WAITING PERIOD FROM DIAGNOSIS FOR MORTGAGE INSURANCE ISSUED TO CANCER SURVIVORS

Antoine Soetewey, Catherine Legrand,
Michel Denuit, Geert Silversmit

REPRINT | 2020 / 39

UCLouvain

ISBA | LIDAM
Louvain Institute of Data Analysis and Modeling in economics and statistics

European Actuarial Journal (2021) 11:135–160
https://doi.org/10.1007/s13385-020-00254-x

ORIGINAL RESEARCH PAPER

Waiting period from diagnosis for mortgage insurance issued to cancer survivors

Antoine Soetewey[1] · **Catherine Legrand**[1] · **Michel Denuit**[1] · **Geert Silversmit**[2]

Received: 10 February 2020 / Revised: 9 September 2020 / Accepted: 3 November 2020 /
Published online: 22 November 2020

Abstract

The Massart (J Cancer Policy 15:70–71, 2018) testimonial illustrates the difficulties faced by patients having survived cancer to access mortgage insurance securing home loan. Data collected by national registries nevertheless suggest that excess mortality due to some types of cancer becomes moderate or even negligible after some waiting period. In relation to the insurance laws passed in France and more recently in Belgium creating a right to be forgotten for cancer survivors, the present study aims to determine the waiting period after which standard premium rates become applicable. Compared to the French and Belgian laws, a waiting period starting at diagnosis (as recorded in national databases) is favored over a waiting period starting at the end of the therapeutic treatment protocol. This aims to avoid disputes when a claim is filed. Since diagnosis is often recorded in the official registry database, as is the case for the Belgian Cancer Registry, its date is reliable and unquestionable in case of claim. Based on 28,994 melanoma and thyroid cancer cases recorded by the Belgian Cancer Registry, the length of the waiting period is assessed with the help of widely-accepted tools from biostatistics, including relative survival models and time-to-cure indicators. It turns out for instance that a waiting period of 4 years after diagnosis is enough for 30-year-old thyroid cancer patients. This appears to be similar to the 3-year period starting at the end of treatment protocol according to the Belgian law in such a case.

Keywords Term insurance · Impaired lives · Cancer · Home loan · Right to be forgotten

✉ Antoine Soetewey
antoine.soetewey@uclouvain.be

[1] Institute of Statistics, Biostatistics and Actuarial Sciences, Louvain Institute of Data Analysis and Modeling in Economics and Statistics, UCLouvain, Louvain-la-Neuve, Belgium

[2] Belgian Cancer Registry, Brussels, Belgium

1 Introduction

Property loans are often accompanied with mortgage insurance that pays the balance of the loan if the mortgagor dies. Coverage is usually awarded in the form of term insurance with decreasing sum insured, with the amount of death benefit diminishing as the debt decreases. This is common practice in Belgium, with about 170,000 new mortgage loans per year, mainly contracted by young adults acquiring their first family house (statistics from the Belgian Central Credit Register indicate that 36% of new mortgage loans in 2017 were contracted by borrowers younger than 35 and about 68% were granted to borrowers younger than 45).

Based on answers to a health questionnaire, insurers evaluate applicant's health status and either impose surcharges in case of impaired lives or refuse to cover the risk. Filling such health questionnaires may create frustration for patients having survived cancer occurred many years ago. Having repeatedly to answer questions related to this disease has psychological consequences and being charged higher premiums or denied coverage generates a feeling of discrimination [45]. This is often felt as a double penalty by cancer survivors.

The restricted access to insurance cover is often regarded as a barrier to property and home ownership (in case of house loan) and to entrepreneurship (in case of professional loan). This lead Belgian authorities to create the *Bureau du suivi de la tarification assurance solde restant dû* (http://www.bureaudusuivi.be)—*Opvolgingsbureau voor de tarifering schuldsaldoverzekering* (http://www.opvolgingsbureau.be) (which could be translated literally as the "Outstanding balance insurance pricing monitoring office") in 2014, in application of the law on insurance. This body reviews health questionnaires used by insurance companies selling mortgage insurance in Belgium and checks whether the proposed premium surcharges or cover denials are justified for impaired lives. In 2017, only 16% of the 454 cases submitted by insurance applicants to the Bureau resulted in improved policy conditions, as it can be read from the annual report published by the Bureau [13].

Faced with a similar situation, France established in 2016 a "*droit à l'oubli*" (translated literally as "right to be forgotten" in the remainder of this text), that is, the right for an insurance applicant not to declare a previous cancer after a period of 10 years starting at the end of the therapeutic protocol. This 10-year waiting period is reduced to 5 years if the applicant suffered cancer before the age of 18. These periods of 10 and 5 years start from the date of the end of the therapeutic treatment, in absence of relapse within this period. The 10-year length of the waiting period is further shortened for several types of cancer (and other non cancer related pathologies), as detailed in the reference grid used in France (known as convention AERAS, see http://www.aeras-infos.fr/cms/sites/aeras/accueil.html) with reduced duration after which survivors have access to the right to be forgotten. After this period, insurance companies cannot take the pathology into account in risk assessment and cannot refuse the insurance, nor impose a premium surcharge because of the pathology.

A similar rule has been introduced in Belgium in a law dated April 4, 2019 (published in the Official Journal on April 18, 2019) for home-related and

professional mortgage insurance. Despite clear similarities, the right to be forgotten established in Belgium differs from its French counterpart in an important way. Cancer survivors must still declare their pathology to insurance companies when applying for mortgage insurance in Belgium but the decision to grant coverage cannot be based on this information. Besides this right to not declare (in France) or to declare but without consequences (in Belgium) a cancer after a given waiting period, the premium surcharges may also be prohibited or limited for some cancer types. This has been implemented in Belgium in a Royal decree published in the Belgian Official Journal on June 14, 2019.

Although the establishment of such a right to be forgotten in Belgium is clearly an improvement for cancer survivors, there is most probably room for further reducing the waiting period for some cancer types. Also, there remains some ambiguity about what is considered as treatment and thus what marks the end of the therapeutic protocol. Since all cancer cases (and date of diagnosis) must be recorded in a national database in Belgium and many other EU countries, defining the start of the waiting period at the recorded date of diagnosis would certainly avoid endless discussions when a claim is filed.

There is abundant literature on cancer survival in biostatistical and medical studies. However, this topic has been the subject of few actuarial papers beyond those dealing with the assessment of extra mortality after Haberman and Renshaw [35] and Renshaw [58], such as Dodd et al. [26]. Let us briefly discuss some of the contributions on pricing life insurance specifically for cancer patients that appeared in the actuarial literature.

Lemaire et al. [42] considered term life pricing in the presence of a family history of breast or ovarian cancer. These authors found that while many women with a family history of breast or ovarian cancer can be accepted at standard rates, women with two family members with cancer or one first-degree relative with cancer at an early age show substantial mortality increases (up to 100%) and can thus probably only be accepted at higher premium rates. Moreover, the authors also found that mortality increases for women with the BRCA1 or BRCA2 gene mutation (a malfunction which results in cells more likely to develop additional genetic alterations that can lead to cancer) reach 150% and can thus possibly only be accepted at a premium rate that incorporates a severe mortality surcharge.

Using 10-year and 20-year term life products, Shang [62] calculated the single premium for breast cancer patients and found that both the average and minimum premium of the sample cancer patients to be much higher than standard premium, with the minimum premium still being close to 40% of the sum insured for the least risky patients. In order to improve the affordability of insurance products for cancer patients, Shang [62] also suggested to set a waiting period during which no death claims will be paid. Since many deaths happen during the first years after diagnosis, a waiting period reduces the mortality risk and thus the net premium. For instance, for a 20-year term life product with sum insured 10,000 contracted by a 40-year-old cancer patient, a waiting period of 1, 2 and 3 years reduces the average premium by, respectively, 24.85%, 41.48% and 51.05% compared to no waiting period, according to the calculations by Shang [62].

The present paper concentrates on the determination of the length of the waiting period embedded in the right to be forgotten, that is, the minimum duration before the applicant can be covered at standard premium rate. To this end, we apply several widely-accepted tools from biostatistics in order to assess excess mortality. We concentrate on melanoma (ICD-10 C43) and thyroid (ICD-10 C73.9) tumors for the sake of illustration, leaving other types of cancer for future research. These two types of cancer (called cancer sites) were selected to get a significant number of incidences occurring before the age of 40 (mortgage insurance applicants being rather young) and as they lead to a fraction of the patients who have a chance of survival close to cancer-free patients. We were thus looking for cancers with a relatively high survival rate or high cured rate. Our analysis is based on 28,994 cancer cases recorded by the Belgian Cancer Registry (BCR): 19,848 melanoma and 9146 thyroid tumors, diagnosed between 2004 and 2016.

Based on survival data recorded by the Belgian Cancer Registry, the present paper aims

- to show that for some types of cancer (with melanoma and thyroid as examples), survivors actually have a survival comparable to that of the general population, that is, excess mortality is negligible.
- to demonstrate that patients having survived long enough to some types of cancer (still with melanoma and thyroid as examples) can access life insurance market at standard insurance rates, contrarily to the common belief within the actuarial community. The technical waiting period appears to be relatively short, and shorter compared to the 10-year period specified in the law.

In addition, we promote a waiting period starting at diagnosis rather than at the end of the therapeutic treatment protocol in order to avoid disputes in case of death. Indeed, diagnosis is recorded in databases maintained by official bodies within the European Union. The results obtained in the present study appear to be particularly encouraging as they suggest a considerable shortening of the 10-year waiting period for some types of cancer.

The remainder of this paper is structured as follows. Section 2 presents the data used to perform the present study. Sections 3 and 4 apply several tools from biostatistics to assess the length of the waiting period, with a focus on the estimation of the survival of cancer patients in Sect. 3 and concentrating rather on the time after which we can consider the patients still alive as "cured" in Sect. 4. Comparisons with standard premium rates based on life tables generally used on the Belgian market are provided in Sect. 5. The final section (Sect. 6) concludes the paper with a discussion.

2 Data sources

2.1 Belgian Cancer Registry (BCR)

The Belgian Cancer Registry (BCR) is a national population-based cancer registry collecting data on all new cancer diagnoses in Belgium since the incidence year 2004. For the execution of this main task, BCR relies on its own specific legislation.

In this study, we restrict our analysis to two cancer sites, as explained in the introductory section. We also limit our analyses to patients from 20 to 69 years old at time of diagnosis since the right to be forgotten mainly concerns young adults and active life. A total of 19,848 cases of melanoma and 9146 cases of thyroid cancer diagnosed between 2004 and 2016 were followed-up until the 1st of July 2018. Follow-up thus varied from 2 years for patients diagnosed in 2016 to 14 years for those diagnosed in 2004. Patients without a national security number (INSZ/NISS) were excluded from our analyses, as we have no vital status on these patients. Patients lost to follow-up (mostly due to moving abroad) and patients still alive at the end of the follow-up period were censored. Table 1 summarizes the percentage of lost to follow-up before the 1st of July 2018 in each remaining subgroup together with the number of included cases. The fraction of patients lost to follow-up per subgroup varied from 1.04% for women with melanoma cancer aged 50–69 to 4.41% for male melanoma cancer patients aged 20–34. The total fraction of patients lost to follow-up cases, regardless of gender, site or age group was 1.87%. The analyses were conducted separately for women and men, representing 65.13% and 34.87% of all cases respectively, and age at diagnosis was included as a covariate, continuous or categorical depending on the approach. Age at diagnosis ranges from 20 to 69 years and 3 age groups were considered: 20–34, 35–49 and 50–69.

Table 1 Numbers of melanoma and thyroid cancer cases diagnosed in Belgium between 2004 and 2016 (BCR data) by gender, site and age group, with percentage of lost to follow-up

Gender	Cancer site	Age at diagnosis	Lost to follow-up (%)	Number of cases included
Women	Melanoma	20–34	3.44	1863
		35–49	1.07	4386
		50–69	1.04	5786
	Thyroid	20–34	3.07	1204
		35–49	2.66	2596
		50–69	1.67	3048
Men	Melanoma	20–34	3.45	725
		35–49	2.29	2360
		50–69	1.84	4728
	Thyroid	20–34	3.30	273
		35–49	2.49	724
		50–69	1.69	1301
Total			1.87	28,994

2.2 General population

In order to estimate excess cancer mortality, mortality in the cancer population must be compared to the expected mortality in the general population. Belgian population life tables, obtained from Statbel (the Belgian statistical office), were used to estimate expected mortality in the general population. Gender-specific mortality rates for the period 2004–2018 (by single year of age) have been smoothed in two dimensions to remove erratic variations. The surface smoothing was performed with the SAS procedure PROC LOESS using local linear polynomials weighted by population size [19–21].

3 Survival of cancer patients

3.1 Overall survival

In this paper, we analyze survival (i.e., time to death) for cancer patients beyond diagnosis according to a number of covariates summarized into the vector $\boldsymbol{Z}$. Specifically, T denotes the remaining lifetime at diagnosis. Given $\boldsymbol{Z} = z$, T has probability density function $f(\cdot|z)$, distribution function $F(\cdot|z)$, survival function $S(\cdot|z) = 1 - F(\cdot|z)$, and hazard rate, or force of mortality $\lambda(\cdot|z) = f(\cdot|z)/S(\cdot|z)$. Contrarily to insurance studies, T denotes the remaining lifetime since diagnosis and age at diagnosis is included in the covariates (attained age is thus obtained by summing age at diagnosis and survival time). This is why we refrain here from complying with the international actuarial notation for survival probabilities and force of mortality (when computing premiums in Sect. 5, we will revert back to the actuarial notation).

The non-parametric Kaplan–Meier [38] estimator is used here to estimate the overall survival function $S(\cdot)$, without distinguishing according to causes of death [5]. Estimated overall survival probabilities according to gender and site (melanoma and thyroid) are detailed in Table 2 and illustrated in Fig. 1. The 10-year overall survival probabilities range from 0.774 (with 95% confidence interval [0.762–0.786]) for men with melanoma cancer to 0.93 (with 95% CI [0.922–0.938]) for women with thyroid cancer (see Table 2). Remember that overall survival probabilities take into account all causes of deaths, that is, both cancer and non-cancer related deaths are considered.

In addition to the two curves for women and men in Figs. 1 and 2, a third curve including both sexes (denoted "Belgium") is also drawn for the simple reason that that since December 21, 2012, the European directive on equality between men and women also applies to outstanding balance insurance, so the gender no longer has an influence on the premiums. Therefore, we believe it is of interest for insurance companies to visualize these survival probabilities for both sexes combined.

Table 2 Estimated overall survival probabilities by gender and site using the non-parametric Kaplan–Meier estimator at 5 and 10 years after diagnosis, $\widehat{S}(t = 5)$ and $\widehat{S}(t = 10)$, with their confidence interval (CI) at 95% level

Sex	Cancer site	$\widehat{S}(t = 5)$	$CI_{95\%}$	$\widehat{S}(t = 10)$	$CI_{95\%}$
Women	Melanoma	0.928	(0.923–0.933)	0.885	(0.878–0.893)
	Thyroid	0.961	(0.956–0.966)	0.930	(0.922–0.938)
Men	Melanoma	0.847	(0.838–0.855)	0.774	(0.762–0.786)
	Thyroid	0.915	(0.903–0.927)	0.839	(0.820–0.860)

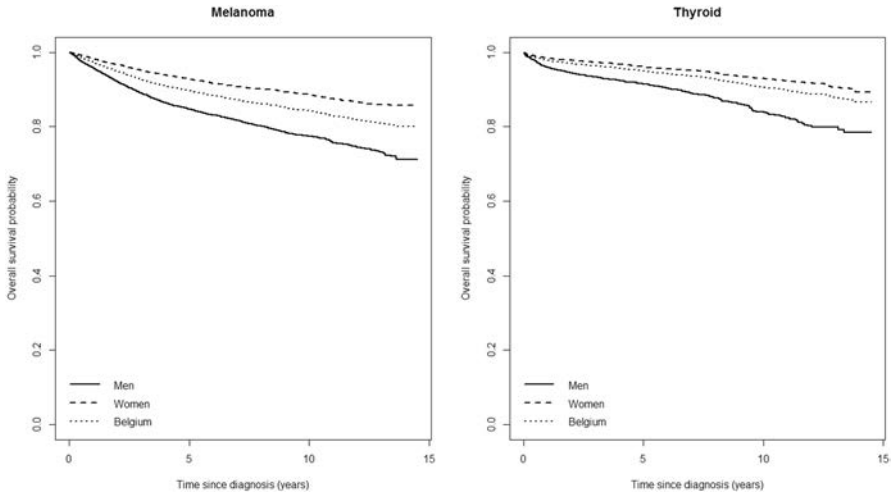

Fig. 1 Estimated overall survival probability by gender and site using the non-parametric Kaplan–Meier ([38]) estimator

3.2 Relative survival

Information on cause of death due to cancer is often unavailable or unreliable and not all deaths can be easily classified as deaths due to the cancer of interest or due to another cause [50]. Relative survival, which does not require information on the cause of death, provides a measure of the excess mortality experienced by cancer patients by comparing the mortality in the cancer population with the mortality in the general population. This led the relative survival to become the standard measure of patient survival for population-based cancer registries, as shown by its prominence around the world in studies related to cancer survival [8, 22, 49, 52, 53, 59].

Relative survival models are divided into two types; (i) additive models and (ii) multiplicative models; see e.g. Pohar and Stare [54]. Despite the wide acceptance of multiplicative specifications within the actuarial community, it turns out that additive models are biologically more plausible in cancer studies and provide a better fit to the data [10, 12, 25, 29, 36]. The additive specification is thus

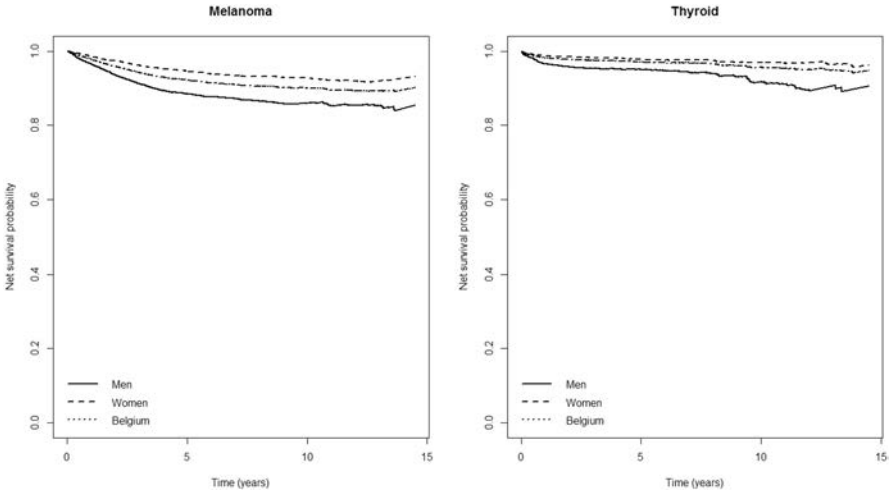

Fig. 2 Net survival by gender and site using the non-parametric Perme et al. [52] estimator

favored here. The hazard rate at time t since diagnosis for cancer patients with covariate vector $\boldsymbol{Z}$, is decomposed into two additive components: the population hazard based on available patient's characteristics $\boldsymbol{Z} = z$, denoted as $\lambda_P(\cdot|z)$, and the excess hazard specific for the disease of interest, denoted as $\lambda_E(\cdot|z)$. Formally,

$$\lambda(t|z) = \lambda_P(t|z) + \lambda_E(t|z). \tag{3.1}$$

From this expression, relative survival model (3.1) can be written as

$$S(t|z) = S_P(t|z)r(t|z)$$

where the relative survival function $r(\cdot|z)$ is defined as

$$r(t|z) = \frac{S(t|z)}{S_P(t|z)}. \tag{3.2}$$

In words, the relative survival function $r(\cdot|z)$ corresponds to the ratio of the survival function of the studied group $S(\cdot|z)$ to the survival function of a comparable group (i.e., with the same characteristics) from the general population $S_P(\cdot|z)$ [25].

In (3.1), $\lambda_P(\cdot|z)$ is usually estimated on the basis of external data such as population life tables, which are usually stratified according to the main factors affecting patient survival such as age, gender, and calendar year. As population life tables take into account all deaths, those due to the cancer of interest are thus also included. However, it is assumed that this does not influence the estimated ratio as mortality for a given cancer represents only a small fraction of the overall mortality. Correcting for this over-representation of the cancer being studied has, in practice, an insignificant effect on estimates of expected survival [30]. Oksanen [47] shows that this holds even for common cancers such as prostate cancer.

Net survival is a measure of patient survival corrected for the effect of other causes of death [25]. It represents the (hypothetical) survival that would be observed if the only possible cause of death was the disease of interest [6, 60]. If we are in a situation where we only have data from (a sample of) our sub-population of interest (e.g., cancer patients) but we have information on the cause of death, then estimation methods for net survival, sometimes referred then as marginal survival [see Geskus 33], can be estimated in a competing risks framework. However, unbiased estimate of this net or marginal (cause-specific) survival can only be obtained if one can assume independence of the censoring due to death from other causes [40, 60]. On the other hand, net survival can also be estimated when the cause of death is unknown by making use of information from the general population. In this different framework, net survival can then be estimated using the relative survival method. Since the BCR does not collect information on the cause of death, we are in this latter situation. The net survival function is thus derived from the excess mortality hazard $\lambda_E(\cdot|z)$. Formally, it is defined as

$$S_n(t|z) = \exp\left(- \int_0^t \lambda_E(u|z)du \right). \tag{3.3}$$

Depending on $\lambda_E(\cdot|z)$, $S_n(\cdot|z)$ may be a proper survival function but this is not necessarily the case.

There are several approaches to estimate net survival of a cohort of patients in a relative survival framework. Danieli et al. [24] showed that only two of them provide unbiased estimates of net survival: (i) the non-parametric Perme et al. [52] estimator and (ii) the excess risk based on an adjusted modeling on the demographic variables of the life tables. We used the non-parametric Perme et al. [52] estimator to estimate net survival as recommended by Danieli et al. [24] for population based studies.

Net survival probabilities by gender and site are illustrated in Fig. 2, while net survival by age group and site is displayed in Fig. 3. Numerical values are listed in Table 3. The net survival functions reach a plateau for both sites and genders, which is indicative for 'cure' of cancer. The estimated survival curves staying practically constant after 5 years since diagnosis indicates that the excess hazard of dying compared to the general population becomes negligible after only a few years after diagnosis. This suggests that a waiting period of moderate length would be enough to apply standard life insurance rates for the cancers under consideration.

Net survival by site and age group (both sexes combined, Fig. 3) follows a quite expected path for melanoma cancer patients (left panel): younger patients have a better net survival compared to older patients. In particular, the 5-year and 10-year net survival probabilities for patients aged 20–34 were respectively 0.959 (with 95% confidence interval [0.951–0.968]) and 0.943 (with 95% confidence interval [0.932–0.955]), meaning that for this age group (which is typically the age at which one starts a loan), patients' survival is close to that of the general population. Net survival by age group for thyroid cancer patients (right panel of Fig. 3) yields even more promising results in the context of mortgage loans as the age group which is most likely to subscribe to such financial products (20–34 years) has a 5-year and

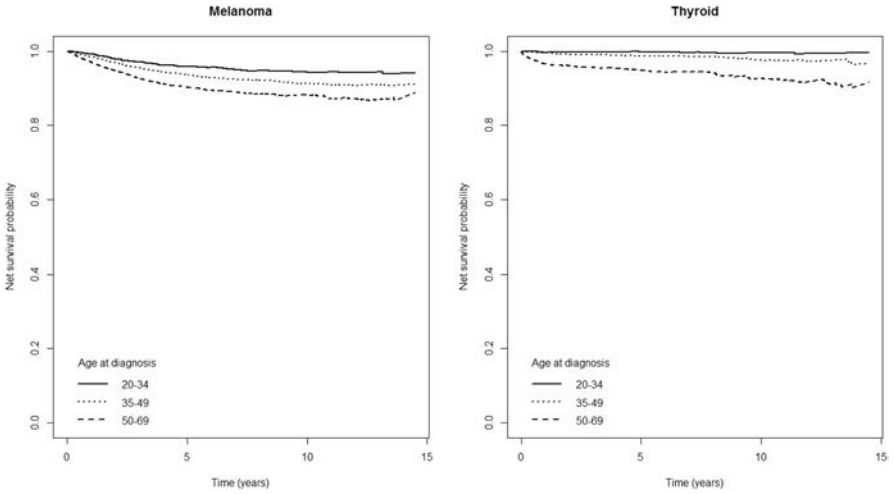

Fig. 3 Net survival by age group and site using the non-parametric Perme et al. [52] estimator

Table 3 Net survival probabilities by gender, site and age group using the non-parametric [52] estimator at 5 and 10 years after diagnosis, $\widehat{S}_n(t = 5)$ and $\widehat{S}_n(t = 10)$, with their confidence intervals at 95% level

Gender	Cancer site	Age at diagnosis	$\widehat{S}_n(t = 5)$	$CI_{95\%}$	$\widehat{S}_n(t = 10)$	$CI_{95\%}$
Women	Melanoma	20–34	0.974	(0.966–0.982)	0.963	(0.951–0.974)
		35–49	0.956	(0.949–0.963)	0.936	(0.925–0.946)
		50–69	0.927	(0.918–0.936)	0.907	(0.893–0.922)
	Thyroid	20-34	0.998	(0.995–1.000)	0.997	(0.990–1.000)
		35–49	0.992	(0.987–0.997)	0.990	(0.981–0.998)
		50–69	0.958	(0.948–0.968)	0.941	(0.924–0.959)
Men	Melanoma	20–34	0.921	(0.901–0.942)	0.894	(0.868-0.921)
		35–49	0.899	(0.886–0.913)	0.871	(0.853–0.890)
		50–69	0.871	(0.859–0.884)	0.849	(0.828–0.870)
	Thyroid	20-34	0.997	(0.987–1.010)	0.989	(0.968–1.010)
		35–49	0.972	(0.957–0.987)	0.932	(0.901–0.964)
		50–69	0.928	(0.908–0.949)	0.895	(0.858–0.933)

10-year net survival probabilities of respectively 0.998 (with 95% confidence interval [0.995–1) and 0.996 (with 95% confidence interval [0.989–1]).

Net survival by gender, site and age group (Table 3) indicates a highly favorable outcome for women and men diagnosed with thyroid cancer and aged 20–34. For both subgroups, the 95% confidence interval for net survival at 10 years after diagnosis indicate that a net survival equal to that of the general population cannot be rejected at the 5% significance level.

Regarding the gap between women and men for the net survival curves (in particular for melanoma cancer patients), we do not have a clear explanation for these

differences in survival probability. However, although stage distribution is very similar across gender, The Belgian Cancer Registry [64], in agreement with Balch et al. [4], showed that compared to men, females have more often melanoma on the arms or legs which has a better prognosis. This partly explains the difference in survival for melanoma cancer patients.

Using data provided by the French network of cancer registries (FRANCIM), Boussari et al. [11] found very similar results for thyroid cancer patients diagnosed between 1995 and 2010. They obtained a 10-year net survival of 0.99 (95% CI [0.99–1.00]) for women aged 15–45, and 0.98 (95% CI [0.96–0.99]) for men of the same age group. Their results are also very close to ours when comparing patients aged 45–55: a 10-year net survival of 0.99 (95% CI [0.98–1.00]) for women, and 0.92 (95% CI [0.88–0.95]) for men. In another study including melanoma cancer patients diagnosed between 1989 and 2004 by the French registries, Jooste et al. [37] obtained a 10-year net survival for men and women aged 15-45 of 0.81 (95% CI [0.78–0.84]) and 0.91 (95% CI [0.89–0.93]), respectively.

3.3 Proportional excess hazards

Previous sections suggested that some patients actually have a survival comparable to that of the general population. In the following sections, we confirm these findings from the point of view of the excess hazard.

Esteve et al. [29] proposed a maximum likelihood method for estimating net survival via the modeling of the excess hazard. The excess hazard $\lambda_E(t|z)$ to be estimated is represented as

$$\log(\lambda_E(t|z)) = (\boldsymbol{\beta}^\top z) + \log\left(\sum_{k=1}^{m} \tau_k I_k(t)\right) \tag{3.4}$$

where $\boldsymbol{\beta}$ is the log hazard ratio corresponding to the covariates, $I_k(t)$ the indicator function for the k^{th} interval (after splitting the follow-up time into short time intervals) and τ_k the net baseline hazard rate in that interval for patients with $z = \mathbf{0}$.

A maximum likelihood approach to estimate parameters of model (3.4) is available in the `relsurv` package [51, 54] in R [56]. Fitting two models to our data (one for each cancer site) with age group (50–69 years old taken as the reference category) and follow-up (with intervals of 5 years) as covariates, we obtain the results presented in Tables 4 and 5. The coefficient estimates $\widehat{\beta}$ and their standard errors are displayed in the second column, p-values are reported in the last column. Note that the significance is largely impacted by the large sample size.

For both cancer sites, being a woman and being younger at the time of diagnosis are good prognostic factors.

3.4 Flexible parametric model

Model (3.4) is based on the assumption of proportional excess hazards, which constrains the hazard ratio to be constant over the follow-up time [34]. Nonetheless, in

Table 4 Results of model (3.4) fitted to melanoma cancer data

Covariates	$\widehat{\beta}$ (S.E.)	p-value
Gender Women	− 0.759 (0.060)	< 0.001
Agegr 20–34	− 0.741 (0.105)	< 0.001
Agegr 35–49	− 0.349 (0.064)	< 0.001
fu [0,5)	− 3.507 (0.045)	< 0.001
fu [5,10)	− 4.590 (0.105)	< 0.001
fu [10,14]	− 5.334 (0.420)	< 0.001

Table 5 Results of model (3.4) fitted to thyroid cancer data

Covariates	$\widehat{\beta}$ (S.E.)	p-value
Gender Women	− 0.754 (0.169)	< 0.001
Agegr 20–34	− 2.864 (0.703)	< 0.001
Agegr 35–49	− 1.252 (0.213)	< 0.001
fu [0,5)	− 4.086 (0.129)	< 0.001
fu [5,10)	− 5.089 (0.314)	< 0.001
fu [10,14]	− 5.276 (0.867)	< 0.001

cancer survival, the effects of prognostic factors often vary with time since diagnosis [25, 55] and it is well known that the linearity assumption of covariates may be too strong and not always verified in practice [46]. As an example with the variable age, the effect on hazards of an increase of one year is often different for patients aged 18 or 60. This is why Remontet et al. [57] and Fauvernier et al. [31, 32] extended the model proposed by Esteve et al. [29] to account for non-linear and non-proportional effects of covariates. This model also allows (i) a flexible modeling of the baseline hazard and (ii) a flexible interaction between several covariates adopting a multidimensional penalized splines approach. This leads to the specification

$$\log(\lambda_E(t|z)) = \sum_{j=1}^{J} g_j(t,z) \tag{3.5}$$

where $g_j(\cdot,\cdot)$ are uni- or multidimensional penalized spline functions and each function $g_j(\cdot,\cdot)$ can be the marginal basis of time, the marginal basis of a covariate or the tensor product of the marginal bases of any number of elements of (t,z) [32]. This model has the advantage that the splines bring the flexibility needed for modeling the hazard and inclusion of penalty terms allows to control this flexibility for smooth estimation [as suggested by Eilers and Marx 27].

Several flexible models to estimate excess hazard were considered in this paper:

- baseline hazard only (BH) model:
  Equation (3.4) without the $(\boldsymbol{\beta}^\top z)$ term, as it considers an excess hazard in each interval and no other covariates
- linear and proportional hazard (LPH) model:

$$\log(\lambda_E(t|z)) = f(t) + age$$

- linear and non-proportional hazard (LNPH) model:
$$\log(\lambda_E(t|z)) = f(t) + age + g(t) \cdot age$$
- non-linear and proportional hazard (NLPH) model:
$$\log(\lambda_E(t|z)) = f(t) + g(age)$$
- non-linear and non-proportional hazard (NLNPH) model:
$$\log(\lambda_E(t|z)) = f(t) + g(age) + g(t) \cdot age$$

with $f(t)$ the flexible parametric function for the baseline/reference hazard as a function of time and $g(z)$ the (non-)linear function of the covariates.

We fitted all these models and compared them based on a likelihood ratio test. The remaining of this section is based on the NLNPH model, deemed the best one according to likelihood ratio test. Excess hazard, assuming non-linear and non-proportional hazard for age at diagnosis, are estimated using the `flexrsurv` package [18] in R.

As previously suggested, for both cancer sites, excess mortality hazard increases with age at diagnosis and decreases with time since diagnosis (Fig. 4). For thyroid cancer patients, excess hazard at the time of diagnosis is approximately 0.15 for patients aged 65 and is close to 0 for patients aged 40 and below. From 4 to 10 years after diagnosis, excess hazard remains constant and is close to 0 for all ages. Beyond 10 years after diagnosis, excess hazard slightly bends up but remains small. This observation can also be an artefact of the spline function, which is "unbounded" at the end (a higher order degree for the spline function can produce this upwards bend). For melanoma cancer patients, excess hazard at the time of diagnosis is approximately 0.035 for patients aged 65 and is close to 0 for the youngest patients. For all ages, excess hazard peaks at 1 year

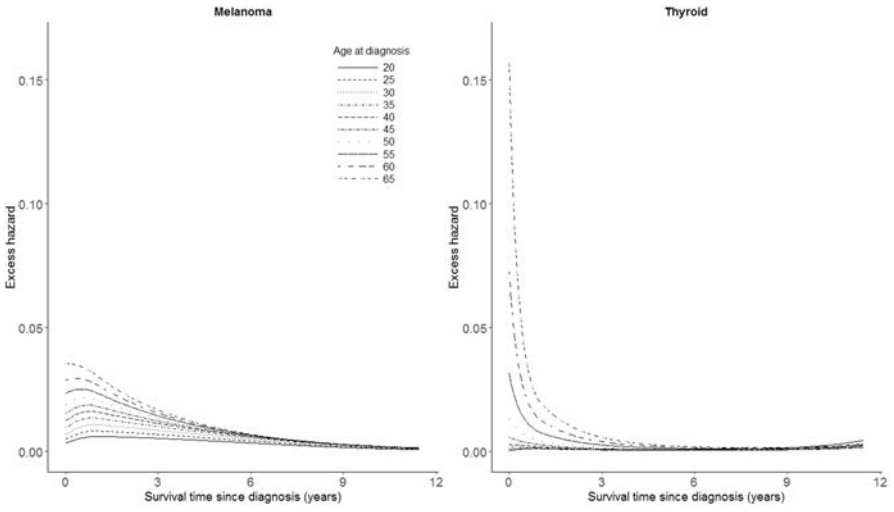

Fig. 4 Excess hazard by age and cancer site estimated with a non-linear and non-proportional hazard model

after diagnosis before decreasing until it becomes negligible, around 8 years after diagnosis.

3.5 Cure models

Cure models are a specific class of survival models which assume that a fraction of the subjects will never develop the event of interest, here death due to cancer. Such models have been used in different fields such as economics (e.g., time until an unemployed person finds a new job), engineering (e.g., time until a machine or device fails), finance (e.g., time until a bank goes bankrupt), marketing (e.g., time until a client buys a new product), for instance. In our context, cure models can be used to determine cancer patients who are considered as "long-term survivors" and those who are not [44, 48]. The long-term survivors are still often referred to as "cured subjects" in the literature as a consequence of the name of this class of models. Cure models are particularly suitable for some cancer sites because if the treatment is successful, the patient will never suffer a relapse of the disease. It is also particularly suitable for less aggressive cancers such as, among others, the two considered in this paper because a considerable percentage of subjects exhibits long-term survival.

There are two main types of cure models in the literature: (i) mixture cure models (the most common type, based on the seminal work by Boag and Berkson and Gage [7, 9]) and (ii) non-mixture cure models [66, 16, 65]; see Amico and Van Keilegom [1] for a recent overview. In the present paper, we consider only the family of mixture cure models. In this approach, the patient population is considered as a mix of two types of patients, that is, long-term survivors who will never die of their cancer and the uncured patients who, if not censored, will die of their cancer [39]. In the global survival setting, the mixture cure model is specified as follows:

$$S(t) = \pi + (1 - \pi)S_u(t),$$

where π is the proportion of patients that are long-term survivors and $S_u(\cdot)$ the survival function of the uncured population. Both π and $S_u(\cdot)$ can then be modeled to depend on covariates. Cure models can be a useful alternative to standard survival models for cancers with a strong medical evidence and a confirmation in the data for the presence of long-term survivors [41]. In the relative survival setting, cure models also allow to determine the proportion of statistically cured cases and survival time of the fatal cases [63].

The estimated proportion of cured cases and mean survival time of fatal cases for the two cancer sites considered (melanoma and thyroid), using a mixture cure model are given in Table 6. Note that NAs for young age groups are due to an insufficient number of cases.

The estimated cured proportion ranges from 84.99% ($CI_{95\%}$ [84.65, 85.33]) for men with melanoma cancer aged 50–69 to 99.86% ($CI_{95\%}$ [99.82, 99.90]) for women with thyroid cancer aged 20-34. The estimated mean survival time of fatal cases ranges from 0.47 years ($CI_{95\%}$ [− 0.18, 1.12]) for women with thyroid cancer aged 20–34 to 10.57 years ($CI_{95\%}$ [1.32, 19.81]) for men with thyroid cancer aged 35–49.

Table 6 Estimated cured fractions (in %) and mean survival time (in year) for the fatal cases by gender, site and age group, with their 95% confidence intervals

Gender	Cancer site	Age at diagnosis	Estimated cured fraction	$CI_{95\%}$	Estimated mean T	$CI_{95\%}$
Women	Melanoma	20–34	96.53	(96.33, 96.72)	3.81	(3.46, 4.22)
		35–49	93.28	(92.94, 93.63)	4.49	(4.04, 5.03)
		50–69	90.01	(89.49, 90.53)	3.93	(3.50, 4.46)
	Thyroid	20–34	99.86	(99.82, 99.90)	0.47	(− 0.18, 1.12)
		35–49	99.03	(98.82, 99.23)	4.87	(3.03, 6.71)
		50–69	95.73	(95.39, 96.07)	1.40	(1.02, 1.78)
Men	Melanoma	20–34	88.77	(88.21, 89.32)	3.54	(3.12, 4.07)
		35–49	86.28	(85.77, 86.78)	3.83	(3.52, 4.20)
		50–69	84.99	(84.65, 85.33)	2.94	(2.79, 3.10)
	Thyroid	20–34	NA	NA	NA	NA
		35–49	91.96	(86.99, 96.93)	10.57	(1.32, 19.81)
		50–69	92.87	(92.34, 93.41)	0.96	(0.67, 1.25)

Estimated mean T = Estimated mean survival time of fatal cases. NAs for young age groups are due to an insufficient number of cases

Moreover, higher age at diagnosis is correlated with lower cured proportions, except for men with thyroid cancer from the oldest age group.

Using data on 818,902 Italian cancer patients diagnosed between 1985 and 2005, Dal Maso et al. [23] also found encouraging results for thyroid cancer patients, with an estimated cured fraction of 99% and 95% for women and men aged 15–45, respectively. Regarding melanoma cancer patients, the results on Italian data are somewhat lower than the estimations presented in this work, with a cured proportion of 85% and 77% for women and men aged 15–45, respectively. The fact that we are using a more recent incidence period may partly explain these improved results.

4 Time-to-cure

The time-to-cure (TTC) is generally referred as the time after which patients can be considered as long-term survivors. Different approaches to define TTC have emerged in the literature. We highlight three of them. Firstly, [15] defined TTC_C as the time at which "almost" (that is, $1 - \epsilon$, with ϵ small enough usually ranging from 0.1 to 0.01) all uncured patients would have died. From that time onwards, the number of deaths attributable to the cancer of interest becomes negligible.

Secondly, Dal Maso et al. [23] define TTC_D as the shortest time after diagnosis at which the 5-year conditional net survival (defined as the ratio between net survival at time $t + 5$ years and net survival at time t) is close to 1.

Thirdly, Boussari et al. [11] defined TTC_B as the shortest time from which the conditional probability of being cured at a given time t after diagnosis knowing

that the patient was alive up to time t is close to 1, that is TTC_B is the smallest value of t such that, for some given (small) value of ϵ

$$\frac{\pi}{S(t)} = \frac{\pi}{\pi + (1-\pi)S_u(t)} \geq 1 - \epsilon, \tag{4.1}$$

where π is the proportion of cured patients and is estimated from the relative survival with the hypothesis of cure.

The main advantage of TTC is that, although the results may depend on the chosen definition, it is a simple and straightforward indicator to set the time after which a patient who had cancer should not be penalized anymore when applying for mortgage insurance.

In this paper, we focus on TTC_B for several reasons. First, TTC_B depends on both the cure proportion and the survival of the uncured, so it is less influenced by high early excess mortality. Second, TTC_B has the advantage of being an increasing function of time, therefore, it is not sensitive to a temporary plateau effect (that is, net survival curve flattening before decreasing again). Third, estimating TTC_B requires a 5-year shorter follow-up than TTC_D, which is a clear advantage of TTC_B over TTC_D [11].

Time-to-cure TTC_B was estimated using the `rstpm2` package [17] in `R`. Table 7 presents estimated TTC_B (with $\epsilon = 0.05$ and $\epsilon = 0.01$) in years and the cure proportion in percentage for each subgroup.

In the following we interpret only TTC_B with $\epsilon = 0.05$. Results of TTC_B with $\epsilon = 0.01$ are still presented for the sake of comparison and completeness. From Table 7 we see that while TTC_B increases with age, ranging from a few days (0.01 year) for the youngest age groups to almost 3.5 years for the oldest age group of male melanoma cancer patients, cure proportion decreases with age, ranging from 99.75% for women aged 20–34 with thyroid cancer to 84.77% for men aged 50–69 with melanoma cancer. With this approach, the subgroups that stand out the most are the ones with a small TTC_B (and confidence intervals as narrow as possible to decrease uncertainty as much as possible) and a large proportion of long-term survivors. Among all subgroups considered, this is the case especially for women with melanoma cancer aged 20–34 ($\widehat{\text{TTC}}_B = 0.01$ with 95% confidence interval [0.00–1.18] and estimated cure proportion = 96.75%).

Figure 5 illustrates the cure proportion obtained from a cure model and the TTC_B ($\epsilon = 0.05$) by age group and gender for both cancer sites. We can easily classify points into two clusters; the ones in the upper left corner with the best possible outcomes in terms of TTC_B and cure proportion and the others. Among melanoma cancer patients, women aged under 50 belong to this group with favorable outcomes, whereas for thyroid cancer patients, women of all ages and men aged under 50 belong to this group.

Similar results for thyroid cancer patients have been obtained by Boussari et al. [11] on FRANCIM data. They also found that women aged 15–65 and men aged 15–45 have highly favorable outcomes, that is, a cure proportion close to 100% and a TTC_B close to 0.

Table 7 Estimated value of time-to-cure ($\widehat{\mathrm{TTC}}_B$ in years, with $\epsilon = 0.05$ and $\epsilon = 0.01$) together with 95% confidence intervals and cure proportion (in %) by cancer site, sex and age group

Gender	Cancer site	Age at diagnosis	$\widehat{\mathrm{TTC}}_B\ \epsilon = 0.05$	$CI_{95\%}\ \epsilon = 0.05$	$\widehat{\mathrm{TTC}}_B\ \epsilon = 0.01$	$CI_{95\%}\ \epsilon = 0.01$	Estimated cure proportion
Women	Melanoma	20–34	0.01	(0.00–1.18)	4.32	(3.30–5.35)	96.75
		35–49	0.81	(0.21–1.42)	5.65	(4.82–6.49)	94.16
		50–69	2.56	(2.04–3.08)	6.39	(5.66–7.12)	90.87
	Thyroid	20-34	0.01	(0.00–23.42)	0.01	(0.00–23.42)	99.75
		35–49	0.01	(0.00–14.45)	0.01	(0.00–14.45)	99.35
		50–69	0.01	(0.00–5.10)	4.82	(3.78–5.86)	94.93
Men	Melanoma	20–34	2.38	(1.92–2.85)	5.64	(4.90—6.39)	89.69
		35–49	2.78	(2.33–3.23)	5.86	(5.13–6.59)	88.12
		50–69	3.42	(3.00–3.85)	6.21	(5.51–6.90)	84.77
	Thyroid	20–34	0.01	(0.00–28.97)	0.11	(0.00–0.78)	98.88
		35–49	0.01	(0.00–16.03)	4.66	(3.19–6.13)	96.46
		50–69	0.91	(0.51–1.31)	7.14	(6.40–7.87)	90.84

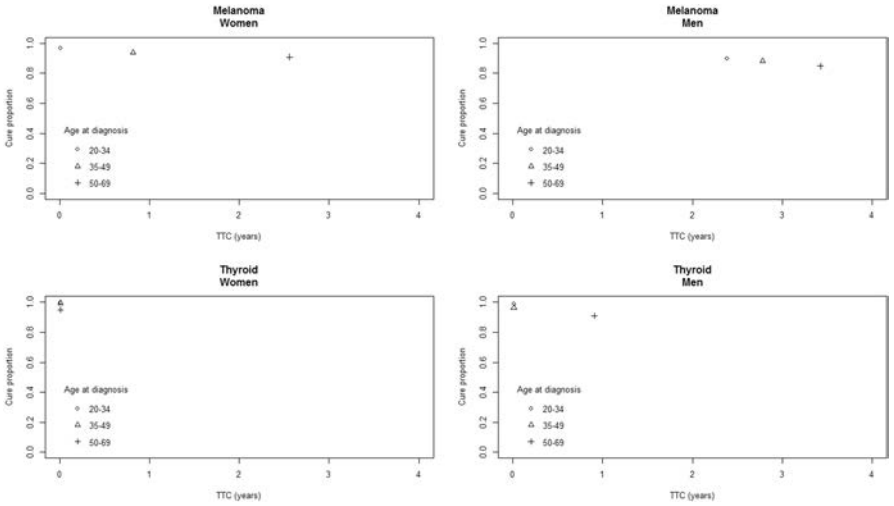

Fig. 5 Cure proportion versus time-to-cure (in years, with $\epsilon = 0.05$) by gender and age group at diagnosis in patients diagnosed with melanoma and thyroid cancer

5 Application to mortgage insurance

All results obtained so far suggest that, for melanoma and thyroid cancer patients, excess mortality becomes negligible after some waiting period. In this section, we determine the length of such a waiting period as the time needed to get back to standard premium rates. Henceforth, standard rates correspond to premiums computed according to life tables commonly used on the Belgian market:

- regulatory life table XK applying to insurance products comprising benefits in case of death (formally, XK defines minimum premium amount for policies with a positive sum at risk). This life table is conservative and generates a relatively high safety loading.
- experience market life table published by the National Bank of Belgium (NBB). These life tables reflect the mortality observed on the market, within portfolios of companies controlled by NBB. There is no safety loading and insurers are only allowed to apply premium rates resulting from NBB tables for relatively short periods of time (rates are subject to revision in case the observed mortality on the market changes over time).

We also include premium rates calculated from general population life table published by Statbel, for the sake of comparison. Term life premiums are smaller for NBB life tables and larger for XK life table, Statbel life tables falling in between because of the higher socio-economic profile of the insured population. Premium rates for cancer patients are computed from the excess mortality hazards estimated with the help of the flexible parametric model discussed in Sect. 3.4, according to the time elapsed since diagnosis.

Consider a mortgage insurance applicant aged x borrowing an amount of 100,000 at interest rate 2% for a duration δ. At time t, the amount of the loan that has not been amortized is denoted as $c(t)$. This loan is secured by mortgage insurance, repaying the lender the amount $c(t)$ in case the policyholder dies at time t. The expected present value (EPV) of benefits paid in case of death is thus equal to

$$\text{EPV} = \int_0^{\delta} c(t)v(0,t)_t p_x \mu_{x+t} dt \tag{5.1}$$

where $c(t)$ is the amount of benefit in case of death at time t (which is in our case, a decreasing sum insured corresponding to the amount of the loan not yet amortized), $v(0,t)$ is the present value at time 0 of a unit payment made at time t (the discount factor), ${}_tp_x$ is the t-year survival probability for a policyholder aged x and μ_{x+t} is the force of mortality at attained age $x+t$.

In case the applicant suffered from cancer, we have to relate ${}_tp_x$ and μ_{x+t} entering the formula to the survival function and hazard rate estimated in the preceding sections. To this end, we assume that the applicant aged x has been diagnosed with cancer at age $x-w$. We then have

$$\begin{aligned} {}_tp_x &= \frac{S(w+t|\text{age at diagnosis} = x-w)}{S(w|\text{age at diagnosis} = x-w)} \\ \mu_{x+t} &= \lambda(w+t|\text{age at diagnosis} = x-w). \end{aligned}$$

with $S(.)$ and $\lambda(.)$ being estimated from the NLNPH model (eq (3.5)) discussed in Sect. 3.4.

This allows us to compute EPV(w) according to (5.1) for each candidate waiting period w after diagnosis and to select the smallest w such that EPV(w) becomes close to the value computed from the XK life table (assuming that cancer patients are priced with the regulatory life table XK and that the market can absorb the extra mortality burden corresponding to the difference between XK and NBB life tables).

Consider a home loan of duration 20 years. A cancer patient aged 30 and another aged 50 apply for mortgage insurance, with technical interest rate of 1% and a term of 20 years. These characteristics have been chosen as they represent a rather standard setting and other scenarios revealed similar patterns when considering different ages between 18 and 50. EPV based on XK, NBB and Statbel life tables have been computed to compare it with EPV(w) for a cancer patient diagnosed with melanoma or thyroid cancer at ages 30-w and 50-w. Results are illustrated in Fig. 6. Notice that EPV are presented for both genders combined, as XK life table applies to both sexes and insurance companies operating in the European Union are not allowed to account for gender in pricing.

Results show that EPV of a 30-year-old patient approaches EPV based on XK life table about 9 years after diagnosis for melanoma cancer patients and after slightly less than 1 year for thyroid cancer patients. For a 30-year-old patient with melanoma cancer, the EPV never reaches within our time period the EPV based on NBB and Statbel life tables. For a 30-year-old patient with thyroid cancer, the EPV goes below the lowest EPV (based on NBB life table) as early as about

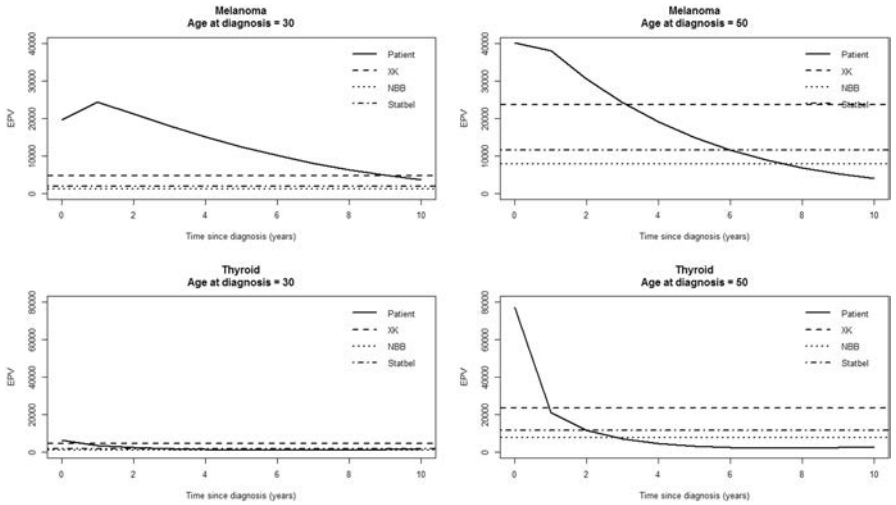

Fig. 6 Expected present value (EPV) of a life insurance contracted by a 30 and 50-year-old cancer patient for a period of 20 years with interest of 1 percent and benefit of 100,000. Horizontal lines correspond to EPV calculated with XK, NBB and Statbel life tables

4 years after diagnosis. For a 50-year-old patient with melanoma cancer, EPV reaches the same level than the one based on XK life table 3 years after diagnosis, and reaches the lowest level (based on NBB life table) less than 8 years after diagnosis. Finally, for a 50-year-old patient with thyroid cancer, EPV reaches XK level 1 year after diagnosis, then stays below the lowest EPV 3 years after diagnosis. This is in line with the reduced waiting periods published in the Royal decree on June 14, 2019, where a duration of 3 years applies in this case. There is however a fundamental difference in the approach because the waiting period starts at diagnosis in the present study whereas it starts at the end of the treatment protocol according to the law.

The improvement for melanoma cancer patients is not as substantial as for thyroid cancer patients since the time period is close to 9 and 3 years for 30 and 50-year-old patients, respectively. Nonetheless, it remains advantageous for patients since the 9 and 3-year period starts from the date of diagnosis and not at the end of the therapeutic treatment. Note that, at first glance there seems to be an advantage for older patients with shorter waiting time, while all the other results seem rather to indicate an advantage for younger patients with shorter time to cure and higher cure rate. This actually comes from the fact that in absolute terms, younger patients have shorter time to cure and thus lower EPV than older patients. As shown in Fig. 6, EPV for a 30-year-old patient is lower than for a 50-year-old patient regardless of the time since diagnosis and for both cancer sites. However, financial burdens (i.e., XK, NBB and Statbel levels) are much lower for younger people than for older people (since young people from the general population have a lower probability of dying than older people).

6 Discussion

Results derived in this paper are in line with the reduced waiting period specified in the Belgian legislation. Furthermore, results are also in line with the reference grid used in France (convention AERAS) as the time after which patients have access to the right to be forgotten according to this convention is relatively short (maximum 6 years after the end of the therapeutic protocol for the two cancers considered in this paper).

All analyzes in this paper are based on the time since diagnosis although the right to be forgotten implemented in Belgium and France is applicable after a certain time after the end of the therapeutic protocol. For the sake of clarity and easiness, an approach based on the time since diagnosis would undeniably be more favorable to patients. A right to be forgotten based on the date of diagnosis would indeed allow patients to know when exactly they can expect to benefit from this right. On the contrary, with the current approach based on treatment end date (which is unknown until the success of the treatment and can even change later in case of relapse), benefiting from this right is subject to a high level of uncertainty as durations of treatments are heterogeneous and unpredictable even within same cancer types and stages. Therefore, patients cannot currently easily estimate when (and if) they will be able to benefit from this right.

One of the main results of this paper is thus to promote the use of the date of diagnostic instead of the end of the therapeutic treatment for defining the waiting period. However, one could argue that the length of medical treatments may have significantly reduced over the period 2004–2018 and if this is the case, both approaches could be now closer than expected. While it could have been interesting to formally compare both approaches, individual data on the type and length of treatment for each case is not reported in the Belgian Cancer Registry and such information is not readily available. Furthermore, the definition of the end of the treatment is in itself debatable and the duration of the treatment can be quite different depending on several factors, which is actually an argument in favor of considering the date of diagnosis. Moreover, durations of treatment are heterogeneous even within the same cancer type, usually unpredictable, and optimal durations are often still open to debates [61]. In any case, a reduction in treatment length due to the progress made in medical treatment of cancer would obviously lead to closer agreement between the two approaches. Since the date of diagnosis, as recorded in national registries, offers the advantage to not be subject to any discussion and to allow the patient to know from the start when the waiting period will end, we think that all parties (actuaries included) will benefit from using the date of diagnosis instead of the end of treatment for more convenience and less uncertainty.

Contrarily to other studies [like Yue et al. 67], calendar time has not been included in the analysis conducted in the present paper, because of the limited amount of cases available (recall that we concentrate on younger ages because of the product under consideration, targeting young adults). That being said, we performed two subanalyses: one for the cohort 2004–2011 and one for the cohort 2012–2018 and compared the results. There are no real changes, so we conclude

that there is no cohort effect. Moreover, simplicity of the system is crucial and given that medical treatments keep improving, the resulting bias of ignoring a potential cohort effect favors insurance providers. Notice that the approaches are dependent on factors such as changing diagnostic criteria and improved diagnostic methods. As these factors may vary over time irrespective of any improvement of the treatment and are different between populations, one can not compare excess risks across different time periods and populations [43]. To illustrate this, suppose that a medical advance allows a cancer to be diagnosed at a less severe stage (cases that are not as fatal as the ones detected with the previous methods) and perhaps also earlier with the consequence that more cases are detected (cases that would not have been detected with the previous detection methods are now detectable). These improvements will yield an increased survival rate, regardless of whether the treatment improved or not. This weakness of the survival rate has been pointed out in the literature extensively [3, 28, among others]. However, although not necessarily the case, earlier diagnoses will in most cases be associated with better efficacy of the treatment.

The melanoma and thyroid cancers may include a variety of types and could be diagnosed at different stages of severity. Moreover, significant gaps are observed between women and men. It is undeniable that including the information on stages of severity and gender in the NLNPH model would refine the analysis. However, these have been ignored on purpose, considering the specific application of the results for insurance practice. Since December 21, 2012, the European directive on equality between men and women also applies to outstanding balance insurance. The judgment of the European Court of Justice indeed stipulates that no discrimination can be made between men and women when establishing such insurance contract, so the gender no longer has an influence on the premiums nor on the coverage conditions of outstanding balance insurances. Therefore, the gender has been omitted not only for the sake of simplicity, but also because it is illegal for insurance companies to use that information. Concerning the stage of severity, it has also been omitted to ensure the simplicity and thus the legal safety of the system. Moreover, not including the stage may bias the analysis but not necessarily in favor of patients diagnosed at an advanced stage (perhaps a shorter waiting period would be indicated for them). In any case, if the method amounts to determining the time to wait before mortality returns to normal, ignoring the stage is actually a conservative approach for the insurer.

To ensure that the coverage cost of cancer patients remains acceptable for the insurance industry, further constrains may be imposed in terms of sum insured, for instance. Also, it could be reasonable to impose that insurers charge a single premium for mortgage insurance to mitigate mortality risk. Last but not least, compensation could be performed at market level to avoid that some insurers face higher costs. Indeed, even if the coverage of mortality risk becomes affordable after a relatively short waiting period, premium rates reflecting the actual mortality of cancer survivors remain sometimes above the market premiums resulting from NBB life tables. This means that when it is the case, this extra cost must be fairly distributed among stakeholders: cancer patients, insurance industry, banking sector (as they sell the loans) and society as a whole.

Notice that cancer stage at diagnosis has not been taken into account in the present study. For patients with more advanced cancer stage, the waiting period will be more conservative because mortality will peak just after diagnosis and before reverting back to the general population level. Note also that *in situ* cancer cases (considered as pre-cancer) have not been included in the present study as they are not classified as cancer per se [14] like any other "regular" cancer cases which were not diagnosed as in situ before. Moreover, cancer is not one disease, but a family of many diverse diseases with different outcomes. Results in the present paper focus on melanoma and thyroid cancer patients, and cannot be applied to other cancer types. A natural extension of this work would be to repeat the analyses for all major cancer types. This would certainly be useful for implementing appropriate market rules but goes beyond the scope of this study which primarily aims to advocate a waiting period starting at diagnosis.

Cancer patient survival has improved over the last few decades, with an increasing proportion of patients being cured for many types of cancer [2, 39]. Providing coverage in case cancer is diagnosed or to long-term cancer survivors is therefore of prime importance, for the society but also for the insurance industry since proper coverage of such risks may well produce attractive returns.

7 Software

Extraction of the data and the surface smoothing for Belgian mortality rates were carried out using SAS version 9.3 (SAS Institute, Cary, North Carolina, USA). All other analyzes and figures were performed with R version 3.6.1 (2019-07-05) [56].

Acknowledgements This work was funded by a UCLouvain grant. The three first authors express their sincere gratitude to the university for making such a research project possible. The support of the Belgian Cancer Registry is gratefully acknowledged for providing access to the data and for research assistance. The authors also thank the two anonymous referees for comments that have been very helpful for revising previous versions of the present work.

References

1. Amico M, Van Keilegom I (2018) Cure models in survival analysis. Annu Rev Stat Appl 5:311–342
2. Andersson TM, Dickman PW, Eloranta S, Lambert PC (2011) Estimating and modelling cure in population-based cancer studies within the framework of flexible parametric survival models. BMC Med Res Methodol 11(1):96
3. Bailar JC III, Smith EM (1986) Progress against cancer? N Engl J Med 314(19):1226–1232
4. Balch CM, Soong S-J, Gershenwald JE, Thompson JF, Reintgen DS, Cascinelli N, Urist M, McMasters KM, Ross MI, Kirkwood JM et al (2001) Prognostic factors analysis of 17,600 melanoma patients: validation of the american joint committee on cancer melanoma staging system. J Clin Oncol 19(16):3622–3634
5. Belot A, Ndiaye A, Luque-Fernandez M-A, Kipourou D-K, Maringe C, Rubio FJ, Rachet B (2019) Summarizing and communicating on survival data according to the audience: a tutorial on different measures illustrated with population-based cancer registry data. Clin Epidemiol 11:53–65
6. Berkson J, Gage RP (1950) Calculation of survival rates for cancer. In: Proceedings of the staff meetings. Mayo Clinic, vol 25, p 270

7. Berkson J, Gage RP (1952) Survival curve for cancer patients following treatment. J Am Stat Assoc 47(259):501–515
8. Berrino F, Capocaccia R, Estève J, Gatta G, Hakulinen T, Micheli A, Sant M, Verdecchia A (1999) Survival of cancer patients in europe: the eurocare-2 study. In: Survival of cancer patients in Europe: the EUROCARE-2 study, pp 1–CD
9. Boag JW (1949) Maximum likelihood estimates of the proportion of patients cured by cancer therapy. J R Stat Soc Ser B (Methodological) 11(1):15–53
10. Bolard P, Quantin C, Esteve J, Faivre J, Abrahamowicz M (2001) Modelling time-dependent hazard ratios in relative survival: application to colon cancer. J Clin Epidemiol 54(10):986–996
11. Boussari O, Romain G, Remontet L, Bossard N, Mounier M, Bouvier A-M, Binquet C, Colonna M, Jooste V (2018) A new approach to estimate time-to-cure from cancer registries data. Cancer Epidemiol 53:72–80
12. Buckley J (1984) Additive and multiplicative models for relative survival rates. Biometrics 40:51–62
13. Bureau du suivi de la tarification (2018) Rapport sur l'activité 2017. https://www.bureaudusuivi.be/images/docs/RapportAnnuel_2017.pdf
14. Chang AE, Ganz PA, Hayes DF, Kinsella T, Pass HI, Schiller JH, Stone RM, Strecher V (2007) Oncology: an evidence-based approach. Springer, Berlin
15. Chauvenet M, Lepage C, Jooste V, Cottet V, Faivre J, Bouvier A-M (2009) Prevalence of patients with colorectal cancer requiring follow-up or active treatment. Eur J Cancer 45(8):1460–1465
16. Chen M-H, Ibrahim JG, Sinha D (1999) A new Bayesian model for survival data with a surviving fraction. J Am Stat Assoc 94(447):909–919
17. Clements M, Liu X-R (2019) rstpm2: Smooth Survival Models, Including Generalized Survival Models. R package version 1(5):1
18. Clerc-Urmès I, Grzebyk M, Hédelin G, CENSUR working survival group (2020) flexrsurv: An R package for relative survival analysis. R package version 1(4):5
19. Cleveland WS, Grosse E (1991) Computational methods for local regression. Stat Comput 1(1):47–62
20. Cleveland WS, Devlin SJ, Grosse E (1988) Regression by local fitting: methods, properties, and computational algorithms. J Econometr 37(1):87–114
21. Cleveland W, Grosse E, Shyu M (1992) A package of c and fortran routines for fitting local regression models. In: Chambers JM (ed) Statistical methods. S. Chapman and Hall Ltd., London
22. Coleman MP, Babb P, Damiecki P, Grosclaude P, Honjo S, Jones J, Knerer G, Pitard A, Quinn M, Sloggett A et al (1999) Cancer survival trends in England and Wales, 1971–1995: deprivation and NHS region. Stationery Office Books
23. Dal Maso L, Guzzinati S, Buzzoni C, Capocaccia R, Serraino D, Caldarella A, Dei Tos A, Falcini F, Autelitano M, Masanotti G et al (2014) Long-term survival, prevalence, and cure of cancer: a population-based estimation for 818 902 italian patients and 26 cancer types. Ann Oncol 25(11):2251–2260
24. Danieli C, Remontet L, Bossard N, Roche L, Belot A (2012) Estimating net survival: the importance of allowing for informative censoring. Stat Med 31(8):775–786
25. Dickman PW, Sloggett A, Hills M, Hakulinen T (2004) Regression models for relative survival. Stat Med 23(1):51–64
26. Dodd EO, Streftaris G, Waters HR, Stott AD (2015) The effect of model uncertainty on the pricing of critical illness insurance. Ann Actuar Sci 9(1):108–133
27. Eilers PH, Marx BD (1996) Flexible smoothing with b-splines and penalties. Stat Sci 11(2):89–102
28. Enstrom JE, Austin DF (1977) Interpreting cancer survival rates. Science 195(4281):847–851
29. Esteve J, Benhamou E, Croasdale M, Raymond L (1990) Relative survival and the estimation of net survival: elements for further discussion. Stat Med 9(5):529–538
30. Esteve J, Benhamou E, Raymond L et al (1994) Statistical methods in cancer research. volume iv. descriptive epidemiology. IARC Sci publ 128(1):302
31. Fauvernier M, Remontet L, Uhry Z, Bossard N, Roche L (2019a) survpen: an r package for hazard and excess hazard modelling with multidimensional penalized splines. J Open Sour Softw 4(40):1434
32. Fauvernier M, Roche L, Uhry Z, Tron L, Bossard N, Remontet L, and in the Estimation of Net Survival Working Survival Group C (2019b) Multi-dimensional penalized hazard model with continuous covariates: applications for studying trends and social inequalities in cancer survival. J R Stat Soc Ser C (Applied Statistics)

33. Geskus RB (2015) Data analysis with competing risks and intermediate states, vol 82. CRC Press, Boca Raton
34. Giorgi R, Payan J, Gouvernet J (2005) Rsurv: a function to perform relative survival analysis with s-plus or r. Comput Methods Progr Biomed 78(2):175–178
35. Haberman S, Renshaw A (1990) Generalised linear models and excess mortality from peptic ulcers. Insur Math Econ 9(1):21–32
36. Hakulinen T, Tenkanen L (1987) Regression analysis of relative survival rates. J R Stat Soc Ser C (Appl Stat) 36(3):309–317
37. Jooste V, Grosclaude P, Remontet L, Launoy G, Baldi I, Molinié F, Arveux P, Bossard N, Bouvier A-M, Colonna M et al (2013) Unbiased estimates of long-term net survival of solid cancers in france. Int J Cancer 132(10):2370–2377
38. Kaplan EL, Meier P (1958) Nonparametric estimation from incomplete observations. J Am Stat Assoc 53(282):457–481
39. Lambert PC, Thompson JR, Weston CL, Dickman PW (2006) Estimating and modeling the cure fraction in population-based cancer survival analysis. Biostatistics 8(3):576–594
40. Latouche A, Allignol A, Beyersmann J, Labopin M, Fine JP (2013) A competing risks analysis should report results on all cause-specific hazards and cumulative incidence functions. J Clin Epidemiol 66(6):648–653
41. Legrand C, Bertrand A (2019) Cure models in oncology clinical trials. Textb Clin Trials Oncol Stat Perspect 1:465–492
42. Lemaire J, Subramanian K, Armstrong K, Asch DA (2000) Pricing term insurance in the presence of a family history of breast or ovarian cancer. N Am Actuar J 4(2):75–87
43. Lenner P (1990) The excess mortality rate: a useful concept in cancer epidemiology. Acta Oncol 29(5):573–576
44. Maller RA, Zhou X (1996) Survival analysis with long-term survivors. Wiley, Oxford
45. Massart M (2018) A long-term survivor's perspective on supportive policy for a better access to insurance, loan and mortgage. J Cancer Policy 15:70–71
46. Mounier M (2015) Apport des méthodes de survie nette dans le pronostic des lymphomes malins non hodgkiniens en population générale. PhD thesis, Université Claude Bernard-Lyon I
47. Oksanen H (1998) Modelling the survival of prostate cancer patients. University of Tampere
48. Othus M, Barlogie B, LeBlanc ML, Crowley JJ (2012) Cure models as a useful statistical tool for analyzing survival. Clin Cancer Res 18(14):3731–3736
49. Pavlič K, Pohar Perme M (2019) Using pseudo-observations for estimation in relative survival. Biostatistics 20(3):384–399
50. Percy C, Stanek E 3rd, Gloeckler L (1981) Accuracy of cancer death certificates and its effect on cancer mortality statistics. Am J Public Health 71(3):242–250
51. Perme MP, Pavlič K (2018) Nonparametric relative survival analysis with the R package relsurv. J Stat Softw 87(8):1–27
52. Perme MP, Stare J, Estève J (2012) On estimation in relative survival. Biometrics 68(1):113–120
53. Perme MP, Estève J, Rachet B (2016) Analysing population-based cancer survival-settling the controversies. BMC Cancer 16(1):1–8
54. Pohar M, Stare J (2006) Relative survival analysis in r. Comput Methods Programs Biomed 81(3):272–278
55. Quantin C, Abrahamowicz M, Moreau T, Bartlett G, MacKenzie T, Adnane Tazi M, Lalonde L, Faivre J (1999) Variation over time of the effects of prognostic factors in a population-based study of colon cancer: comparison of statistical models. Am J Epidemiol 150(11):1188–1200
56. R Core Team (2017) R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria
57. Remontet L, Uhry Z, Bossard N, Iwaz J, Belot A, Danieli C, Charvat H, Roche L, Group C W S (2019) Flexible and structured survival model for a simultaneous estimation of non-linear and non-proportional effects and complex interactions between continuous variables: Performance of this multidimensional penalized spline approach in net survival trend analysis. Stat Methods Med Res 28(8):2368–2384
58. Renshaw AE (1988) Modelling excess mortality using glim. J Inst Actuar 115(2):299–315
59. Ries L, Eisner M, Kosary C, Hankey B, Miller B, Clegg L, Edwards B (2002) Seer cancer statistics review, 1973–1999. bethesda, md: National cancer institute. https://seer.cancer.gov/csr/1973-1999

60. Schaffar R, Rachet B, Belot A, Woods LM (2017) Estimation of net survival for cancer patients: relative survival setting more robust to some assumption violations than cause-specific setting, a sensitivity analysis on empirical data. Eur J Cancer 72:78–83
61. Schvartsman G, Taranto P, Glitza IC, Agarwala SS, Atkins MB, Buzaid AC (2019) Management of metastatic cutaneous melanoma: updates in clinical practice. Ther Adv Med Oncol 11:1758835919851663
62. Shang K (2019) Individual Cancer Mortality Prediction. Fundaciòn Mapfre
63. Silversmit G, Jegou D, Vaes E, Van Hoof E, Goetghebeur E, Van Eycken L (2017) Cure of cancer for seven cancer sites in the Flemish region. Int J Cancer 140(5):1102–1110
64. The Belgian Cancer Registry (2012) Cancer survival in belgium. https://kankerregister.org/media/docs/publications/CancerSurvivalinBelgium.PDF
65. Tsodikov A, Ibrahim J, Yakovlev A (2003) Estimating cure rates from survival data: an alternative to two-component mixture models. J Am Stat Assoc 98(464):1063–1078
66. Tsodikov AD, Yakovlev AY, Asselain B (1996) Stochastic models of tumor latency and their biostatistical applications, vol 1. World Scientific, Singapore
67. Yue JC, Wang H-C, Leong Y-Y, Su W-P (2018) Using Taiwan national health insurance database to model cancer incidence and mortality rates. Insur Math Econ 78:316–324