Cancer Epidemiology 37 (2013) 857–863

Contents lists available at ScienceDirect

Cancer Epidemiology

The International Journal of Cancer Epidemiology, Detection, and Prevention

journal homepage: www.cancerepidemiology.net

Probabilities of dying from cancer and other causes in French cancer patients based on an unbiased estimator of net survival: A study of five common cancers☆

H. Charvat [a,b,c,*], N. Bossard [a,b,c], L. Daubisse [d,e], F. Binder [f,g], A. Belot [a,b,c,h], L. Remontet [a,b,c]

[a] Université Claude Bernard Lyon 1, 43 Boulevard du 11 Novembre 1918, F-69100 Villeurbanne, France
[b] CNRS, UMR5558, Laboratoire de Biométrie et Biologie Evolutive, Equipe Biostatistique-Santé, 16 rue Raphaël Dubois, F-69100 Villeurbanne, France
[c] Hospices Civils de Lyon, Service de Biostatistique, 162 avenue Lacassagne, F-69003 Lyon, France
[d] Registre des cancers du Tarn, BP 37, F-81001 Albi, France
[e] CRLCC Institut Claudius Regaud, 20-24 rue du Pont Saint-Pierre, F-31052 Toulouse, France
[f] Registre des cancers du Bas-Rhin, Laboratoire d'Epidemiologie et de Sante Publique, EA 3430 Faculte de Medecine, Universite de Strasbourg, 4 rue Kirschleger, F-67085 Strasbourg, France
[g] Reseau des Registres Francais des Cancers FRANCIM, F-31073 Toulouse, France
[h] Institut de Veille Sanitaire, Département des maladies chroniques et traumatismes, F-94410 Saint-Maurice, France

ARTICLE INFO

Article history:
Received 18 January 2013
Received in revised form 10 June 2013
Accepted 13 August 2013
Available online 22 September 2013

Keywords:
Cancer registries
Competing risks
Crude probability
Net survival

ABSTRACT

Background: Net survival is the survival that would be observed if cancer were the only possible cause of death. Although it is an important epidemiological tool allowing temporal or geographical comparisons, it cannot inform on the "crude" probability of death of cancer patients; *i.e.*, when taking into account other possible causes of deaths.
Methods: In this work, we provide estimates of the crude probabilities of death from cancer and from other causes as well as the probability of being alive up to ten years after cancer diagnosis according to the age and year of diagnosis. Based on a flexible excess hazard model providing unbiased estimates of net survival, our methodology avoids the pitfalls associated with the use of the cause of death. We used data from FRANCIM, the French network of cancer registries, and studied five common cancer sites: head and neck, breast, prostate, lung, and colorectal cancers.
Results: For breast, prostate, and colorectal cancers, the impact of the other causes on the total probability of death increased with the age at diagnosis whereas it remained negligible for lung and head and neck cancers whatever the age. For breast, prostate, and colorectal cancer, the more recently was the cancer diagnosed, the less was the probability of death from cancer.
Conclusion: The crude probability of death is an intuitive concept that may prove particularly useful in choosing an appropriate treatment, or refining the indication of a screening strategy by allowing the clinician to estimate the proportion of cancer patients who will die specifically from cancer.

1. Introduction

The analysis of cancer-registry data allows epidemiologists to gain some insight into the survival of patients diagnosed with a particular type of tumour. This is important to assess the benefit of a treatment or a screening campaign [1]. However, a correct assessment of the risk of death from cancer should consider that cancer patients may die from other causes too [2]. Indeed, in elderly patients, the probability of dying from other causes may be greater than that of dying from cancer. Thus, it would be interesting to estimate the probability of death associated with each of these two mutually exclusive competing risks of death.

This probability of death from one cause (*e.g.*, cancer) in the presence of other causes is called the "crude" probability of death [3–9]. In competing risk settings, this probability is also known as the cumulative incidence function [10,11]. In the presence of cause of death information, this function can be obtained using competitive risk methods [12] by estimating the two (or more) cause-specific hazards and combining them with the overall survival function. However, the cause of death is not always known and, when known, it is not always reliable [13–15]; besides, in many situations it is very difficult to determine the real cause of death, especially in elderly patients with various comorbidities and deaths indirectly attributable to cancer (*e.g.*, pulmonary embolism after tumour surgery).

☆ This work was carried out at Hospices Civils de Lyon, Service de Biostatistique, Lyon, France.

* Corresponding author at: Laboratoire Biostatistique-Santé, Centre Hospitalier Lyon Sud, Pavillon 4D, 165, chemin du Grand Revoyet, F-69495 Pierre-Bénite Cedex, France. Tel.: +33 4 78 86 57 75; fax: +33 4 78 86 57 74.
E-mail addresses: hadrien.charvat@chu-lyon.fr, chadrien@ncc.go.jp (H. Charvat).

1877-7821/$ – see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.canep.2013.08.006

The pitfalls of using the cause of death have long been known and motivated the design of "relative survival" methods [16–18]. Without requiring the cause of death, these methods provide an estimation of the "net survival"; that is, the survival that would be observed in cancer patients if cancer were the only cause of death [18–20]. Relative survival methods rely mainly on the assumption that, in cancer patients, death from other causes can be correctly obtained from the all-cause mortality of the general population. Once the net survival estimated, one minus that estimate can be interpreted as the "net" probability of death from cancer.

The net survival setting is particularly appropriate to derive estimates of the crude probability of death without requiring cause of death information. In 2000, Cronin and Feuer developed a method based on life-tables to estimate the crude probability of death from cancer [4] and, in 2010, Lambert et al. [7] showed that the crude probability of death can be derived from an excess-hazard model (or "relative survival model").

Using the general principle of Lambert's method and the flexible excess-hazard model of Remontet et al. [21], the objective of the present study was to provide estimates of the crude probability of death due – directly or indirectly – to cancer for the five major cancer sites plaguing the French population; *i.e.*, head and neck, breast, lung, prostate, and colorectal cancers [22].

2. Material and methods

2.1. Data description

The dataset used here was provided by the French network of cancer registries (FRANCIM). Eight cancer registries – covering about 9% of the French population – were considered; those of Départements Calvados, Côte-d'Or, Doubs, Isère, Bas-Rhin, Saône-et-Loire, Somme, and Tarn. The breast, prostate, lung, head and neck, and colorectal cancers of patients diagnosed between 1 January 1989 and 31 December 2004 were classified according to the International Classification of Diseases for Oncology, Third Edition (ICD-O-3). The administrative censoring date was 1 January 2008. The number of cancer cases by site and sex were as follows: 46,166 women with breast cancer, 26,059 men with colorectal cancer, 21,399 women with colorectal cancer, 12,261 men with head and neck cancer, 1727 women with head and neck cancer, 24,963 men with lung cancer, 4482 women with lung cancer, and 36,607 men with prostate cancer.

2.2. Net survival estimation

The first step in estimating the crude probability of death was to obtain an unbiased estimate of net survival through the use of a flexible fully parametric excess hazard survival model [21]. The basic idea of this model is that the total hazard, λ_{Tot}, observed in the study population may be considered as the sum of a disease-specific hazard (the excess hazard), λ_C, and an other-cause-specific hazard, $\lambda_{\bar{C}}$:

$$\lambda_{Tot}(t) = \lambda_C(t) + \lambda_{\bar{C}}(t) \tag{1}$$

$\lambda_{\bar{C}}$ was considered known and obtained from population life tables by year, age, sex, and Département assuming that the all-cause mortality in the general population is a good estimate of the other-cause mortality in cancer patients.

The logarithm of the baseline excess hazard was modelled using a cubic spline with two knots at one and five years after diagnosis. The model included also covariates that influence additively the logarithm of the baseline excess hazard. Age at diagnosis was introduced into the model as a continuous variable and its non-linear effect was described by a cubic spline with a knot at the mean age of the study population. Moreover, the effect of age at diagnosis was allowed to vary over time; that is, the hazard for age was modelled as non-proportional. The effect of variable "Département" was modelled using a categorical variable. This constituted the basic model. Besides, the potential effect of the year of diagnosis was tested through fitting two more models derived from the basic one through: (i) the use of a quadratic spline with a knot at year 1997; and, (ii) allowing this effect to be non-proportional. For each cancer site and each sex, the selection of the final model was made by comparing the Akaike information criteria of the three above-described models. A more detailed description of the model-building strategy is provided in Appendix 1. The final model included all the covariates that define the population life table and have a significant effect on the excess hazard, thus providing unbiased estimates of net survival [19,20].

This procedure allowed us to estimate, for each cancer site and sex, the net survival associated with cancer, S_C, according to the age at diagnosis, the Département, and the year of diagnosis:

$$S_C(t, \text{age}, \text{dept}, \text{year}) = \exp\left[-\int_0^t \lambda_C(u, \text{age}, \text{dept}, \text{year})\, du\right] \tag{2}$$

Conversely, the net survival associated with the other causes of death, $S_{\bar{C}}$, is given by:

$$S_{\bar{C}}(t, \text{age}, \text{dept}, \text{year}) = \exp\left[-\int_0^t \lambda_{\bar{C}}(u, \text{age}, \text{dept}, \text{year})\, du\right] \tag{3}$$

Formula (1) implies that the overall survival, S_{Tot}, is the product of the two above-defined survivals:

$$S_{Tot}(t, \text{age}, \text{dept}, \text{year}) = S_C(t, \text{age}, \text{dept}, \text{year}) \cdot S_{\bar{C}}(t, \text{age}, \text{dept}, \text{year}) \tag{4}$$

2.3. Estimation of the probability of death from cancer

The crude probability of death from cancer, Cr_C (respectively, from other causes, $Cr_{\bar{C}}$), was estimated by considering the cumulative effect of the excess hazard, λ_C (respectively, $\lambda_{\bar{C}}$) on the proportion of patients still alive (that is, the overall survival) at each time point [4,7]:

$$Cr_C(t, X) = \int_0^t \lambda_C(u, X) \cdot S_C(u, X) \cdot S_{\bar{C}}(u, X)\, du \tag{5}$$

where X stands for the vector of covariates (age, dept, year). The crude probability of death from other causes $Cr_{\bar{C}}$ is defined as in formula (5) by replacing λ_C by $\lambda_{\bar{C}}$.

In the following section, we present age- and year of diagnosis-specific crude probability of death estimates. The method used to obtain these estimates is described in Appendix 2.

2.4. Interpretation of the results

The administrative censoring date being the 1st of January 2008, patients diagnosed after that date were not followed-up for ten years: this means that the effect of the year of diagnosis for year 2004 (for example) is estimated from individuals followed between 0 and 4 years. Using the model to estimate the crude probabilities of death ten years after diagnosis for patients diagnosed in 2004 can thus be seen as an extrapolation (indicated by asterisks in the tables and dashed lines in the figures). This extrapolation is mainly based on the assumption that the effect of the year of diagnosis remains the same between 4 and 10 years (in the case of a proportional effect of year of diagnosis) or on the assumption that the non-proportional effect of year of diagnosis