Institute
and Faculty
of Actuaries

ORIGINAL RESEARCH PAPER

Cancer insurance pricing under different scenarios associated with diagnosis and treatment

Ayşe Arık[1,2] ⓘ, Andrew J. G. Cairns[2], Erengul Dodd[3], Angus S. Macdonald[2] ⓘ,
Adam Shao[4] and George Streftaris[2]

[1]School of Risk and Actuarial Studies, University of New South Wales, Sydney, Australia; [2]Department of Actuarial
Mathematics and Statistics, Heriot-Watt University, and the Maxwell Institute for Mathematical Sciences, Edinburgh, UK;
[3]Mathematical Sciences and S3RI, University of Southampton, Southampton, UK; and [4]Biometric Risk Modelling Chapter,
SCOR, Singapore, Singapore
Corresponding author: Ayşe Arık; Email: a.arik@unsw.edu.au

Abstract

We consider pricing of a specialised critical illness and life insurance contract for breast cancer (BC) risk. We compare (a) an industry-based Markov model with (b) a recently developed semi-Markov model, which accounts for unobserved BC cases and progression through clinical stages of BC, and (c) an alternative Markov model derived from (b). All models are calibrated using population data in England and data from the medical literature. We show that the semi-Markov model aligns best with empirical evidence. We then consider net premiums of specialized life insurance products under various scenarios of cancer diagnosis and treatment. The results show strong dependence on the time spent with diagnosed or undiagnosed pre-metastatic BC. This proves to be significant for refining cancer survival estimates and accurately estimating related age dependence by cancer stage. In contrast, the industry-based model, by overlooking this critical factor, is more sensitive to the model assumptions, underscoring its limitations in cancer estimates.

Keywords: Breast cancer; model risk; multiple state models; pricing; semi-Markov model

1. Introduction

Critical illness insurance (CII) is one of the most significant insurance products and provides coverage against a wide range of diseases, including cancer. Cancer remains one of the most common causes of morbidity and mortality worldwide with an estimated 19.3 million new cases and almost 10 million deaths occurred in 2020 (Sung et al., 2021). Meanwhile, in the aftermath of the COVID-19 pandemic, cancer has been one of the diseases impacted the most due to the dramatic changes in relevant healthcare pathways introduced as a response to the pandemic (Alagoz et al. 2021; Lai et al., 2020; Maringe et al., 2020; Sud et al., 2020).

Our focus is on breast cancer (BC). This is one of the most common cancers diagnosed in women and one of the leading causes of death for women in several countries (American Cancer Society, 2021; McDonald et al., 2008). Besides, BC is one of the most common causes of CII claims, e.g., accounting for 44% of female CII claims in the UK in 2014 (CMI, 2011; Aviva, 2015). Also, BC is one of the cancer types with significantly higher cancer survival as compared to other cancer types, e.g., lung cancer. This, combined with a better understanding of BC risk, has led to changes in insurance practice, such as the "right to be forgotten" initiative in Europe, with more and more BC survivors being insurable and with new insurance options being available to women

with medical history (iamInsured, 2023; Insurance Europe, 2021; SCOR, 2020; The Insurance Surgery, 2023).

It is also important to note that BC is one of the key cancer types for which screening is often available. The availability of screening is crucial for early diagnosis of cancer, which is also a main determinant of cancer survival. Unfortunately, national-level lockdowns, introduced as a response to the COVID-19 pandemic, had significant consequences leading, for instance, to suspension of cancer screening programs and treatments between March and June 2020 in the UK (CRUK, 2021).

In this paper, we build on and extend earlier work (Arık et al., 2023, 2024), to derive BC and other cause mortality at a wide range of age groups, and provide pricing for a specialized accelerated CII contract and a life insurance contract, for both healthy and post-cancer women. Our modeling allows the inclusion of cancer stage and results are obtained under various scenarios relating to cancer diagnosis and treatment. We also consider a commonly used modeling framework for pricing (Baione & Levantesi, 2018; CMI, 1991) and provide detailed comparisons between these approaches. The comparisons showcase the diverse implications arising from varying modeling assumptions in the context of life insurance.

Specifically, we consider two main multiple state models in a continuous time framework for modeling BC, e.g., see Soetewey et al. (2022, 2023), Baione & Levantesi (2018), Ozkok Dodd et al. (2014). First, we consider an industry-based Markov model (Reynolds & Faye, 2016) with 4 states, also discussed in Baione & Levantesi (2018), as our baseline model. This model is similar to the 3-state (healthy-ill-dead) model used by Soetewey et al. (2022), but the latter is within a semi-Markov framework. The difference between these models stems from how death is defined. In the industry-based model, death is divided into two categories: due to critical illness or other causes. Since we are interested in pricing both CII and traditional life insurance contracts, we focus on the 4-state model rather than the 3-state model. Moreover, these earlier literature models account for cancer incidence in a single state. An exception is the study by Soetewey et al. (2023). They introduce cancer stage into the model, while retaining all-cause mortality for death.

As an alternative to the existing models, we propose to use a semi-Markov model as the second approach that differentiates between life histories on the basis of cancer stage and whether or not cancer diagnosis is made while maintaining two distinct causes of death (Arık et al. 2024). The model presented in Arık et al. (2024) is further developed here in the following ways: (i) we consider a collection of semi-Markov models for a wider age range starting from age 30 and (ii) we apply generalised additive models to define transition intensities to account for changes in rates of transition over time for a given cohort. We are particularly interested in understanding how combining specific events, like BC registrations across different cancer stages at the time of diagnosis as assumed by the industry-based model, affects insurance cash flows, especially when there is not enough data available. Thus, we show how both modeling frameworks can be implemented to quantify net insurance premiums for specialized insurance products for women aged 35 to 60 years at the time of purchase. At the same time, we address differences raised by different model assumptions. This includes clarifying the consequences of relying on the more compact model and ignoring duration dependence in this specific case. Additionally, we highlight the potential benefits of incorporating duration dependence by examining a special case of the proposed semi-Markov model. This is achieved through comparisons of insurance premiums calculated under the semi-Markov model and its special case. This is relevant to ongoing discussions around as-if-Markov modeling as a substitute for semi-Markov processes for more refined actuarial (reserve) calculations (Christiansen, 2021).

Calibrating BC risk in the absence of comprehensive data is complex, but we demonstrate a suitable approach based on a pragmatic combination of publicly available cancer registrations and deaths data in England from the Office for National Statistics (ONS), and published clinical studies. Our study discusses shortcomings of the earlier literature (CMI 1991; Baione & Levantesi, 2018; Reynolds & Faye, 2016) based on comparisons of estimated BC incidence and mortality with

Table 1. Multiple state models used in the numerical results

M0	Industry-based Markov model
M1	Semi-Markov model
M2	Markov model, as a special case of the semi-Markov model
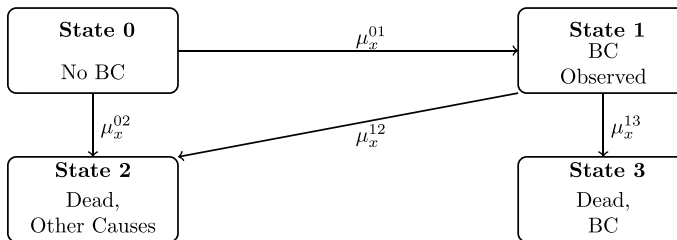
Figure 1. Industry-based 4-state Markov model, M0, as in Baione & Levantesi (2018). Intensities μ are functions of age x.

observed rates. We also clarify the definition of a certain method, known as k_x-method, practised by the insurance industry to determine other causes of death through critical illness deaths. Overall, we conclude that the earlier literature model (which we call the industry-based model) is more sensitive to model assumptions and should be approached with caution. We later show differences in net single premiums for a specialized CII contract and a life insurance contract tailored for women at ages 35 to 60 under different considerations associated with diagnosis and treatment. Our calculations highlight each model's capacity to adapt to shifts in insurance practice. We demonstrate that our approach could be useful for underwriting more inclusive life insurance products whilst capturing important features of BC risk, in contrast to the industry-based model.

The remainder of this paper is organized as follows. In Section 2, we introduce the multiple state models used in this study. In Section 3, we describe available data. In Section 4, we explain how to calibrate, implement, and validate the models. In Section 5, we describe specialized insurance contracts and present estimated net insurance premiums under various scenarios linked to BC diagnosis and treatment. In Section 6, we examine the sensitivity of the industry-based model using an alternative calibration. In Section 7, we discuss our main findings and their implications for the insurance industry.

2. Multiple state models

This section introduces the three models used in the paper to define the life history of an individual exposed to BC risk (Table 1). In Section 2.1, we first present an industry-based 4-state Markov model (M0), while in Section 2.2 we introduce a semi-Markov multiple state model (M1) and a related simplified Markov version (M2). While Model 0 is based on an earlier study by Baione & Levantesi (2018), Model 1 is a recently developed framework from Arık et al. (2024). Models M1 and M2 are further discussed and extended in later sections.

2.1 An industry-based multiple state model

We consider a multiple state model with four states as shown in Fig. 1. This model is used as a CII industry model in Reynolds & Faye (2016) and is similar to the multiple state models discussed in CMI (1991) and Baione & Levantesi (2018).

Fig. 1 demonstrates the life history of an insured individual in a continuous time setting. Transition intensities from state i to state j at age x are denoted by μ_x^{ij}.
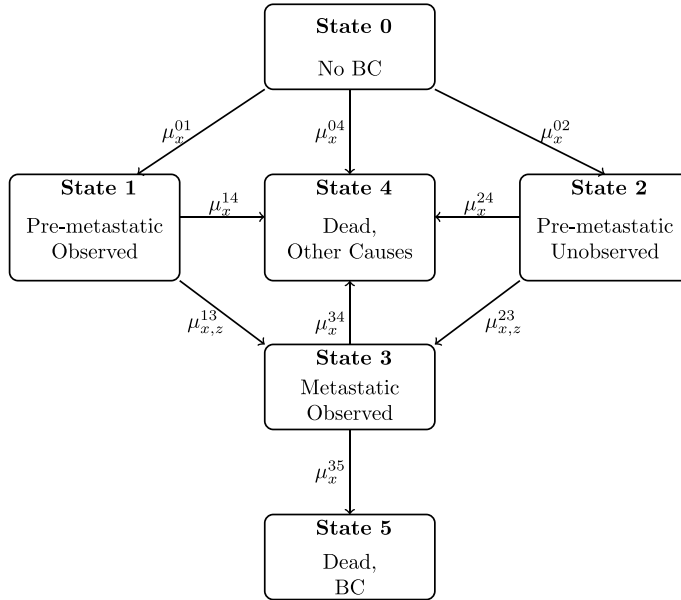
Figure 2. A breast cancer semi-Markov model, M1. Intensities μ are functions of age x and/or duration z.

While four states allow us to price a contract paying a benefit on death specifically from other causes (see Section 5), we would need additional states to describe different stages of BC, or observed and unobserved cases. Such extensions are easily accommodated if necessary by expanding the Kolmogorov equations for Model M0 provided in Appendix A.

2.2 Semi-Markov multiple state model

Fig. 2 displays a continuous time semi-Markov model with 6 states for the life history of a policyholder. This model, recently developed by Arık et al. (2024), presents an enhanced modeling approach for BC risk by offering an improved accuracy in related estimates. It describes the development of a single policy depending on age-specific transition intensities from state i to state j at age x, denoted by μ_x^{ij}, and age- and duration-dependent transition intensities at age x and duration z from state i to state j, denoted by $\mu_{x,z}^{ij}$.

In this model, State 0 represents individuals free of BC. Fig. 2 distinguishes between observed and unobserved BC cases. In particular, State 1 and State 3 show observed BC cases, whereas State 2 shows unobserved BC cases. Note that observed and unobserved BC are distinguished with respect to the availability of BC diagnosis. Incorporating unobserved BC cases as a separate state allows for a more realistic health trajectory for each individual. This approach also enables us to develop scenarios concerning BC diagnosis and the availability of BC treatment (Section 5.2). States 4 and 5 correspond to death from other causes and BC, respectively.

Based on this model, the onset of BC, that is, new cancer cases at a given age x, can be determined by the sum of the rates of transition from State 0 to State 1 and to State 2 such that

$$\mu_x^{01} + \mu_x^{02} = \mu_x^{*}, \tag{1}$$

where μ_x^{*} corresponds to all new BC cases. (1) leads to a convenient parametrisation such that

$$\mu_x^{01} = \alpha_x \, \mu_x^{*}, \qquad \mu_x^{02} = (1 - \alpha_x) \, \mu_x^{*}, \tag{2}$$

where $0 < \alpha_x < 1$ quantifies a proportional relationship between μ_x^{01} and μ_x^{02}. See Arık et al. (2024) for a detailed discussion about the model assumptions.

As part of the model, duration dependence is considered only in certain states, e.g., from pre-metastatic BC to metastatic BC by following the related medical literature (e.g. Colzani et al., 2014). Besides, it is assumed that individuals in State 1 may have treatments for BC, whereas this would not be possible for individuals in State 2. Thus, we assume a lower rate of transition to metastatic BC from State 1, compared to State 2, such that

$$\mu_{x,z}^{13} = \beta_{x,z}\, \mu_{x,z}^{23}, \tag{3}$$

where $0 < \beta_{x,z} < 1$.

Model M1 in Fig. 2 is characterised by a set of modified Kolmogorov equations where a system of integral-differential equations is involved due to the existence of duration dependence from States 1 and 2 to State 3. These equations are explicitly provided in Appendix B. A fourth-order Runge–Kutta scheme is applied to numerically solve the (modified) Kolmogorov equations under consideration.

2.3 A special case of the semi-Markov multiple state model

A Markov model (M2) is also considered as a special case of the semi-Markov model (M1) in Fig. 2. This assumes a simplification of the semi-Markov model, where no duration dependence is considered in any of the model states. In particular, (3) is modified as follows:

$$\mu_x^{13} = \beta_x\, \mu_x^{23},$$

where $0 < \beta_x < 1$.

Model M2 is introduced with the purpose of demonstrating differences in BC rates and related insurance prices caused by duration dependence (e.g. Section 4). The relevant differences are reported through a comparison between the semi-Markov and Markov models.

3. Data

In this section, we describe available data. The data were also used in Arık et al. (2024), which focused on older age groups (65+). Here, our primary interest lies in the insured age groups, specifically ages 30–60. However, the model calibration spans ages 30–89, as detailed in Section 4.

Specifically, the data consist of numbers of deaths by cause-of-death and cancer incidence registrations in England for the following age groups: 30–49, 50–54, 55–59, ..., 85–89. The cause-specific number of deaths data are available up to age 89 from 2001 to 2022, whereas the cancer registrations data are available from 2001 to 2020.

Cancer registrations are stratified by five-year age-at-diagnosis groups, type of tumor, single calendar year, and gender, for the years between 2001 and 2020. These data are provided by the ONS up to 2017 and later by the Health and Social Care Information Centre of the National Health Service (NHS) of England, also known as NHS Digital. Cause-specific death numbers, provided by the ONS, have similar granularity, where the data are split by five-year age-at-death groups, causes of death defined based on ICD 10 classification, single calendar year, and gender, between 2001 and 2022. Mid-year population estimates up to 2022 are also available from the ONS.

Fig. 3 shows age-specific BC incidence and mortality along with mortality from other causes over time. This builds on a similar figure presented in Arık et al. (2024), incorporating more recent data in the current version. Here, BC incidence is defined as new BC registrations at a certain age-at-diagnosis group in a given year, divided by the related mid-year population estimates. Mortality from other causes is defined similarly, with deaths from BC being excluded.
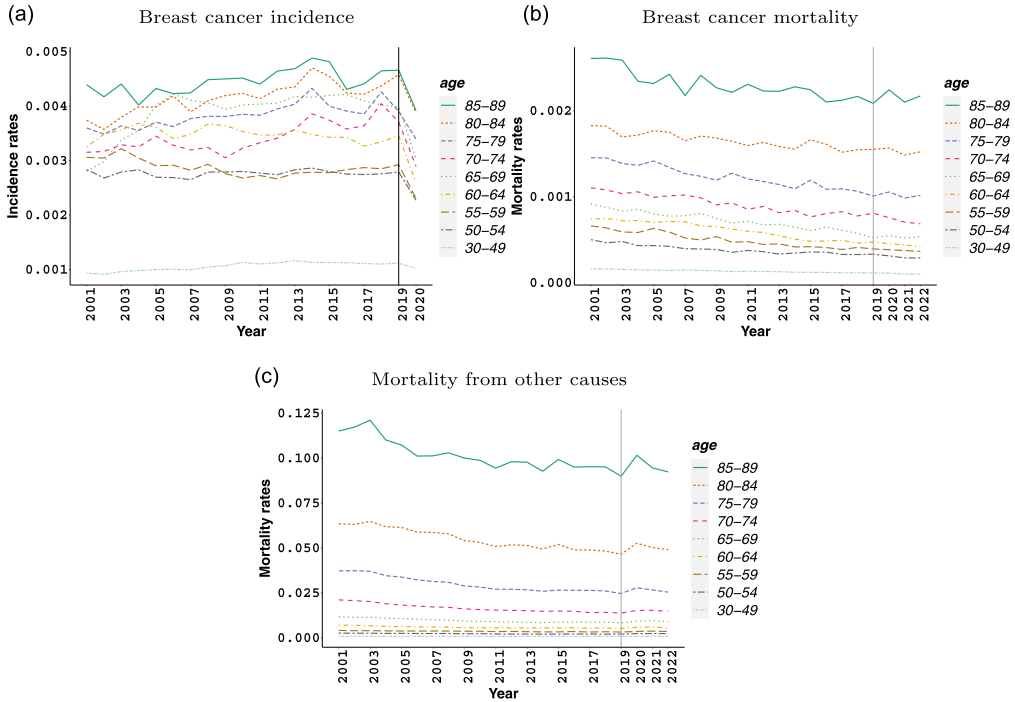
(a) Breast cancer incidence

(b) Breast cancer mortality

(c) Mortality from other causes

Figure 3. Breast cancer incidence, mortality, and mortality from other causes (excluding breast cancer).

Fig. 3a demonstrates that BC is a greater risk for women from age 50 onward. Also, BC incidence mostly exhibits an increasing trend at different ages over calendar years apart from 2020, where we observe a sharp decline, as low as 25% at ages 60–64 (see also Fig. 4a). This is more evident for women older than 30–49, reflecting a reduction in cancer registrations due to the changes in availability of cancer services, e.g., a halt in cancer screening from late March 2020 till June 2020, as a result of national lockdowns introduced as a response to the COVID-19 pandemic (CRUK, 2021).

Fig. 3b and c show a generally decreasing trend in BC and other-cause mortality. This trend was interrupted by COVID-19 in 2020, with increasing mortality from other causes at all considered ages by 10–13% as compared to 2019 (Fig. 4c). BC mortality increased in 2020, up around 7%, for older ages, while it decreased for younger ages (Fig. 4b).

4. Model calibration and validation

In this section, we first present key transition intensities used to calibrate the models described in Section 2. We show our main findings based on different modeling assumptions, summarised in Table 1. Each model across M0–M2 is calibrated for women aged 30 to 89 years, partly using the population data of England (see Section 3). However, only M1 is accounting for duration dependence, and the other two models (M0 and M2) do not involve duration dependence. We exclude the calendar years 2020–2022, while calibrating the models, as the COVID-19 pandemic has had a major impact not only on cancer diagnoses and treatments but also on health-seeking behavior in general. The models are then used to estimate age-specific occupancy probabilities in future years, starting from 1 January 2020 onward. The occupancy probabilities are subsequently used to derive net cancer survival at different ages of diagnosis.
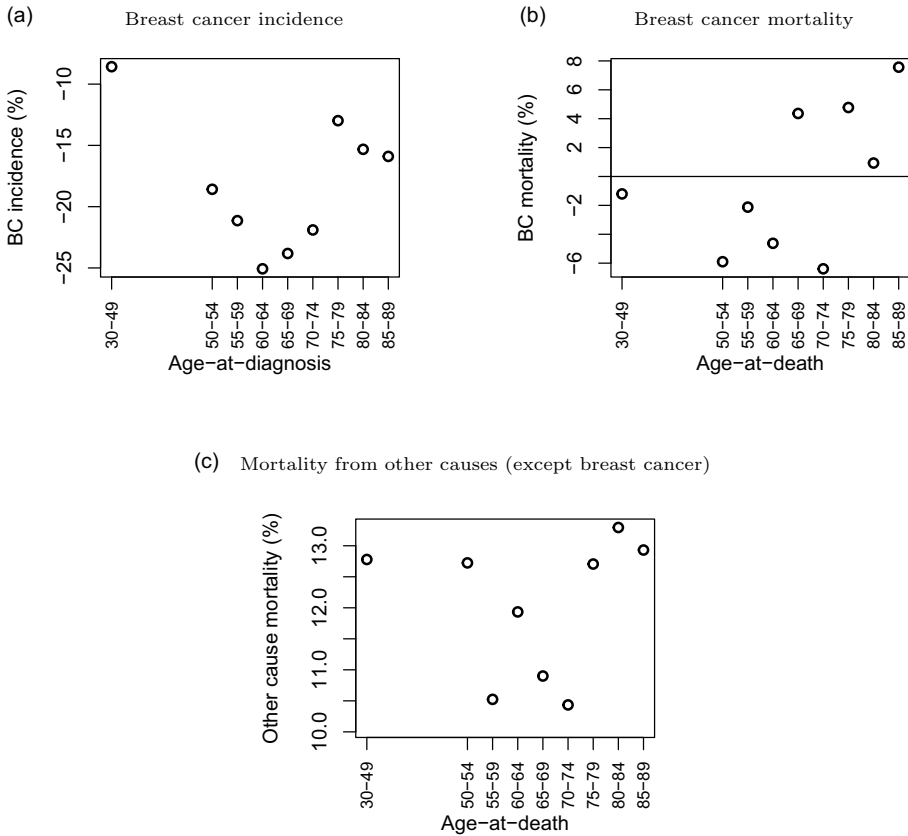
(a) Breast cancer incidence

(b) Breast cancer mortality

(c) Mortality from other causes (except breast cancer)

Figure 4. Change (%) in breast cancer incidence and mortality, and in mortality from other causes, in 2020 as compared to 2019.

4.1 Transition intensities

Key transition intensities used in this paper are summarised in Table 2. In the models tabulated in Table 1, we use available data as introduced in Section 3, up to 2019. To be specific, we use average cancer incidence and type-specific mortality between 2001 and 2019 to determine age-specific rates of transition in the models under consideration, and by doing so we ignore the time trend. This is mainly for easiness of computation. We then fit generalised additive models to these rates, see Section 4.2, while obtaining insurance premiums to account for changes in incidence and mortality rates at different ages.

Hereby, we use average cancer incidence in England to determine age-specific rates of transition from State 0 to State 1, μ_x^{01}. We note that BC registrations by stage in England are available between years of diagnosis 2012–2015 (ONS, 2016). Following Table 1 in ONS (2016b), 81% of new cancer registrations between 2001 and 2019 is used to represent Stages 1–3 BC, i.e. transitions to State 1 in M1 and M2.

Furthermore, in the models summarised in Table 1, we use average mortality from other causes between 2001 and 2019 to determine age-specific rates of transition from State 0 to death from other causes, i.e., μ_x^{02} in M0 and μ_x^{04} in M1 and M2. Survival probabilities from other causes for BC patients could be different than the ones for women with no BC (Cho et al., 2013). Particularly, BC patients could be more susceptible to die from other causes due to additional risk factor (cancer incidence) (Andersen & Keiding, 2012). However, in the absence of comprehensive data, we

Table 2. Age-specific transition intensities for the BC models M0–M2 based on available data and medical literature

Age	μ_x^{01} in M0	μ_x^{01} in M1&M2	μ_x^{02} in M0 μ_x^{04} in M1&M2	μ_x^{13} in M0 μ_x^{35} in M1&M2
30–49	0.00106	0.00086	0.00084	0.16739
50–54	0.00277	0.00224	0.00228	0.24005
55–59	0.00287	0.00233	0.00363	0.24005
60–64	0.00349	0.00282	0.00588	0.28060
65–69	0.00393	0.00318	0.00952	0.28060
70–74	0.00345	0.00280	0.01643	0.36002
75–79	0.00384	0.00311	0.02987	0.40000
80–84	0.00417	0.00338	0.05496	0.49711
85–89	0.00447	0.00362	0.10112	0.50000

assume that transitions to death due to other causes from all live states are equal to the transition to death due to other causes from State 0, i.e.,

$$\mu_x^{12} = \mu_x^{02},$$

in M0, and

$$\mu_x^{14} = \mu_x^{24} = \mu_x^{34} = \mu_x^{04},$$

in M1 and M2.

We determine rates of transition to death from BC denoted by μ_x^{13}, in M0, and by μ_x^{35} in M1 and M2, using the information reported in Zhao et al. (2020). Moreover, we refer the study of Colzani et al. (2014) to determine rates of transition for developing metastatic BC after being diagnosed with pre-metastatic BC, denoted as $\mu_{x,z}^{13}$ in M1. See Arık et al. (2024) for an in-depth discussion about the appropriateness of these sources for model calibration purposes. Note that we assume $\mu_x^{13} = 0.0194$ in M2, which is a special case of M1. This value is consistent with first distant metastasis rates based on Table 1 in Colzani et al. (2014).

Rates of transition for developing metastatic BC in the absence of cancer diagnosis and treatment, i.e., $\mu_{x,z}^{23}$, are determined based on $\mu_{x,z}^{13}$ (see (3)). Thus, under M1, $\mu_{x,z}^{23}$ accounts for duration dependence, whereas duration z is ignored under M2, with $\mu_x^{13} = 0.0194$.

There is no reliable data to support a particular assumption regarding the choice of α_x and $\beta_{x,z}$ (see (2)–(3)). Thus, an extensive scenario analysis is carried out by considering a range of values of $\alpha_x = \alpha = \{0.1, 0.2, \ldots, 0.9\}$ and $\beta_{x,z} = \beta = \left\{ \frac{1}{2}, \frac{1}{3}, \ldots, \frac{1}{10} \right\}$ in the context of life insurance premiums (see Section 5.2). Furthermore, following Arık et al. (2023), we assume $\alpha_x = \alpha = 0.6$ and $\beta_{x,z} = \beta = 1/7$, in M1 and M2, while providing comprehensive discussions regarding BC risk and related insurance products in Sections 4–5.

4.2 Generalized additive models for key transition intensities

We consider Markov (M0 and M2), and semi-Markov (M1) models, and we model transition intensities as a smooth function of age by using a generalized additive model. Hereby, we apply generalized additive models of the following form to the observed transition intensities in Table 2:

$$g(E(y_i)) = \alpha + \sum_p s_p(x_{ip}), \qquad (4)$$

where α is the intercept; $g(.)$ is a smooth monotonic link function used to transform the expectation of the outcome y; and y is modeled as the sum of smooth functions, $s(.)$, of covariates x (Wood, 2017). We use cubic splines as basis functions.
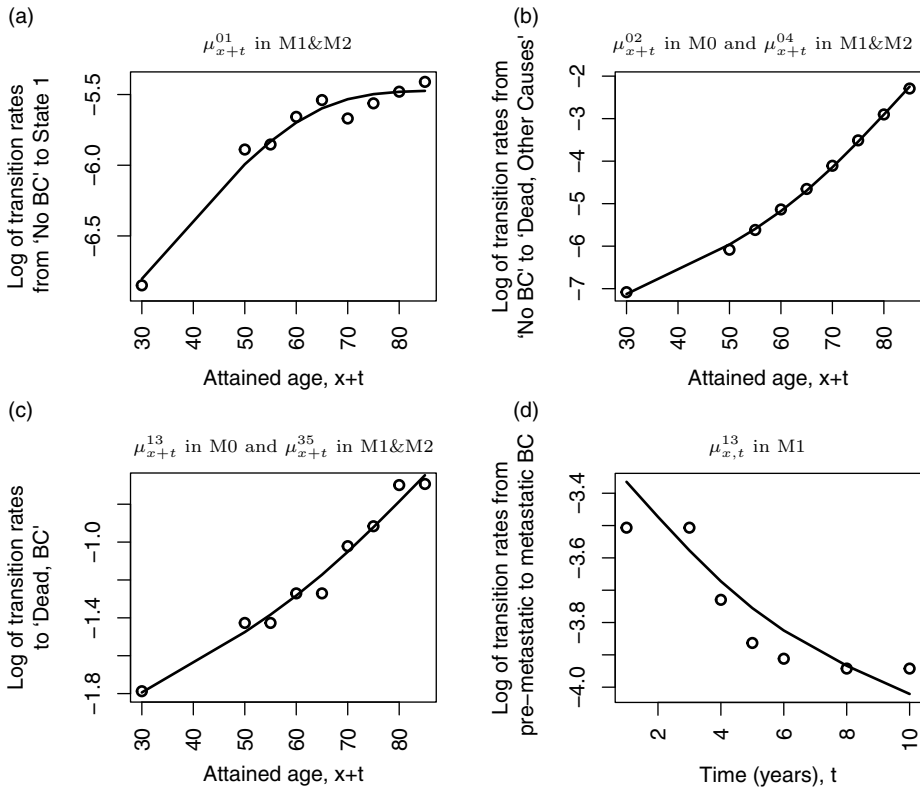
Figure 5. Key transition rates, in the models listed in Table 1, as functions of attained age, $x + t$, where circles show observed values taken from Table 2, and lines show fitted values from the relevant generalized additive models.

Fig. 5 displays observed, provided in Table 2, and fitted values of key transition intensities, as functions of (attained) age based on the additive models in (4). The plots in Fig. 5 demonstrate smooth fitted rates across different ages. Fig. 5a shows, once again, a significant difference in BC morbidity before and after age 50, where the increase in BC incidence seems to slow down from age 65 onward. Fig. 5b and c exhibit increasing mortality rates from other causes and BC, respectively, for higher ages. Fig. 5d suggests a decrease in developing a higher stage of BC, i.e., a more progressed BC, after being diagnosed with BC over time. Note that this is taken as fixed for each age.

4.3 Occupancy probabilities for different insured ages

In Markov-type models, the occupancy probability, $_tp_x^{ij}$, is defined to be the probability that an individual in state i at age x will be in state j at age $x + t$ (Macdonald et al., 2018).

Fig. 6 presents occupancy probabilities for an individual aged between 30 and 60 years, with no BC at the beginning of the insurance contract, based on M0. Note that these probabilities can be associated with a policyholder at these ages at the time of purchase of a particular insurance product. Hereby, in the figure, we assume that time spent on the contract depends on the policyholder's age at purchase. We also assume that the maximum age for these insurance contracts is 90. This means that the policy purchased by a person who is aged 50 years could be in-force for at most 40 years, whereas policy purchased by a 60 year old could be in-force for at most 30 years. Accordingly, in each subplot, we see shorter time periods for higher ages on x axis.
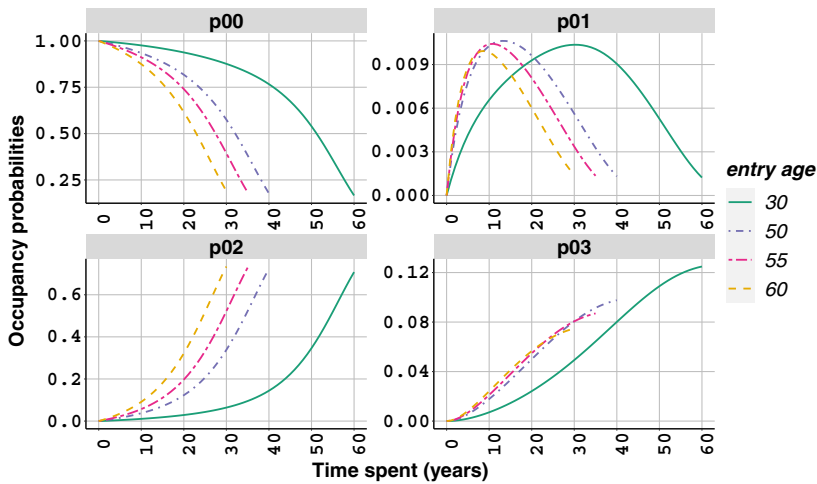
Figure 6. Occupancy probabilities for a policyholder with no breast cancer, at different contract entry ages, based on the industry-based model M0.
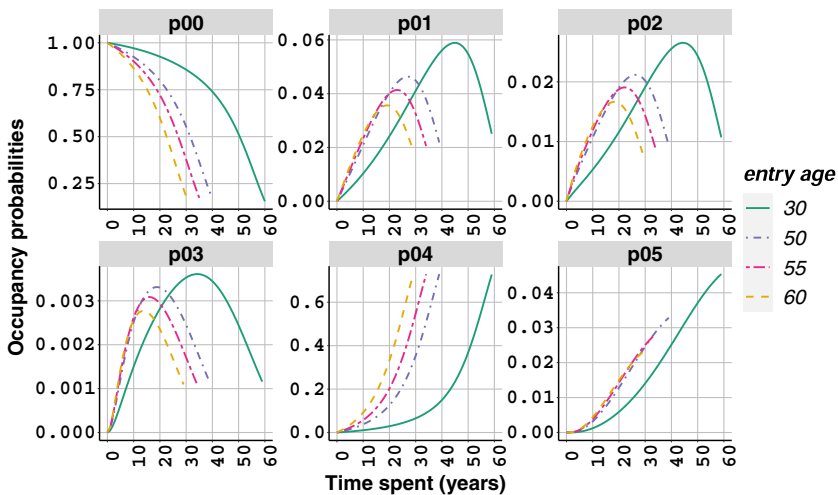

Figure 7. Occupancy probabilities for a policyholder with no breast cancer, at different contract entry ages, based on the semi-Markov model M1.

Similarly, Fig. 7 displays occupancy probabilities for women with different ages ranging from 30 to 60, with no BC at time zero, based on M1.

We note that the death probabilities from BC for a healthy woman under M0, $_tp_x^{03}$, are estimated as significantly higher than those under the semi-Markov model, M1, $_tp_x^{05}$, for the same t. This is related to the assumption about the risk of dying from BC after receiving a BC diagnosis, that is μ_x^{13} under M0. In the absence of relevant data, we use a similar assumption for μ_x^{35} under M1, leading to high deaths from BC (Table 2). See Section 6 for further discussion related to this assumption.

4.4 Model validation: Breast cancer survival

We calculate age-specific net cancer survival, following the ONS definition (ONS, 2019b), with the aim of comparing our results with the ONS findings. This measures survival from a given
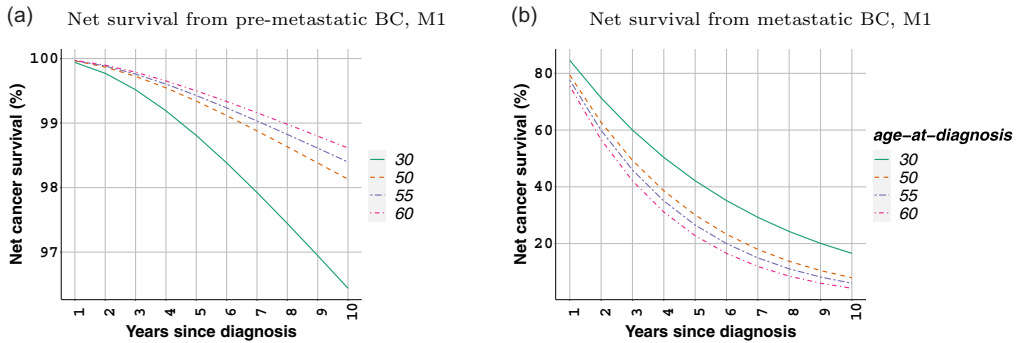
Figure 8. Estimated net cancer survival at different ages under the semi-Markov model M1, for a woman diagnosed with: (a) pre-metastatic breast cancer and (b) metastatic breast cancer.

cancer type after receiving a diagnosis, assuming that this cancer can be the only cause of death (Mariotto et al., 2014; Swaminathan and Brenner, 2011). As an example, we can consider a woman diagnosed with pre-metastatic BC at age x. BC survival of this woman in t years can be obtained, based on M1, as follows:

$$\frac{1 - {}_tp_x^{14} - {}_tp_x^{15}}{1 - {}_tp_x^{14}}, \tag{5}$$

where ${}_tp_x^{14}$ represents mortality from other causes, and ${}_tp_x^{15}$ represents mortality from BC at age x in t years time.

Fig. 8 shows net cancer survival from pre-metastatic and metastatic BC for a woman diagnosed with the related BC at different ages, based on M1. The relationship between BC survival with a pre-metastatic BC and age, Fig. 8a, demonstrates an inverse pattern, where survival seems to increase from age 30 to 60. These estimates are aligned with BC statistics in England. For instance, 5-year cancer survival for women aged 15–39 years at diagnosis between 2010 and 2014, followed up to 2015, is reported to be 85.5%, whereas the same survival for women aged 40–69 years is between 90.7% and 93.0%. Note that these figures do not distinguish cancer stages. Higher survival at older ages may be linked to the availability of BC screening for women aged 50–70 years at the time (ONS, 2016). Another reason might be the type of BC, which is more likely to be a more aggressive cancer for younger women as compared to older women (ONS, 2019a).

At the same time, Fig. 8b points out significantly lower cancer survival estimates for women with metastatic BC, as compared to those with pre-metastatic BC (in Fig. 8a), where higher age is associated with lower cancer survival. One- and 5-year age standardised survival rates for women diagnosed with Stage 4 BC in 2012 to 2016, followed up to 2017, are reported to be 66% and 27.9%, respectively, while even 5-year survival estimates remain very high for women diagnosed with Stages 1–3 BC, which is above 96.5% (ONS, 2019a). It is important to highlight that while the age-standardized rates reported by the ONS are based on ages from 15 to 99 years, our estimates specifically focus on insured ages ranging from 30 to 60 years.

Note that survival estimates obtained under M0 are identical to the ones in Fig. 8b due to the assumptions about rates of transition to death from BC and other causes, respectively (see Table 2).

4.5 Model validation: The proportion of breast cancer deaths over all deaths

Model M0 in Fig. 1 is widely applied in CII by the insurance industry. Also, in the absence of reliable cause of death data, a particular approach, known as k_x-method, is employed based on this model (Dash & Grimshaw, 1990; SAS, 2011; Institute and Faculty of Actuaries, 2014;

Reynolds & Faye, 2016). Specifically, this approach aims to identify the proportion of deaths from other causes in a given year depending on an input: k_x. This input, k_x, is defined to be the proportion of deaths from a certain critical illness condition over all deaths at a given age x. Appendix F provides a detailed account of derivation of k_x as outlined, e.g., in Reynolds & Faye (2016). Note that the related derivation shall be read in conjunction with the assumptions made under M0 (see Section 4.1).

We understand that the motivation behind k_x-method is to find a reasonable and simple approach to differentiate between critical illness and other causes of death in the absence of a comprehensive dataset. Particularly, k_x values are used to indirectly estimate other causes of deaths in practice. Here, we attempt to define model-based k_x values under each model considered as part of this study. Afterward, the model-based k_x values are compared with BC statistics in England to further validate model results.

Particularly, we define the proportion of model-based BC deaths at attained age x as

$$\hat{k}_x = \frac{{}_x p_0^{01} \mu_x^{13}}{{}_x p_0^{00} \mu_x^{02} + {}_x p_0^{01} \mu_x^{12} + {}_x p_0^{01} \mu_x^{13}},$$

based on M0; and

$$\hat{k}_x = \frac{{}_x p_0^{03} \mu_x^{35}}{{}_x p_0^{00} \mu_x^{04} + {}_x p_0^{01} \mu_x^{14} + {}_x p_0^{02} \mu_x^{24} + {}_x p_0^{03} \mu_x^{34} + {}_x p_0^{03} \mu_x^{35}},$$

based on M1 and M2. Note that, for simplicity, the formulas above are defined for a cohort starting in State 0 at age 0.

Fig. 9 exhibits model based and observed values of k_x. Specifically, Fig. 9a shows estimated BC deaths over all deaths for policyholders aged 30 years, with no BC (State 0), at time zero. Note that the models are calibrated as described in Sections 4.1–4.2 with a boundary condition ${}_0 p_{30}^{00} = 1$ (Appendices A and B). Meanwhile, Fig. 9b displays the proportion of BC deaths over all deaths, based on the population data of England in 2001 and 2019. We see significantly higher estimates based on M0 as compared to the observed values, such as around 48%, for a woman aged 30 years, after staying for 10 years in State 0. The same value is estimated to be 18% and 25% based on M1 and M2, respectively, which are more aligned with the observed value at ages 40–44. The main reason for this difference is that the rates of transition to death from BC after a diagnosis under M0, have been defined using the mortality of women with metastatic BC. It is important to note that this assumption, made in the absence of relevant data, results in higher than expected BC deaths in that modeling setting. Nonetheless, this indicates high sensitivity to model assumptions under M0, urging to be cautious with the related assumptions. Provided that we rely on average rates of transition in 2001–2019, we see that the estimates based on M1 (primarily) and M2 in Fig. 9a are broadly aligned with the empirical values shown in Fig. 9b, with an obvious discrepancy at age 30 due to the boundary condition.

5. Net premium rates: Definitions and results

In this section, we first introduce two main types of insurance contracts and explain calculation of related net insurance premiums. Then, we examine the insurance premiums for the CII contract described in (7) and the life insurance contract described in (9) under various scenarios linked to BC diagnosis and treatment. Last, we present and compare net insurance premiums for a wider range of life insurance contracts for women across varying ages under different modeling assumptions.
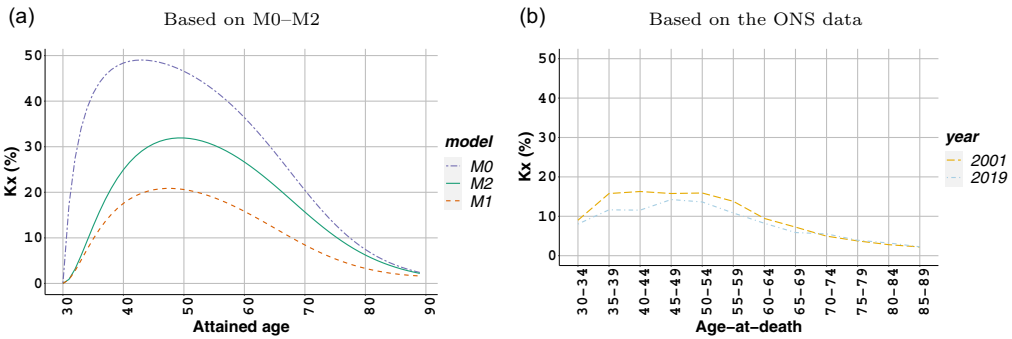
Figure 9. $\hat{k}_x$ values (a) based on different models, and observed k_x values (b) based on the ONS data.

5.1 Net single premiums for different insurance contracts

CII is a popular insurance contract that covers cancer as one of the core diseases, also including heart attack, stroke, and so on. Alongside, BC life insurance has recently been attracting more attention in the life insurance industry. This is possibly linked to increasing demands from people with existing conditions or, perhaps, with a related medical history. However, premium rates for the latter contract are noted to vary significantly from one insurance company to another, as a result of several factors impacting the underwriting process (iamInsured, 2023). For instance, age, a common risk factor, can be a strong determinant of a BC premium, due to the distinctive age-specific curve in BC incidence resulting from very large changes in estrogen level of women after age 50 (Bray et al., 2004; Henderson et al., 1988). At the same time, provided high survival rates from BC, increasing chances of preventing recurrence with more advanced medical technology and a better understanding of the disease over time, it would be possible to tailor a contract for an individual surviving BC. In that case, other factors, such as cancer stage and time since the end of treatment, can become the main determinants of the insurance premium (The Insurance Surgery, 2023).

We consider two different insurance contracts here, noting that our focus is on the impact of different modeling assumptions. Following the study of Baione & Levantesi (2018), the first is a special accelerated CII contract, where a single benefit is paid when the insured

(i) is either diagnosed with BC for the first time; or

(ii) dies from other causes before being diagnosed with BC.

The net single premium for this contract is an accelerated death benefit along with a benefit paid at the time of diagnosis, and this can be calculated based on the industry-based model, M0, in Fig. 1 as

$$\text{CI,1}\bar{A}_x = \int_0^\infty e^{-\delta t} {}_t p_x^{00} \left(\mu_{x+t}^{01} + \mu_{x+t}^{02} \right) dt. \tag{6}$$

The net single premium for the same contract, based on the semi-Markov model, M1, in Fig. 2, can be determined as

$$\text{CI,2}\bar{A}_x = \int_0^\infty e^{-\delta t} {}_t p_x^{00} \left(\mu_{x+t}^{01} + \mu_{x+t}^{04} \right) dt \tag{7}$$
$$+ \int_0^\infty e^{-\delta u} {}_u p_x^{00} \mu_{x+u}^{02} \int_0^\infty e^{-\delta t} {}_t p_{[x+u]}^{22} \left(\mu_{[x+u]+t}^{23} + \mu_{[x+u]+t}^{24} \right) dt \; du.$$

Here, δ is the instantaneous constant force of interest rate, and we assume that

> **(i)** the insured purchases the product before being diagnosed with BC, i.e., in State 0 "No BC"; and

> **(ii)** there is no waiting time between cancer diagnosis and the insurance payment.

Note that, in order to make the formulae clearer, we have slightly modified earlier notation and have used actuarial selection notation. For instance, $\mu_{x,z}^{23}$ is presented based on select attained age $[x]$ with duration z, such that $\mu_{x,z}^{23} = \mu_{[x]+z}^{23}$.

The second contract under consideration is a life insurance contract that can also be purchased with an existing BC condition and provides a single death benefit at the time of death. The net single premium for this contract, that is a death benefit from any cause for an insured person with no BC at the time of purchase, can be expressed based on the industry-based model as

$$_{\text{LI},1}\bar{A}_x = \int_0^\infty e^{-\delta t}\left({}_tp_x^{00}\mu_{x+t}^{02} + {}_tp_x^{01}\left(\mu_{x+t}^{12} + \mu_{x+t}^{13}\right)\right)dt, \tag{8}$$

while under the semi-Markov model it can be written as

$$_{\text{LI},2}\bar{A}_x = \int_0^\infty e^{-\delta t}\left({}_tp_x^{00}\mu_{x+t}^{04} + {}_tp_x^{01}\mu_{x+t}^{14} + {}_tp_x^{02}\mu_{x+t}^{24}\right)dt \tag{9}$$

$$+ \int_0^\infty e^{-\delta u}{}_up_x^{01}\mu_{x+u}^{13}\int_0^\infty e^{-\delta t}{}_tp_{[x+u]}^{33}\left(\mu_{[x+u]+t}^{34} + \mu_{[x+u]+t}^{35}\right)dt\,du$$

$$+ \int_0^\infty e^{-\delta u}{}_up_x^{02}\mu_{x+u}^{23}\int_0^\infty e^{-\delta t}{}_tp_{[x+u]}^{33}\left(\mu_{[x+u]+t}^{34} + \mu_{[x+u]+t}^{35}\right)dt\,du.$$

At the same time, the net single premium for the same contract, for an insured individual with BC at the time of purchase, can be determined based on the industry-based model as

$$_{\text{LI},3}\bar{A}_x = \int_0^\infty e^{-\delta t}{}_tp_x^{11}\left(\mu_{x+t}^{12} + \mu_{x+t}^{13}\right)dt, \tag{10}$$

and the premium for an insured person with pre-metastatic BC, is written based on the semi-Markov model as

$$_{\text{LI},4}\bar{A}_x = \int_0^\infty e^{-\delta t}{}_tp_x^{11}\mu_{x+t}^{14}dt \tag{11}$$

$$+ \int_0^\infty e^{-\delta u}{}_up_x^{11}\mu_{x+u}^{13}\int_0^\infty e^{-\delta t}{}_tp_{[x+u]}^{33}\left(\mu_{[x+u]+t}^{34} + \mu_{[x+u]+t}^{35}\right)dt\,du.$$

Note that the net single premiums of the contracts under consideration, defined in (6)–(11), are expressed for whole life insurance contracts. The upper bound of related integrals in each formula would change in the case of term insurance contracts. We also note that the net single premiums for these contracts under the special case of the semi-Markov model, M2, can be found in Appendix E.

5.2 *Different diagnosis and treatment considerations*

In this section, we examine estimated net insurance premiums under various considerations relating to BC diagnosis and treatment in the semi-Markov model (M1). This is done by choosing a range of values for parameters α and β, such that $\alpha = \{0.1, 0.2, \ldots, 0.9\}$ and $\beta = \left\{\frac{1}{2}, \frac{1}{3}, \ldots, \frac{1}{10}\right\}$.

Particularly, the value of parameter α is associated with BC diagnosis, allowing us to distinguish between diagnosed and undiagnosed BC cases, by considering the proportion of BC diagnoses out
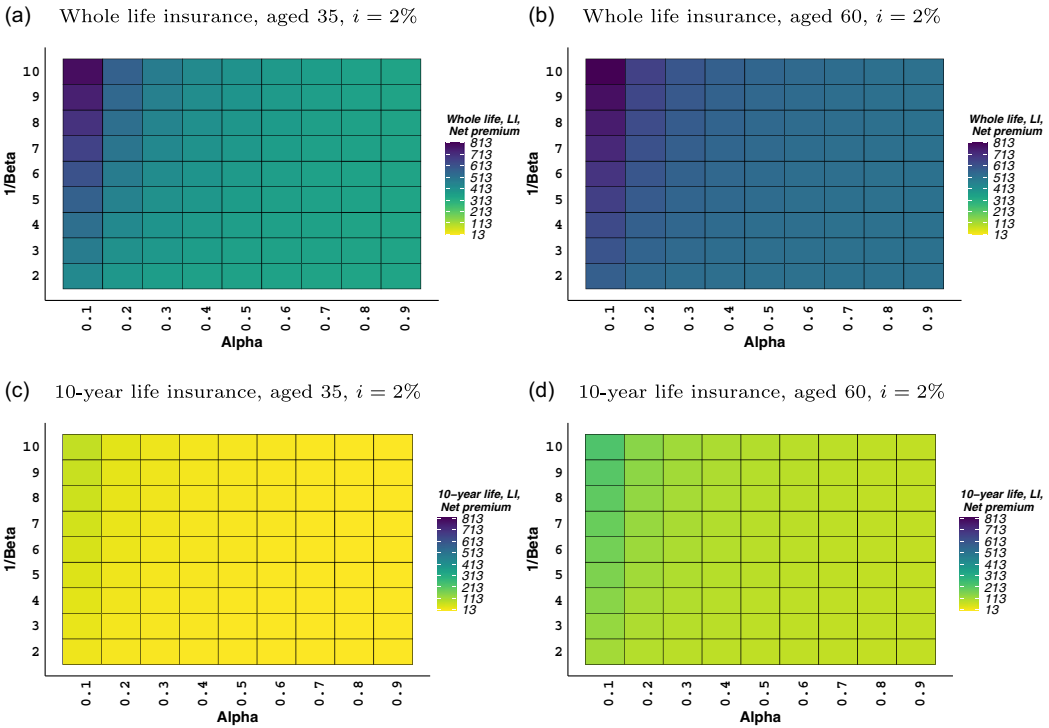
Figure 10. Net single premium rates for a specialized life insurance contract, (9), for policyholders without breast cancer at the time of purchase, £1,000 benefit, payable at the time of death, based on the semi-Markov model M1.

of all BC cases. We consider a range of α values changing from 0.1 to 0.9, while all other model quantities are calibrated as before. Note that α is defined within the range $0 < \alpha < 1$ (see (2)), where $\alpha = 0.1$ is associated with, e.g., a health system with notably poor BC diagnosis (representing 10% of all BC cases) while $\alpha = 0.9$ corresponds to high BC diagnosis (representing 90% of all BC cases).

The value of parameter β is used to differentiate between rates of transition from pre-metastatic BC, whether observed or unobserved, to metastatic BC. In other words, this parameter can be considered as a proxy to the rate of BC metastasis in the absence of cancer treatment since BC treatment would not be accessible to the women with undiagnosed cancer. We consider a set of β values from $\frac{1}{2}$ to $\frac{1}{10}$ within the range $0 < \beta < 1$ (see (3)). Here, $\beta = \frac{1}{2}$ is associated with a higher-level access to BC treatment, whereas $\beta = \frac{1}{10}$ is relating to a lower-level access to BC treatment.

Our analysis suggests that changes in the values of α and β have no impact on life insurance premiums, described in (11) and (12), for women with (pre-metastatic) BC (see (2)–(3) and Table 2 for the definitions of related transition intensities). However, any adjustment in the value of α or β does impact insurance premiums for a healthy woman. Thus, we examine changes in net insurance premiums for the accelerated life insurance and CII contracts, in (9) and (7), respectively, for a woman with no BC medical record at the time of purchase based on various values of α and β in Figs. 10 and 11.

Specifically, Fig. 10 presents the net single premiums for a whole life and 10-year life insurance contract for a woman aged 35 years, Fig. 10a and c, respectively, and for a woman aged 60 years, Fig. 10b and d, where the effective rate of interest is 2% per year. As expected, for a given α and β, higher net insurance premiums are shown as a result of a higher insured age. Our
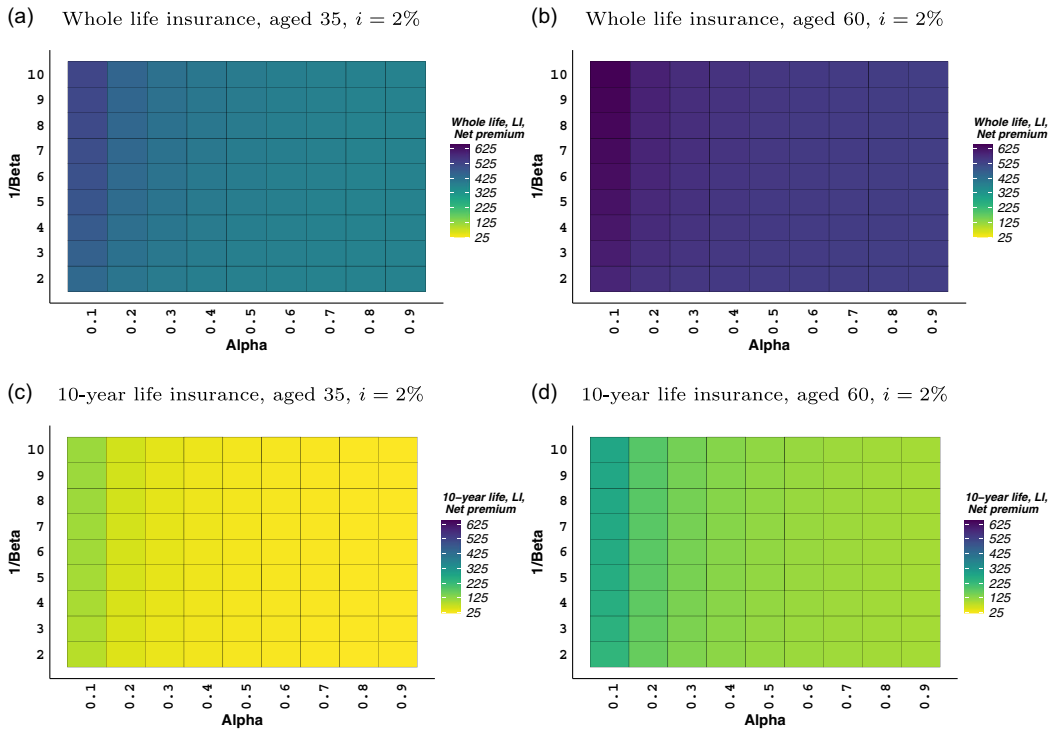
Figure 11. Net single premium rates for a specialized CII contract, (7), for policyholders without breast cancer at the time of purchase, £1,000 benefit, payable at the time of event, based on the semi-Markov model M1.

results demonstrate sensitivity to the choice of α and β parameters, particularly in extreme cases. Importantly, a poor BC diagnosis, where $\alpha < 0.4$, leads to greater sensitivity in the model results, especially in the context of limited access to BC treatment (lower values of β). However, when BC diagnosis is reasonably high, e.g., $\alpha > 0.5$, more consistent net premium rates are obtained for different purchasing ages and maturities.

Fig. 11 shows the estimated net single insurance premiums for the accelerated CII insurance contract at selected maturities for women aged 35 years, Fig. 11a and c, and aged 60 years, Fig. 11b and d. Similar to the estimates in Fig. 10, an α value associated with a poor BC diagnosis seems to be the main reason of greater sensitivity in our findings.

Our findings are also calculated at a different effective rate of interest, 4%, which corroborates the conclusions drawn from the results presented in Figs. 10 and 11 (see Appendix G).

5.3 Net single premium results

Fig. 12 displays the estimated net single premiums for the accelerated CII contracts, in (6) and (7), and the life insurance contracts, in (8)–(11), calculated based on M0–M2 at varying maturities, across different policyholders aged between 35–60 for $\alpha = 0.6$ and $\beta = 1/7$. We use 2% and 4% effective rates of interest, respectively. The figure mainly compares the estimated net single premiums for a woman with no BC, with those for a woman with pre-metastatic BC diagnosis at the time of purchase or 5 years before the time of purchase. As expected, a higher age-at-purchase, a longer maturity, or a lower rate of interest lead to higher net single premiums.

We note that additional analysis is carried out by examining specific values of α, namely 0.4 and 0.8, while keeping all other model parameters calibrated consistently, and assuming β to be $\frac{1}{7}$.

Similarly, we explore selected values of β, specifically $\frac{1}{5}$ and $\frac{1}{10}$, again with other model parameters calibrated consistently, and assuming α to be 0.6. Details can be found in Appendix H.

Differences in net premiums across different models. Pricing differences across different models can be justified with respect to the following points: (i) number of departures from State 0 and (ii) definition of State 1. Specifically, under M1 and M2, we allow three departures from State 0. One of these three departures, to "Pre-metastatic Unobserved," is not considered under M0. Meanwhile, one of the other two departures, to State 1, linked to cancer registrations, is defined differently (see Section 4.1). These have two main consequences. First, the occupancy probability $_{t}p_x^{00}$, which is crucial for pricing purposes, is estimated to be lower under M1 and M2, as compared to M0, because of the higher number of departures in the former model(s). Second, State 1 in M1 and M2 is defined to be a state involving pre-metastatic BC cases, i.e., Stages 1–3 BC, whereas State 1 in M0 combines all BC registrations into a single state without accounting for cancer stage information. This then implies that we have lower rates of transition to State 1, μ_x^{01}, in M1 and M2, in comparison to the rates used in M0 (Table 2).

Differences between CII and life insurance premiums. We can see smaller net single premiums in Fig. 12 for the life insurance contracts for a woman with no BC, as compared to those calculated for the CII contracts. Note that the differences between related premiums are much smaller under M0 due to the definition of State 1 and the assumption linked to BC deaths. The main reason for observing smaller life insurance premiums in general is that BC morbidity is a bigger risk than BC mortality for a healthy woman, and both BC diagnosis or death from other causes lead to a benefit payment under the CII contract. Meanwhile, only death from any cause leads to a benefit payment under the life insurance contract.

Differences in life insurance premiums with BC diagnosis. We observe considerably higher premium estimates for a woman with BC diagnosis at the time of purchase under M0, in comparison to the estimates for a woman with pre-metastatic BC diagnosis under M1–M2. This highlights the importance of our more detailed modeling in M1 and M2.

Our findings for women with pre-metastatic BC under M1 and M2 point toward the significance of modeling for the time spent with observed and unobserved pre-metastatic BC. The combined impact on the insurance premiums leads to higher premiums under M1, as compared to M2. This is more evident for a contract with shorter maturity. We can see, e.g., in a whole life insurance contract under M1, that the differences in the estimated premiums for women with no BC or pre-metastatic BC diagnosis at the time of purchase get smaller with higher age-at-purchase or age-at-diagnosis and a longer time to maturity (Fig. 12a and b). This can be because the risk of developing a metastatic cancer, after being diagnosed with pre-metastatic BC, becomes fairly stable after about 5 years with a peak at about the first 2 years (also see Fig. 1 in Colzani et al., 2014), where BC risk becomes a relatively lower risk as opposed to other risk factors.

Life insurance premiums 5-year after pre-metastatic BC diagnosis. We have estimated net single premiums for a life insurance contract for a woman diagnosed with pre-metastatic BC 5 years ago at time of purchase, under M1. This aims to determining the impact of eliminating the high-risk years of developing metastatic BC on the net insurance premiums, where we would expect lower insurance premiums by allowing the insurance contract to be tailored in a way to be more inclusive. This calculation can be carried out by modifying (11) as follows:

$$
_{\text{LI},4}\bar{A}_x = \int_0^\infty e^{-\delta t}\,_{t}p_{[x-5]+5}^{11}\mu_{x+t}^{14}dt \tag{12}
$$

$$
+ \int_0^\infty e^{-\delta u}\,_{u}p_{[x-5]+5}^{11}\mu_{[x-5]+5+u}^{13}\int_0^\infty e^{-\delta t}\,_{t}p_{x+u}^{33}\left(\mu_{x+u+t}^{34}+\mu_{x+u+t}^{35}\right)dt\ du,
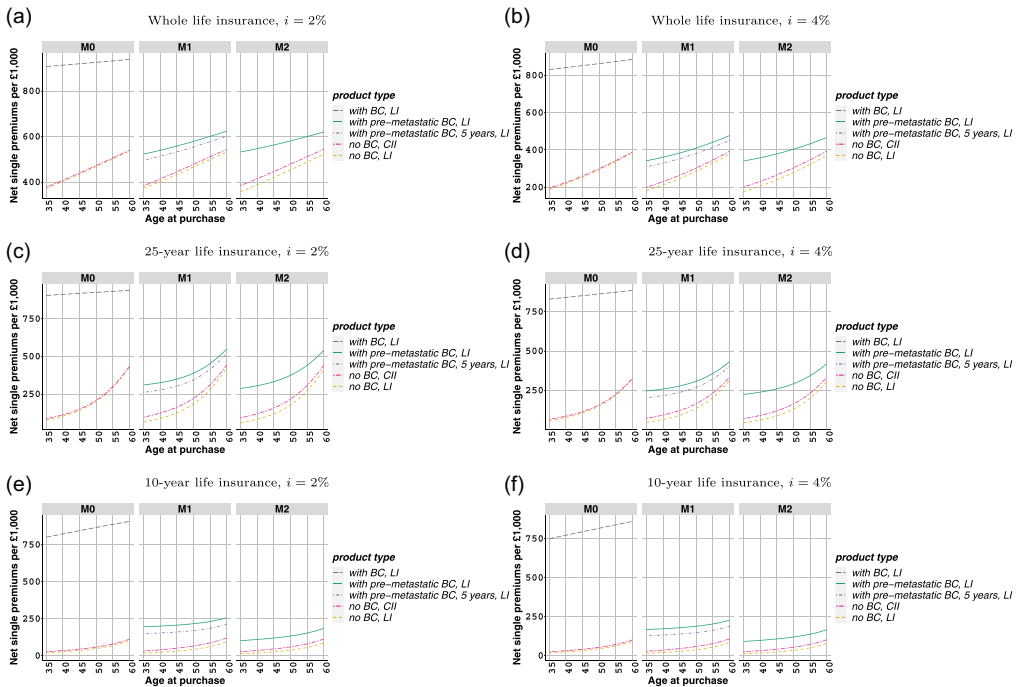$$

Figure 12. Net single premium rates for specialized critical illness, (6)–(7), and life insurance contracts, (8)–(11), and (12), for policyholders with or without breast cancer at the time of purchase, £1,000 benefit, payable at the time of event, based on M0–M2 in Table 1, when $\alpha = 0.6$ and $\beta = 1/7$.

where the upper boundary in the integrals would be set differently for a term life insurance contract.

Our findings suggest lower premiums for a woman after 5 years of BC diagnosis, where smaller differences are observed with an increasing time to maturity. Our results demonstrate intuitive outcomes, aligned with medical observations. Yet, these premiums can be higher than expected by considering, for instance, the premiums for a woman after 5 years of BC diagnosis under "right to be forgotten" initiative (Insurance Europe, 2021). The difference between the estimated premiums under M1 and the premiums, e.g., under the "right to be forgotten" initiative could be linked to two assumptions that we maintained during our calculations: (i) the woman with pre-metastatic BC is assumed to be in State 1 at the time of purchase, where the implicit assumption is that the woman is not free of BC, even after 5 years, but the risk of developing metastatic BC has been considerably reduced and (ii) we have not accounted for time trend in any of the transition rates in our modeling, including the trend in the risk of developing metastatic BC, which may have improved significantly as a result of medical advances (Colzani et al., 2014).

We note that a similar calculation as (12) under M2 would lead to the same results for a woman with pre-metastatic BC diagnosis due to the lack of duration dependence assumption in M2. Also, such calculation would not be possible under M0 since this model does not distinguish between different cancer stages, and it does not account for duration dependence in transition intensities after BC diagnosis.

6. Post-cancer Mortality from breast cancer under Model M0

Under M0, in the absence of data, we have assumed that the rates of transition from BC to death could be determined using the rates based on the risk of death from BC for women with metastatic

Table 3. Modified rates of transition from State 1 to State 3 at different ages in the industry-based model M0

Age	30–49	50–54	55–59	60–64	65–69	70–74	75–79	80–84	85–89
μ_x^{13}	0.00041	0.00115	0.00150	0.00182	0.00214	0.00274	0.00369	0.00497	0.00687

BC (Section 4.1). In other words, in Section 5, these transition rates have been calibrated by considering the risk of death from BC for women with metastatic BC (Zhao et al., 2020). However, State 1 in M0 involves all BC registrations, where rates of transition from State 1 to death from BC, μ_x^{13}, at a given age x, should be determined by considering the risk of death from BC for women with any type of BC.

Although we can assume that death from BC without metastatic BC is rare enough to ignore (Redig & McAllister, 2013), this assumption suggests that everyone in State 1, under M0, has the same risk of death from BC, similar to a woman with metastatic BC. Thus, this leads to unrealistic net premiums, as compared to M1. We therefore consider the sensitivity of the main findings to a different set of μ_x^{13} values.

Giannakeas et al. (2020) estimate the risk of death from BC for women diagnosed with primary ductual carcinoma in situ (DCIS) between 1995 and 2014, based on the SEER database. DCIS is the appearance of cancer cells or tumors within the breast area without showing any presence beyond that area. Hereby, this type of cancer can be associated with early stage BC, such as Stage 1 BC. The risk of dying from BC has been found to be approximately threefold greater than that for women at the same age with no BC in the general population (Giannakeas et al., 2020). Following this study, we can determine μ_x^{13} by using the population mortality from BC in England (Fig. 3b). In particular, we use average mortality from BC between 2001 and 2019, with this being increased by a factor of 3 for all ages (Table 3). Note that the rates in Table 3 accept that everybody in State 1 would be exposed to the same risk of death from BC similar to a woman with DCIS at the time of diagnosis. Although this would lead to lower estimates than expected, we could consider these rates to be, perhaps, a lower boundary to define μ_x^{13}.

Changing rates of transition to death from BC has a significant impact on $\hat{k}_x$ values. As a result of using the rates of transition in Table 3, we have found that the proportion of estimated BC deaths over all deaths has dramatically declined to a level that is lower than that obtained under M1 (Fig. I22 in Appendix I). This result is aligned with the medical literature. The risk of death from BC with DCIS within 20 years is very low, such that approximately 3% of the women with DCIS would be expected to die from BC (Narod et al., 2015). This clearly demonstrates that the rates of transition to death from BC, μ_x^{13}, under M0 have a crucial role in identifying k_x values, and they should be determined with caution.

It is also important to highlight the impact of the change in μ_x^{13} values, on net cancer survival (Fig. I21). As expected, setting μ_x^{13} with respect to the women with DCIS leads to higher cancer survival rates, which is more in line with net cancer survival from a pre-metastatic BC. However, the model seems not to be able to capture the adverse pattern in age-specific cancer survival observed in Fig. 8a.

The impact on life insurance premiums is also relevant, with premiums obtained with μ_x^{13} in Table 3 under M0 being shown in Fig. I23, along with the related insurance premiums under M1 and M2, when $\alpha = 0.6$ and $\beta = 1/7$. As a result of having considerably fewer BC deaths due to the change in μ_x^{13} (see Table 3), the estimated premiums for a life insurance contract, with or without BC diagnosis at the time of purchase, based on M0, are estimated to be considerably lower. For example, they are lower than the premiums for a woman after 5 years of pre-metastatic BC diagnosis under M1. Note that the impact on CII insurance pricing is not discussed here, since the definition of μ_x^{13} is not relevant to the pricing of the CII contract under M0 (see, e.g., (6)).

7. Discussion

We have examined actuarial net premiums for two important insurance products by considering three related models: an industry-based Markov model (M0), a semi-Markov model (M1), and a Markov model (M2), which represents a simplified case of the semi-Markov model. We have obtained net single premiums for a specialized CII contract for healthy women and have compared the estimated premiums under these models. Alongside, we have examined net single premiums for a specialized life insurance contract for women with and without (pre-metastatic) BC. The differences in premiums for a given insurance contract across different models seem to be reduced with increasing age and longer time to maturity.

Our findings under the semi-Markov model are broadly in agreement with the empirical evidence related to net cancer survival from pre-metastatic or metastatic BC, and the proportion of BC deaths over all deaths (Figs. 9, H17 and H20). Furthermore, our work shows that the semi-Markov model has demonstrated insightful results by combining important information, such as cancer stage and the availability of BC diagnostic and treatment services, in a pragmatic way. Our results also show the significance of assuming duration dependence in the modeling, i.e., accounting for the time spent with pre-metastatic BC.

We also note that the overall impact of the COVID-19 pandemic years on insurance premiums is considered by including observations from 2020 to 2022 whilst defining the rates of transition from State 0 to State 1 and State 4, i.e., μ_x^{01} and μ_x^{04}, under M1 and M2 (Section 3). We have found that net single premiums would only differ by less than 1% in all cases, as compared to the earlier results based on the dataset between 2001 and 2019.

There are three major strengths of this study. First, we have demonstrated how to implement a recently developed semi-Markov modeling approach (Arık et al. 2024) in the context of life insurance. Second, we have provided comparisons between this modeling approach and an industry-based model in terms of pricing of BC risk. Last, we have provided separate estimates of net insurance premiums associated with availability of cancer diagnosis and treatment in the semi-Markov model. Our findings point out sensitivity to the model parameters associated with BC diagnosis and treatment only in extreme cases, such as significantly lower BC diagnosis, e.g., 10%, among all BC cases. While overinterpretation of the absolute numerical results should be avoided, our findings suggest that the industry-based model should be approached with caution, since it is particularly sensitive to model assumptions, such as the risk of death from BC for women with BC diagnosis (i.e., definition of μ_x^{13}). Besides, the industry-based model is not able to capture the relationship between age and BC survival in general. Thus, the semi-Markov model, together with our findings, can help life insurers understand the impact of different modeling assumptions on insurance cash flows and pricing calculations. This is relevant when considering new insurance products developed to meet the needs of individuals with medical history of BC.

There are ongoing efforts to develop new medical technologies, such as liquid biopsy, or increase the availability of – and access to –—cancer screening programs in order to improve early cancer diagnosis. In this context, the semi-Markov model can serve as a valuable tool, for instance in quantifying the effect of early BC diagnosis on BC survival. The model could also be extended to include a separate state representing eligibility for a cancer screening program, with clinical diagnosis determined in the model as a separate state (e.g., see Bhatt et al., 2024). This approach would help to estimate the age of onset and quantify the impact of screening on cancer risk. Such extension would be particularly relevant to medical underwriting for related insurance contracts. In particular, the semi-Markov model can help insurers implement a more inclusive approach for underwriting related life insurance as demonstrated in Section 5.3.

Our study has certain limitations. First, in the absence of comprehensive data, we have grouped ages 30–49 to be a single age group, and Stages 1–3 BC to represent pre-metastatic BC. These groupings may be too broad given the nature of BC. Second, we note that our modeling has not taken into account BC potential recovery from pre-metastatic BC, which could have an impact on

exposure in State 0 "No BC." This implies that the exposure in the initial state used in the modeling may have been lower than expected. Furthermore, we have not accounted for a time trend in cancer incidence, type-specific mortality, or the risk of developing metastatic BC. This is particularly relevant in understanding the impact of the COVID-19 pandemic on insurance pricing, as the pandemic significantly affected overall mortality and indirectly influenced cancer incidence in 2020 and 2021. Also, it will be useful to consider competing risks when assessing COVID-19 and its broader implications on the insurance prices. Thus, further research needs to be carried out for incorporating a possible time trend in these aspects of the models. Additionally, we calibrated the models primarily using general population data from England. This may introduce additional risk, as BC risk could differ in insured populations. Insured individuals may belong to higher socioeconomic groups, which are typically associated with better access to healthcare services. This could lead to higher cancer awareness and earlier diagnoses, together with better access to cancer treatment, potentially influencing BC incidence and mortality.

We further highlight the challenge posed by the lack of comprehensive data in determining transition rates within the models under consideration. As a result of this constraint, we rely on point estimates derived from publicly available data in England and medical literature. It is important to acknowledge the uncertainty surrounding transition intensities, which contributes to the overall uncertainty in calculating net single premium rates for insurance contracts. To address uncertainties surrounding BC diagnoses and availability of any treatment, and their implications for insurance premiums, we employ a scenario-based approach. While scenario-driven methods have inherent limitations, we have reinforced our analysis by grounding assumptions in the existing literature and comparing results with available empirical evidence, such as net cancer survival rates provided by the ONS. Last, we have used a constant interest rate when determining net premium rates, rather than variable rates. This is driven by the need to focus on the impact of uncertainty linked to incidence and mortality rates across different models, together with varying modeling assumptions.

References

Alagoz, O., **Lowry**, K., **Kurian**, A., & **Mandelblatt**, J. (2021). Impact of the COVID-19 pandemic on breast cancer mortality in the US: Estimates from collaborative simulation modeling. *Journal of the National Cancer Institute*, **113**(11), 1484–1494.

American Cancer Society (2021). https://www.cancer.org/healthy/cancer-facts/cancer-facts-for-women.html. Cancer facts for women.

Andersen, P., & **Keiding**, N. (2012). Interpretability and importance of functionals in competing risks and multistate models. *Statistics in Medicine*, **31**(11-12), 1074–1088.

Arık, A., **Cairns**, A., **Dodd**, E., **Macdonald**, A., & **Streftaris**, G. (2023). Estimating the impact of the COVID-19 pandemic on breast cancer deaths among older women. In Living to 100 Research Symposium, Society of Actuaries.

Arık, A., **Cairns**, A., **Dodd**, E., **Macdonald**, A., & **Streftaris**, G. (2024). The effect of the COVID-19 health disruptions on breast cancer mortality for older women: A semi-Markov modelling approach. *Scandinavian Actuarial Journal*, **2024**(8), 848–879. doi: 10.1080/03461238.2024.2340964.

Aviva. (2015). UK: Breast cancer accounts for 44% of female critical illness claims. https://www.aviva.com/newsroom/news-releases/2015/09/uk-breast-cancer-accounts-for-44-of-female-critical-illness-claims-17538/.

Baione, F., & **Levantesi**, S. (2018). Pricing critical illness insurance from prevalence rates: Gompertz versus Weibull. *North American Actuarial Journal*, **22**(2), 270–288.

Bhatt, R., **van den Hout**, A., & **Antoniou**, A. C. (2024). Estimation of age of onset and progression of breast cancer by absolute risk dependent on polygenic risk score and other risk factors. *Cancer*, **130**(9), 1590–1599.

Bray, F., **McCarron**, P., & **Parkin**, D. (2004). The changing global patterns of female breast cancer incidence and mortality. *Breast Cancer Research*, **6**(06).

Cho, H., **Mariotto**, A., **Mann**, B., **Klabunde**, C., & **Feuer**, E. (2013). Assessing non-cancer-related health status of US cancer patients: Other-cause survival and comorbidity prevalence. *American Journal of Epidemiology*, **178**(3), 339–349.

Christiansen, M. (2021). On the calculation of prospective and retrospective reserves in non-Markov models. *European Actuarial Journal*, **11**(2), 441–462.

CMI. (1991). Report no. 12, The analysis of permanent health insurance data. Technical report, Faculty of Actuaries and Institute of Actuaries, Edinburgh and London. https://www.actuaries.org.uk/system/files/documents/pdf/cmir12all.pdf.

CMI. (2011). Report no. 52, Cause-specific CMI critical illness diagnosis rates for accelerated business, 2003-2006. Technical report, Faculty of Actuaries and Institute of Actuaries, Edinburgh and London. https://www.actuaries.org.uk/system/files/documents/pdf/cmiwp-52.pdf.

Colzani, E., **Johansson**, A., **Liljegren**, A., & **Foukakis**, T. (2014). Time-dependent risk of developing distant metastasis in breast cancer patients according to treatment, age and tumour characteristics. *British Journal of Cancer*, **110**(5), 1378–1384.

CRUK. (2021a). Survival for all stages of lung cancer. https://www.cancerresearchuk.org/about-cancer/lung-cancer/survival.

CRUK. (2021b). Evidence of the impact of COVID-19 across the cancer pathway: Key stats. Technical report, Cancer Research UK.

Dash, A., & **Grimshaw**, D. (1990). Dread disease cover: An actuarial perspective. Technical report, Continuous Mortality Investigation.

Giannakeas, V., **Sopik**, V., & **Narod**, S. (2020). Association of a diagnosis of ductual carcinoma in situ with death from breast cancer. *JAMA Network Open*, **3**(9).

Henderson, B., **Ross**, R., & **Bernstein**, L. (1988). Estrogens as a cause of human cancer: the Richard and Hinda Rosenthal Foundation Award Lecture. *Cancer Research*, **48**(2), 246–253.

iamInsured. (2023). Breast cancer life insurance. https://iaminsured.co.uk/conditions/breast-cancer-life-insurance/#link-7.

Institute and Faculty of Actuaries (2014). Extending the critical path. https://www.actuaries.org.uk/system/files/documents/pdf/20140217-ifoa-meeting-extending-critical-path-v4.pdf.

Insurance Europe (2021). Beating cancer: Right to be forgotten must account for how insurance works. https://www.insuranceeurope.eu/news/2478/right-to-be-forgotten-must-account-for-how-insurance-works/.

Lai, A., **Pasea**, L., **Banerjee**, A. (2020). Estimated impact of the COVID-19 pandemic on cancer services and excess 1-year mortality in people with cancer and multimorbidity: near real-time data on cancer care, cancer deaths and a population-based cohort study. *BMJ Open*, **10**(11), e043828.

Macdonald, A., **Richards**, S., & **Currie**, I. (2018). *Modelling mortality with actuarial applications*. Cambridge.

Maringe, C., **Spicer**, J., **Morris**, M., **Purushotham**, A., **Nolte**, E., **Sullivan**, R. (2020). The impact of the COVID-19 pandemic on cancer deaths due to delays in diagnosis in England, UK: A national, population-based, modelling study. *The LANCET Oncology*, **21**(8), 1023–1034.

Mariotto, A., **Noone**, A., **Howlader**, N., & **Cho**, H. e. a. (2014). Cancer survival: An overview of measures, uses, and interpretation. *Journal of the National Cancer Institute. Monographs*, **49**(49), 145–186.

McDonald, M., **Hertz**, R., & **Pitman Lowenthal**, S. (2008). The burden of cancer in Asia. https://cdn.pfizer.com/pfizercom/products/cancer_in_asia.pdf.

Narod, S. A., **Iqbal**, J., **Giannakeas**, V., **Sopik**, V., & **Sun**, P. (2015). Breast cancer mortality after a diagnosis of ductal carcinoma in situ. *JAMA Oncology*, **1**(7), 888–896.

ONS. (2016a). Cancer survival in England: Patients diagnosed between 2010 and 2014 and followed up to 2015. Technical report, Office for National Statistics.

ONS. (2016b). One-year net cancer survival for bladder, breast, colorectal, kidney, lung, melanoma, ovary, prostate and uterus, by stage at diagnosis. https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/conditionsanddiseases/datasets/oneyearnetcancersurvivalforbladderbreastcolorectalkidneylungmelanomaovaryprostateanduterusbystageatdiagnosis.

ONS. (2016c). Cancer survival by stage at diagnosis for England (experimental statistics): Adults diagnosed 2012, 2013 and 2014 and followed up to 2015. Technical report, Office for National Statistics.

ONS. (2019a). Cancer survival in England: national estimates for patients followed up to 2017. Technical report, Office for National Statistics.

ONS. (2019b). Cancer survival statistical bulletins QMI. Technical report, Office for National Statistics.

Ozkok Dodd, E., **Streftaris**, G., **Waters**, H., & **Stott**, A. (2014). The effect of model uncertainty on the pricing of critical illness insurance. *Annals of Actuarial Science*, **9**(1), 108–133.

Redig, A., & **McAllister**, S. (2013). Breast cancer as a systemic disease: A view of metastasis. *Journal of Internal Medicine*, **274**(2), 113–126.

Reynolds, **C.**, & **Faye**, **M.** (2016). CI pricing detectives. Technical report, Partner Reviews. https://partnerre.com/wp-content/uploads/2017/08/CI_Pricing_Detectives.pdf.

SAS. (2011). Singapore insured lives: Mortality investigation 2004-2008. https://www.actuaries.org.sg/files/library/other/Other%20Reports/MortalityInvestigationReport%2020111103v18.pdf.

SCOR. (2020). Insuring more cancer survivors through inclusive underwriting. https://www.scor.com/en/expert-views/insuring-more-cancer-survivors-through-inclusive-underwriting.

Soetewey, **A.**, **Legrand**, **C.**, **Denuit**, **M.**, & **Silversmit**, **G.** (2022). Semi-markov modeling for cancer insurance. *European Actuarial Journal*, **12**(2), 813–837. doi: 10.1007/s13385-022-00308-2.

Soetewey, **A.**, **Legrand**, **C.**, **Denuit**, **M.**, & **Silversmit**, **G.** (2023). Right to be forgotten for mortgage insurance issued to cancer survivors: critical assessment and new proposal. Universite Catholique de Louvain, Institute of Statistics, Biostatistics and Actuarial Sciences.

Sud, **A.**, **Torr**, **B.**, **Jones**, **M.** (2020). Effect of delays in the 2-week-wait cancer referral pathway during the COVID-19 pandemic on cancer survival in the UK: A modelling study. The LANCET: Oncology.

Sung, **H.**, **Ferlay**, **J.**, **Siegel**, **R.**, **Laversanne**, **M.**, **Soerjomataram**, **I.**, **Jemal**, **A.**, & **Bray**, **F.** (2021). Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: A Cancer Journal for Clinicians*, **71**(3), 209–249.

Swaminathan, **R.**, & **Brenner**, **H.** (2011). Statistical methods for cancer survival analysis. https://survcan.iarc.fr/survival/chap2.pdf.

The Insurance Surgery (2023). https://www.the-insurance-surgery.co.uk/medical-conditions-life-insurance/breast-cancer-life-insurance/.Breast cancer life insurance.

Wood, **S.** (2017). *Generalized additive models: An introduction with R*. Chapman and Hall/CRC.

Zhao, **Y.**, **Xu**, **G.**, **Guo**, **X.** (2020). Early death incidence and prediction in Stage IV breast cancer. *Medical Science Monitor*, **26**.

Appendix

A. Kolmogorov equations in the alternative model

Kolmogorov equations for the industry-based Markov model are given as follows:

$$\frac{d}{dt}{}_tp_x^{00} = -{}_tp_x^{00}\left(\mu_{x+t}^{01} + \mu_{x+t}^{02}\right)$$

$$\frac{d}{dt}{}_tp_x^{01} = {}_tp_x^{00}\mu_{x+t}^{01} - {}_tp_x^{01}\left(\mu_{x+t}^{12} + \mu_{x+t}^{13}\right)$$

$$\frac{d}{dt}{}_tp_x^{02} = {}_tp_x^{00}\mu_{x+t}^{02} + {}_tp_x^{01}\mu_{x+t}^{12}$$

$$\frac{d}{dt}{}_tp_x^{03} = {}_tp_x^{01}\mu_{x+t}^{13}$$

$$\frac{d}{dt}{}_tp_x^{11} = -{}_tp_x^{11}\left(\mu_{x+t}^{12} + \mu_{x+t}^{13}\right)$$

$$\frac{d}{dt}{}_tp_x^{12} = {}_tp_x^{11}\mu_{x+t}^{12}$$

$$\frac{d}{dt}{}_tp_x^{13} = {}_tp_x^{11}\mu_{x+t}^{13}$$

The appropriate boundary conditions are ${}_0p_x^{00} = {}_0p_x^{11} = 1$; ${}_0p_x^{01} = {}_0p_x^{02} = {}_0p_x^{03} = 0$; and ${}_0p_x^{12} = {}_0p_x^{13} = 0$.

B. Modified Kolmogorov equations with duration dependence in the Semi-Markov model

Modified Kolmogorov equations for the semi-Markov BC model are given as below. Note that more details can be found in CMI (1991), based on a 3-state multiple model, allowing recovery from the disease under inspection along with duration dependence. Here, in order to make

integrals clearer, we introduce actuarial selection notation. For instance, $\mu_{x,t}^{13}$ is shown based on select attained age $[x]$ with duration t, specifically $\mu_{x,t}^{13} = \mu_{[x]+t}^{13}$.

$$\frac{d}{dt}{}_tp_x^{00} = -{}_tp_x^{00}\left(\mu_{x+t}^{01} + \mu_{x+t}^{02} + \mu_{x+t}^{04}\right)$$

$$\frac{d}{dt}{}_tp_x^{01} = {}_tp_x^{00}\mu_{x+t}^{01} - {}_tp_x^{01}\ \mu_{x+t}^{14} - \int_{u=0}^{t} {}_up_x^{00}\ \mu_{x+u}^{01}\ {}_{t-u}p_{[x+u]}^{11}\ \mu_{[x+u]+t-u}^{13}\ du$$

$$\frac{d}{dt}{}_tp_x^{02} = {}_tp_x^{00}\mu_{x+t}^{02} - {}_tp_x^{02}\mu_{x+t}^{24} - \int_{u=0}^{t} {}_up_x^{00}\ \mu_{x+u}^{02}\ {}_{t-u}p_{[x+u]}^{22}\ \mu_{[x+u]+t-u}^{23}\ du$$

$$\frac{d}{dt}{}_tp_x^{03} = \int_{u=0}^{t} {}_up_x^{00}\ \mu_{x+u}^{01}\ {}_{t-u}p_{[x+u]}^{11}\ \mu_{[x+u]+t-u}^{13}\ du+$$

$$\int_{u=0}^{t} {}_up_x^{00}\ \mu_{x+u}^{02}\ {}_{t-u}p_{[x+u]}^{22}\ \mu_{[x+u]+t-u}^{23}\ du - {}_tp_x^{03}\left(\mu_{x+t}^{34} + \mu_{x+t}^{35}\right)$$

$$\frac{d}{dt}{}_tp_x^{04} = {}_tp_x^{00}\mu_{x+t}^{04} + {}_tp_x^{01}\mu_{x+t}^{14} + {}_tp_x^{02}\mu_{x+t}^{24} + {}_tp_x^{03}\mu_{x+t}^{34}$$

$$\frac{d}{dt}{}_tp_x^{05} = {}_tp_x^{03}\mu_{x+t}^{35}$$

We note that the select notation on age $[x]$ is kept in the equations below, where this is based on the assumption of being in the relevant initial state.

$$\frac{d}{dt}{}_tp_{[x]}^{11} = -{}_tp_{[x]}^{11}\left(\mu_{[x]+t}^{13} + \mu_{[x]+t}^{14}\right)$$

$$\frac{d}{dt}{}_tp_{[x]}^{13} = {}_tp_{[x]}^{11}\ \mu_{[x]+t}^{13} - {}_tp_{[x]}^{13}\mu_{[x]+t}^{34} - {}_tp_{[x]}^{13}\mu_{[x]+t}^{35}$$

$$\frac{d}{dt}{}_tp_{[x]}^{14} = {}_tp_{[x]}^{11}\ \mu_{[x]+t}^{14} + {}_tp_{[x]}^{13}\mu_{[x]+t}^{34}$$

$$\frac{d}{dt}{}_tp_{[x]}^{15} = {}_tp_{[x]}^{13}\mu_{[x]+t}^{35}$$

$$\frac{d}{dt}{}_tp_{[x]}^{22} = -{}_tp_{[x]}^{22}\left(\mu_{[x]+t}^{23} + \mu_{[x]+t}^{24}\right)$$

$$\frac{d}{dt}{}_tp_{[x]}^{23} = {}_tp_{[x]}^{22}\ \mu_{[x]+t}^{23} - {}_tp_{[x]}^{23}\left(\mu_{[x]+t}^{34} + \mu_{[x]+t}^{35}\right)$$

$$\frac{d}{dt}{}_tp_{[x]}^{24} = {}_tp_{[x]}^{22}\ \mu_{[x]+t}^{24} + {}_tp_{[x]}^{23}\ \mu_{[x]+t}^{34}$$

$$\frac{d}{dt}{}_tp_{[x]}^{25} = {}_tp_{[x]}^{23}\ \mu_{[x]+t}^{35}$$

$$\frac{d}{dt}{}_tp_{[x]}^{33} = -{}_tp_{[x]}^{33}\left(\mu_{[x]+t}^{34} + \mu_{[x]+t}^{35}\right)$$

$$\frac{d}{dt}{}_tp_{[x]}^{34} = {}_tp_{[x]}^{33}\ \mu_{[x]+t}^{34}$$

$$\frac{d}{dt}{}_tp_{[x]}^{35} = {}_tp_{[x]}^{33}\ \mu_{[x]+t}^{35}$$

The appropriate boundary conditions are ${}_0p_x^{00} = {}_0p_{[x]}^{11} = {}_0p_{[x]}^{22} = {}_0p_{[x]}^{33} = 1$; ${}_0p_x^{01} = {}_0p_x^{02} = {}_0p_x^{03} = {}_0p_x^{04} = {}_0p_x^{05} = 0$; ${}_0p_{[x]}^{13} = {}_0p_{[x]}^{14} = {}_0p_{[x]}^{15} = 0$; ${}_0p_{[x]}^{23} = {}_0p_{[x]}^{24} = {}_0p_{[x]}^{25} = 0$; and ${}_0p_{[x]}^{34} = {}_0p_{[x]}^{35} = 0$.

C. Runge–Kutta method for breast cancer model in Fig. 2

Runge–Kutta methods estimate function values in a given small interval, and then use those values to obtain a better estimate of the function under inspection. A fourth-order Runge–Kutta scheme is based on four recursive estimates of the increment in the function value per time step (Macdonald et al., 2018).

We have a 6-state model in Fig. 2, and hence, in full, a 6×6 matrix of occupancy probabilities denoted by ${}_h p_x^{ij} \equiv y_t$ as

$$
y_t = \begin{bmatrix} {}_t p_x^{00} \\ {}_t p_x^{01} \\ {}_t p_x^{02} \\ \vdots \\ {}_t p_x^{55} \end{bmatrix}, \quad \frac{d}{dt} y_t = \begin{bmatrix} \frac{d}{dt} {}_t p_x^{00} \\ \frac{d}{dt} {}_t p_x^{01} \\ \frac{d}{dt} {}_t p_x^{02} \\ \vdots \\ \frac{d}{dt} {}_t p_x^{55} \end{bmatrix} = f(t, y_t).
$$

Now, suppose we would like to solve $\frac{d}{dt} y_t = f(t, y_t)$, $y_{t_0} = y_0$. Then, we could write

$$
y_{t_{n+1}} = y_{t_n} + \frac{h}{6}\left(k_1 + 2k_2 + 2k_3 + k_4 \right),
$$

for $t_{n+1} = t_n + h$ and

$$
k_1 = f(t_n, y_{t_n})
$$
$$
k_2 = f\left(t_n + \frac{h}{2}, y_{t_n} + h\frac{k_1}{2} \right)
$$
$$
k_3 = f\left(t_n + \frac{h}{2}, y_{t_n} + h\frac{k_2}{2} \right)
$$
$$
k_4 = f\left(t_n + h, y_{t_n} + hk_3 \right).
$$

Here, the four intermediate steps, denoted by k_1, k_2, k_3, and k_4, are also vector quantities such that

$$
k_1 = \begin{bmatrix} k_1^{00} \\ k_1^{01} \\ k_1^{02} \\ \vdots \\ k_1^{55} \end{bmatrix}, \quad k_2 = \begin{bmatrix} k_2^{00} \\ k_2^{01} \\ k_2^{02} \\ \vdots \\ k_2^{55} \end{bmatrix}, \quad k_3 = \begin{bmatrix} k_3^{00} \\ k_3^{01} \\ k_3^{02} \\ \vdots \\ k_3^{55} \end{bmatrix}, \quad k_4 = \begin{bmatrix} k_4^{00} \\ k_4^{01} \\ k_4^{02} \\ \vdots \\ k_4^{55} \end{bmatrix}.
$$

D. Computation of insurance premiums in the Semi-Markov model

We take the net single premium for the CII contract, without a death rider, described in (7), as an example, that is

$$
{}_{CI,2}\bar{A}_x = \int_0^\infty e^{-\delta t} {}_t p_x^{00} \mu_{x+t}^{01} dt
$$
$$
+ \int_0^\infty e^{-\delta u} {}_u p_x^{00} \mu_{x+u}^{02} \int_0^\infty e^{-\delta t} {}_t p_{[x+u]}^{22} \mu_{[x+u]+t}^{23} dt \, du.
$$

The first component in the equation above is straightforward to carry out. That is why we focus on the second component between time zero and nh which is

$$
\int_{u=0}^{T=nh} e^{-\delta u} {}_u p_x^{00} \mu_{x+u}^{02} \int_{t=0}^{nh-u} e^{-\delta t} {}_t p_{[x+u]}^{22} \mu_{[x+u]+t}^{23} dt \, du,
$$

where we can simplify this expression by denoting the inner integral by $g(x + u, T - u)$ as

$$\int_{u=0}^{T=nh} e^{-\delta u} {}_u p_x^{00} \mu_{x+u}^{02} \ g(x + u, T - u)du,$$

and we can approximately calculate this part of the premium by choosing each time step to be equal to h, e.g., based on trapezoidal rule, as follows:

$$\int_{u=0}^{T=nh} e^{-\delta u} {}_u p_x^{00} \mu_{x+u}^{02} \ g(x + u, T - u)du \approx \frac{h}{2} \left(e^{-\delta 0} {}_0 p_x^{00} \ \mu_x^{02} \ g(x, nh) \right.$$

$$+ 2e^{-\delta h} {}_h p_x^{00} \ \mu_{x+h}^{02} \ g(x + h, nh - h)$$

$$+ 2e^{-\delta 2h} {}_{2h} p_x^{00} \ \mu_{x+2h}^{02} \ g(x + 2h, nh - 2h) + \dots$$

$$\left. + e^{-\delta nh} {}_{nh} p_x^{00} \ \mu_{x+nh}^{02} \ g(x + nh, nh - nh) \right).$$

Here, the important thing to notice is that $g(x + u, T - u)$ shows different net single premiums to be paid at the time of diagnosis for policyholders with unobserved BC at selected ages $[x + u]$.

E. Net single premiums for different insurance contracts under M2

Section 5.1 explains how to calculate the net single premiums for a special CII contract and a couple life insurance contracts based on M0 and M1. This section further explains how to calculate net single premiums for these contracts under M2.

The net single premium for the CII contract, formulated in (7), can be redefined under M2 as

$$_{CI,2}\bar{A}_x = \int_0^\infty e^{-\delta t} {}_t p_x^{00} \left(\mu_{x+t}^{01} + \mu_{x+t}^{04} \right) dt \tag{E1}$$

$$+ \int_0^\infty e^{-\delta t} {}_t p_x^{02} \left(\mu_{x+t}^{23} + \mu_{x+t}^{24} \right) dt.$$

The net single premiums for the life insurance contracts for a woman without BC, (9), and with pre-metastatic BC, (11), can be redefined under M2, as well, where (E2) shows the premium for a healthy woman at the time of purchase as

$$_{LI,2}\bar{A}_x = \int_0^\infty e^{-\delta t} {}_t p_x^{00} \mu_{x+t}^{04} + {}_t p_x^{01} \mu_{x+t}^{14} + {}_t p_x^{02} \mu_{x+t}^{24} + {}_t p_x^{03} \left(\mu_{x+t}^{34} + \mu_{x+t}^{35} \right) dt, \tag{E2}$$

whereas (E3) shows the premium for a woman with a pre-metastatic BC diagnosis at the time of purchase as

$$_{LI,4}\bar{A}_x = \int_0^\infty e^{-\delta t} {}_t p_x^{11} \mu_{x+t}^{14} dt \tag{E3}$$

$$+ \int_0^\infty e^{-\delta t} {}_t p_x^{13} \left(\mu_{x+t}^{34} + \mu_{x+t}^{35} \right) dt.$$

F. Deriving an expression for μ_x^{02} in the absence of cause of death data

All-cause mortality, denoted by $\hat{\mu}_x$, can be estimated as follows:

$$\hat{\mu}_x = \frac{D_x}{E_x^0 + E_x^1},$$

with $D_x = D_x^{02} + D_x^{12} + D_x^{13}$ based on M0. Let us accept that deaths from BC at age x, D_x^{13}, are a percentage ($k_x\%$) of all deaths, such that $D_x^{13} = k_x D_x$. Thus,

$$D_x^{02} + D_x^{12} = D_x - k_x D_x.$$

Provided our assumption under M0 in Section 4.1, we can write

$$\hat{\mu}_x^{12} = \hat{\mu}_x^{02} \Rightarrow \frac{D_x^{12}}{E_x^1} = \frac{D_x^{02}}{E_x^0} \Rightarrow D_x^{12} = D_x^{02} \frac{E_x^1}{E_x^0},$$

and

$$D_x^{02} + D_x^{12} = D_x - k_x D_x \Rightarrow D_x^{02} + D_x^{02} \frac{E_x^1}{E_x^0} = (1 - k_x) D_x$$

$$\frac{D_x^{02}}{E_x^0} = (1 - k_x) \frac{D_x}{E_x^0 + E_x^1},$$

that leads to

$$\hat{\mu}_x^{02} = (1 - k_x) \hat{\mu}_x,$$

where rates of transition to death from other causes, $\hat{\mu}_x^{02}$, are indirectly determined by using all-cause mortality, $\hat{\mu}_x$, and k_x.

G. Different diagnosis and treatment considerations when $i = 4\%$
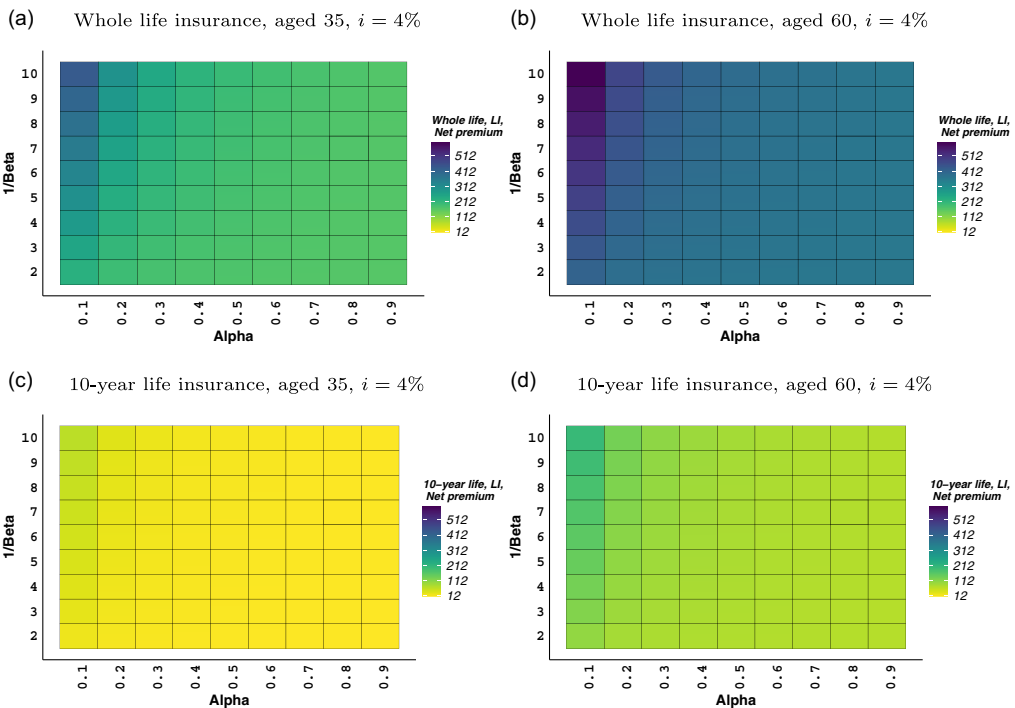
Figure G13. Net single premium rates for a specialized life insurance contract, (9), for policyholders without breast cancer at the time of purchase, £1,000 benefit, payable at the time of death, based on the semi-Markov model M1 with $i = 4\%$.
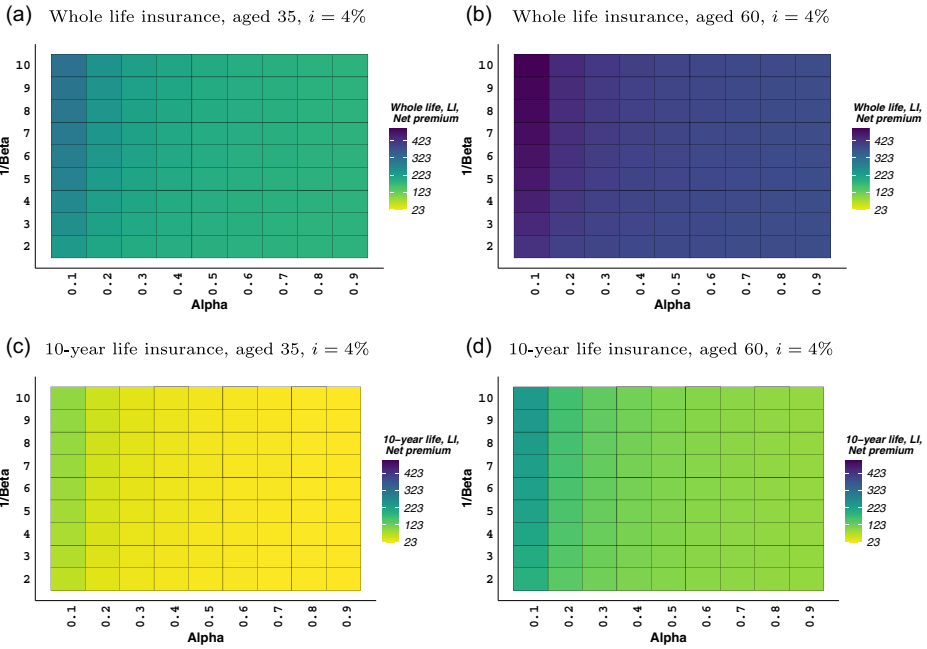
Figure G14. Net single premium rates for a specialized CII contract, (7), for policyholders without breast cancer at the time of purchase, £1,000 benefit, payable at the time of event, based on the semi-Markov model M1 with $i = 4\%$.
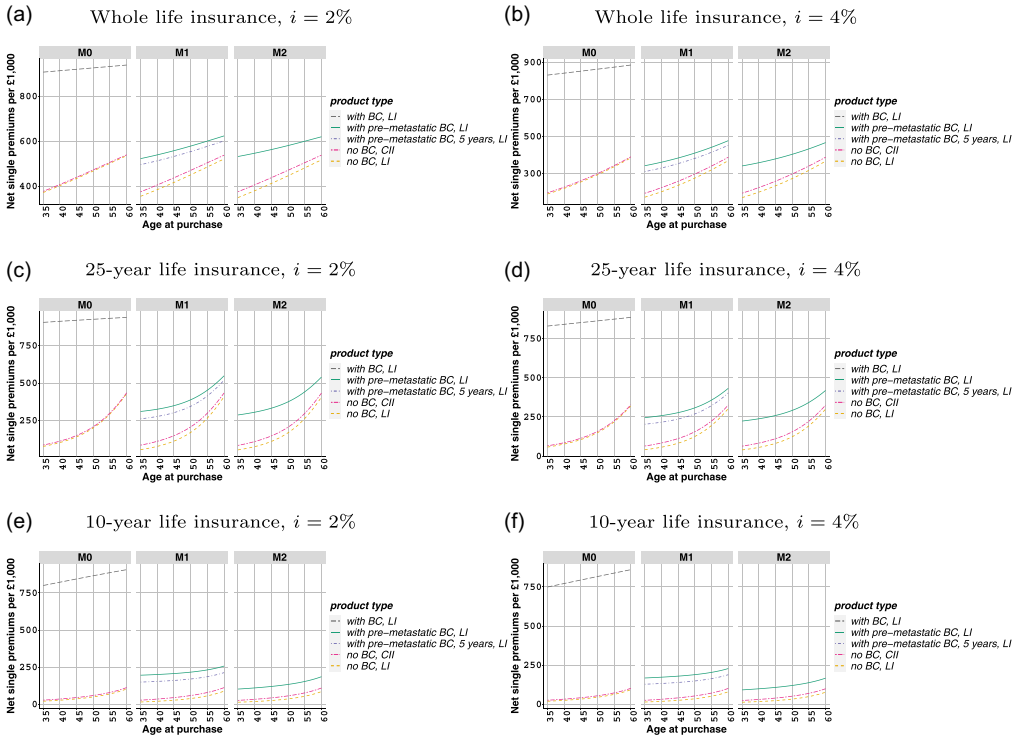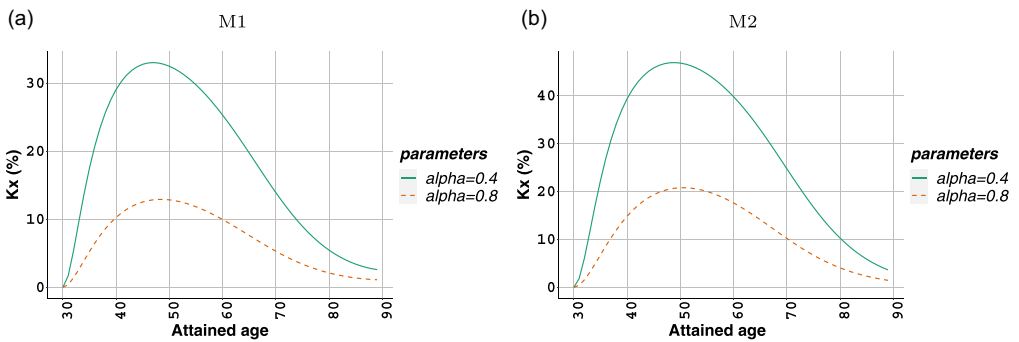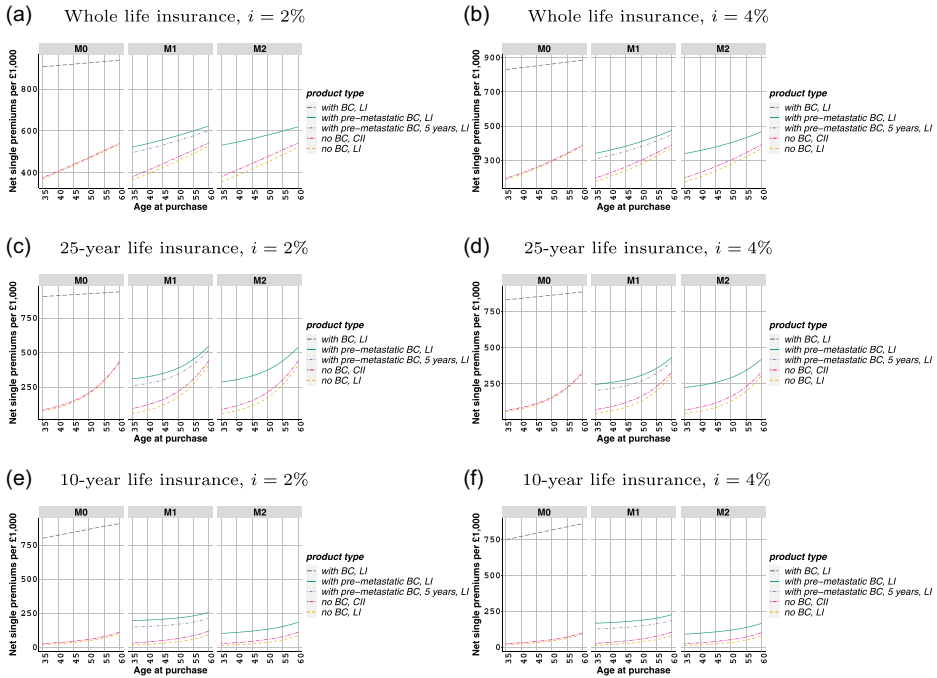
H. Further findings in Section 5.3
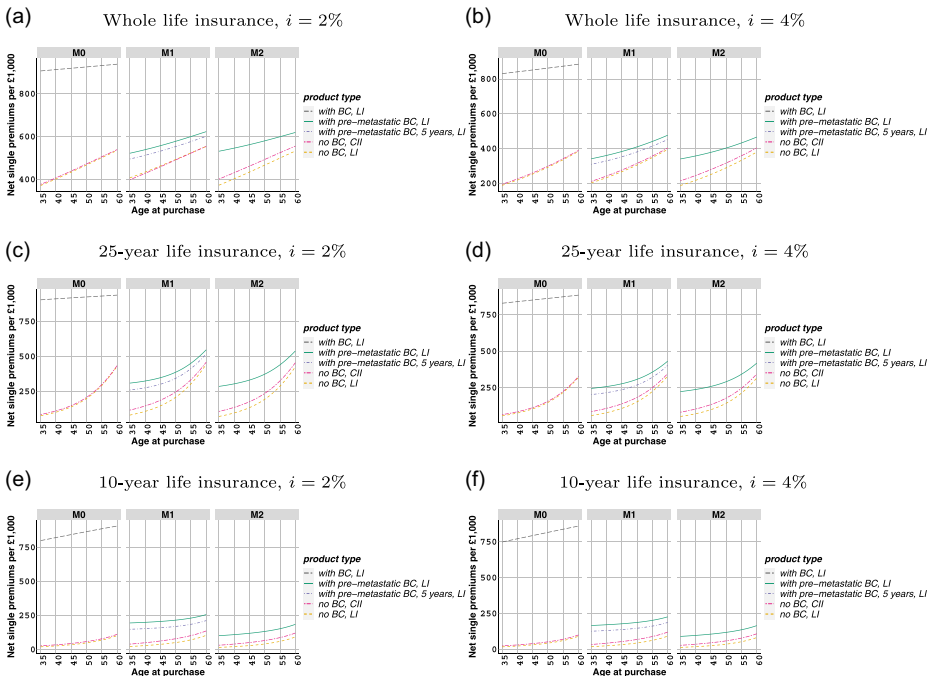
Figure H15. Net single premium rates for specialized critical illness, (6)–(7), and life insurance contracts, (8)–(11), and (12), for policyholders with or without breast cancer at the time of purchase, £1,000 benefit, payable at the time of event, based on M0–M2 in Table 1, when $\alpha = 0.4$ and $\beta = 1/7$.
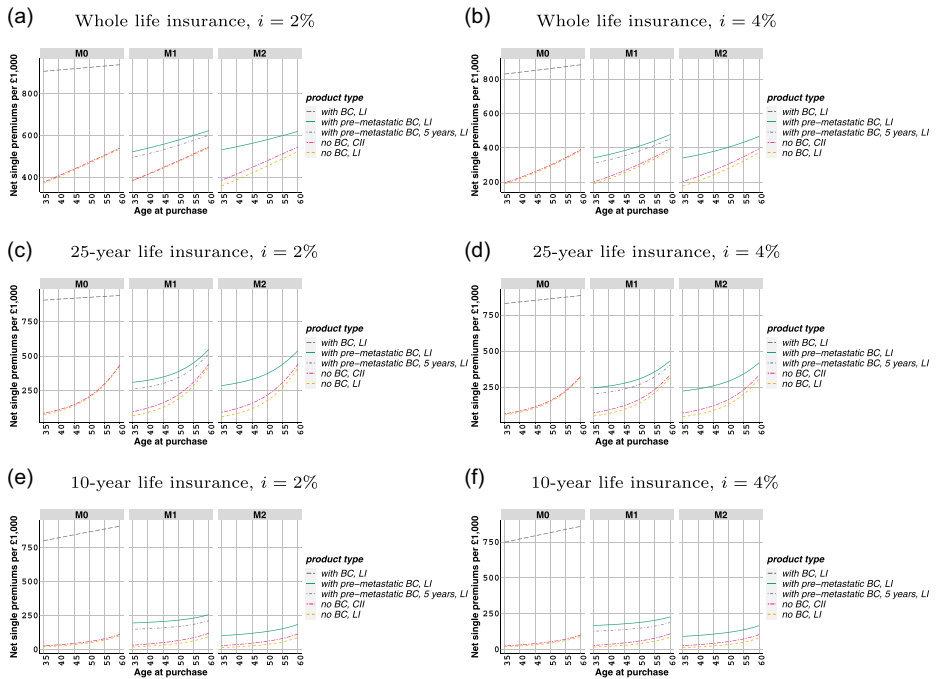
Figure H16. Net single premium rates of specialized critical illness, (6)–(7), and life insurance contracts, (8)–(11), and (12), for policyholders with or without breast cancer at the time of purchase, £1,000 benefit, payable at the time of event, based on M0–M2 in Table 1, when $\alpha = 0.8$ and $\beta = 1/7$.
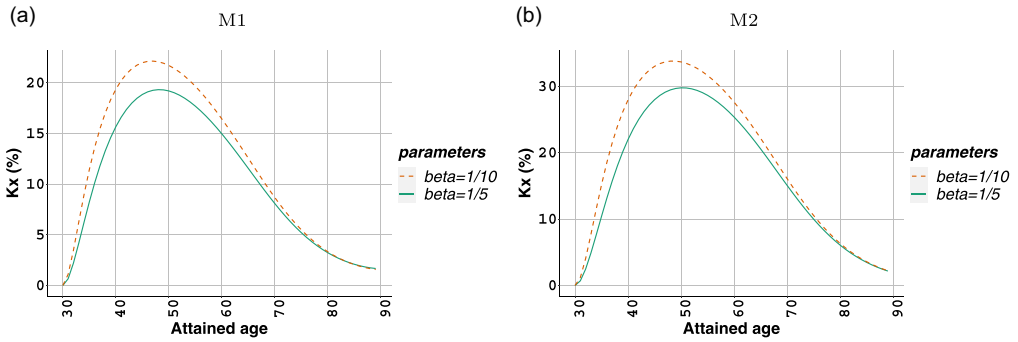

Note: Attained age = Age-at-entry + Time

Figure H17. Estimated $\hat{k}_x$ values for a policyholder aged 30 years, with no breast cancer, at time zero, based on M1 and M2, when $\alpha = 0.4$ or $\alpha = 0.8$ and $\beta = 1/7$.

Figure H18. Net single premium rates for specialized critical illness, (6)–(7), and life insurance contracts, (8)–(11), and (12), for policyholders with or without breast cancer at the time of purchase, £1,000 benefit, payable at the time of event, based on M0–M2 in Table 1, when $\alpha = 0.6$ and $\beta = 1/5$.

Figure H19. Net single premium rates for specialized critical illness, (6)–(7), and life insurance contracts, (8)–(11), and (12), for policyholders with or without breast cancer at the time of purchase, £1,000 benefit, payable at the time of event, based on M0–M2 in Table 1, when $\alpha = 0.6$ and $\beta = 1/10$.

(a) M1

(b) M2

Note: Attained age = Age-at-entry + Time

Figure H20. Estimated $\hat{k}_x$ values for a policyholder aged 30 years, with no breast cancer, at time zero, based on M1 and M2, when $\beta = 1/5$ or $\beta = 1/10$ and $\alpha = 0.6$.
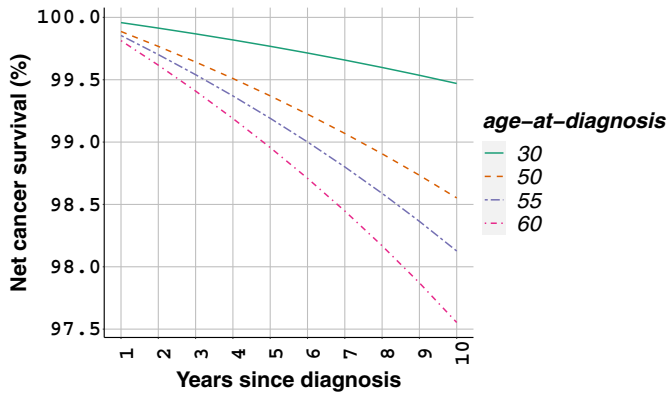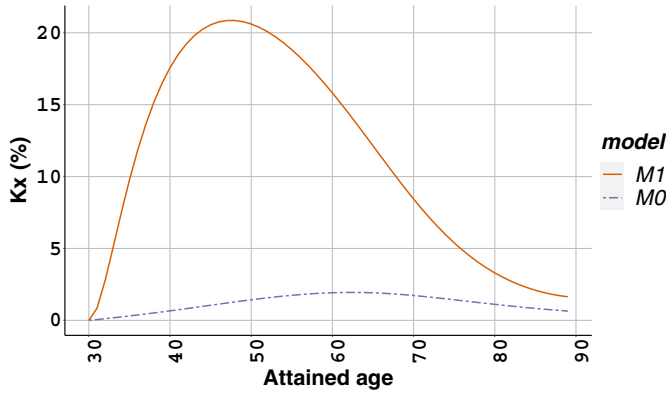
I. Further findings in Section 6

Figure I21. Estimated net cancer survival for a woman diagnosed with breast cancer at different ages under M0.

Note: Attained age = Age-at-entry + Time

Figure I22. Estimated $\hat{k}_x$ values based on M0 and M1, when $\alpha = 0.6$ and $\beta = 1/7$.
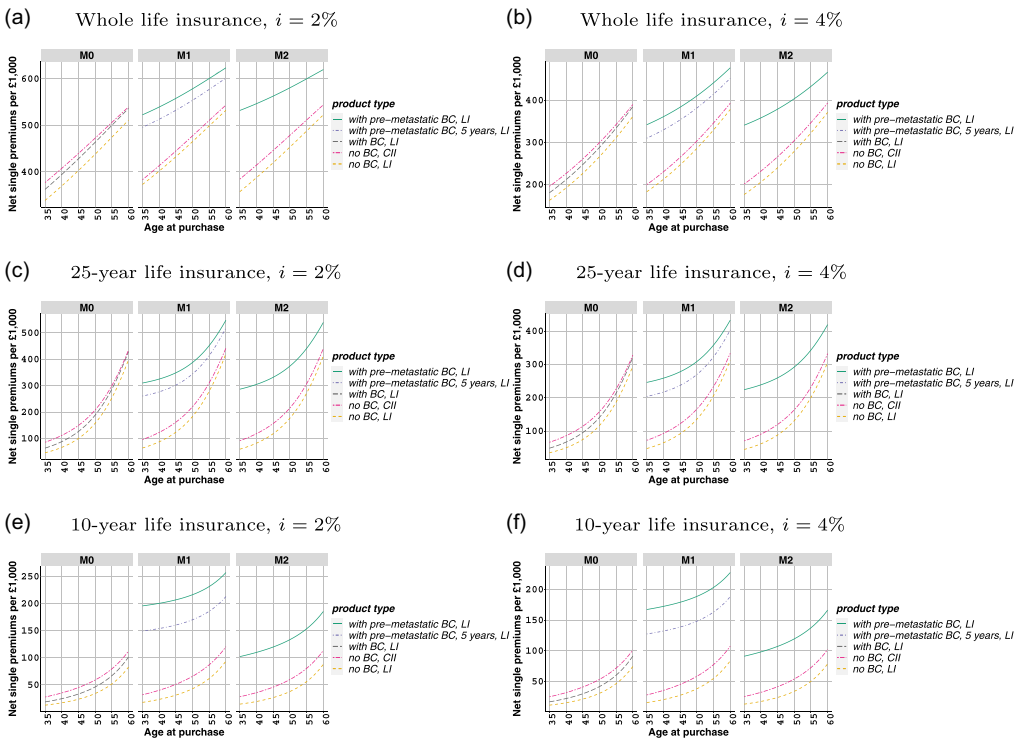

Figure I23. Net single premium rates for specialized critical illness, (6)–(7), and life insurance contracts, (8)–(11), and (12), for policyholders with or without breast cancer at the time of purchase, £1,000 benefit, payable at the time of event, based on M0–M2 in Table 1, when $\alpha = 0.6$ and $\beta = 1/7$, and μ_x^{13} under M0 based on Table 3.