Cancer registry data based estimation of regional cancer incidence: application to breast and colorectal cancer in French administrative regions

Marc Colonna, Pascale Grosclaude, Jean Faivre, Ali Revzani, Patrick Arveux, Gilles Chaplain, Brigitte Tretarre, Guy Launoy, Josette Macé Lesec'h, Nicole Raverdy, Paul Schaffler, Antoine Buémi, François Ménégoz, Roger J Black

Abstract

Study objective—In many countries, cancer registries cover only a small part of the national population. Cancer incidence for the rest of the country has therefore to be estimated. This can be done from mortality data using the relation between incidence and mortality observed in the cancer registry areas. Such an approach was used to study geographical variation and trend of colorectal and breast cancer incidence in France where 10% of the national population is covered by cancer registries.

Design—This study applies the incidence/mortality ratios of cancer registry areas to regional mortality data to obtain an estimation of cancer incidence at a given point in time. Age and period effects are included in the statistical models.

Main results—The incidence estimations are given for 21 administrative regions and three time points (1985, 1990, 1995). The European standardised incidence rates for breast cancer ranged from 86.8 to 128.8. For colorectal cancer, these rates ranged from 48.2 to 79.6 for men, and from 32.5 to 48.8 for women. Breast cancer incidence has increased considerably between 1985 and 1995 with a higher increase in the north than in the south of France. The incidence of colorectal cancer has also increased, albeit to a lesser extent.

Conclusion—The incidence estimation method proposed leads to regional incidence rates that are useful for planning health care services on a regional basis and may also be used to study regional differences in incidence. This method is useful when only partial incidence data are available.

(J Epidemiol Community Health 1999;53:558–564)

Cancer is responsible for over 130 000 deaths per annum in France, representing about a quarter of the total mortality. Although mortality statistics give an overview of the extent of cancer, they do not provide information on the incidence of the different cancer sites, being unaware of curable cancer types and of other mortality causes of incidents and prevalent cancer cases. Only cancer registries, which record exhaustively incident cases in a given geographical area, can establish reliable morbidity statistics that are essential to health care planning, aetiological or trend studies. As cancer registries do not cover the whole country in many countries, incidence rates have often to be estimated. An estimation method that has been previously applied in France to estimate national incidence used mortality data from an area without a cancer registry and the incidence and mortality relation observed in the regional cancer registries. In this paper, this method has been extended to provide regional incidence estimates. Results of breast and colorectal cancer incidence estimation are provided and clearly illustrate the geographical heterogeneity that exists in cancer incidence.

Methods

Data collection


The data on mortality in each region, year of death, age group, sex and site of the tumour were available for the 1975–1992 period. They are provided by the Institut National de la Santé et de Recherche Médicale (INSERM). As practically no incident cases are observed before the 20–24 years age group for both breast and colorectal cancer, the age range starts at 20 years.

The population data estimated with respect to age, year, sex, and region were supplied by the Institut National de la Statistique et des Etudes Economiques (INSEE). Table 1 gives the breakdown of the population by sex and region for 1992. In table 1, the percentage of persons aged 75 years and over is also mentioned. It can be seen that there was considerable diversity in the age structures of the regional populations.

Study design

Modélisation

The previous studies applied the incidence/mortality ratios of cancer registry areas to national (for example, France) mortality data to obtain an estimation of cancer incidence at a given point in time. In a recent
Table 1  Breakdown of total population by sex and regions in 1992. Percentages of men and women over 75 years old

<table>
<thead>
<tr>
<th>Region</th>
<th>Whole population (men)</th>
<th>Whole population (women)</th>
<th>% 75+ (men)</th>
<th>% 75+ (women)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alsace</td>
<td>805 028</td>
<td>844 141</td>
<td>3.4</td>
<td>7.3</td>
</tr>
<tr>
<td>Aquitaine</td>
<td>1 366 365</td>
<td>1 461 691</td>
<td>5.7</td>
<td>9.9</td>
</tr>
<tr>
<td>Auvergne</td>
<td>641 997</td>
<td>676 429</td>
<td>5.6</td>
<td>10.1</td>
</tr>
<tr>
<td>Basse Normandie</td>
<td>681 460</td>
<td>717 718</td>
<td>4.1</td>
<td>8.2</td>
</tr>
<tr>
<td>Bourgogne</td>
<td>789 167</td>
<td>826 975</td>
<td>5.7</td>
<td>10.0</td>
</tr>
<tr>
<td>Bretagne</td>
<td>1 366 441</td>
<td>1 446 451</td>
<td>4.5</td>
<td>8.9</td>
</tr>
<tr>
<td>Centre</td>
<td>1 174 379</td>
<td>1 224 403</td>
<td>5.6</td>
<td>9.3</td>
</tr>
<tr>
<td>Champagne</td>
<td>662 033</td>
<td>686 686</td>
<td>4.3</td>
<td>8.0</td>
</tr>
<tr>
<td>Franche Comté</td>
<td>544 828</td>
<td>558 154</td>
<td>4.4</td>
<td>8.0</td>
</tr>
<tr>
<td>Haute Normandie</td>
<td>857 736</td>
<td>898 086</td>
<td>3.8</td>
<td>7.3</td>
</tr>
<tr>
<td>Languedoc R</td>
<td>1 045 453</td>
<td>1 113 210</td>
<td>6.1</td>
<td>9.7</td>
</tr>
<tr>
<td>Limousin</td>
<td>349 105</td>
<td>372 436</td>
<td>7.6</td>
<td>12.7</td>
</tr>
<tr>
<td>Lorraine</td>
<td>1 130 878</td>
<td>1 175 224</td>
<td>3.7</td>
<td>7.3</td>
</tr>
<tr>
<td>Midi Pyrénées</td>
<td>1 201 232</td>
<td>1 260 054</td>
<td>6.2</td>
<td>10.1</td>
</tr>
<tr>
<td>Nord</td>
<td>1 929 205</td>
<td>2 045 053</td>
<td>3.2</td>
<td>6.8</td>
</tr>
<tr>
<td>Pays de Loire</td>
<td>1 506 920</td>
<td>1 580 311</td>
<td>4.5</td>
<td>8.3</td>
</tr>
<tr>
<td>Picardie</td>
<td>902 701</td>
<td>927 298</td>
<td>3.7</td>
<td>7.0</td>
</tr>
<tr>
<td>Poitou Charente</td>
<td>784 460</td>
<td>820 529</td>
<td>6.3</td>
<td>10.1</td>
</tr>
<tr>
<td>PACA</td>
<td>2 088 994</td>
<td>2 246 857</td>
<td>5.6</td>
<td>9.5</td>
</tr>
<tr>
<td>Ile de France</td>
<td>5 242 685</td>
<td>5 561 017</td>
<td>3.4</td>
<td>6.8</td>
</tr>
<tr>
<td>Rhône Alpes</td>
<td>2 668 829</td>
<td>2 782 482</td>
<td>4.0</td>
<td>7.7</td>
</tr>
</tbody>
</table>

Figure 1  Plot of observed incidence and mortality European age standardised ratio against time for the French départements covered by a cancer registry—breast cancer.

study, the mortality trend, the trends in mortality and in incidence/mortality ratios in the total registered population were used to estimate national incidence rates over a 20 year period and therefore an estimate of incidence trend. This method has been extended to regional incidence. The procedure was divided into four main stages:

1. For each combinations of sex and site, we fitted a log-linear model for the age (subscript $i$), population (subscript $j$) and period (subscript $t$) specific mortality rates $\lambda_{ijt}$ in the national and pooled registry populations expressed by the following relation:

$$\log(\lambda_{ijt}) = a_i + \beta_j + \gamma_f t + \eta_{ijt}$$

(1)

where $a_i$, $\beta_j$, $\gamma_f$ and $\eta_{ijt}$ are categorical variables coding for age groups (20–24, 25–29,..., 80–84 and 85+) and populations (national and registry), $t$ is coding for the single years 1975–92. The model was fitted to obtain smoothed values for national and pooled registries mortalities. The model allows rates to differ in both populations,

subject to the constraint of a common rate ratio for all age groups and identical age specific time trends. The effect of the larger national population is to smooth the sparse age specific data of the registry populations. The inclusion of second degree polynomial term allows curvature in the age specific time trends. This seemed to be adequate for representing the slight and slowly changing trends observed in cancer mortality in large populations.

2. For each combination of sex and site, we fitted log-linear models for the age (subscript $i$), population (subscript $j$) and period (subscript $t$) specific mortality rates $\lambda_{ijt}$ in the 21 French areas, expressed by the relation:

$$\log(\lambda_{ijt}) = a_i + \beta_j + \gamma_f t + \theta_{ijt} + \xi_{ijt}$$

(2)

where the parameters have the same meaning as in model (1) except that $j$ codes now for the 21 different French areas. The model was fitted to obtain smoothed values for the local mortalities on the same assumptions as in model (1) with a possible area specific time trend.

3. The observed incidence and mortality data from the pooled registries seemed unsuitable for estimating age and year specific incidence/mortality ratios as these observations may be sparse for some age groups. Models were therefore fitted to estimate trends in the age specific incidence/mortality ratios in the pooled registry populations, expressed by the relation:

$$\log(\lambda_{ijt}) = a_i + \beta_j + \gamma_f t$$

(3)

where $c_i$ is the number of observed incident cases in the pooled registry populations at age $i$ and year $t$ and $\hat{d}_i$ is the estimated deaths in these populations from model (1). The model assumes a log-linear relation between incidence and mortality. A quadratic term may be included for some particular sites when sudden changes in diagnostic or treatment techniques appeared, as it has been shown for prostatic carcinoma.

4. Finally, for each cancer site, local incidence estimates were obtained by applying the estimated incidence/mortality ratios from model (3):

$$\frac{\hat{c}_i}{\hat{d}_i}$$

(4)

to the fitted values for local mortality from model (2). Fitted values for the age and year specific local area mortality are required instead of the observed values to avoid extreme variability because of the small number of observed deaths in age groups for some sites. Estimated number of incident cases $\hat{c}_i$ in the age group $i$, area $j$ and year $t$ are:

$$\hat{c}_i = \hat{d}_i \frac{\hat{c}_i}{\hat{d}_j}$$

(5)

where $\hat{d}_i$ is the number of estimated mortality cases from model (2).
Assessment of the reliability of the method

To check the reliability of the estimations, we have used some form of cross validation. We have dropped each cancer registry data from the analysis in turn and estimated its incidence from the others. Then we have compared observed and estimated incidence from each of the nine modelisations we have performed.

Variance estimation of the estimates

To estimate the variance of the estimated number of cases, we used the formulation (5). We have computed the variance of $\hat{c}_{ijt}$ using the classic variance of predicted value; the variance of $d_{ij}$ was obtained from model (2) and of the variance of the ratio $\hat{c}_{ijt}/d_{ij}$ was obtained using a combination of the variance of the predicted values $d_{ij}$ from model (2) and of the variance of the ratio $\hat{c}_{ijt}$. The variance of $d_{ij}$ was obtained using the classic variance of predicted value,

$$\text{var}(\hat{c}_{ij}) = \frac{(\text{var}(\hat{c}_{ij}))^2 + (d_{ij} \text{var}(\hat{c}_{ij}) + \text{var}(d_{ij}))}{d_{ij}^2}$$

The GLIM statistical package was used for calculations and to set up the models. European age standardised incidence rates were computed to take into account the differences in age structure between French areas.

Results

INCIDENCE/MORTALITY

Our estimations of incidence are based essentially on the incidence/mortality ratios. Figures 1 and 2 show the plots of the components of these ratios over the period 1980–1992 for the two cancer sites and for the nine French départements covered by a cancer registry. For each cancer site, the comparison between age standardised mortality rates and age standardised incidence rates shows a definite increase in the incidence/mortality ratio over this 12 year period. The most distinct increase was for breast cancer. The underlying hypothesis of a log-linear trend of the observed incidence/mortality thus appeared to be verified for the two studied sites.

REGIONAL INCIDENCE OF BREAST AND COLORECTAL CANCER IN 1992

Tables 2, 3, and 4 list the estimated number of cases of breast cancer in women and of colorectal cancer in both men and women for 1992 for each age group in each region. The standard deviation of estimated number of cases for each

*Estimated number of new cases. EASIR: European Age Standardised Incidence Rate. EASMR: European Age Standardised Mortality Rate. Numbers in parentheses are estimated standard deviation of total number of estimated incidence cases.
Table 3  Number of cases of colorectal cancer in women in the different age groups European standardised incidence and mortality rates in 1992.

<table>
<thead>
<tr>
<th>Year</th>
<th>20–44*</th>
<th>45–64*</th>
<th>65–74*</th>
<th>75–84*</th>
<th>85+*</th>
<th>all ages*</th>
<th>EASIR</th>
<th>EASMR</th>
<th>all ages*</th>
<th>EASIR</th>
<th>all ages*</th>
<th>EASIR</th>
</tr>
</thead>
<tbody>
<tr>
<td>1992</td>
<td>104</td>
<td>125</td>
<td>129</td>
<td>57</td>
<td>430</td>
<td>89.9</td>
<td>40.7</td>
<td>16.4</td>
<td>378</td>
<td>40.7</td>
<td>415</td>
<td>45.0</td>
</tr>
<tr>
<td>1985</td>
<td>104</td>
<td>125</td>
<td>129</td>
<td>57</td>
<td>430</td>
<td>89.9</td>
<td>40.7</td>
<td>16.4</td>
<td>378</td>
<td>40.7</td>
<td>415</td>
<td>45.0</td>
</tr>
<tr>
<td>1990</td>
<td>104</td>
<td>125</td>
<td>129</td>
<td>57</td>
<td>430</td>
<td>89.9</td>
<td>40.7</td>
<td>16.4</td>
<td>378</td>
<td>40.7</td>
<td>415</td>
<td>45.0</td>
</tr>
</tbody>
</table>

Estimated number of new cases. EASIR : European Age Standardised Incidence Rate. EASMR : European Age Standardised Mortality Rate. Numbers in parentheses are estimated standard deviation of total number of estimated incidence cases.

Table 4  Number of cases of colorectal cancer in men in the different age groups European standardised incidence and mortality rates in 1992.

<table>
<thead>
<tr>
<th>Year</th>
<th>20–44*</th>
<th>45–64*</th>
<th>65–74*</th>
<th>75–84*</th>
<th>85+*</th>
<th>all ages*</th>
<th>EASIR</th>
<th>EASMR</th>
<th>all ages*</th>
<th>EASIR</th>
<th>all ages*</th>
<th>EASIR</th>
</tr>
</thead>
<tbody>
<tr>
<td>1992</td>
<td>104</td>
<td>125</td>
<td>129</td>
<td>57</td>
<td>430</td>
<td>89.9</td>
<td>40.7</td>
<td>16.4</td>
<td>378</td>
<td>40.7</td>
<td>415</td>
<td>45.0</td>
</tr>
<tr>
<td>1985</td>
<td>104</td>
<td>125</td>
<td>129</td>
<td>57</td>
<td>430</td>
<td>89.9</td>
<td>40.7</td>
<td>16.4</td>
<td>378</td>
<td>40.7</td>
<td>415</td>
<td>45.0</td>
</tr>
<tr>
<td>1990</td>
<td>104</td>
<td>125</td>
<td>129</td>
<td>57</td>
<td>430</td>
<td>89.9</td>
<td>40.7</td>
<td>16.4</td>
<td>378</td>
<td>40.7</td>
<td>415</td>
<td>45.0</td>
</tr>
</tbody>
</table>

Estimated number of new cases. EASIR : European Age Standardised Incidence Rate. EASMR : European Age Standardised Mortality Rate. Numbers in parentheses are estimated standard deviation of total number of estimated incidence cases. 

There was some disparity in incidence between regions for breast cancer. An incidence above national average was noted in the Nord département, while Bretagne and Midi-Pyrenees had below average incidences. Overall, the estimated number of new cases in 1992 was a slightly over 32,000. The proportion of cases diagnosed before 45 years of age is 15%, 44% between 45 and 64 years, 23% between 65 and 74 years, 13% between 75 and 84 years, and 5% after 84 years. These variations varied little from region to region. The European standardised mortality rates for breast cancer were almost four times lower than the European standardised incidence rates.

Table 4  Number of cases of colorectal cancer in men in the different age groups European standardised incidence and mortality rates in 1992.

<table>
<thead>
<tr>
<th>Year</th>
<th>20–44*</th>
<th>45–64*</th>
<th>65–74*</th>
<th>75–84*</th>
<th>85+*</th>
<th>all ages*</th>
<th>EASIR</th>
<th>EASMR</th>
<th>all ages*</th>
<th>EASIR</th>
<th>all ages*</th>
<th>EASIR</th>
</tr>
</thead>
<tbody>
<tr>
<td>1992</td>
<td>104</td>
<td>125</td>
<td>129</td>
<td>57</td>
<td>430</td>
<td>89.9</td>
<td>40.7</td>
<td>16.4</td>
<td>378</td>
<td>40.7</td>
<td>415</td>
<td>45.0</td>
</tr>
<tr>
<td>1985</td>
<td>104</td>
<td>125</td>
<td>129</td>
<td>57</td>
<td>430</td>
<td>89.9</td>
<td>40.7</td>
<td>16.4</td>
<td>378</td>
<td>40.7</td>
<td>415</td>
<td>45.0</td>
</tr>
<tr>
<td>1990</td>
<td>104</td>
<td>125</td>
<td>129</td>
<td>57</td>
<td>430</td>
<td>89.9</td>
<td>40.7</td>
<td>16.4</td>
<td>378</td>
<td>40.7</td>
<td>415</td>
<td>45.0</td>
</tr>
</tbody>
</table>

Estimated number of new cases. EASIR : European Age Standardised Incidence Rate. EASMR : European Age Standardised Mortality Rate. Numbers in parentheses are estimated standard deviation of total number of estimated incidence cases. 

There was less regional disparity in incidence of colorectal cancer for women than for men. In this last group, the incidence rates were 10% above national average in Alsace, Nord, Lorraine and Bretagne, but 10% under in Provence Alpes Côte d’Azur, Midi-Pyrenees, Languedoc Roussillon and Franche Comté.

The incidence of colorectal cancer was slightly higher in men than in women (16 401 estimated cases versus 14 836 in 1992). The sex ratio calculated from the incidence rates ranged from 1.3 to 2 depending on the region. Colorectal cancer was sparse before 45 years with 3% of cases in both men and women. The proportion of cases diagnosed at 85 years and over was 7% in 15% the women.

TRENDS OF REGIONAL INCIDENCE OF COLORECTAL AND BREAST CANCER

In an attempt to evaluate the trends in cancer incidence, we have indicated in tables 2, 3, and 4 the number of estimated cases and
In 10 years, the incidence of breast cancer has increased considerably but with regional differences. Increases of more than a third were noted in Aquitaine, Auvergne, Basse-Normandie, Champagne, Limousin, Lorraine, Nord and Picardie. Smaller increases were observed in the other regions.

The increases in incidence of colorectal cancer in both men and women have been relatively modest, with a mean increase in incidence of 20%. Incidence remained stable over this decade in: Franche-Comté, Haute-Normandie, Limousin and Provence-Alpes-Côte-d'Azur for men, and in Alsace, Franche-Comté and Pays-de-Loire for women.

**RELIABILITY ASSESSMENT**

In tables 5, 6, and 7, the number of observed and estimated cases in each cancer registry are presented. Each cancer registry estimations are computed without taking into account its own observations. The confidence intervals of observed cases are provided. The ratio number of observed cases/number of estimated cases are computed. Except for cancer registry number 8 for the three cancer sites and for cancer registry number 6 for breast cancer, these ratio lie between 0.9 and 1.1. Apart from those two previous situations, the number of estimated cases are inside the corresponding confidence intervals computed according to the number of observed cases.

**Discussion**

We provide in this study the first ever estimation of the incidence of cancers in different regions in France. The incidences of breast and colorectal cancer were used as typical examples as they are common sites.

The quality of the estimation of incidence depends on the fit of the models, which in turn relies on the underlying hypotheses. A feature of our methodology was the procedure for smoothing mortality data in the départements covered by a cancer registry. With the assumption of a parallelism in the trend of adjusted mortality rates (model (1)), the data in pooled registries areas were considered as being representative of national data. This actually may not be the case because of geographical differences in the catchment areas of the regional cancer registries. This can be resolved by modelling only the mortality data of the départements covered by a regional cancer registry in the model (1), although it may not be valid when mortality is relatively low.

Using this latter method for breast and colorectal cancer, we obtained similar estimations. Model (2) contained parameters for modelling the spatio-temporal trends in regional mortality.

Model (3) could be improved by the inclusion of a second degree term in the time variable. This correction is not statistically useful for the cancer sites studied here, although it could be used for other sites, such as the prostate, which has presented rapid changes in incidence over the past decades. As each of the estimated models leads to a good statistical fit, we considered that this methodology is one possible approach for the estimation of regional incidence rates.

Our reliability assessments confirm that the method leads to reasonably good results. An internal study of French cancer registry seems to conclude that the discrepancy showed for cancer registry number 8 is attributable to a cancer registration problem. For breast cancer in cancer registry number 6, we found that incidence level is the highest of the nine French cancer registries and cancer mortality is at the same level as in other cancer registries so that no clear reason can be given to explain the discrepancy between observation and estimation in that area.
The data presented in table 1 show that the estimations involved regions with considerable differences in population size. This means that the variability in the estimations is not constant from region to region, which may lead to a bias in classifying the least populated regions. Further analysis using Bayesian methods would perhaps aid comparison between regions differing in population size. It may be useful to give confidence intervals of estimated numbers of cases and standardised rate estimates. These estimates were obtained from successive models so that the analytic expression of their confidence intervals are not readily defined. To evaluate the variability of our estimations, we have provided a variance estimation of the total number of estimated cases in each areas.

The use of incidence/mortality ratios to estimate regional incidence is based on the assumption of comparable survival rates in the different regions. This assumption is largely validated by the data from the EUROCAR study. The survival rates in the French départements that participated in this study were comparable and have provided the overall data for France. One should bear in mind, however, that targeted public health measures such as screening may affect the validity of this assumption in the future. The fact that the estimates of regional incidence rates for cancer rely on certain simplifying assumptions means that the values obtained should be viewed as orders of magnitude. The extrapolations to 1995 for the regions considered here produced values that were slightly higher than those of a previous national evaluation. This was attributed to the fact that the data in this study were obtained over a longer period of time, especially for the Bas-Rhin département. The extrapolations should thus be interpreted with due caution and may need regular updating.

Estimations of incidence are crucial for health care planning. Knowledge of the expected number of cases in each region, the available data on care facilities, taking the age factor into account, can provide an estimate of needs for surgical beds or of the number of people likely to require chemotherapy and radiotherapy. Estimations of regional incidence produce objective data on which health care needs in oncology may be relied on.

The estimations of incidence also highlight the limitations of mortality statistics for determining the magnitude of the problems represented by the two types of cancer studied here. For breast cancer, the incidence rates were four times higher than mortality rates, and even for colorectal cancer, incidence rates were double those of mortality. The statistics of incidence and mortality were also complementary. The increase in the incidence/mortality ratio is indicative of an improved prognosis for the two types of cancer studied. The increase was more pronounced for breast than for colorectal cancer. This trend has also been reported recently in Sweden for breast cancer and in the USA for colorectal cancer. The improved prognosis may stem from an earlier diagnosis or the advent of more effective treatments, or both. Further analysis of registry data could provide more information on this point.

The results of this study indicate that there were some regional variations in the incidence of breast cancer in France, the incidence ratio of the highest risk region to the lowest is only 1.5. In general, the incidence of breast cancer is lower than average in the south and above average in the north. A same ratio is estimated for colorectal cancer for women. For men, we found a ratio of highest to lowest risk incidences of 1.7. Incidence was particularly high in Alsace, and above average in Lorraine and Nord regions. The lowest incidences were found in Provence Alpes Côte d’Azur. The role of the Mediterranean diet has been proposed to account for the lower incidence in the south than in the north of France. A diet rich in fruit and vegetables has been shown to reduce the risks of both types of cancer studied.

The analysis of time trends is expected to provide valuable information, although more detailed analysis will require data collected over a longer period of time. Already it seems that there are regional differences in such trends. The changes in incidence of breast cancer were found to be more pronounced in the north than in the south of France. The incidence of colorectal cancer has also increased but to a smaller extent than for breast cancer. In certain regions the incidence has remained stable.

Our study shows that there is a heterogeneity both in spatial distribution and in time trends of breast and colorectal cancer incidence. The results of this study are therefore helpful to plan regional health care needs when no national cancer incidence registration exists. The method we propose uses the empirical relation between incidence and mortality for each period and age class. This method is useful for countries lacking nationwide coverage by cancer registries when no great heterogeneity exists between areas in survival. It is an interesting alternative method to that used by Verdecchia et al when no reliable survival data are available. Of course, our method may be improved. Firstly, our method uses observed correlation between incidence and mortality in areas covered by cancer registries at each time point where observations are available so that no time lag between incidence and mortality is taken into account; it may be interesting to evaluate the effect of including such a lag. Secondly, our mortality smoothing does not include any spatial dependence; perhaps such information may improve smoothed mortality rates. Lastly, no mortality variance is included in our model (3), where each estimated mortality case is an offset.


Conflicts of interest: none.