea, bronchus and lung was the 10th leading cause of death in the world in 1990, being the third in the developed regions. Smoking was estimated to be responsible for another 20% of all cancer deaths, all of which are preventable *(2)*. While the need for reliable estimates of cancer burden is clear, much more work is still needed to improve their reliability. Parallel to the development of national systems of death registration, there is a need to develop new methodologies to help improve the accuracy of the current estimates, based on existing data. In this paper, we outline an approach to measuring cancer mortality and incidence based on existing sources.

While vital registration of causes of death and national cancer registries are perhaps the best source of data on cancer disease burden, mortality data are still scarce, poor or even unavailable for some regions of the world (see Section 2). Innovative methods will thus continue to be needed to exploit available data. Estimating mortality from morbidity and, especially, morbidity from mortality was a common practice in the 70's and 80's *(9;10)*. More recently, some authors have also used information on incidence and survival to estimate cancer death *(2;6)*, but by means of a different methodology. Still others have made use of vital statistics and cancer incidence data to predict the number of new cancer cases and deaths for the US in the subsequent year *(11)*.

Globocan 2000 estimates *(6)* for global cancer incidence and mortality are shown in Table 1. The mortality estimates are based on vital registration data, where available, and for other regions, on mortality estimates derived from survival models using estimates of cancer incidence derived from available cancer registry data in each region. As described in Section 2, the Global Burden of Disease 2000 project has also estimated total global cancer mortality as part of its detailed analysis of all-cause mortality levels, and cause of death distributions, for 191 WHO Member States. The GBD 2000 estimate for global cancer deaths is 11% higher

than the Globocan 2000 estimates, and is substantially higher for Africa and South East Asia. It is quite likely that cancer registry data in these regions systematically underestimates both incidence and mortality. The GBD 2000 deals with this problem by estimating total cancer mortality for each Member State, starting from an analysis of the overall mortality envelope, in order to ensure that the cause-specific estimates add to the total all cause mortality by age and sex, and that there is not systematic underestimation or double counting of deaths (see Section 2). For countries and regions where information on the distribution of cancer deaths is not available, a similar approach has been taken to that used in Globocan 2000, of using available incidence distributions by site, together with estimates of site-specific survival, to estimate the distribution of cancer deaths by site.

Table 1. Globocan 2000 estimates of global cancer incidence and mortality, 2000

Site	Incidence	Mortality
Mouth and oropharynx cancers	462,979	250,900
Oesophagus cancer	386,612	350,841
Stomach cancer	950,319	714,452
Colon and rectum cancers	944,677	510,021
Liver cancer	554,344	536,904
Pancreas cancer	201,506	200,865
Trachea, bronchus and lung cancers	1,211,804	1,089,258
Melanoma	131,469	37,654 [a]
Breast cancer	1,017,207	371,680
Cervix uteri cancer	472,387	232,153
Corpus uteri cancer	185,951	44,359
Ovary cancer	188,482	114,488
Prostate cancer	536,279	202,201
Bladder cancer	326,523	131,681
Lymphomas and multiple myeloma	405,995	236,494
Leukaemia	255,932	209,328
Other sites	1,678,413	1,027,317 [b]
Total	**9,910,878**	**6,260,596**

Source: GLOBOCAN 2000 (6).

a Does not include other skin cancers

b Includes unknown primary site and Kaposi's sarcoma

In this paper, we present a detailed model to estimate cancer survival in different parts of the world as a key input to estimate the distribution of cancer deaths by site. Cancer sites for which survival was calculated were mouth and pharynx (ICD-9 140-149), oesophagus (ICD-9 150), stomach (ICD-9 151), colon and rectum (ICD-9 153, 154), liver (ICD-9 155), pancreas (ICD-9 157), lung (ICD-9 162), melanoma of skin (ICD-9 172), female breast (ICD-9 174), cervix uterine (ICD-9 180), corpus uteri (ICD-9 182), ovary (ICD-9 183), prostate (ICD-9 185), bladder (ICD-9 188), lymphomas (ICD-9 200-203), leukemia (ICD-9 204-208), and other cancer (balance of ICD-9 140-208). On the basis of available published information on age-, sex-, and site-specific cancer incidence and survival, we developed an algorithm to estimate region-specific cancer incidence, survival and death distributions, rates and absolute numbers of cases for the year 2000.

These data have been used to estimate the global burden of cancer as part of the Global Burden of Disease 2000 project (GD 2000) *(12)*. Version 1 estimates of cancer burden in DALYs were published in the World Health Report 2001 *(1)* and more detailed estimates by

site, age and sex for GBD 2000 subregions are available in a Discussion Paper *(12)* and on the WHO website at www.who.int/evidence. The methods for estimation of disease burden are described elsewhere *(13)* and will be revised to take into account new information on survival, incidence and long-term sequelae for the World Health Report 2002.

Some characteristics of cancer epidemiology and of its natural history, make it relatively simple to calculate estimates of mortality. Cancer incidence is reasonably stable over time. However, as procedures of detection vary over time, incidence may rise abruptly, which is artifactual, due only to increased detection. For some cancer sites, incidence increased in earlier years and has recently started to decline. An example of this is prostate cancer *(14;15)*. Increases in the incidence of cancer of the brain have also been the focus of debate in the literature *(16;17)*, but, as opposed to prostate cancer, its increase seems to be less affected by artifacts than that of prostate cancer. Survival, which is itself basically dependent on the development of new techniques of detection as well as of new treatment, changes relatively slowly.

Sankaranarayanan et al *(18)* have published detailed data on cancer survival for selected sites in the late 1980s for nine cancer registries in developing countries (see Table 2). There are substantial variations in relative 5-year survival (all ages) for some sites; these variations are even larger, and fluctuate substantially with age, when the age-sex specific survival estimates are examined. In some cases, survival rates are higher than those reported for developed countries. This may reflect incomplete follow-up and case finding in some instances, and also

Table 2. Relative 5-year survival (%) by cancer site for registries in some developing regions of the world.

Sex Site	China Qidong	China Shanghai	India Bangalore	India Bombay	India Madras	Philippine s Rizal	Thailand Chiang Mai	Thailand Khon Kaen	Cuba
	1982-91	1988-91	1982-91	1988-92	1982-96	1987	1983-92	1985-92	1988-91
Males									
Oesophagus	4.2	10.5			6.8		2.2	33.0	
Stomach	15.1	24.8			7.7	18.3	9.2	14.9	
Colorectal	27.6	42.3		i		34.6	33.6	31.1	36.9
Liver	1.8	4.3				13.3	0.0	8.5	
Pancreas	5.8	6.9			7.2		4.3	4.5	
Lung	3.4	12.1			7.2	7.0	3.0	10.3	10
Melanoma		42.5					43.8	57.4	
Prostate		40.1				21.3	42.3	41.1	45.1
Bladder	43.7	64.1			25.2		39.7	61.5	
Leukemias	6.1	15.1			20.2	18.8	10.2	22.0	22.3
Females									
Oesophagus	4.0	12.7			6.1		6.3	22.5	
Stomach	13.0	22.3			9.2	4.9	7.7	23.3	
Colorectal	25.3	44.1				31.2	30.1	39.2	41.6
Liver	2.7	4.8				19.0	1.1	8.3	
Pancreas	5.1	5.1			0.0		3.0	5.1	
Lung	4.1	11.3			10.2	7.9	3.1	9.5	12.6
Melanoma		48.9					44.3	45.3	
Breast	55.7	72.0	45.1	55.1	49.5	45.6	63.7	47.1	60.8
Cervix	33.6	51.9	40.4	50.7	60.0	29.0	68.2	57.5	55.9
Corpus uteri		76.8					69.5	78.7	60.9
Ovary		44.2					44.9	35.6	43.3
Bladder	21.3	51.2			15.0		35.2	39.0	
Leukemias	3.2	15.8	26.4		23.5	16.3	10.6	19.2	20

a Adapted from Sankaranarayanan et al, *(18)*.

the effects of random variation with small numbers of cases. To deal with these issues, and to ensure that site-specific cancer incidence and mortality estimates vary smoothly and appropriately across age groups, and to ensure that all available evidence, including historical trends in survival in developed countries, is taken into account, we have developed an age-period-cohort survival model which enables us to estimate relative survival by site, age and sex for all regions of the world.

For regions where detailed data on the distribution of cancer deaths by site is not available, we have used incidence estimates (drawn to a large extent from the comprehensive estimates undertaken for Globocan 2000 supplemented by some other incidence studies) together with cancer survival data from all regions of the world to construct a detailed model to estimate cancer survival in different parts of the world as a key input to estimate the distribution of cancer deaths by site. These distributions were then used, where necessary, to distribute total cancer deaths (estimated as described in Section 2) to various sites. In the following Section 3, we describe the cancer survival model. The resulting estimates of cancer deaths by site are compared with the Globocan estimates in Section 4. The use of the survival model to estimate cancer incidence is then described in Section 5.

2. Global cancer mortality in the year 2000

In this Section, we describe the Global Burden of Disease 2000 approach to the estimation of global cancer mortality and compare it with the Globocan 2000 estimates made by the International Agency on Research in Cancer (IARC) *(6)*.

The GBD 2000 study has estimated the all-cause age-specific death rates, by sex, for all 191 WHO Member States for the year 2000 *(19)*. The importance of this approach for disease-specific mortality estimates cannot be overemphasized. The number of deaths, by age and sex, provides an essential "envelope" which constrains individual disease and injury estimates of deaths. Competing claims for the magnitude of deaths from various causes must be reconciled within this envelope. The sum of deaths from all specific causes for any sex-age group must sum to the total number of deaths for that age-sex group estimated via the data sources and methods described below.

Complete or incomplete vital registration data together with sample registration systems now cover 74% of global mortality in 128 countries. Survey data and indirect demographic techniques provide information on levels of child and adult mortality for the remaining 26% of estimated global mortality. The available sources of mortality data for the 14 mortality subregions of the GBD 2000 are summarised in Table 3. Methods used to estimate global all-cause mortality from these data are described elsewhere *(12)*.

Causes of death for the WHO subregions and the world have been estimated based on data from national vital registration systems that capture about 17 million deaths annually. In addition, information from sample registration systems, population laboratories and epidemiological analyses of specific conditions have been used to improve estimates of the cause of death patterns *(12)*. Cause of death data have been carefully analysed to take into account incomplete coverage of vital registration in countries and the likely differences in cause of death patterns that would be expected in the uncovered and often poorer sub-populations. Techniques to undertake this analysis have been developed based on the global burden of disease study *(20)* and further refined using a much more extensive database and more robust modelling techniques *(21)*.

Table 3. Mortality data sources (number of Member States with recent deaths coverage) by WHO subregion for the GBD2000

Subregion	Complete vital statistics (coverage of 95%+)	Incomplete vital statistics	Sample registration and surveillance systems	Surveys and indirect demographic methods	No recent data	Total Member States
Afro D	2	2	0	18	4	26
Afro E	0	2	1	13	4	20
Amro A	3	0	0	0	0	3
Amro B	17	9	0	0	0	26
Amro D	0	4	0	1	1	6
Emro B	4	4	0	5	0	13
Emro D	0	2	0	5	2	9
Euro A	26	0	0	0	0	26
Euro B	7	9	0	0	0	16
Euro C	8	1	0	0	0	9
Searo B	1	1	0	1	0	3
Searo D	0	2	2	1	2	7
Wpro A	4	1	0	0	0	5
Wpro B	3	12	1	6	0	22
Total	**75**	**49**	**4**	**50**	**13**	**191**

Source (12)

As a general rule, vital registration data, suitably corrected for ill-defined coding and probable systematic biases in certifying deaths to non-specific vascular, cancer and injury codes were used to estimate the cause of death pattern. Vital registration data to do so was available for 65 countries. In a further 28 countries, cause of death models were used to correct vital registration data by age and sex to yield more plausible patterns across Groups I, II and III. The distribution of specific causes within groups was then based on the recorded cause of death patterns from vital registration data. The resulting estimates were then systematically corrected on the basis of other epidemiological evidence from registries, community studies and disease surveillance systems.

For China and India, cause patterns of mortality were based on existing mortality registration systems, namely the Disease Surveillance Points system (DSP) and the Vital Registration System of the Ministry of Health in China, and the Medical Certificate of Cause of Death (MCCD) for urban India and the Annual Survey of Causes of Death (SCD)) for rural areas of India. For all other countries lacking vital registration data, cause of death models were used to firstly estimate the maximum likelihood distribution of deaths across the broad categories of communicable, non-communicable and injuries, based on estimated total mortality rates and income *(21)*. A regional model pattern of specific causes of death was then constructed based on local vital registration and *verbal autopsy* data and this proportionate distribution was then applied within each broad cause group. Finally, the resulting estimates were then adjusted based on other epidemiological evidence from specific disease studies.

Table 4 shows the resulting regional estimates of total cancer mortality (all sites) for the GBD 2000 and compares it with regional estimates from Globocan 2000 *(6)*. The Globocan estimates have been adjusted to exclude Karposi's sarcoma deaths and the proportion of NHL due to HIV/AIDS (see Section 4). These two sets of estimates are also compared in Figure 1. Overall, the GBD 2000 estimate for global cancer deaths is 11% higher than the GLOBOCAN 2000 estimate. This difference is predominantly due to the very large difference in the AFRO

region (GBD estimate is almost double that of GLOBOCAN) and the SEARO region (where the GBD estimate is one third higher than the GLOBOCAN estimate).

The Globocan estimates shown in Table 4 have been adjusted to exclude cancer deaths attributable to HIV/AIDS (included under HIV/AIDS deaths in the GBD 2000) but they have not been adjusted to include a proportion of deaths coded to ill-defined causes in vital registration data. The GBD 2000 redistributes these deaths pro-rata among Group 1 and Group 2 causes (communicable, maternal, perinatal, and non-communicable diseases). For this reason, we would expect GBD estimates of cancer deaths to be higher than GLOBOCAN estimates in regions with good vital registration data. In other regions, a more fundamental reason for the differences between the two sets of estimates relates to the methods used. The GLOBOCAN estimates are based on either cancer incidence data from cancer registries in the region (with a survival model used to estimate deaths) or on mortality data collected by regional cancer registries or other sources. Both these sources of data are likely to be incomplete and to result in underestimation of cancer deaths.

Table 4. GBD 2000 total cancer deaths by WHO region and comparison with GLOBOCAN 2000 estimated cancer deaths[a] by WHO region.

	Estimated cancer deaths ('000)						
	AFRO	**AMRO**	**EMRO**	**EURO**	**SEARO**	**WPRO**	**World**
GBD 2000	533	1,074	242	1,882	1,103	2,096	6,930
GLOBOCAN 2000	278	1,089	253	1,811	831	1,954	6,216
% difference (GBD – GLOBOCAN)	***92***	***-1***	***-4***	***4***	***33***	***7***	***11***

a Globocan estimates have been adjusted to exclude Karposi's sarcoma deaths and the proportion of NHL due to HIV/AIDS.

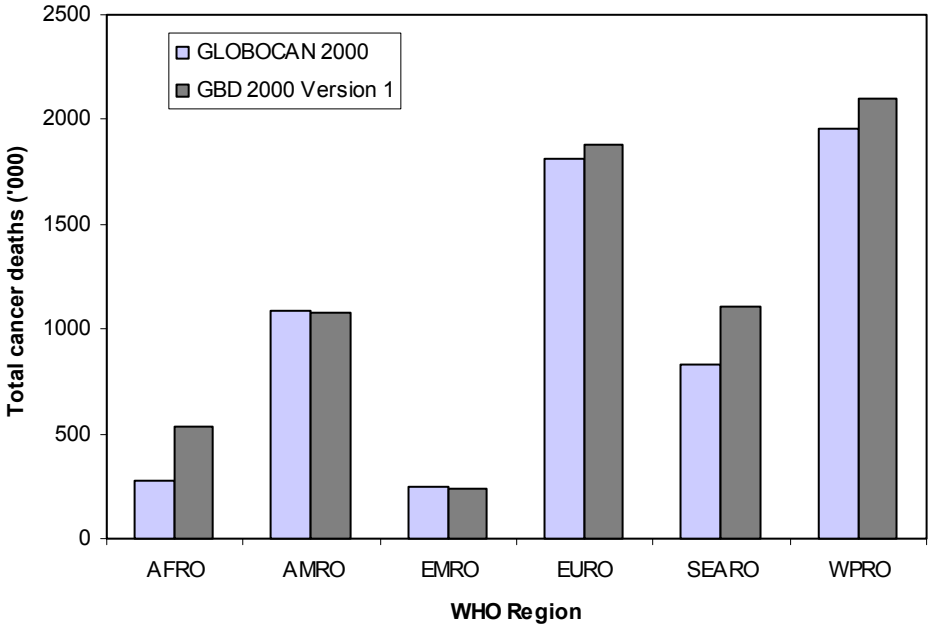

Figure 1. Total cancer deaths by WHO region, GBD 2000 and GLOBOCAN 2000 estimates

On the other hand, the GBD 2000 starts with data on the level of all-cause mortality, and uses available cause of death data and cause of death models, where such data is not available, to estimate the distribution of major cause groups, including malignant neoplasms (cancers). It is possible that these methods result in an overestimate of total cancer deaths in some regions, and work is underway to obtain additional data from these regions in order to check the validity of these estimates, and where appropriate, to improve them.

3. The cancer survival model

3.1 Data Sources

The data sources used to develop the cancer survival model were the National Cancer Institute Surveillance, Epidemiology, and End Results (SEER) statistical program (SEER*Stat), the Connecticut survival data from Cancer in Connecticut – Survival Experience, 1935-1962 *(22;23)* and the US vital statistics.

The SEER program is considered as the standard for quality among cancer registries around the world, being the most authoritative source of information on cancer incidence and survival in the United States. It includes data from population-based cancer registries, which collect cancer data on a routine basis, and covers approximately 14% of the US population *(22)*. SEER*Stat was created for the analysis of SEER and other cancer databases, and produces frequencies, rates, and survival statistics. We obtained cancer incidence and survival data from SEER*Stat to build our survival model.

The Connecticut State Department of Health published Cancer in Connecticut – Survival Experience *(23)*, which focused on the survival experience of patients from Connecticut only. Its data were based on the Connecticut Tumor Registry, which collects information on all cases of cancer diagnosed in the state of Connecticut since 1935, and carries out a lifetime follow-up of each of these patients in order to access survival. Relative survival rates for 1- 3-, 5-, and 10-year were available for some selected sites for the periods 1935-44, 1945-54, and 1955-63. We have used this source of data to obtain the relative survival data for the 30's, 40's, and 50's.

3.2 Multiplicative model for the relative interval survival.

In order to estimate cancer death distribution for the regions where no mortality data is available, we made use of incidence and survival data – component measures of our outcome. We will define survival here as it is done in SEER*Stat: observed interval survival rate (OIS), expected interval survival rates (EIS), and relative interval survival rates (RIS). OIS is "the probability of surviving a specified time interval as calculated from the cohort of cancer cases". EIS is "the probability of surviving the specified time interval in the general US population. It has been generated from the US population and matched to the cohort cases by race, sex, age, and date at which age was coded". RIS is "the observed survival probability for the specified time interval adjusted for the expected survival. Such adjustment accounts for the general survival rate of the US population for race, sex, age, and date at which the age was coded". Cumulative survival rates (CS) can be obtained by simply multiplying consecutive interval survival rates.

Cancer patients are at risk of dying from both cancer and other causes of death, and the observed survival (OIS) is influenced by both. Expected survival (EIS) is the survival

experience of a comparable group of individuals who are at risk of death from causes other than the cancer under study. Because the relative survival is adjusted for the expected survival, based on the general mortality experience of the population, the relative interval survival (RIS) was chosen to be modelled. Mathematically, it can be defined as: $RIS = OIS / EIS$. *RIS* was directly obtained from the SEER database within SEER*Stat for every age group, sex, and cancer site.

The basic model was developed as a three-dimension age-period-cohort model, separately for each cancer site. To simplify notation below, we suppress the subscript s for cancer site on all quantities, but the model description should be read as referring to a specific cancer site. To incorporate all three time dimensions, we have taken into account the relative survival for every 5-year age group from 0 up to 85+ years of age, for time since cancer diagnosis (survival time) from 1- up to 15-year survival, and for calendar year (cohort) from 1981 to 1995. Because the SEER data do not provide survival beyond the 10th year, we calculated RIS from the 11th to the 15th year of survival by means of a linear regression model, using survival data from year 1 to 10, as follows:

$$Y_\tau = \kappa + \beta * \tau$$

where

Y_τ is the estimated RIS for time τ since diagnosis (in years),

κ and β are the regression coefficients, and

τ = time since diagnosis (in years)

After obtaining the time-specific survival data, we have then further indexed all the age, time, and calendar year survival information to the first year interval survival for each sex, and cancer site. The first year of survival was chosen because, for most if not all cancer sites, it is the most critical year concerning cancer survival experience. After the first year of survival, the relative survival curve usually increases and then flattens smoothly. Indexing was done by dividing each of the time-specific RIS by the survival at 1-year interval. The age-time-dimension was estimated for each age by assuming that the same RIS of the 5-year age group applied for each single age year.

We then obtained RIS' – our model estimated relative interval survival – from the following basic multiplicative three-dimensional time survival model (age-, time-, and calendar year-specific RIS), by calculating:

$$RIS'_{\alpha,t,\tau} = 1 - (1 - RIS_1) * A_\alpha * T_t * Y_\tau$$

where

$RIS'_{\alpha,t,\tau}$	is the estimated relative interval survival for age α, calendar year t across the interval τ–1 to τ where τ is time since diagnosis in years
$RIS_1 = 1 - RIS_{.,1973-95,1}$	is the relative probability of death after 1 year for all ages, averaged across the calendar years 1975 to 1995
$A_\alpha = \frac{1 - RIS_{\alpha,1973-95,1}}{1 - RIS_1}$	is the ratio of the relative probability of death after 1 year at age α to the relative probability of death after 1 year for all ages, averaged across the calendar years 1975 to 1995

$T_t = \frac{1 - RIS_{.,t,1}}{1 - RIS_1}$ is the ratio of the relative probability of death after 1 year for all ages in calendar year t to the relative probability of death after 1 year for all ages, averaged across the calendar years 1975 to 1995

$Y_\tau = \frac{1 - RIS_{.,1973-95,\tau}}{1 - RIS_1}$ is the ratio of the relative probability of death after τ years for all ages, averaged across the calendar years 1975 to 1995, to the relative probability of death after 1 year for all ages, averaged across the calendar years 1975 to 1995

Calculations were performed for 18 age groups (α = 1 to 18), from 0-4 to 85+ years of age; for 23 calendar years (t = 1 to 23), from 1973 to 1995; and for 15 years of survival (τ = 1 to 15).

3.3 Cancer death distribution.

The modelled cancer death distributions were calculated from SEER's age-specific incidence data from 1981 to 1995, and from the described modelled $RIS'_{\alpha,t,\tau}$. We assumed that incidence was constant for every single year of age within its corresponding 5-year age group. Based on each cohort age- and year- survival experience, from 1981 up to 1995, we calculated $RIS''_{\alpha,\tau} = RIS'_{\alpha,1995,\tau}$ for t = 1995, the 15th year of survival. The double quotes are used to indicate calendar year 1995 in the following equations to simplify notation.

To obtain the number of deaths and, from them, our final outcome of interest – cancer death distribution, we needed to estimate the number of individuals who survived up to 1995 by age and time of survival as well as their corresponding probability of death during this year.

The number of surviving individuals at age α in 1995 was calculated by multiplying incidence at age α in year 1995- τ by $OIS''_{\alpha,\tau}$, the observed interval survival for τ years since diagnosis for individuals aged α in 1995, and summing over τ. We first estimated the relative cumulative survival ($RCS''_{\alpha,\tau}$) for every single age (α = 0 to 89) and year of survival (τ = 1 to 15) for 1995 to enable us to estimate $OIS''_{\alpha,\tau}$. $RCS''_{\alpha,\tau}$ was calculated by multiplying $RIS''_{\alpha,\tau}$ over the years of survival. Next, by using a standard life table, and age- and time-specific $RCS''_{\alpha,\tau}$, we estimated $OIS''_{\alpha,\tau}$ for 1995 by single age and time of survival:

$$OIS''_{\alpha,\tau} = RCS''_{\alpha,\tau} * (l_{\alpha+1} / l_{\alpha-\tau+1})$$

where l_x is the number of individuals surviving at exact age x in the life table.

For ease of calculation in a spreadsheet, and to facilitate calculation of the probability of dying, this equation can be rewritten:

$$OIS''_{\alpha,\tau} = exp\left(-\left(-ln\left(RCS''_{\alpha,\tau}\right) + \sum_{x=\alpha-\tau+1}^{\alpha} h_x \right) \right)$$

where

$h_x = ln(l_{x+1} / l_x)$

α is single year of age (0 to 89), and

τ *is* time since diagnosis (1 to 15)

The number of individuals $S''_{\alpha,\tau}$ who had survived up to 1995 was then estimated, for every year of age α and time of survival τ, by multiplying incidence and observed interval survival for the corresponding year of age and survival time:

$$S''_{\alpha,\tau} = Inc_{\alpha-\tau,1995-\tau} * OIS''_{\alpha,\tau}$$

where

$Inc_{\alpha,t}$ is the incidence at age α in calendar year t

For example, the number of individuals who were 7 years of age ($\alpha = 7$) in 1995, and who had survived cancer for 4 years ($\tau = 4$) in 1995 was calculated by multiplying the incidence of cancer for the cohort of individuals who were 3 years of age ($\alpha-\tau = 3$) in 1991 ($=1995-\tau$) (year of diagnosis) by the $OIS''_{\alpha,\tau}$ calculated for a 7 year old person who had survived 4 years since cancer diagnosis.

The probability of dying in 1995, due to cancer hazard, for each single age, and year of survival was calculated as follows:

$$PD''_{\alpha,\tau} = \left(1 - exp\left(-\left(-ln\left(RIS''_{\alpha,\tau}\right) + h_\alpha\right)\right)\right) * \left(-ln\left(RIS''_{\alpha,\tau}\right) / \left(-ln\left(RIS''_{\alpha,\tau}\right) + h_\alpha\right)\right)$$

In order to obtain the number of cancer deaths estimated to occur in 1995 among those individuals aged α years, and who had survived cancer for τ years, we multiplied the number of survivors $S''_{\alpha,\tau}$ by the relevant probability of dying in 1995 due to cancer hazard $PD''_{\alpha,\tau}$:

$$D''_{\alpha,\tau} = S''_{\alpha,\tau} * PD''_{\alpha,\tau}$$

and then, to obtain total cancer deaths in 1995 at age α years, we summed over all survival times τ:

$$D''_{\alpha} = \sum_{\tau=0}^{Min(15,\alpha)} S''_{\alpha,\tau} * PD''_{\alpha,\tau}$$

3.4 Model Validation.

In order to check the performance of the model, we have graphically compared our estimated $RIS'_{\alpha,t,\tau}$ for $\tau = 1$ to 10 years individuals diagnosed with cancer in 1986 with the SEER $RIS_{\alpha,t,\tau}$ for $\tau = 1$ to 10 years for the same cohort of individuals. We show the results obtained for males and females 55-59 years old, and for every cancer site in Figure 2. From these figures, we can observe that the model predicts very well the relative interval survivals.

For those cancer sites with greater number of cases, such as colon, lung, breast, corpus uteri, and prostate cancer, the model fits very well. For those with smaller numbers, the estimated RIS' smoothes the curves for the observed RIS, also showing a very good fit.

Figure 2: Comparison between predicted and observed relative interval survival for 55-59 year olds with year of diagnosis, 15 cancer sites, by sex, 1986.

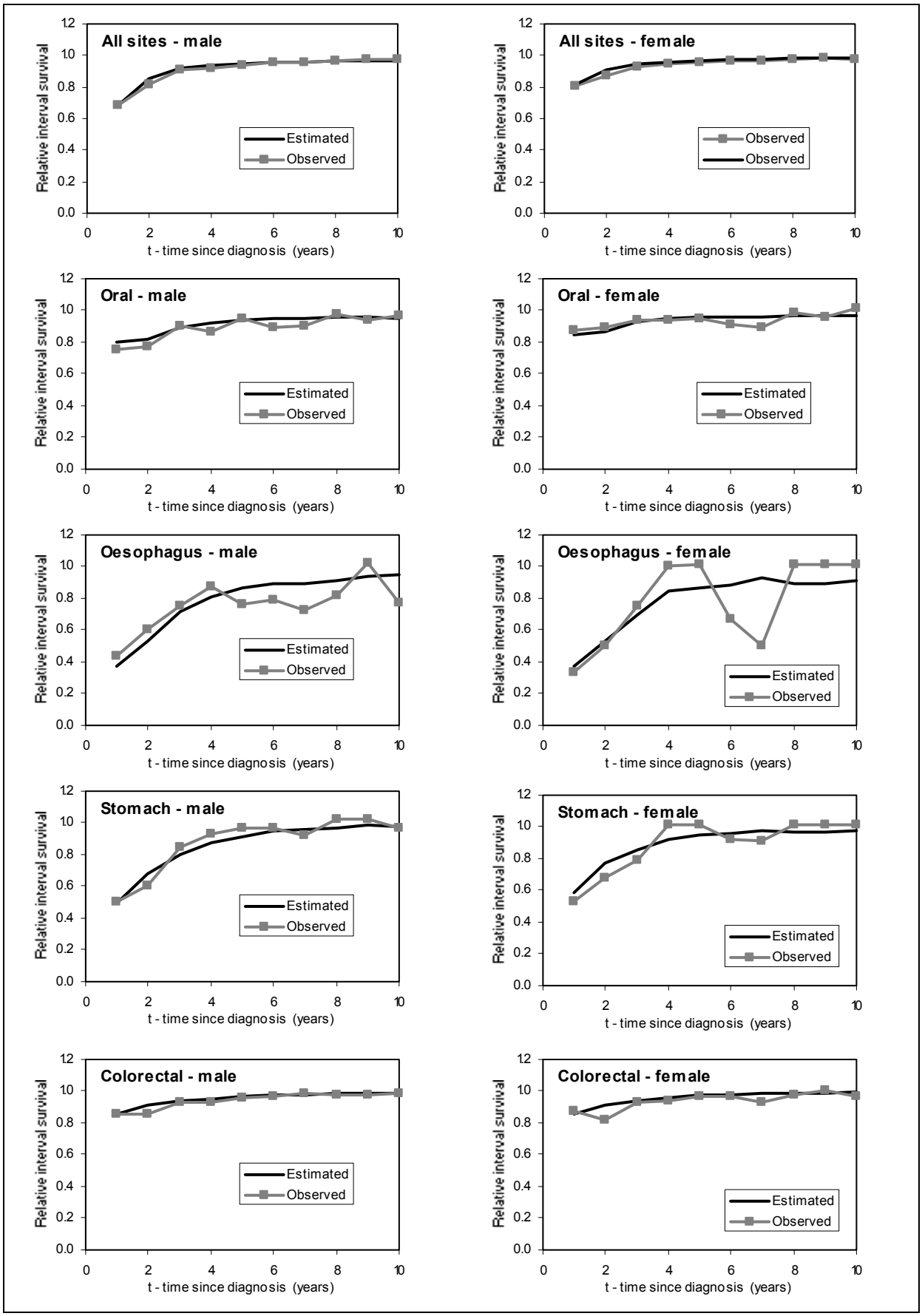

Figure 2 (continued): Comparison between predicted and observed relative interval survival for 55-59 year olds with year of diagnosis, 15 cancer sites, by sex, 1986.

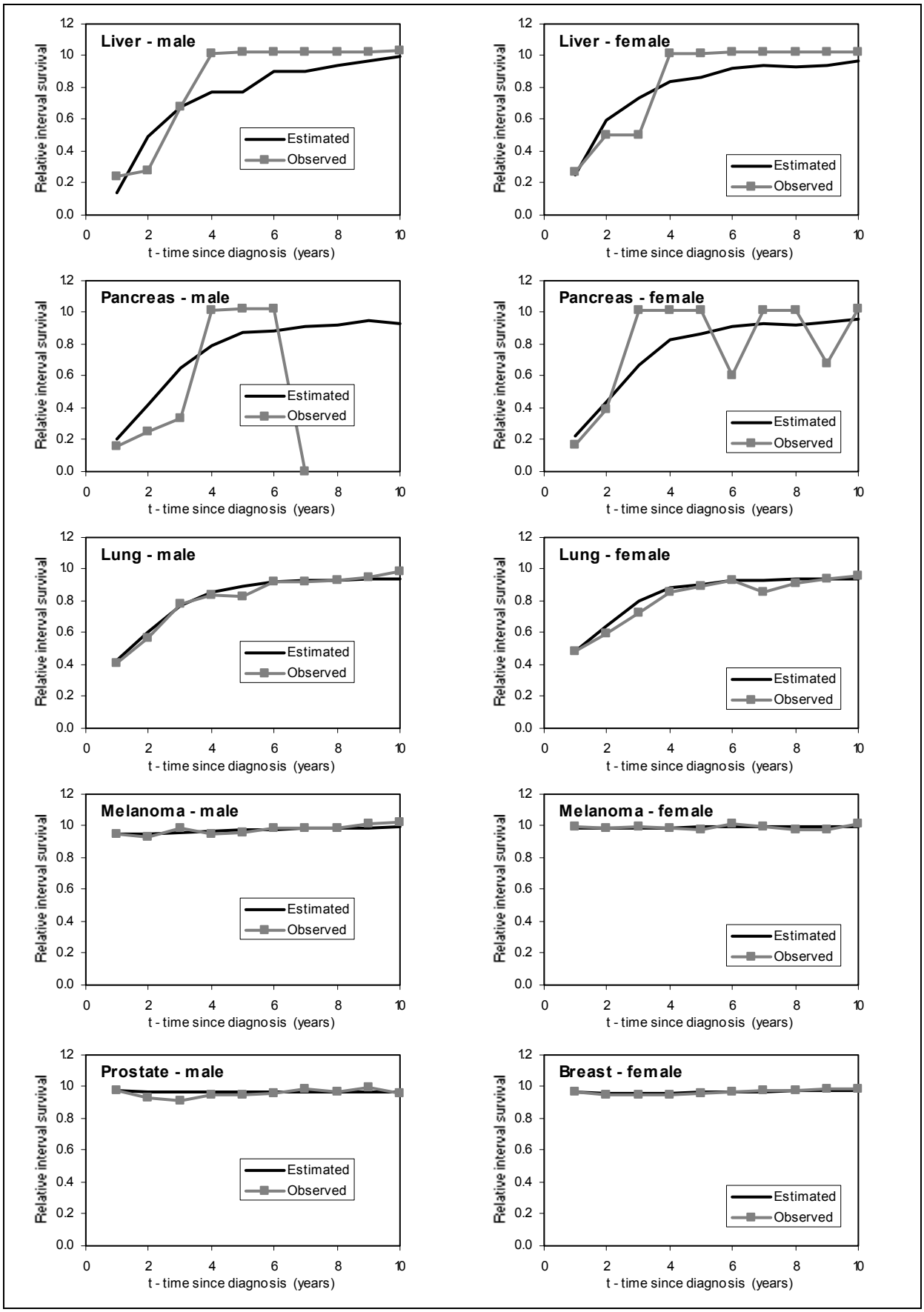

Figure 2 (continued): Comparison between predicted and observed relative interval survival for 55-59 year olds with year of diagnosis, 15 cancer sites, by sex, 1986.

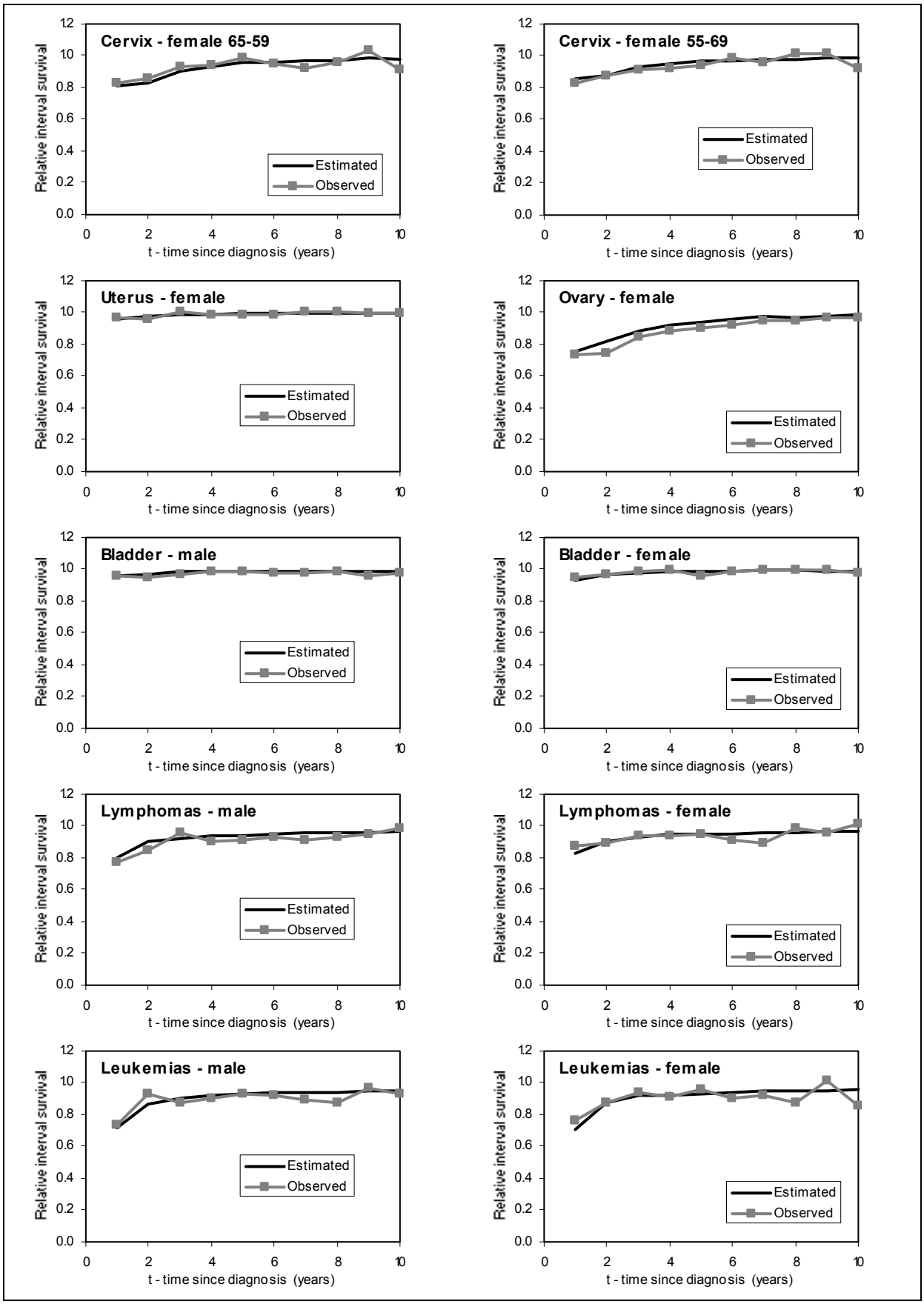

Table 5: Cancer death ratios SEER / US vital statistics by site, age groups, and sex. 1990-1995.

	All cancers		Oral		Oesophagus		Stomach		Colorectal	
Age	Male	Female	Male	Female	Male	Female	Male	Female	Male	Female
0-4	1.41	1.48	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
5-9	1.24	1.45	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
10-14	1.48	1.57	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
15-19	1.16	1.32	1.00	1.00	1.00	1.00	1.00	1.00	1.21	0.28
20-24	1.08	1.26	1.61	1.00	1.00	1.00	1.27	0.12	0.79	0.77
25-29	2.34	1.03	4.72	0.70	1.00	0.72	0.85	1.41	1.37	1.05
30-34	1.68	1.07	4.72	1.36	0.50	1.00	5.21	1.19	1.42	0.76
35-39	1.82	0.90	4.02	6.21	2.04	0.70	1.83	1.16	1.40	0.92
40-44	1.53	0.99	3.65	1.17	0.67	2.63	1.74	0.88	0.96	0.86
45-49	1.46	1.06	3.16	1.89	1.56	1.06	1.41	0.86	1.23	1.25
50-54	1.44	1.13	1.84	1.80	1.44	1.41	1.54	1.09	0.97	1.08
55-59	1.39	1.16	2.17	1.66	1.08	1.32	1.28	1.24	0.98	1.06
60-64	1.28	1.11	2.32	1.92	0.83	2.00	1.28	1.19	0.96	1.00
65-69	1.27	1.12	2.54	2.36	1.00	1.13	1.39	1.68	0.99	1.11
70-74	1.17	1.14	2.58	2.60	0.77	1.47	1.28	1.05	1.00	1.18
75-79	1.10	1.15	2.31	2.05	0.90	0.94	1.48	1.15	1.27	1.23
80-84	0.97	1.14	2.30	1.82	0.67	0.93	1.04	1.13	1.27	1.22
85+	0.75	1.01	1.50	1.65	0.70	0.74	0.92	0.99	1.14	1.10

	Liver		Pancreas		Lung		Bladder		Lymphoma	
Age	Male	Female	Male	Female	Male	Female	Male	Female	Male	Female
0-4	1.00	0.67	1.00	1.00	1.00	1.00	1.00	1.00	0.51	1.00
5-9	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	3.52	1.00
10-14	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	2.99	2.85
15-19	0.08	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.24	0.78
20-24	0.91	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.51	1.92
25-29	0.37	1.55	1.00	1.57	1.00	0.46	1.00	1.00	2.34	2.28
30-34	4.55	0.83	1.70	1.25	1.41	1.82	1.00	1.00	1.67	2.19
35-39	1.16	1.26	1.15	0.73	1.74	1.21	0.29	1.67	2.92	1.18
40-44	1.49	1.50	0.97	1.19	1.34	1.15	1.68	3.20	2.18	1.29
45-49	1.11	0.90	0.90	1.00	1.10	1.06	1.21	2.23	2.20	1.31
50-54	1.22	0.99	1.30	1.15	1.18	1.12	2.67	1.58	2.05	1.16
55-59	1.28	0.99	1.19	1.29	1.18	1.19	1.17	1.08	1.30	1.01
60-64	1.07	0.74	0.95	1.03	1.01	1.15	1.60	1.81	0.97	1.12
65-69	1.13	0.94	0.96	1.21	1.15	1.13	1.20	1.52	0.94	0.92
70-74	0.89	0.95	1.03	1.08	1.09	1.07	1.43	1.84	0.79	1.08
75-79	0.88	0.78	0.99	0.99	1.08	1.13	1.33	1.48	0.88	1.06
80-84	0.59	0.74	0.86	0.93	1.01	1.03	1.68	1.57	0.82	1.09
85+	0.90	0.67	0.86	0.96	0.84	0.97	1.26	1.06	0.64	1.03

	Leukemia		Melanoma		Breast	Cervix	Uterus	Ovary	Prostate
Age	Male	Female	Male	Female	Female	Female	Female	Female	Male
0-4	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
5-9	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
10-14	1.34	1.05	1.00	1.00	1.00	1.00	1.00	1.00	1.00
15-19	2.07	1.59	1.16	1.48	1.00	0.28	1.00	2.50	1.00
20-24	1.21	0.79	0.80	1.19	0.35	1.96	1.00	2.05	1.00
25-29	1.76	1.03	5.94	1.71	0.72	0.60	1.00	5.25	1.00
30-34	1.42	0.76	0.75	2.17	0.67	1.05	1.00	3.49	1.00
35-39	1.85	2.47	1.58	1.88	0.61	1.15	1.85	1.37	1.00
40-44	1.20	1.69	1.24	1.54	0.63	1.05	2.56	1.74	0.27
45-49	2.22	1.67	1.58	1.46	0.75	1.28	1.09	1.05	0.77
50-54	1.25	1.77	1.43	1.71	1.11	1.52	0.82	1.04	0.86
55-59	1.37	1.29	1.16	1.15	1.13	1.40	1.32	0.87	0.66
60-64	1.22	0.87	1.59	1.44	1.13	1.15	1.03	0.95	0.62
65-69	1.10	1.01	1.42	1.78	1.15	1.43	0.94	0.89	0.60
70-74	1.13	0.73	1.18	1.72	1.12	1.42	1.23	0.95	0.55
75-79	1.02	0.83	1.28	1.53	1.18	1.42	1.36	0.95	0.91
80-84	0.92	0.75	1.69	1.09	1.26	1.11	1.28	0.90	0.99
85+	0.87	0.61	1.18	1.43	1.11	1.26	1.17	0.98	0.62

To exemplify this, let us take the case of liver cancer. There were 49 cases of liver cancer in males, at the start of follow-up. Among them, 37 individuals died during the first year of follow-up. After that, the numbers became very small in every interval. The observed relative survival increased from the second year on and went beyond one from the forth year of survival on, period during which the only two individuals who had survived the forth year, remained alive. A similar phenomena is seen among females, for whom there were only 15 cases at the start of follow-up, and among those individuals with pancreatic cancer. Of the 83 males diagnosed with pancreatic cancer in 1986, 70 died during the first year of follow-up; all individuals had died by the end of the seventh year. In such cases, our model has smoothed the survival curves.

3.5 Application to the US vital statistics data.

We have compared the estimated age-, sex-, and site-specific cancer deaths to those reported by the US vital statistics for the same areas covered by the SEER program (see Appendix 1). In order to do so, we calculated the ratios between our estimates and the observed deaths reported by the US vital statistics by sex and age-group. The data corresponded to deaths between 1990 and 1995. The ideal situation would be to obtain ratios close to 1, in which case, deaths estimated by the model would be similar to those reported by the US vital statistics. These ratios are presented in Table 5. Ratios vary considerably for young ages (up to 25 years old) because there were few or no deaths at these ages for most cancer sites for both SEER-based estimates and the US vital statistics (exceptions were all cancers, lymphomas, and leukemia).

We observe that, among those 45 years of age and older - age groups for which cancer incidence and mortality start to increase and are more stable, the ratios were closer to one (bounds 0.75 and 1.33), for all cancers (1.01 to 1.16), lymphomas (0.92 to 1.31), and cancers of the breast (0.75 to 1.26), and of ovary (0.87 to 1.05) among females. In males and females, such bounds held for cancers of colon and rectum (0.96 to 1.27; 1.00 to 1.25, respectively), pancreas (0.86 to 1.30; 0.93 to 1.29, respectively), and lung (0.84 to 1.18; 0.97 to 1.19, respectively).

Ratios did not go beyond 0.50 or 2.00, a somewhat wider range, for all cancers (0.75 to 1.46) and prostate cancer (0.55 to 0.99) among males, and for leukemias (0.61 to 1.77), cervical (1.11 to 1.52) and uterine (0.82 to 1.36) cancers for females. For males and females, those were the bounds for cancer of oesophagus (0.67 to 1.56; 0.74 to 2.00, respectively), stomach (0.92 to 1.54; 0.86 to 1.68, respectively), liver (0.59 to 1.28; 0.67 to 0.99, respectively), and melanoma of skin (1.16 to 1.69; 1.09 to 1.78, respectively). Poor consistency (very wide bounds) was observed for oral and bladder cancer among males and females, and for lymphomas and leukemias among males.

In the GBD 1990, deaths coded to ICD-9 195–199, (malignant neoplasm of other and unspecified sites including those whose point of origin cannot be determined, secondary and unspecified neoplasm) were redistributed pro-rata across all malignant neoplasm categories within each age–sex group, so that the category ‘Other malignant neoplasms’ includes only malignant neoplasms of other specified sites. The comparison of the predicted deaths from the survival model with those reported in US Vital Statistics was used to identify four sites where there did not appear to be any significant coding of cancer deaths to the ‘garbage codes’ ICD-9 195–199 (see Table 6). So the cancer garbage code redistribution algorithm was revised for the GBD 2000 to redistribute cancer garbage code deaths pro-rata across only the included sites listed on the left side of Table 6.

Table 6. Sites included in the redistribution of deaths coded to cancer garbage codes, GBD 2000

Included	Excluded
Mouth and oropharynx cancers	Liver cancer
Oesophagus cancer	Pancreas cancer
Stomach cancer	Trachea, bronchus and lung cancers
Colon and rectum cancers	Ovary cancer
Melanoma and other skin cancers	
Breast cancer	
Cervix uteri cancer	
Corpus uteri cancer	
Ovary cancer	
Prostate cancer	
Bladder cancer	
Lymphomas and multiple myeloma	
Leukaemia	
Other malignant neoplasms (excluding garbage codes)*	

* ICD-9 195-199

4. Estimation of cancer mortality by site and region

We have applied the multiplicative survival model to 7 regions/subregions for which the mortality data were either scarce or non existent at level of specific cancer sites: AFRO (D and E), EMRO (B and D), SEARO (B and D), AMRO (B and D), and Wpro B (see Murray et al (ref) for definitions of the subregions). For doing so, we needed estimates of the period survival factor T_r by site for each of the regions *r*, and estimated incidence distributions by site for each of these regions/subregions.

4.1 Survival data for developing regions

To estimate survival for developing regions, where little or no data is available, based on the SEER survival patterns by site, age and sex, we need to estimate the "equivalent" calendar year survival term T_r for each region/subregion. T_r is the ratio of the relative probability of death after 1 year for all ages in the relevant region to the relative probability of death after 1 year for all ages in the SEER data, averaged across the calendar years 1975 to 1995. In this way, we obtain a new calendar year survival term for the model.

Equivalent period survival terms were estimated for each region by examining the relationship between period survival terms and gross domestic product per capita (measured in purchasing power parity dollars or international dollars) using the following data

(1) SEER survival data for the USA for the years 1973 to 1995 *(22)*

(2) Connecticut survival data for the years 1950 and 1958 *(23)*

(3) Survival data for the late 1980s from cancer registries in 5 developing countries (see Table 2) *(18)*,

(4) Survival data for four Eastern European countries (Poland, Estonia, Slovenia, Slovakia) for the late 1980s *(24)*.

Calendar year survival terms (T_t) for each cancer site were calculated as described in Section 3 for those years of the series for which SEER survival data were available. For the other data sources, available survival data were also used to estimate T_t as follows.

Survivorship functions were estimated from the relative survival data by fitting a Weibull survival distribution function to the all-ages data. To allow for a proportion who are cured and never die from the cancer, we modify the usual Weibull model as follows:

$$S(t) = \alpha + (1-\alpha)\exp\left(-(\lambda t)^{\gamma}\right)$$

where α is the proportion who never die from the cancer, λ is the location parameter ($1/\lambda$ is the time at which 50% of those will die have died) and γ is the shape parameter. We use the 10 year relative survival S_{10} as an estimate of the proportion who never die from the cancer. This is an approximation to avoid the need for iterative solution of an equation which cannot be solved analytically. Empirical test suggest that this does not introduce significant error in the mean survival time estimates, but in future revision of these estimates, numerical methods for obtaining exact solutions will be further explored.

For survival data sets where S_{10} is not available, we estimate it from S_5 using the latest SEER data from the USA on the ratio of 10 to 5 year survival by site, age and sex as follows:

$$S_{10} = S_5 \times \left(\frac{S_{10}}{S_5}\right)_{SEER}$$

We use 1, 3 and 5 year relative survival rates to fit the Weibull distribution as follows:

$$\sigma_1 = \frac{S_1 - S_{10}}{1 - S_{10}}$$

$$\sigma_3 = \frac{S_3 - S_{10}}{1 - S_{10}}$$

$$\gamma = ln\left(\frac{ln\,\sigma_3}{ln\,\sigma_1}\right) \Big/ ln\,3$$

$$\lambda = \left[-ln\,\sigma_1\right]^{1/\gamma}$$

To check the goodness of fit of the resulting survival curve, we computed S_5 using these parameters, and compared with the observed S_5. Good fits were obtained in all cases.

The T factors for all the available survival data were plotted against GDP per capita (international dollars) for each site and sex as shown in Figure 3, and trend lines fitted. In each plot, the data points above $17,000 per capita are the SEER survival factors, the two points between $10,000 and $15,000 per capita are the factors for the 1950 and 1958 Connecticut data, and the other points below $10,000 per capita are for the Eastern European and developing country data.

Based on the trend lines for each site and sex, and the estimated GDP per capita in international dollars for each region in 1997, T factors were estimated for each site and sex for each GBD 2000 region. The results are shown in Table 7. An example is shown for breast cancer in Table 7: knowing that GDP per capita in AFRO D was $1,536 in 1997, this corresponded to an indexed calendar year-specific T_t = 3.231. This was then the value used in the age-period-cohort survival model for breast cancer in the AFRO D region. A similar process was applied to the other regions, and for other cancer sites.

The main advantage of this approach to estimating regional survival distributions by cancer site for developing regions is that it correctly estimates survival and smooths it in regions where good data are provided, and it ensures that regional survival estimates are consistent

with trends in survival across all regions, where the numbers for some cancer sites are small and, consequently, 'noisy' for that region.

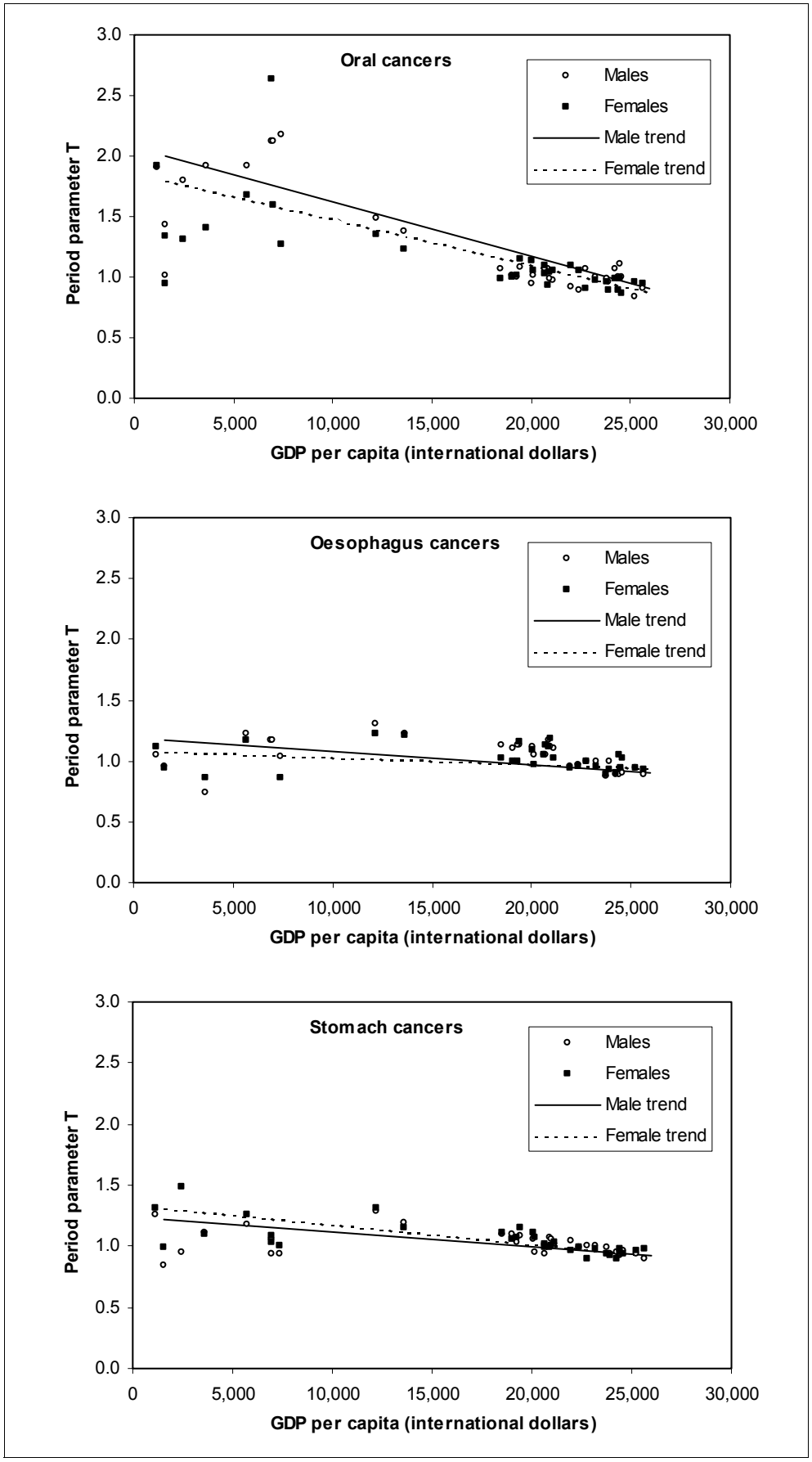

Figure 3. Survival T factor versus GDP per capita, USA and developing countries

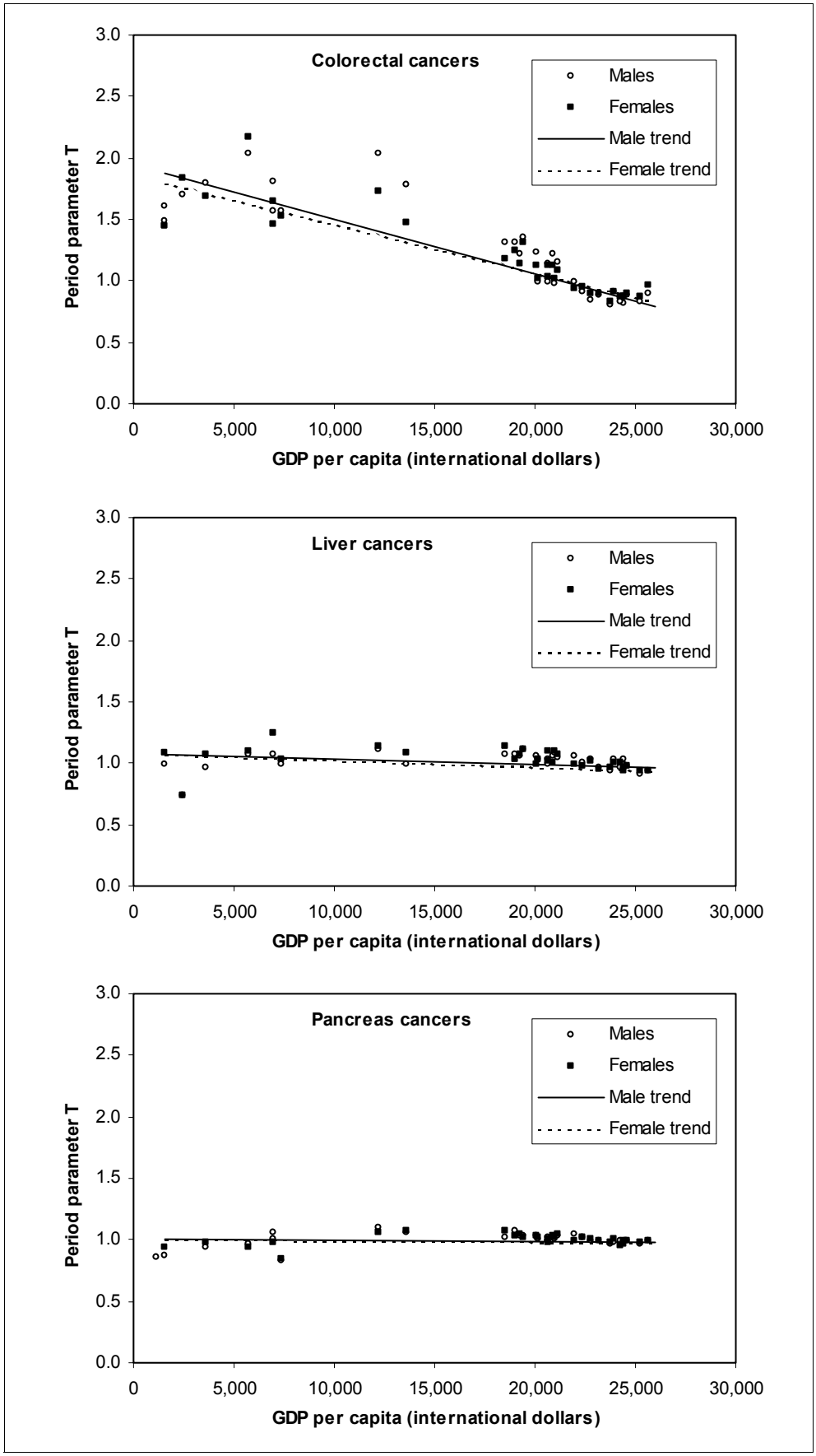

Figure 3 (continued). Survival T factor versus GDP per capita, USA and developing countries

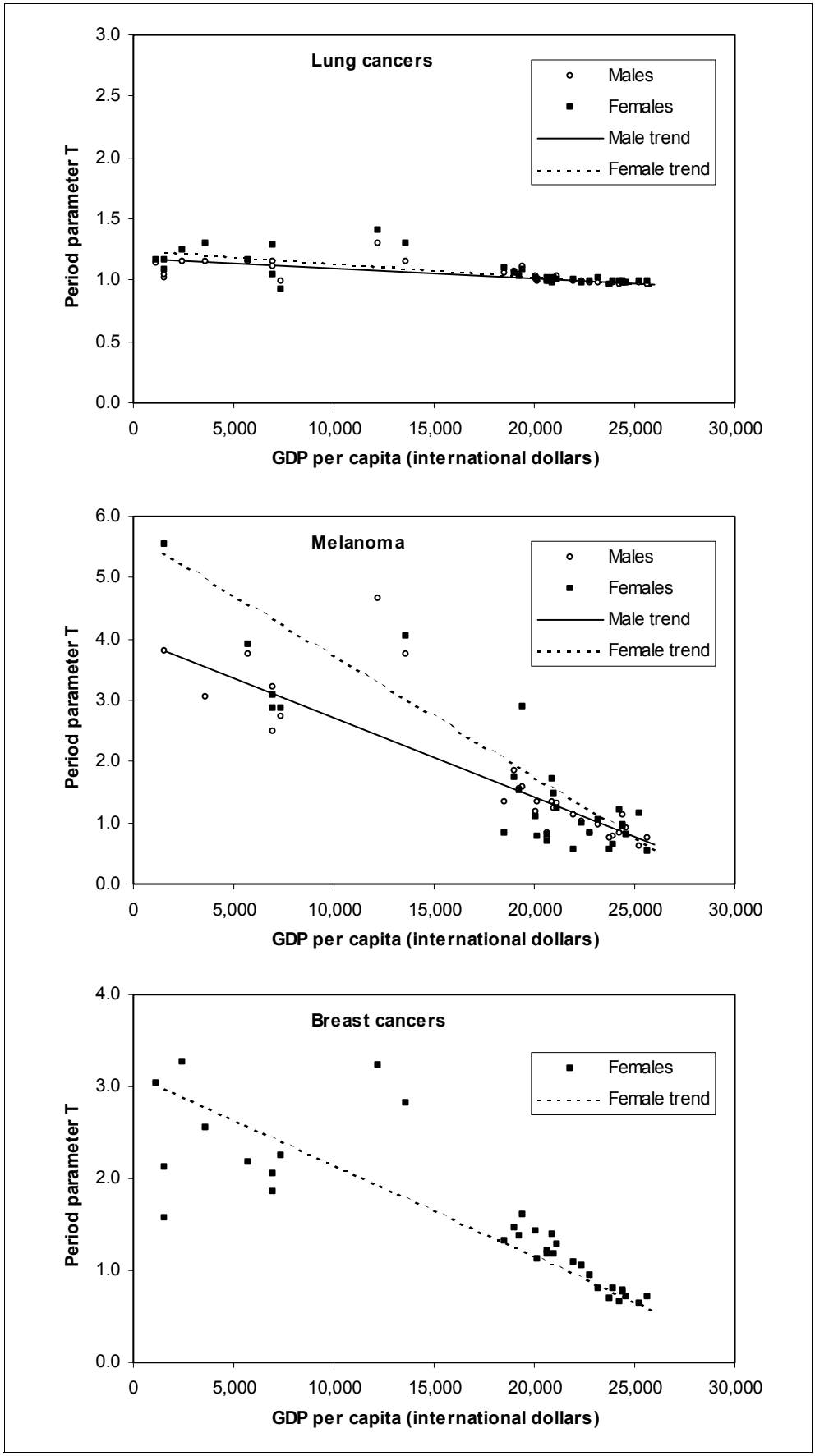

Figure 3 (continued). Survival T factor versus GDP per capita, USA and developing countries

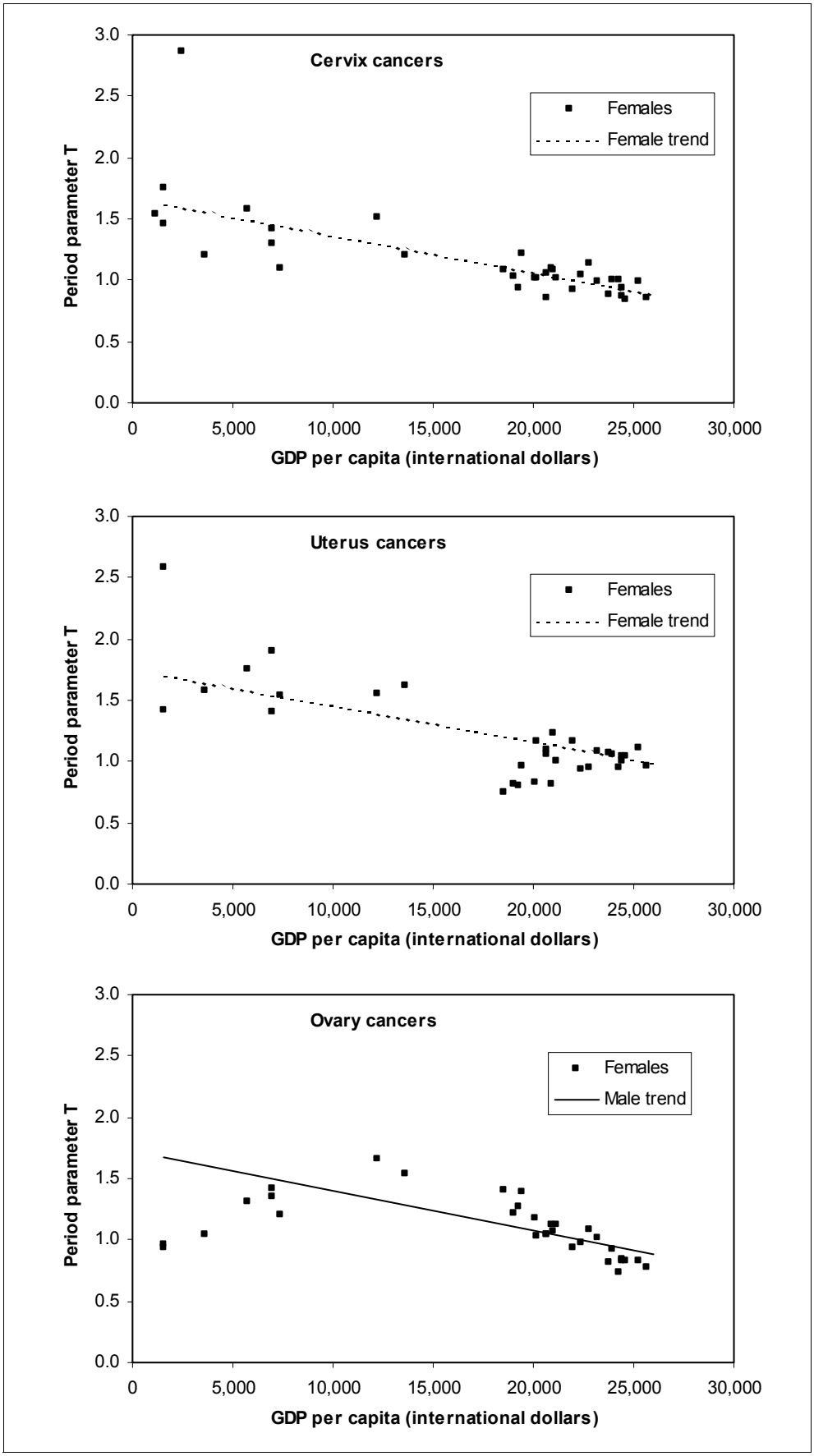

Figure 3 (continued). Survival T factor versus GDP per capita, USA and developing countries

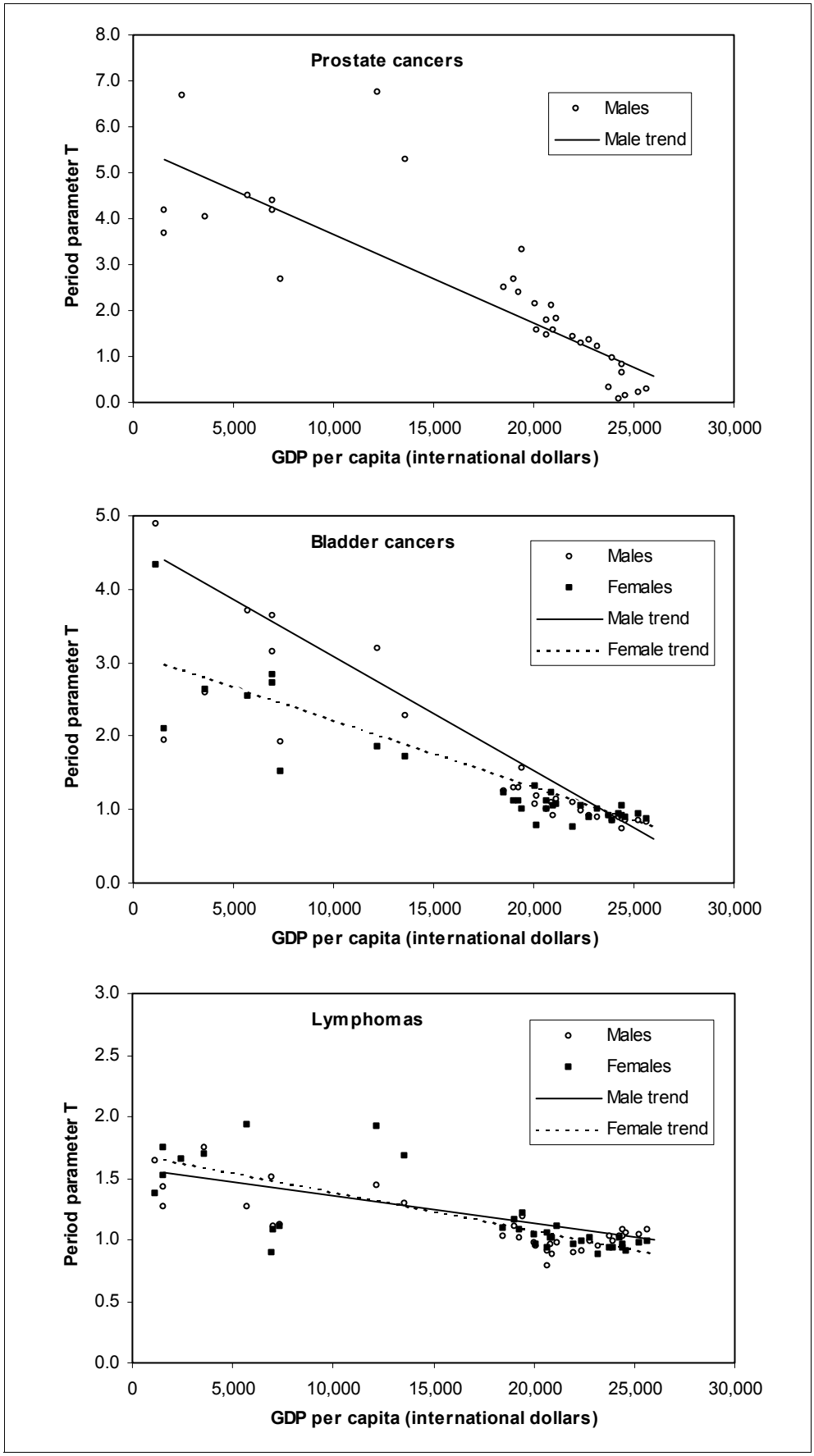

Figure 3 (continued). Survival T factor versus GDP per capita, USA and developing countries

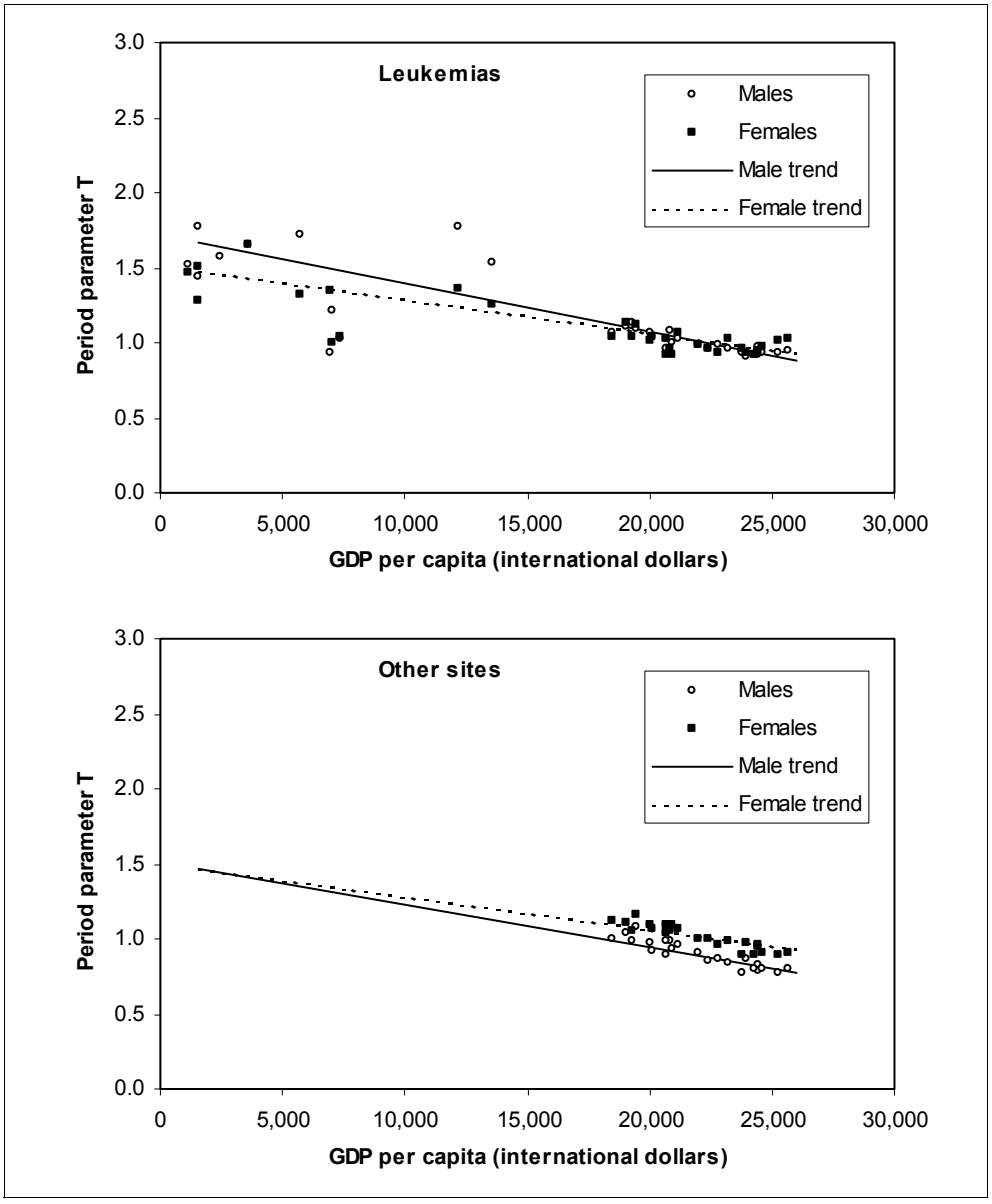

Figure 3 (continued). Survival T factor versus GDP per capita, USA and developing countries

As can be seen in Figure 3, cancer registry survival estimates for some sites in some developing countries are better than recent US experience, or significantly below the trend line with GDP per capita, suggesting that survival may have been overestimated due to small numbers or incomplete case followup. In these cases, the survival model provides survival estimates more consistent with the complete body of evidence.

4.2 Incidence data

In order to apply the survival model to estimate the distribution of cancer deaths in each developing region, we need to have estimates of the cancer incidence distribution by site for each region. Cancer Incidence in Five Continents *(25)* is one of the periodic publications from the International Agency for Research on Cancer (IARC). It consists of a series of monographs, which are published every five years, with incidence data from registries all over the world. Data are collected, coded and analyzed in a standard way and are only

Table 7: Estimated regional survival parameters T_R for age-period-cohort survival model.

Region*	Oral		Oesophagus		Stomach		Colorectal		Liver	
R	Male	Female	Male	Female	Male	Female	Male	Female	Male	Female
AFRO D	2.000	1.800	1.165	1.075	1.225	1.313	1.875	1.800	1.075	1.138
AFRO E	1.980	1.780	1.160	1.075	1.220	1.300	1.850	1.780	1.075	1.130
AMRO A	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
AMRO B	1.800	1.656	1.145	1.075	1.163	1.238	1.653	1.626	1.075	1.101
AMRO D	1.858	1.720	1.160	1.075	1.205	1.277	1.806	1.736	1.075	1.128
EMRO B	1.800	1.670	1.150	1.075	1.175	1.250	1.700	1.650	1.075	1.110
EMRO D	1.980	1.780	1.165	1.075	1.220	1.290	1.845	1.775	1.075	1.130
EURO A	1.000	1.000	1.003	0.870	0.900	0.920	1.290	1.290	1.000	1.000
EURO B1	2.000	2.000	1.038	1.000	1.000	1.009	1.568	1.533	1.000	1.034
EURO B2	1.800	1.635	1.157	1.054	1.084	1.136	1.633	1.553	1.074	1.173
EURO C	1.800	1.600	1.164	1.033	1.000	1.033	1.568	1.463	1.074	1.240
SEARO B	1.800	1.700	1.160	1.075	1.200	1.275	1.800	1.700	1.075	1.130
SEARO D	1.980	1.780	1.165	1.075	1.225	1.315	1.845	1.775	1.075	1.130
Japan	1.000	1.000	0.860	0.760	0.570	0.560	1.040	1.110	0.830	0.940
Australia/NZ	0.950	0.950	0.950	0.950	0.860	0.910	0.950	0.900	1.000	1.000
WPRO B1	1.800	1.700	1.160	1.075	1.210	1.280	1.835	1.775	1.075	1.135
WPRO B2	2.000	1.900	1.170	1.075	1.225	1.320	1.900	1.825	1.075	1.138
WPRO C	1.900	1.820	1.150	1.075	1.215	1.285	1.850	1.800	1.075	1.125

	Pancreas		Lung		Bladder		Lymphoma		Leukemia	
Age	Male	Female	Male	Female	Male	Female	Male	Female	Male	Female
AFRO D	1.000	1.000	1.170	1.230	4.400	3.000	1.545	1.650	1.670	1.485
AFRO E	1.000	1.000	1.168	1.220	4.310	2.900	1.540	1.635	1.650	1.470
AMRO A	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
AMRO B	1.000	1.000	1.122	1.175	3.559	2.475	1.396	1.501	1.480	1.354
AMRO D	1.000	1.000	1.152	1.203	4.083	2.666	1.488	1.588	1.610	1.424
EMRO B	1.000	1.000	1.130	1.180	3.700	2.475	1.420	1.525	1.520	1.375
EMRO D	1.000	1.000	1.165	1.220	4.300	2.750	1.530	1.600	1.650	1.440
EURO A	1.000	1.000	1.200	1.200	1.600	1.500	1.000	1.000	1.100	1.100
EURO B1	1.000	1.000	1.000	1.000	1.910	1.525	1.115	1.105	1.032	1.035
EURO B2	1.000	1.000	1.139	1.113	3.664	2.654	1.460	1.235	1.293	1.364
EURO C	1.002	1.000	1.148	1.049	3.629	2.846	1.501	1.000	1.100	1.352
SEARO B	1.000	1.000	1.146	1.190	4.000	2.475	1.470	1.575	1.605	1.420
SEARO D	1.000	1.000	1.165	1.220	4.300	2.950	1.550	1.600	1.650	1.440
Japan	0.950	0.910	1.000	1.100	1.090	1.060	1.430	1.280	1.520	1.360
Australia/NZ	0.750	0.760	1.000	1.000	1.400	1.280	0.750	0.840	0.930	0.890
WPRO B1	1.000	1.000	1.155	1.210	4.150	2.850	1.500	1.625	1.625	1.440
WPRO B2		1.000	1.175	1.225	4.500	3.150	1.570	1.650	1.680	1.475
WPRO C		1.000	1.155	1.210	4.300	2.850	1.520	1.600	1.630	1.450

	Melanoma		Miscellaneous		Breast	Cervix	Uterus	Ovary	Prostate
Age	Male	Female	Male	Female	Female	Female	Female	Female	Male
AFRO D	3.800	5.400	1.470	1.470	2.975	1.625	1.700	1.330	5.300
AFRO E	3.700	5.175	1.450	1.450	2.950	1.600	1.675	1.330	5.200
AMRO A	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
AMRO B	2.988	3.694	1.336	1.400	2.419	1.493	1.551	1.330	4.311
AMRO D	3.457	4.762	1.423	1.428	2.785	1.587	1.638	1.330	4.935
EMRO B	3.120	4.000	1.360	1.400	2.525	1.519	1.575	1.330	4.500
EMRO D	3.650	5.175	1.470	1.470	2.900	1.600	1.650	1.330	5.150
EURO A	1.700	1.600	1.100	1.100	1.400	1.200	1.500	1.200	2.700
EURO B1	2.733	2.867	1.400	1.400	2.249	1.103	1.547	1.200	2.664
EURO B2	2.795	3.516	1.385	1.375	2.273	1.466	1.729	1.371	4.453
EURO C	2.504	3.091	1.410	1.350	2.046	1.414	1.899	1.414	4.406
SEARO B	3.400	4.650	1.410	1.400	2.750	1.575	1.625	1.330	4.900
SEARO D	3.650	5.175	1.470	1.470	2.900	1.600	1.650	1.330	5.150
Japan	1.000	1.000	1.300	1.300	0.800	1.000	2.100	1.110	3.100
Australia/NZ	0.500	0.500	0.950	0.950	0.770	0.760	1.120	1.000	0.930
WPRO B1	3.500	4.875	1.420	1.420	2.850	1.625	1.675	1.330	5.000
WPRO B2	3.850	5.400	1.470	1.450	3.000	1.650	1.725	1.330	5.350
WPRO C	3.600	5.150	1.450	1.440	2.900	1.600	1.650	1.330	5.150

included after an assessment of quality is performed. It has become the reference source of data on international cancer incidence. Volume VII – the most recent of these monographs which were published, includes data from 50 different countries for the years 1988-1992. These data have been used to project cancer incidence rates to the year 2000 for the Globocan 2000 database *(6)*.

In order to estimate regional incidence patterns for those regions where the model was to be applied, we carefully examined the methods used to estimate country-specific incidence data in Globocan 2000, to ensure that for all the regions where we required incidence estimates, the Globocan estimates were based on cancer registry incidence data, and not modelled from mortality data using assumptions about survival (which would then result in circularity in our mortality estimation process for regions without good mortality data by cancer site). Incidence estimates used in Globocan 2000 were based on data from national or local cancer registries from 69 countries as listed in Table 8.

Table 8. Cancer registry data used to estimate incidence patterns by region*, GBD 2000

Region	Cancer registry data used
AFRO D	Algeria (Oran), Burkina Faso (Ouagadougou), Guinea (Conakry), Mali (Bamako), Gambia, Niger (Niamey)
AFRO E	Malawi (Blantyre), Rwanda, South Africa, Uganda (Kyadondo County), Zimbabwe (Harare – black population), Rep. of the Congo (Brazzaville), Swaziland, Côte d'Ivoire (Abidjan)
AMRO A	Canada (national, Ontario), USA (SEER)
AMRO B	Argentina (Concordia), Brazil (Belém, Goiânia, Porto Alegre), Colombia (Cali), Costa Rica, Cuba, Jamaica, Paraguay (Ascunsion), Uruguay, France (Guadeloupe, La Martinique), Puerto Rico
AMRO D	Bolivia (La Paz), Peru (Lima, Trujillo), Ecuador (Quito)
EMRO B	Kuwait, Jordan, Iran, Oman, Tunisia (North and Suisse), Israel (non-Jews), Saudi Arabia
EMRO D	Egypt, Iraq
EURO A	Austria, Croatia, Czech Rep., Denmark, Finland, France, Germany, Iceland, Ireland, Israel, Italy, Malta, Netherlands, Norway, Slovenia, Spain, Sweden, Switzerland, UK
EURO B	Turkey (Izmir), Poland, Slovakia, Yugoslavia (Vojvodina)
EURO C	Belarus, Estonia, Latvia, Lithuania
SEARO B	Thailand (Chiang Mai, Khon Kaen, Songkhla), Singapore, Philippines (Manila, Rizal)
SEARO D	Singapore (Indian), India (Bangalore, Bombay, Karunagappally, Madras, Trivandrum), Pakistan (South Karachi)
WPRO A	Australia, New Zealand
WPRO B1	China (Qidong, Shanghai, Tianjin, Hong Kong), Mongolia, Korea (Kangwha County)
WPRO B2	Viet Nam (Hanoi)
WPRO B3	Fiji, Guam, Samoa, French Polynesia

* Globocan 2000 *(6)* estimates of incidence are based, in part, on incidence data from the cancer registries listed. Published papers describe cancer registry data for some of these sources *(26-45)*

Globocan 2000 estimates of cancer incidence by site for countries differ from those required for the GBD 2000 in three major respects:

- Globocan 2000 estimates include Karposi's sarcoma and non-Hodgkin's lymphomas (NHL) caused by HIV/AIDS. The GBD 2000 includes these cases among AIDS sequela and their burden is included with the HIV/AIDS burden *(46-49)*.
- Globocan 2000 estimates exclude incidence of skin cancers other than melanoma.
- Globocan 2000 estimates include cancers of unknown primary with cancers of other specified sites. The GBD 2000 attributes these cancers back to specific sites as described in Section 3.5.

Globocan 2000 incidence estimates by age, sex, site and country were adjusted for these differences as follows. Firstly, unpublished data on the incidence of Karposi's sarcoma for countries in Africa were provided by IARC and used to adjust incidence of 'other cancers' to

remove Karposi's sarcoma. Secondly, the fraction of NHL incidence attributable to HIV/AIDS was estimated for the Globocan regions of Africa based on the fraction for Zimbabwe *(27;50;51)* and using the ratio of Karposi's sarcoma for the region to that for Zimbabwe. The resulting attributable fractions shown in Table 9 were used to adjust the incidence estimates for NHL.

Table 9. Non-Hodgkin's lymphoma: estimated attributable fractions for HIV/AIDS by Globocan region*

	Male					Female				
Country/Region	0-14	15-44	45-54	55-64	65+	0-14	15-44	45-54	55-64	65+
Uganda	0.75	0.33	0.33	0.30	0.30	0.87	0.43	0.43	0.30	0.30
Zimbabwe	0.00	0.50	0.50	0.30	0.30	0.00	0.50	0.50	0.30	0.30
Eastern Africa	0.58	0.38	0.38	0.30	0.30	0.79	0.32	0.32	0.30	0.30
Middle Africa	0.40	0.20	0.20	0.10	0.10	0.50	0.25	0.25	0.10	0.10
Northern Africa	0.00	0.01	0.01	0.01	0.01	0.00	0.01	0.01	0.01	0.01
Southern Africa	0.04	0.02	0.02	0.02	0.02	0.03	0.02	0.02	0.02	0.02
Western Africa	0.06	0.03	0.03	0.03	0.03	0.04	0.02	0.02	0.02	0.02
Other regions	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

*Globocan 2000 regions are defined in Ferlay et al. *(6)*.

Finally, incidence estimates for cancers of unknown primary site were redistributed among specific sites (including the 'Other sites' category) using the GBD 2000 algorithm described in Section 3.5. The proportion of the Globocan 'Other sites' category corresponding to unknown primary sites was estimated from published data on the distribution of cancer incidence by site which included unknown primary as a specific category *(26-29;33;36-40;44;45)*

After adjusting the Globocan incidence estimates for each country as described above, these estimates were summed for the countries in each GBD 2000 region, resulting in estimated incidence distributions by site, age and sex for each region.

Table 10. Estimated proportion of Globocan 'Other sites' incidence attributable to unknown primary site

GBD Region	Unknown primary as proportion of Other sites
AFRO D, E	0.4
AMRO A	0.3
AMRO B, D	0.45
EMRO B, D	0.3
EURO A, B, C	0.3
SEARO B, D	0.5
WPRO A	0.3
WPRO B	0.5

4.3 Estimation of mortality distributions

The site-specific distributions of cancer mortality were estimated directly from vital registration data for countries in the A regions (Amro A, Euro A and Wpro A) and for countries in Euro B and Euro C. Vital registration data for Amro B did not include codes to identify pancreas and ovary cancer. For these two cancers in Amro B, and for all sites in the other regions of the world, we used the estimated incidence distribution by site for each region (described above) in the survival model to calculate the mortality distribution by site for the year 2000. This distribution was then used to disagreggate the estimated total cancer deaths by age and sex for each region, estimated as described in Section 2.

To apply the survival model for a region, we replaced the SEER incidence data in the survival model by the total incidence estimates for each region for the year 2000. We assumed incidence rates to be constant over the years and we then estimated the region-specific number of new cases for 1985 to 1999, by applying these age-specific incidence rates to the annual population.

The GBD 2000 uses the latest population estimates for WHO Member States prepared by the UN Population Division *(52)*. In order to obtain incidence from 1985 to 2000, we estimated the age-specific population by sex for each of these years, using growth rates also from the United Nation's data. Based on these region-specific estimated incidence and survival levels, cancer deaths were finally calculated by means of the multiplicative survival model for each region by age group and sex. The final results were then used to estimate the distribution, but not the magnitude, of cancer by site, sex, and age-group for 1999.

4.4 Results

Table 11 shows the resulting estimate cancer deaths by site and WHO region for Version 1 of the GBD 2000 *(12)*. Table 12 compares the mortality fractions for each site (site-specific cancer deaths as a fraction of total cancer deaths) for the GBD 2000 and Globocan 2000 by region. The ratio of mortality fractions is bolded in the table if it is 0.7 or less or 1.3 or more. There is good agreement between the GBD 2000 and Globocan estimates for most sites in most regions, with the exception of melanoma and other skin cancers and cancer of the uterus, where GBD 2000 estimated mortality fractions are about 40% higher. In both these cases, there is a difference in the definition of the site category. The GBD 2000 category includes deaths from other skin cancers in the category 'melanoma and other skin cancers' whereas Globocan 2000 excludes incidence from other skin cancers and includes their mortality in the 'Other sites' category'. For the regions with good vital registration data, the GBD 2000 category includes all skin cancer mortality. For the other regions, where the Globocan incidence data have been used in the survival model, the resulting distributions will underestimate total skin cancer mortality as they will miss fatal non-melanoma skin cancers.

The GBD 2000 category for cancer of the uterus includes 'cancer of the uterus, part unknown', whereas Globocan 2000 includes on corpus uteri cancers.

Figure 4 compares GBD 2000 global mortality estimates for specific cancer sites with those from Globocan 2000.

More detailed estimates of cancer mortality by age, sex, site and region are available at the WHO website http://www.who.int/evidence/bod (select GBD 2000 Version 1 results).

Table 11. GBD 2000 Version 1 estimated cancer deaths[a] by site and WHO region, and comparison with GLOBOCAN 2000 total cancer deaths by region

	GBD 2000 estimated cancer deaths ('000)							%total cancer
Site	**AFRO**	**AMRO**	**EMRO**	**EURO**	**SEARO**	**WPRO**	**World**	**deaths**
Mouth and oropharynx cancers	33	23	22	52	169	40	340	4.9
Oesophagus cancer	26	30	13	55	71	217	413	6.0
Stomach cancer	36	71	18	186	64	370	744	10.7
Colon and rectum cancers	25	105	12	237	55	144	579	8.4
Liver cancer	63	32	11	67	52	400	626	9.0
Pancreas cancer	8	48	3	93	16	46	214	3.1
Trachea, bronchus and lung cancers	23	232	31	373	153	401	1,213	17.5
Melanoma and other skin cancers[b]	9	18	2	28	3	5	65	0.9
Breast cancer	38	87	16	155	104	59	459	6.6
Cervix uteri cancer	59	29	19	29	116	35	288	4.2
Corpus uteri cancer[c]	3	20	1	35	4	13	76	1.1
Ovary cancer	10	23	4	48	24	20	128	1.9
Prostate cancer	44	74	6	94	21	19	258	3.7
Bladder cancer	14	23	11	65	21	23	157	2.3
Lymphomas and multiple myeloma	38	65	17	76	54	41	291	4.2
Leukaemia	20	48	16	60	50	71	265	3.8
Other sites	83	145	39	231	125	192	814	11.8
Total GBD 2000	**533**	**1,074**	**242**	**1,882**	**1,103**	**2,096**	**6,930**	**100.0**
Total GLOBOCAN 2000	**278**	**1,089**	**253**	**1,811**	**831**	**1,954**	**6,216**	
% difference (GBD – GLOBOCAN)	***92***	***-1***	***-4***	***4***	***33***	***7***	***11***	

a Globocan estimates have been adjusted to exclude Karposi's sarcoma deaths and the proportion of NHL due to HIV/AIDS and to redistribute a proportion of 'Other and unknown' sites to known sites using same algorithm as for GBD mortality estimates (see Section 4.2).

Table 12. Ratio of GBD deaths as % of total cancer deaths to Globocan deaths as % of total cancer deaths[a],by site and WHO region

	Ratio of GBD mortality fraction to Globocan mortality fraction						
Site	**AFRO**	**AMRO**	**EMRO**	**EURO**	**SEARO**	**WPRO**	**World**
Mouth and oropharynx cancers	**1.3**	1.0	1.0	1.0	1.0	0.8	1.1
Oesophagus cancer	1.0	1.0	**0.7**	1.0	0.8	1.0	1.0
Stomach cancer	0.8	0.9	**1.3**	0.9	0.9	0.8	0.9
Colon and rectum cancers	1.0	1.0	1.0	1.0	1.0	0.8	0.9
Liver cancer	0.9	1.1	1.0	1.1	0.9	1.1	1.0
Pancreas cancer	0.9	0.9	0.9	1.1	0.9	0.9	1.0
Trachea, bronchus and lung cancers	1.1	1.0	**1.4**	1.0	**1.3**	1.0	1.0
Melanoma[b]	1.2	**1.4**	**1.5**	**1.4**	1.1	1.1	1.4
Breast cancer	**0.9**	1.0	0.6	1.0	1.2	1.0	1.0
Cervix uteri cancer	0.9	**0.7**	**2.1**	0.8	0.9	1.0	1.0
Corpus uteri cancer[c]	0.9	**1.5**	**0.6**	**1.5**	**0.7**	**2.1**	**1.4**
Ovary cancer	0.9	1.0	0.8	1.1	0.9	1.0	1.0
Prostate cancer	**1.4**	1.0	1.2	1.0	1.0	1.0	1.0
Bladder cancer	1.1	1.0	0.7	1.1	0.8	1.0	1.0
Lymphomas and multiple myeloma	0.9	1.0	0.9	1.0	1.0	1.1	1.0
Leukaemia	1.0	1.0	1.1	0.9	1.1	1.0	1.0
Other sites	1.0	1.1	1.0	1.0	1.0	**1.7**	1.1

a Globocan estimates have been adjusted to exclude Karposi's sarcoma deaths and the proportion of NHL due to HIV/AIDS and to redistribute a proportion of 'Other and unknown' sites to known sites using same algorithm as for GBD mortality estimates (see Section 4.2).

b GBD estimates include deaths due to other skin cancers

c GBD estimates include cancer of uterus, part unknown

5. Estimation of cancer incidence by site and region

The cancer survival model was also used to calculate incidence to mortality ratios by age and sex for each cancer site in all regions of the world for the year 2000. These incidence to mortality ratios were then applied to the mortality estimates in order to estimate cancer incidence by age and sex for each site and region.

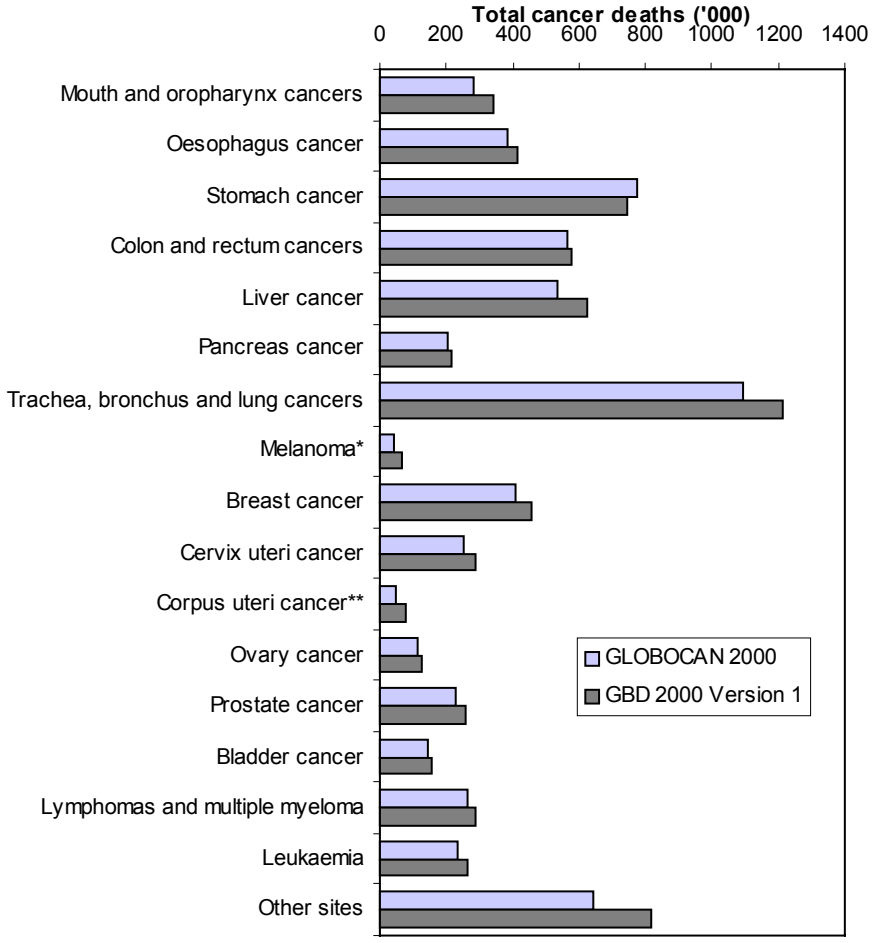

Figure 4. Estimated total cancer deaths by site, GBD 2000 and Globocan 2000.

For the countries in the Wpro A and Euro A regions, country-specific survival data were used to calculate country-specific T factors for use in the survival model. Survival data for European countries *(24)* and for Australia *(36;37)* and Japan *(35)* were used in this analysis.

As shown in Table 13, the resulting GBD 2000 global incidence estimate for all sites is almost identical to that from Globocan 2000. However, there are some differences across regions, with the GBD 2000 estimates being higher for AFRO and SEARO, reflecting the higher mortality estimates, and somewhat lower for AMRO and EMRO. The estimates for these two

regions are being reviewed in more detail as part of the revision of the cancer burden estimates for Version 2 of the GBD 2000.

Table 13. GBD 2000 Version 1 estimated cancer incidence[a] by site and WHO region, and comparison with GLOBOCAN 2000 total cancer incidence by region

	GBD 2000 estimated cancer incidence ('000)							
Site	AFRO	AMRO	EMRO	EURO	SEARO	WPRO	World	%total cancer incidence
Mouth and oropharynx cancers	41	33	30	94	221	65	485	4.7
Oesophagus cancer	27	32	14	58	75	229	434	4.2
Stomach cancer	38	79	20	220	73	470	900	8.7
Colon and rectum cancers	35	160	19	377	84	242	917	8.9
Liver cancer	63	33	12	70	55	418	651	6.3
Pancreas cancer	8	49	3	97	17	48	222	2.2
Trachea, bronchus and lung cancers	22	255	31	399	163	431	1,302	12.7
Melanoma and other skin cancers[b]	13	69	3	83	5	23	197	1.9
Breast cancer	55	234	29	367	179	138	1,002	9.7
Cervix uteri cancer	104	64	38	57	267	74	604	5.9
Corpus uteri cancer[c]	10	105	8	129	18	56	327	3.2
Ovary cancer	16	39	8	70	47	37	217	2.1
Prostate cancer	57	177	9	192	29	40	505	4.9
Bladder cancer	20	58	19	168	30	47	341	3.3
Lymphomas and multiple myeloma	54	104	28	119	83	63	452	4.4
Leukaemia	25	70	22	79	72	99	367	3.6
Other sites	129	264	71	380	204	316	1,364	13.3
Total GBD 2000	**719**	**1,825**	**363**	**2,959**	**1,624**	**2,795**	**10,286**	**100.0**
Total GLOBOCAN 2000	**504**	**2,208**	**401**	**2,958**	**1,259**	**2,704**	**10,032**	**100.0**
% difference (GBD – GLOBOCAN)	***43***	***-17***	***-9***	***0***	***29***	***3***	***3***	

a Globocan estimates have been adjusted to exclude Karposi's sarcoma incidence and the proportion of NHL due to HIV/AIDS and to redistribute a proportion of 'Other and unknown' sites to known sites using same algorithm as for GBD mortality estimates (see Section 4.2).

6. Discussion and conclusions

One important advantage of this multiplicative survival model is that it takes into account time in its three dimensions: time since cancer (survival), age, and calendar year (cohort). One other advantage, due to the availability of data, was the possibility of comparing the outcomes of the model to the data reported by the US vital statistics. This has given us the opportunity to evaluate our model and the data available. However, its main advantage is to correctly estimate survival and smooth it in regions where good data are provided, and where the numbers for some cancer sites are small and, consequently, 'noisy'.

The main limitations for applying this model were the relative lack of region-specific survival data and very few, and probably not always representative, regional cancer incidence data for some developing regions. We assumed that cancer incidence reported by a few countries of one region/subregion would represent the incidence of the whole area, which may not always be the case.

The model uses the available published population-based survival data from developing countries such as those published by Sankaranarayanan et al *(18)*. However, although population-based estimates, they may not to be representative of the whole countries they should represent. Such estimates are restricted to some geographic areas, and based on cancer

registries' data; consequently, also related to better health care and surveillance. Furthermore, several developing regions of the world were not included in these estimates, and the need to produce other estimates would persist.

Because of the poor quality and sparseness of survival data for the developing regions of the world, we decided to use all the available data, including lengthy time series data from the USA, to establish trends in survival with GDP per capita and then to use latest estimates of GDP per capita for developing regions, in order to estimate survival by site. This approach takes into account, through increases in average GDP per capita for regions, the likely improvements in survival over the periods since those for which developing country survival data are available.

As can be seen from Figure 3, the modelled survival estimates for developing regions are, for some sites, in some cases higher and in some cases lower than those presented by Sankaranarayanan et al *(18)*. There is a wide range of variation within countries in the survival estimates presented by Sankaranarayanan et al *(18)*. Important differences are shown for cancers of the oesophagus in Thailand, for those of bladder and for leukaemia in both Thailand and China, and for cancers of the breast, and of cervix uterine in all three countries compared – Thailand, China, and India. Despite all possible differences in the handling of the disease, and in host- or tumour-related factors, at least part of the variation observed is likely to be due to differences in the mechanisms of data collection. Death registration in developing countries is often scarce, incomplete or non-existent. It is possible, then, that even if the cancer cases were thoroughly followed up, the use of generic life tables could have underestimated the background mortality. Consequently, the expected survival would be overestimated, giving rise to a lower relative survival estimate in such cases. However, poor data quality tends to increase the estimates of observed survival, in general.

As shown in Figure 2, there was good consistency between our estimated number of deaths and those reported by the US vital statistics for some important cancer sites in the US, such as colon and rectum, lung, and breast. Consistency was reasonable for a number of other cancer sites (prostate, cervix uterine, ovary, esophagus, stomach, liver, and melanoma of the skin), and poor for others (oral, bladder, lymphomas, and leukaemias). There are a number of possible explanations for such inconsistency. The first one is that the model may not have been adequate to estimate these cancer sites, due to one of its three time dimensions-related behavior. Different age patterns is one of them. Lymphomas and leukaemias, for example, have a different age pattern from all other cancer sites, and it is possible that the model has not been able to adequately capture and estimate such bimodal age pattern.

The other time dimensions that affect the model are time since cancer and calendar year. A recent change in incidence or survival could not have been captured by the model. However, as neither incidence nor survival changes are usually abrupt, it is unlikely that this has happened, unless it was due to some artefact related to increasing screening/incidence and to its consequent lead time (survival) bias. This could have been the case of prostate cancer, for example. Although due in part to increased detection (introduction of prostate-specific antigen – PSA), prostate cancer incidence has risen steeply from 1986 to 1991 in the US *(14)*, especially in younger age groups. We have assumed that incidence was constant over the last few years. This fact, together with a plausible but, at least in part, artifactual increased survival, could have influenced our results and be a partial explanation for the differences. This could then explain the ratios smaller than 1 that were observed, especially in the young age groups.

A second possibility is that SEER incidence/survival data may be over or under estimating the number and/or the duration of some cancers. Third, as mortality data can be affected by both the accuracy of cancer diagnosis and its certification as the underlying cause of death on death

certificates, it is possible that the US vital statistics could also be either over- or under-reporting certain site-specific cancer deaths. For example, as shown by SEER data, between 1990 and 1995, only about 44% of the women registered as having incident cervical cancer by SEER, and who died in this period, had cervical cancer assigned as their underlying cause of death. This could explain in part the ranges for the ratios SEER / US vital statistic for cervical cancer being between 1.11 and 1.52.

In summary, the survival model presented offers a new approach to the calculation of the number and distribution of deaths for areas where mortality data are either scarce or unavailable. It can also be applied in areas with good quality data, but where there are small numbers of some site-specific cancers. In our future work, we will attempt to gather incidence information from more individual countries, as well as further information on survival in order to improve our estimates, with more precise inputs for the model.

The analyses reported here have built extensively on IARC work to synthesise and estimate cancer incidence distributions by site for all regions of the world. They have also contributed to better harmonizing the GBD 2000 and IARC estimates of cancer incidence and mortality by site for most regions of the world. Where there are differences, the potential factors contributing to these differences are now better understood, in terms of the differences in the methods used. For Africa and South-East Asia (mainly India), where differences between the two sets of estimates are greatest, further work is underway to check and refine the GBD 2000 estimates. In addition, further work will be undertaken to check regional differences between the GBD 2000 and Globocan incidence estimates.

This discussion paper has summarised the analysis of cancer incidence and mortality for the Global Burden of Disease 2000 project. These have been used as a basis for the Version 1 estimates of cancer burden published in the World Health Report 2001. Over the next 12 months, work will continue on the revision of YLD and YLL estimates for cancer with a particular emphasis on improving incidence and mortality estimates for Africa and South East Asia, on improving estimates of average durations in the various cancer stages, and on improving the estimates of YLD associated with long term sequelae in cancer survivors.

Acknowledgements

Many people are contributing to the analysis of cancer incidence and mortality for the GBD 2000 both inside and outside WHO. We wish to particularly acknowledge the contributions of staff within EIP/GPE who have contributed to the estimation of total cancer deaths for each WHO Member State in the year 2000: Majid Ezzati, Brodie Ferguson, Mie Inoue, Rafael Lozano, Doris Ma Fat, Joshua Salomon and Lana Tomaskovic. We also thank staff of the International Agency for Research on Cancer (IARC) for provision of data, advice on survival analyses carried out by IARC and methods used to estimate cancer incidence and mortality for GLOBOCAN 2000, particularly Max Parkin, Jacques Ferlay, Paola Pisani and Fred Bray.

References

(1) World Health Organization. World Health Report 2001. Mental Health: New Understanding, New Hope. Geneva: WHO, 2001.

(2) Pisani P, Parkin DM, Bray F, Ferlay J. Estimates of the worldwide mortality from 25 cancers in 1990. International Journal of Cancer 1999; 83(1):18-29.

(3) Murray CJL, Lopez AD. The Global Burden of Disease: a comprehensive assessment of mortality and disability from diseases, injuries and risk factors in 1990 and projected to 2020. 1 ed. Cambridge: Harvard University Press, 1996.

(4) Parkin DM. The global burden of cancer. Seminars in Cancer Biology 1998; 8(4):219-235.

(5) Parkin DM, Pisani P, Ferlay J. Estimates of the worldwide incidence of 25 major cancers in 1990. International Journal of Cancer 1999; 80(6):827-841.

(6) Ferlay J, Bray F, Pisani P, Parkin DM. Globocan 2000: Cancer Incidence, Mortality and Prevalence Worldwide, Version 1.0. IARC CancerBase No. 5. 2001. Lyon, IARCPress.

(7) Doll D, Peto R. The causes of cancer. Quantitative estimates of avoidable risks of cancer in the United States today. Oxford: Oxford University Press, 1981.

(8) Parkin DM, Pisani P, Munoz N, Ferlay J. The global health burden of infection associated cancers. Cancer Surveys 1999; 33:5-33.

(9) Lundberg O. Methods of estimating morbidity and prevalence of disablement by use of mortality statistics. Acta Psychiatrica Scandinavica 1973; 49(3):324-331.

(10) Damiani P, Masse H, Aubenque M. Evaluation of morbidity from mortality. Biomedicine & Pharmacotherapy 1983; 37(3):105-106.

(11) Wingo PA, Landis S, Parker S, Bolden S, Heath Jr CW. Using cancer registry and vital statistcs data to estimate the number of new cancer cases and deaths in the United States for the upcoming year. J Reg Management 1998; 25:43-51.

(12) Murray CJL, Lopez AD, Mathers CD, Stein C. The Global Burden of Disease 2000 project: aims, methods and data sources. GPE Discussion Paper No. 36. 2001. Geneva, WHO.

(13) Mathers CD, Boschi-Pinto C. Global burden of cancer in the year 2000: Version 1 estimates. 2001. Geneva, World Health Organization. GBD 2000 Draft Methods Paper.

(14) Potosky AL, Miller BA, Albertsen PC, Kramer BS. The role of increasing detection in the rising incidence of prostate cancer. JAMA 1995; 273(7):548-552.

(15) Hsing AW, Tsao L, Devesa SS. International trends and patterns of prostate cancer incidence and mortality. International Journal of Cancer 2000; 85(1):60-67.

(16) Polednak AP. Interpretation of secular increases in incidence rates for primary brain cancer in Connecticut adults, 1965-1988. Neuroepidemiology 1996; 15(1):51-56.

(17) Lowry JK, Snyder JJ, Lowry PW. Brain tumors in the elderly: recent trends in a Minnesota cohort study. Archives of Neurology 1998; 55(7):922-928.

(18) Sankaranarayanan R, Black RJ, Parkin DM. Cancer survival in developing countries. IARC Scientific Publications No. 145. Lyon, France: International Agency for Research on Cancer, 1998.

(19) Lopez AD, Ahmad O, Guillot M, Inoue M, Ferguson B. Life tables for 191 countries for 2000: data, methods, results. GPE Discussion Paper No. 40. 2001. Geneva, WHO.

(20) Murray CJL, Lopez AD. The Global Burden of Disease: a comprehensive assessment of mortality and disability from diseases, injuries and risk factors in 1990 and projected to 2020. 1, 211. 1996. Cambridge, Harvard University Press. Global Burden of Disease and Injury Series.

(21) Salomon JA, Murray CJL. Compositional models for mortality by age, sex and cause. GPE Discussion Paper No. 11. 2001. Geneva, WHO.

(22) Ries LAG, Kosary CL, Hankey BF, Miller BA, Harras A, Edwards BK. SEER cancer statistics review, 1973-1994. NIH Pub. No. 97-2789. Bethesda, MD: National Cancer Institute, 1997.

(23) Eisenberg H, Sullivan PD, Connelly RR. Cancer in Connecticut. Survival experience. Hartford: Connecticut State Department of Health, 1968.

(24) Berrino F, Capocaccia R, Estève J, Gatta G, Hakulinen T, Micheli A et al. Survival of cancer patients in Europe: the EUROCARE-2 study. IARC Scientific Publications No. 151. Lyon, France: International Agency for Research on Cancer, 1999.

(25) Parkin DM, Whelan SL, Ferlay J, Raymond L, Young J. Cancer incidence in five continents. IARC Scientific Publications No. 143. Lyon, France: International Agency for Research on Cancer, 1997.

(26) Sitas F, Madhoo J, Wessie J. Cancer in South Africa, 1993-1995. Johannesburg: National Cancer Registry of South Africa, South African Institute for Medical Research, 1998.

(27) Chokunonga E, Levy LM, Bassett MT, Mauchaza BG, Thomas DB, Parkin DM. Cancer incidence in the African population of Harare, Zimbabwe: second results from the cancer registry 1993-1995. International Journal of Cancer 2000; 85(1):54-59.

(28) Wabinga HR, Parkin DM, Wabwire-Mangen F, Mugerwa JW. Cancer in Kampala, Uganda, in 1989-91: changes in incidence in the era of AIDS. International Journal of Cancer 1993; 54(1):26-36.

(29) Newton R, Ngilimana PJ, Grulich A, Beral V, Sindikubwabo B, Nganyira A et al. Cancer in Rwanda. International Journal of Cancer 1996; 66(1):75-81.

(30) Koulibaly M, Kabba IS, Cisse A, Diallo SB, Diallo MB, Keita N et al. Cancer incidence in Conakry, Guinea: first results from the Cancer Registry 1992-1995. International Journal of Cancer 1997; 70(1):39-45.

(31) Bayo S, Parkin DM, Koumare AK, Diallo AN, Ba T, Soumare S et al. Cancer in Mali, 1987-1988. International Journal of Cancer 1990; 45(4):679-684.

(32) Echimane AK, Ahnoux AA, Adoubi I, Hien S, M'Bra K, D'Horpock A et al. Cancer incidence in Abidjan, Ivory Coast: first results from the cancer registry, 1995-1997. Cancer 2000; 89(3):653-663.

(33) Bah E, Hall AJ, Inskip HM. The first 2 years of the Gambian National Cancer Registry. British Journal of Cancer 1990; 62(4):647-650.

(34) Abdallah MB, Zehani S. Registre des cancers Nord-Tunisie 1994. 2000. Tunisia, Ministere de la Sante Publique, Institut Salah Azaiz, Institut National de la Sante Publique.

(35) Oshima A, Tsukuma H, Ajiki W, Kitano M, Kitagawa T, Tanaka A et al. Survival rates of cancer patients in Osaka, 1975-89. Osaka, Department of Cancer Control and Statistics, Osaka Medical Center for Cancer and Cardiovascular Diseases.

(36) Australian Institute of Health and Welfare (AIHW), Australasian Association of Cancer Registries (AACR). Cancer survival in Australia, 2001. Part 1: National summary statistics. 2001. Canberra, AIHW.

(37) Australian Institute of Health and Welfare (AIHW), Australasian Association of Cancer Registries (AACR). Cancer survival in Australia, 2001. Part 2: Statistical tables. 2001. Canberra, AIHW.

(38) Martin AA, Galan YH, Rodriguez AJ, Graupera M, Lorenzo-Luaces P, Fernandez LM et al. The Cuban National Cancer Registry: 1986-1990. European Journal of Epidemiology 1998; 14(3):287-297.

(39) Brooks SE, Hanchard B, Wolff C, Samuels E, Allen J. Age-specific incidence of cancer in Kingston and St. Andrew, Jamaica, 1988-1992. West Indian Medical Journal 1995; 44(3):102-105.

(40) Adib SM, Mufarrij AA, Shamseddine AI, Kahwaji SG, Issa P, el Saghir NS. Cancer in Lebanon: an epidemiological review of the American University of Beirut Medical Center Tumor Registry (1983-1994). Annals of Epidemiology 1998; 8(1):46-51.

(41) Bhurgri Y, Bhurgri A, Hassan SH, Zaidi SH, Rahim A, Sankaranarayanan R et al. Cancer incidence in Karachi, Pakistan: first results from Karachi Cancer Registry. International Journal of Cancer 2000; 85(3):325-329.

(42) Jin F, Devesa SS, Chow WH, Zheng W, Ji BT, Fraumeni JF, Jr. et al. Cancer incidence trends in urban shanghai, 1972-1994: an update. International Journal of Cancer 1999; 83(4):435-440.

(43) Nguyen MQ, Nguyen CH, Parkin DM. Cancer incidence in Ho Chi Minh City, Viet Nam, 1995-1996. International Journal of Cancer 1998; 76(4):472-479.

(44) Paksoy N, Bouchardy C, Parkin DM. Cancer incidence in Western Samoa. International Journal of Epidemiology 1991; 20(3):634-641.

(45) Sitas F, Blaauw D, Terblanche M, Madhoo J, Carrara H. Cancer in South Africa, 1992. Johannesburg: National Cancer Registry of South Africa, South African Institute for Medical Research, 1997.

(46) Analo HI, Akanmu AS, Akinsete I, Njoku OS, Okany CC. Seroprevalence study of HTLV-1 and HIV infection in blood donors and patients with lymphoid malignancies in Lagos, Nigeria. Central African Journal of Medicine 1998; 44(5):130-134.

(47) Sitas F, Bezwoda WR, Levin V, Ruff P, Kew MC, Hale MJ et al. Association between human immunodeficiency virus type 1 infection and cancer in the black population of Johannesburg and Soweto, South Africa. British Journal of Cancer 1997; 75(11):1704-1707.

(48) Mueller N. Overview of the epidemiology of malignancy in immune deficiency. Journal of Acquired Immune Deficiency Syndromes 1999; 21 Suppl 1:S5-10.

(49) Smith C, Lilly S, Mann KP, Livingston E, Myers S, Lyerly HK et al. AIDS-related malignancies. Annals of Medicine 1998; 30(4):323-344.

(50) Chitsike I, Siziya S. Seroprevalence of human immunodeficiency virus type 1 infection in childhood malignancy in Zimbabwe. Central African Journal of Medicine 1998; 44(10):242-245.

(51) Chokunonga E, Levy LM, Bassett MT, Borok MZ, Mauchaza BG, Chirenje MZ et al. Aids and cancer in Africa: the evolving epidemic in Zimbabwe. AIDS 1999; 13(18):2583-2588.

(52) United Nations. World Population Prospects - The 1998 revision Volume III: Analytical Report. 2000. New York.

APPENDIX A: Areas covered by SEER program and compared to the US vital statistics

County Registries

San Francisco-Oakland SMSA.	Alameda County (001), Contra Costa County (013), Marin County (041), San Francisco County (075), San Mateo County (081).
Connecticut.	Fairfield County (001), Hartford County (003), Litchfield County (005), Middlesex County (007), New Haven County (009), New London County (011), Tolland County (013), Windham County (015), Unknown (999).
Detroit (Metropolitan).	Macomb County (099), Oakland County (125), Wayne County (163).
Hawaii.	Hawaii County (001), Honolulu County (003), Kalawao County (005), Kauai County (007), Maui County (009), Hawaii (900) - Populations Only, Unknown (999).
Iowa.	Adair County (001), Adams County (003), Allamakee County (005), Appanoose County (007), Audubon County (009), Benton County (011), Black Hawk County (013), Boone County (015), Bremer County (017), Buchanan County (019), Buena Vista County (021), Butler County (023), Calhoun County (025), Carroll County (027), Cass County (029), Cedar County (031), Cerro Gordo County (033), Cherokee County (035), Chickasaw County (037), Clarke County (039), Clay County (041), Clayton County (043), Clinton County (045), Crawford County (047), Dallas County (049), Davis County (051), Decatur County (053), Delaware County (055), Des Moines County (057), Dickinson County (059), Dubuque County (061), Emmet County (063), Fayette County (065), Floyd County (067), Franklin County (069), Fremont County (071), Greene County (073), Grundy County (075), Guthrie County (077), Hamilton County (079), Hancock County (081), Hardin County (083), Harrison County (085), Henry County (087), Howard County (089), Humboldt County (091), Ida County (093), Iowa County (095), Jackson County (097), Jasper County (099), Jefferson County (101), Johnson County (103), Jones County (105), Keokuk County (107), Kossuth County (109), Lee County (111), Linn County (113), Louisa County (115), Lucas County (117), Lyon County (119), Madison County (121), Mahaska County (123), Marion County (125), Marshall County (127), Mills County (129), Mitchell County (131), Monona County (133), Monroe County (135), Montgomery County (137), Muscatine County (139), O'Brien County (141), Osceola County (143),

	Page County (145), Palo Alto County (147), Plymouth County (149), Pocahontas County (151), Polk County (153), Pottawattamie County (155), Poweshiek County (157), Ringgold County (159), Sac County (161), Scott County (163), Shelby County (165), Sioux County (167), Story County (169), Tama County (171), Taylor County (173), Union County (175), Van Buren County (177), Wapello County (179), Warren County (181), Washington County (183), Wayne County (185), Webster County (187), Winnebago County (189), Winneshiek County (191), Woodbury County (193), Worth County (195), Wright County (197), Unknown (999).
New Mexico.	Bernalillo County (001), Catron County (003), Chaves County (005), Cibola County (006), Colfax County (007), Curry County (009), DeBaca County (011), Dona Ana County (013), Eddy County (015), Grant County (017), Guadalupe County (019), Harding County (021), Hidalgo County (023), Lea County (025), Lincoln County (027), Los Alamos County (028), Luna County (029), McKinley County (031), Mora County (033), Otero County (035), Quay County (037), Rio Arriba County (039), Roosevelt County (041), Sandoval County (043), San Juan County (045), San Miguel County (047), Santa Fe County (049), Sierra County (051), Socorro County (053), Taos County (055), Torrance County (057), Union County (059), Valencia County (061), Cibola + Valencia (910) - 1969-1981.
Seattle (Puget Sound).	Clallam County (009), Grays Harbor County (027), Island County (029), Jefferson County (031), King County (033), Kitsap County (035), Mason County (045), Pierce County (053), San Juan County (055), Skagit County (057), Snohomish County (061), Thurston County (067), Whatcom County (073).
Utah.	Beaver County (001), Box Elder County (003), Cache County (005), Carbon County (007), Daggett County (009), Davis County (011), Duchesne County (013), Emery County (015), Garfield County (017), Grand County (019), Iron County (021), Juab County (023), Kane County (025), Millard County (027), Morgan County (029), Piute County (031), Rich County (033), Salt Lake County (035), San Juan County (037), Sanpete County (039), Sevier County (041), Summit County (043), Tooele County (045), Uintah County (047), Utah County (049), Wasatch County (051), Washington County (053), Wayne County (055), Weber County (057).
Atlanta (Metropolitan).	Clayton County (063), Cobb County (067), DeKalb County (089), Fulton County (121), Gwinnett County (135).

APPENDIX B: Estimated parameters for the age-period-cohort survival model

Table B.1: Estimated age parameters A_α, age-period-cohort survival model based on SEER data 1981-1995.

	Oral		Oesophagus		Stomach		Colorectal		Liver	
Age α	Male	Female	Male	Female	Male	Female	Male	Female	Male	Female
0-4	0.870	1.093	0.957	0.944	0.864	0.814	1.169	0.288	0.330	0.388
5-9	0.429	0.222	0.957	0.944	0.864	0.814	1.169	0.288	0.430	0.600
10-14	0.216	0.258	0.957	0.944	0.864	0.814	1.169	0.288	0.527	0.315
15-19	0.185	0.200	0.957	0.944	0.864	0.814	1.169	0.688	0.225	0.585
20-24	0.337	0.099	0.957	0.944	0.864	0.895	1.169	0.467	0.580	0.638
25-29	0.710	0.177	0.957	0.944	0.864	0.888	1.169	0.820	0.818	0.689
30-34	0.913	0.314	0.957	0.944	0.864	0.685	1.169	0.590	0.949	0.773
35-39	0.761	0.436	0.957	0.888	0.864	0.811	0.966	0.700	0.960	0.752
40-44	0.818	0.526	0.957	0.985	0.864	0.746	0.765	0.739	1.033	0.780
45-49	0.883	0.631	0.965	0.977	0.931	0.791	0.787	0.687	0.974	0.819
50-54	0.908	0.894	0.969	1.036	0.899	0.851	0.784	0.671	0.980	0.969
55-59	0.936	0.939	0.948	0.996	0.877	0.816	0.797	0.739	1.010	0.962
60-64	1.017	0.884	0.967	0.959	0.940	0.863	0.812	0.749	1.023	0.998
65-69	1.048	1.086	1.029	0.949	0.972	0.923	0.861	0.857	1.010	1.025
70-74	1.145	1.103	1.038	0.975	0.990	0.982	0.979	0.898	1.051	1.034
75-79	1.141	1.304	1.063	0.969	1.082	1.055	1.134	1.069	1.032	1.082
80-84	1.220	1.354	1.096	1.066	1.169	1.121	1.368	1.275	1.070	1.126
85+	1.358	1.993	1.152	1.210	1.259	1.279	1.793	1.619	1.039	1.126

	Pancreas		Lung		Bladder		Lymphoma		Leukemia	
Age	Male	Female	Male	Female	Male	Female	Male	Female	Male	Female
0-4	0.824	0.408	0.711	0.849	0.107	0.141	0.875	0.829	0.356	0.354
5-9	0.824	0.408	0.711	0.849	0.107	0.141	0.574	0.470	0.280	0.224
10-14	0.824	0.613	0.711	0.849	0.107	0.141	0.519	0.354	0.483	0.525
15-19	0.824	0.204	0.711	0.849	0.107	0.141	0.367	0.235	0.703	0.684
20-24	0.824	0.204	0.711	0.849	0.107	0.141	0.382	0.233	0.865	0.815
25-29	0.824	0.531	0.711	0.849	0.107	0.159	0.657	0.269	0.883	0.936
30-34	0.824	0.497	0.926	0.849	0.182	0.393	1.005	0.367	0.814	0.796
35-39	0.893	0.712	0.939	0.849	0.157	0.533	1.069	0.419	0.866	0.872
40-44	0.907	0.745	0.922	0.890	0.344	0.459	0.987	0.491	0.702	0.801
45-49	0.927	0.902	0.909	0.891	0.443	0.452	0.817	0.541	0.770	0.882
50-54	0.936	0.922	0.920	0.891	0.459	0.405	0.746	0.657	0.780	0.830
55-59	0.977	0.952	0.932	0.918	0.550	0.522	0.743	0.689	0.811	0.829
60-64	0.986	0.970	0.954	0.937	0.756	0.637	0.881	0.838	0.899	0.880
65-69	1.015	0.989	0.986	0.979	0.751	0.764	0.966	0.935	0.991	0.959
70-74	1.016	1.020	1.039	1.038	1.069	0.878	1.168	1.139	1.192	1.068
75-79	1.016	1.033	1.090	1.038	1.305	1.199	1.376	1.484	1.404	1.185
80-84	1.061	1.058	1.169	1.038	1.740	1.520	1.683	1.484	1.501	1.337
85+	1.082	1.084	1.219	1.038	2.453	1.956	2.004	1.484	1.563	1.535

	Melanoma		Miscellaneous		Breast	Cervix	Uterus	Ovary	Prostate
Age	Male	Female	Male	Female	Female	Female	Female	Female	Male
0-4	1.000	3.573	0.514	0.430	0.874	0.574	0.322	0.247	0.548
5-9	1.000	1.000	0.444	0.405	0.874	0.574	0.322	0.247	0.548
10-14	0.324	0.706	0.395	0.299	0.874	0.574	0.322	0.247	0.548
15-19	0.959	0.937	0.377	0.222	0.874	0.574	0.322	0.202	0.548
20-24	0.670	0.531	0.279	0.147	0.874	0.425	0.322	0.241	0.548
25-29	0.764	0.455	0.229	0.154	0.679	0.329	0.553	0.173	0.548
30-34	0.564	0.493	0.261	0.207	0.733	0.406	0.382	0.236	0.548
35-39	0.902	0.542	0.375	0.270	0.651	0.457	0.353	0.336	0.548
40-44	0.857	0.769	0.527	0.400	0.569	0.508	0.458	0.433	0.548
45-49	0.871	0.710	0.709	0.498	0.564	0.739	0.411	0.496	0.548
50-54	0.815	1.073	0.850	0.620	0.790	0.994	0.473	0.614	0.548
55-59	1.110	0.769	0.957	0.741	0.954	1.025	0.610	0.808	0.548
60-64	1.087	1.297	1.072	0.852	0.933	1.218	0.701	0.955	0.551
65-69	1.093	1.150	1.226	0.954	1.067	1.366	0.901	1.089	0.508
70-74	0.988	1.507	1.375	1.076	1.092	1.740	1.219	1.369	0.508
75-79	1.425	2.241	1.523	1.219	1.215	2.259	1.629	1.650	1.083
80-84	1.620	1.909	1.714	1.372	1.513	2.729	2.570	1.650	1.876
85+	2.008	2.738	1.914	1.555	2.438	3.358	3.705	1.650	1.876

Table B.2: Estimated time since diagnosis parameters Y_τ, age-period-cohort survival model based on SEER data 1981-1995.

τ	Oral		Oesophagus		Stomach		Colorectal		Liver	
(years)	Male	Female	Male	Female	Male	Female	Male	Female	Male	Female
1	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
2	0.877	0.855	0.752	0.753	0.640	0.564	0.609	0.572	0.591	0.542
3	0.542	0.472	0.462	0.483	0.387	0.349	0.443	0.388	0.368	0.352
4	0.377	0.329	0.313	0.251	0.251	0.186	0.345	0.267	0.265	0.219
5	0.300	0.268	0.214	0.216	0.173	0.117	0.240	0.189	0.268	0.174
6	0.235	0.282	0.175	0.181	0.109	0.099	0.185	0.134	0.116	0.104
7	0.241	0.268	0.173	0.109	0.087	0.065	0.144	0.114	0.118	0.083
8	0.217	0.224	0.145	0.168	0.063	0.084	0.094	0.085	0.073	0.093
9	0.211	0.241	0.095	0.168	0.035	0.068	0.079	0.068	0.041	0.075
10	0.236	0.213	0.084	0.138	0.048	0.046	0.068	0.059	0.000	0.049
11	0.219	0.197	0.057	0.145	0.016	0.041	0.024	0.033	0.000	0.045
12	0.217	0.180	0.030	0.142	0.000	0.031	0.000	0.014	0.000	0.034
13	0.214	0.164	0.004	0.139	0.000	0.021	0.000	0.000	0.000	0.022
14	0.211	0.147	0.000	0.137	0.000	0.010	0.000	0.000	0.000	0.010
15	0.208	0.131	0.000	0.134	0.000	0.000	0.000	0.000	0.000	0.000

	Pancreas		Lung		Bladder		Lymphoma		Leukemia	
Age	Male	Female	Male	Female	Male	Female	Male	Female	Male	Female
1	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
2	0.721	0.720	0.683	0.693	0.621	0.478	0.506	0.551	0.466	0.438
3	0.434	0.430	0.402	0.384	0.391	0.267	0.380	0.381	0.334	0.280
4	0.257	0.226	0.254	0.235	0.313	0.191	0.320	0.321	0.282	0.259
5	0.156	0.170	0.181	0.186	0.277	0.141	0.292	0.304	0.251	0.236
6	0.150	0.108	0.145	0.144	0.251	0.141	0.255	0.291	0.225	0.206
7	0.116	0.092	0.123	0.134	0.249	0.131	0.233	0.265	0.225	0.189
8	0.100	0.104	0.120	0.117	0.231	0.108	0.194	0.230	0.215	0.172
9	0.062	0.078	0.113	0.116	0.252	0.139	0.196	0.197	0.194	0.165
10	0.085	0.051	0.113	0.114	0.257	0.165	0.159	0.170	0.168	0.153
11	0.047	0.048	0.101	0.102	0.253	0.154	0.139	0.137	0.162	0.139
12	0.029	0.036	0.093	0.094	0.255	0.160	0.116	0.106	0.147	0.126
13	0.010	0.023	0.085	0.087	0.256	0.165	0.093	0.075	0.133	0.113
14	0.000	0.010	0.078	0.079	0.258	0.171	0.070	0.044	0.118	0.100
15	0.000	0.000	0.070	0.071	0.260	0.177	0.047	0.013	0.104	0.088

	Melanoma		Miscellaneous		Breast	Cervix	Uterus	Ovary	Prostate
Age	Male	Female	Male	Female	Female	Female	Female	Female	Male
1	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
2	0.967	0.972	0.423	0.361	1.205	0.869	0.656	0.726	1.108
3	0.801	0.962	0.215	0.194	1.297	0.507	0.418	0.461	1.142
4	0.622	0.741	0.152	0.127	1.154	0.337	0.233	0.315	1.065
5	0.459	0.514	0.109	0.091	1.051	0.232	0.141	0.232	1.051
6	0.392	0.531	0.097	0.079	0.926	0.196	0.118	0.169	1.038
7	0.280	0.339	0.082	0.067	0.879	0.162	0.070	0.110	1.113
8	0.239	0.329	0.070	0.056	0.744	0.146	0.066	0.113	1.099
9	0.220	0.248	0.069	0.052	0.756	0.095	0.048	0.089	1.022
10	0.085	0.217	0.066	0.050	0.705	0.119	0.038	0.061	1.013
11	0.101	0.117	0.054	0.039	0.633	0.078	0.014	0.038	1.015
12	0.026	0.045	0.047	0.032	0.576	0.056	0.000	0.014	1.001
13	0.000	0.000	0.039	0.025	0.520	0.034	0.000	0.000	0.987
14	0.000	0.000	0.032	0.017	0.464	0.012	0.000	0.000	0.973
15	0.000	0.000	0.024	0.010	0.407	0.000	0.000	0.000	0.959

Table B.3: Estimated period (calendar year) parameters T_t, age-period-cohort survival model based on SEER data 1981-1995.

Period	Oral		Oesophagus		Stomach		Colorectal		Liver	
t	Male	Female	Male	Female	Male	Female	Male	Female	Male	Female
1981	0.991	1.035	1.121	1.180	1.060	1.003	0.983	1.013	1.092	1.092
1982	1.019	1.048	1.054	0.971	0.949	1.066	0.986	1.017	1.031	1.036
1983	1.065	1.095	1.046	1.046	0.939	0.992	0.989	1.026	1.019	1.033
1984	0.925	1.090	0.959	0.944	1.051	0.967	0.998	0.938	0.991	1.053
1985	0.895	1.053	0.975	0.957	0.993	0.994	0.904	0.949	0.974	1.000
1986	1.070	0.910	0.997	0.995	1.007	0.900	0.850	0.893	1.015	1.028
1987	0.994	0.979	0.996	0.954	1.002	0.976	0.883	0.904	0.952	0.966
1988	0.954	0.887	0.998	0.925	0.940	0.918	0.914	0.917	1.009	1.030
1989	1.104	0.984	0.933	0.943	0.919	0.980	0.870	0.876	0.944	1.029
1990	1.000	0.896	0.886	1.049	0.937	0.956	0.824	0.867	0.959	0.985
1991	0.991	0.959	0.880	0.893	0.998	0.932	0.810	0.824	0.969	0.944
1992	1.070	0.992	0.889	0.902	0.952	0.903	0.832	0.871	1.011	0.959
1993	0.998	0.865	0.904	1.017	0.966	0.944	0.888	0.894	0.972	0.982
1994	0.842	0.965	0.936	0.938	0.935	0.966	0.833	0.875	0.933	0.914
1995	0.899	0.942	0.887	0.927	0.900	0.981	0.896	0.971	0.942	0.943

	Pancreas		Lung		Bladder		Lymphoma		Leukemia	
Age	Male	Female	Male	Female	Male	Female	Male	Female	Male	Female
1981	1.004	1.022	1.020	1.017	0.916	1.052	0.886	1.019	1.004	0.916
1982	0.998	1.014	0.994	1.000	1.179	0.775	0.945	0.964	1.042	1.043
1983	0.997	0.980	0.988	0.996	0.998	1.005	0.790	0.930	1.033	1.030
1984	1.039	0.989	0.994	1.009	1.090	0.762	0.895	0.965	0.981	0.985
1985	1.018	1.022	0.997	0.972	0.972	1.039	0.910	0.986	0.962	0.963
1986	0.988	1.000	0.976	0.986	0.925	0.883	0.992	1.018	0.982	0.936
1987	0.988	0.987	0.977	1.015	0.885	0.995	0.945	0.881	0.959	1.023
1988	0.976	1.008	0.980	0.995	0.851	0.859	0.980	0.934	0.905	0.937
1989	0.984	0.967	0.982	0.981	0.746	0.908	1.030	0.960	0.977	0.946
1990	0.972	0.986	0.982	0.985	0.869	1.040	1.078	0.934	0.918	0.930
1991	0.962	0.974	0.968	0.971	0.918	0.907	1.021	0.940	0.930	0.959
1992	0.993	0.949	0.964	0.986	0.894	0.942	1.030	1.015	0.920	0.925
1993	0.994	0.985	0.979	0.977	0.844	0.897	1.056	0.908	0.927	0.972
1994	0.964	0.983	0.975	0.990	0.844	0.935	1.046	0.971	0.927	1.018
1995	0.995	0.987	0.968	0.990	0.815	0.878	1.080	0.981	0.952	1.026

	Melanoma		Miscellaneous		Breast	Cervix	Uterus	Ovary	Prostate
Age	Male	Female	Male	Female	Female	Female	Female	Female	Male
1981	1.230	1.467	0.928	1.089	1.182	1.091	1.229	1.076	1.589
1982	1.332	0.765	0.921	1.060	1.131	1.020	1.160	1.036	1.556
1983	0.830	0.689	0.894	1.044	1.174	0.862	1.104	1.039	1.454
1984	1.124	0.567	0.902	0.996	1.095	0.921	1.160	0.931	1.411
1985	1.014	0.979	0.855	0.994	1.056	1.041	0.932	0.972	1.301
1986	0.838	0.830	0.862	0.966	0.949	1.132	0.953	1.084	1.358
1987	0.971	1.038	0.843	0.982	0.810	0.993	1.089	1.015	1.210
1988	0.784	0.633	0.871	0.977	0.810	1.002	1.056	0.926	0.976
1989	1.124	0.969	0.823	0.954	0.764	0.870	1.047	0.849	0.833
1990	0.925	0.945	0.784	0.941	0.779	0.941	1.001	0.833	0.651
1991	0.757	0.574	0.770	0.894	0.695	0.881	1.075	0.817	0.309
1992	0.838	1.197	0.806	0.897	0.667	1.002	0.955	0.738	0.078
1993	0.902	0.813	0.803	0.910	0.723	0.848	1.042	0.827	0.126
1994	0.610	1.163	0.777	0.891	0.649	0.985	1.115	0.836	0.199
1995	0.757	0.543	0.805	0.910	0.721	0.862	0.968	0.779	0.290

Table B.4: Estimated relative probability of death after 1 year, RIS_1, age-period-cohort survival model based on SEER data 1981-1995.

Site	Males	Females
Oral	0.2023	0.1795
Oesophagus	0.6601	0.6351
Stomach	0.5737	0.5583
Colorectal	0.2069	0.2231
Liver	0.8305	0.7687
Pancreas	0.8282	0.8157
Lung	0.631	0.5677
Melanoma	0.0518	0.0286
Breast		0.039
Cervix		0.1257
Uterus		0.0730
Ovary		0.2865
Prostate	0.0372	
Bladder	0.0905	0.153
Lymphoma	0.2713	0.2446
Leukemia	0.3598	0.376
Miscellaneous	0.3481	0.3859