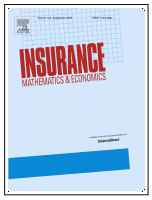
Insurance Mathematics and Economics 126 (2026) 103194

Contents lists available at ScienceDirect

Insurance: Mathematics and Economics

journal homepage: www.elsevier.com/locate/ime



An age–period–cohort model in a Dirichlet framework: A coherent causes of death estimation

Rebecca Graziani [a], Andrea Nigri [b,*]

[a] Department of Decision Sciences, Dondena Center and Bidsa, Bocconi University, Milan, Italy
[b] Department of Social Sciences, University of Foggia, Foggia, Italy

ARTICLE INFO

JEL classification:
C10
C11
J11
I14

Keywords:
Cause of death
Compositional data
Age–period–cohort model
Mixed model

ABSTRACT

Though pivotal in longevity studies, multi-outcome modelling is largely neglected in the associated statistical literature. Here, we focus on the case of compositional data, especially relevant in longevity analysis, where overall mortality can be described as the composition of several causes of death. We propose an age–period–cohort model within the Dirichlet framework with a specific interest in its use for modelling longevity with multiple causes of death. We introduce a flexible approach to incorporating the Dirichlet distribution into the age–period–cohort framework. Then, using US cause-specific mortality data, we provide a comprehensive discussion and comparison of alternative modelling approaches.

1. Introduction

The rise in longevity resulting from medical and social progress represents an important achievement of modern society. Nevertheless, longevity comes with substantial challenges because of the increased number of pensioners compared to taxpayers and the increased need for disease care, among other factors. Longevity modelling thus plays a central role in shaping policy decisions. Detecting trends and changes, by age and time, in cause-specific death data is an important input into public health policies. Understanding whether a cause of death is increasing or decreasing lets policy-makers better target healthcare services. In demographic studies, as stressed by Preston (2013), recognising cause-of-death patterns helps in identifying the determinants of population mortality structures. In particular, cause-of-death data can be used to detect changes in mortality trends that might otherwise be hidden at the population level. For a broader discussion of the advantages and limitations of decomposing mortality by cause of death, see, among others, Booth and Tickle (2008) and Richards (2009).

Causes of death are defined to be mutually exclusive and exhaustive but, because the true (net) dependence structure is impossible to identify (Tsiatis, 1975), since the early work of Chiang (1968), it has been common practice to separately and independently model and forecast causes of death (see e.g. Peltonen and Asplund, 1996; Knorr-Held and Rainer, 2001; Quinn et al., 2003; Di Cesare and Murphy, 2009; Mathers and Loncar, 2006). Such independent modelling, especially as the basis for a forecasting exercise, might produce an implausible cause-specific mortality estimation, ultimately inconsistent with the total mortality and leading to more pessimistic results than an all-causes-combined estimate (Wilmoth, 1995).

There have been some attempts to jointly model and forecast cause-of-death mortality data and account for their competing risk nature. In Manton and Poss (1979) and Arnold and Sherris (2013), correlation between cause-specific mortality rates is induced by their joint dependence on the same individual risk factors; Foreman et al. (2017) use a Bayesian hierarchical model with conditional autoregressive (CAR) components to jointly forecast death rates from a parsimonious and collectively exhaustive set of causes of death at the subnational level; Manton et al. (1986) employ frailty models to induce dependence between the various causes; Kaishev et al. (2007) and Dimitrova et al. (2013) jointly model causes of death through copulas. More recently, Lanfiuti Baldi et al. (2025) proposed a Skellam-based approach to model differences in death counts between two major causes of death.

In this paper, we jointly model cause-of-death data by drawing on their compositional nature. Cause-of-death data are compositional because death counts or proportions are positive and subject to a sum constraint, due to the competing risks between causes. Alai et al. (2015)

* Corresponding author.
E-mail addresses: rebecca.graziani@unibocconi.it (R. Graziani), andrea.nigri@unifg.it (A. Nigri).

https://doi.org/10.1016/j.insmatheco.2025.103194
Received 6 November 2024; Received in revised form 17 November 2025; Accepted 19 November 2025
Available online 24 November 2025

works along these lines by suggesting a multinomial logistic regression to jointly model death counts for different causes; in Stefanucci and Mazzuco (2022), a compositional functional analysis is suggested to recover cause-specific mortality trends; Oeppen et al. (2008), Bergeron-Boucher et al. (2017), Kjærgaard et al. (2019), and Cardillo et al. (2023) use life table deaths to forecast cause-specific mortality based on compositional data analysis. In the last set of works, following Aitchison (1986), a centred log transformation is applied to life table deaths to overcome the sum constraint and apply Lee–Carter-like forecasting procedures.

Our proposal is to avoid any transformation and directly model cause-specific death proportions that are easier to interpret. In this paper, we propose jointly modelling cause-of-death proportions through a multivariate Dirichlet distribution. In compositional data analysis, Dirichlet regression models are used as an alternative to classical multivariate regression analysis on log transformations; (see e.g. Gueorguieva et al., 2008; Hijazi and Jernigan, 2009; Van der Merwe, 2019). In particular, Grunwald et al. (1993), Brunsdon and Smith (1998), and Mills (2010) rely on the Dirichlet distribution to model time series of continuous proportions.

Our aim is to disentangle the effect that the three time dimensions, age, period, and cohort, might separately have on the joint dynamics of cause-of-death proportions. The age–period–cohort (APC) approach is widely used to model total mortality, starting from the classical APC model discussed in Hobcraft et al. (1985) (see e.g. Renshaw and Haberman, 2003, 2006; Currie, 2006; Cairns et al., 2009; Plat, 2009; Hunt and Blake, 2014; Richards et al., 2019; Billari and Graziani, 2024). However, to the extent of our knowledge, this is the first time an APC model has been used in a Dirichlet regression framework to jointly model cause-of-death data. Undoubtedly, the cohort effect holds significant importance in the analysis of factors influencing mortality. As highlighted by Kjærgaard et al. (2019), the integration of cohort effects presents challenges due to the intricate interplay between age, time, and cohort. Notably, the prevailing methodologies on simultaneous cause-of-death analysis do not incorporate cohort. Integrating cohort components could enhance both the accuracy of fitting models and the precision of forecasts within certain populations.

In light of that, the main contributions of our study are twofold. First, we contribute to the literature on the analysis of cause-specific mortality data by outlining an approach that draws on their compositional nature to model them jointly, and, at the same time, to recover their age, period, and cohort dynamics. Second, we explore the performance of the hierarchical or mixed approach in APC modelling which, with the well-known limitations that we discuss later the paper, turns out to be easy to implement through user-friendly routines. The remainder of the paper is organized as follows: Section 2 introduces the Dirichlet modelling. Section 4 describes the motivating dataset. Section 5 presents the implementation and evaluation of the proposed models. Finally, Section 6 concludes the paper.

2. The model

In this section, we outline an hierarchical Bayesian Dirichlet regression model to jointly explore the age, period, and cohort dynamics of cause-of-death proportions.

2.1. The Dirichlet likelihood

We investigate cause-specific mortality proportions by age a and time t; these proportions are defined as the ratio between the number of deaths by age a and time t due to cause d to the number of all deaths by age a and time t. These are therefore related to the total (non-cause-specific) mortality by

$$m_{a,t,d} = \frac{D_{a,t,d}}{D_{a,t}}$$

For a given age and period, the vector of cause-specific death proportions has positive components summing to 1, due to the competing-risks nature of the causes of death. We follow Grunwald et al. (1993) and assign a Dirichlet distribution to the vector of such compositions.

A random vector of nonnegative random variables $\boldsymbol{Y} = (Y_1, \dots, Y_D)^T$ with unit sum has a Dirichlet distribution with parameter $\boldsymbol{\alpha} = (\alpha_1, \dots, \alpha_D)^T$ if it has the following joint density:

$$p(y_1, \dots, y_{D-1} \mid \boldsymbol{\alpha}) = \frac{1}{B(\boldsymbol{\alpha})} \prod_{d=1}^{D} y_d^{\alpha_d - 1} \qquad \sum_{d=1}^{D} y_d = 1, \quad y_d \geqslant 0 \tag{1}$$

where $\boldsymbol{\alpha} = (\alpha_1, \dots, \alpha_D)^T \in \mathbb{R}^D_{>0}$ is the vector of shape (or concentration) parameters for each category, and

$$B(\boldsymbol{\alpha}) = \frac{\prod_{d=1}^{D} \Gamma(\alpha_d)}{\Gamma\left(\sum_{d=1}^{D} \alpha_d\right)} \tag{2}$$

is the Dirichlet function, a d-dimensional analogue to the beta function, acting as the normalizing constant.

If we set $\alpha_0 = \sum_{d=1}^{D} \alpha_d$, we have

$$\mathrm{E}(Y_d) = \alpha_d/\alpha_0, \ \mathrm{Var}(Y_d) = \frac{\alpha_d(\alpha_0 - \alpha_d)}{\alpha_0^2(\alpha_0 + 1)}$$

so that α_0 is often interpreted as a precision parameter. The shape parameter $\boldsymbol{\alpha}$ of the Dirichlet distribution then controls both the location and dispersion of the distribution. We rely on the alternative parameterization of the Dirichlet distribution, as in Grunwald et al. (1993), in terms of a location parameter $\boldsymbol{\theta}$, for the distribution of $\mathbf{Y}$ in the simplex S^d of dimension d, thus accounting for the constrained nature of the simplex, and in terms of scale, the $\boldsymbol{\tau}$ parameter, by setting $\boldsymbol{\theta} = \boldsymbol{\alpha}/\boldsymbol{\tau}$, and $\tau = \alpha_0 = \sum_{d=1}^{D} \alpha_d$. In this way, the expected value and variance of the Dirichlet distribution turn out to be, respectively,

$$E[\mathbf{Y} \mid \boldsymbol{\theta}, \tau] = \boldsymbol{\theta} \text{ and } \mathrm{Var}[\mathbf{Y} \mid \boldsymbol{\theta}, \tau] = \frac{\boldsymbol{\theta}\boldsymbol{\theta}^T}{\tau + 1}.$$

This reparameterization, by separately specifying location and scale, lets us easily incorporate a regression structure into a generalized linear model. Indeed, a multinomial logit link can be used to relate the location parameter θ to the linear predictor, while a model for the precision parameter τ can be specified through the log-link functions.

For a given age a and time t, we assign a Dirichlet distribution to the vector of cause-specific deaths proportion $\mathbf{m}_{a,t} = (m_{a,t,d_1}, \dots, m_{a,t,d_D})$ and model the location parameter as a function of age, period, and cohort factors, as follows:

$$\log\left(\frac{\theta_{a,t,d}}{\theta_{a,t,D}}\right) = \beta_{a,d} + \beta_{t,d} + \beta_{k,d}, \tag{3}$$

where we define $a = 1, \dots, A$ ages, $t = 1, \dots, T$ periods, and $k = 1, \dots, A + T - 1$ cohorts and take cause of death D as the reference category. It is worth stressing that, through (3), we model the log odds of mortality by cause d rather than cause D as a function of age, period, and cohort components treated as factors: $\beta_{a,d}$ models the departure of the log odds from the overall mean due to the age component, $\beta_{t,d}$ the departure due to the period component, and $\beta_{k,d}$ the departure due to the cohort component.

2.2. The mixed age–period–cohort regression model

As is well known, APC models suffer from a lack of identification, due to the linear dependency among the three components: when two out of the three components are known, the value of the third is also determined. Linear trends in age, period, and cohort are not identifiable since they are linearly dependent. Several solutions have been proposed; see Smith and Wakefield (2016) for a detailed review on APC models, Hunt and Blake (2014) for a discussion of the identification issue in APC mortality models, and Smith (2020) and Bell (2020) for recent discussions

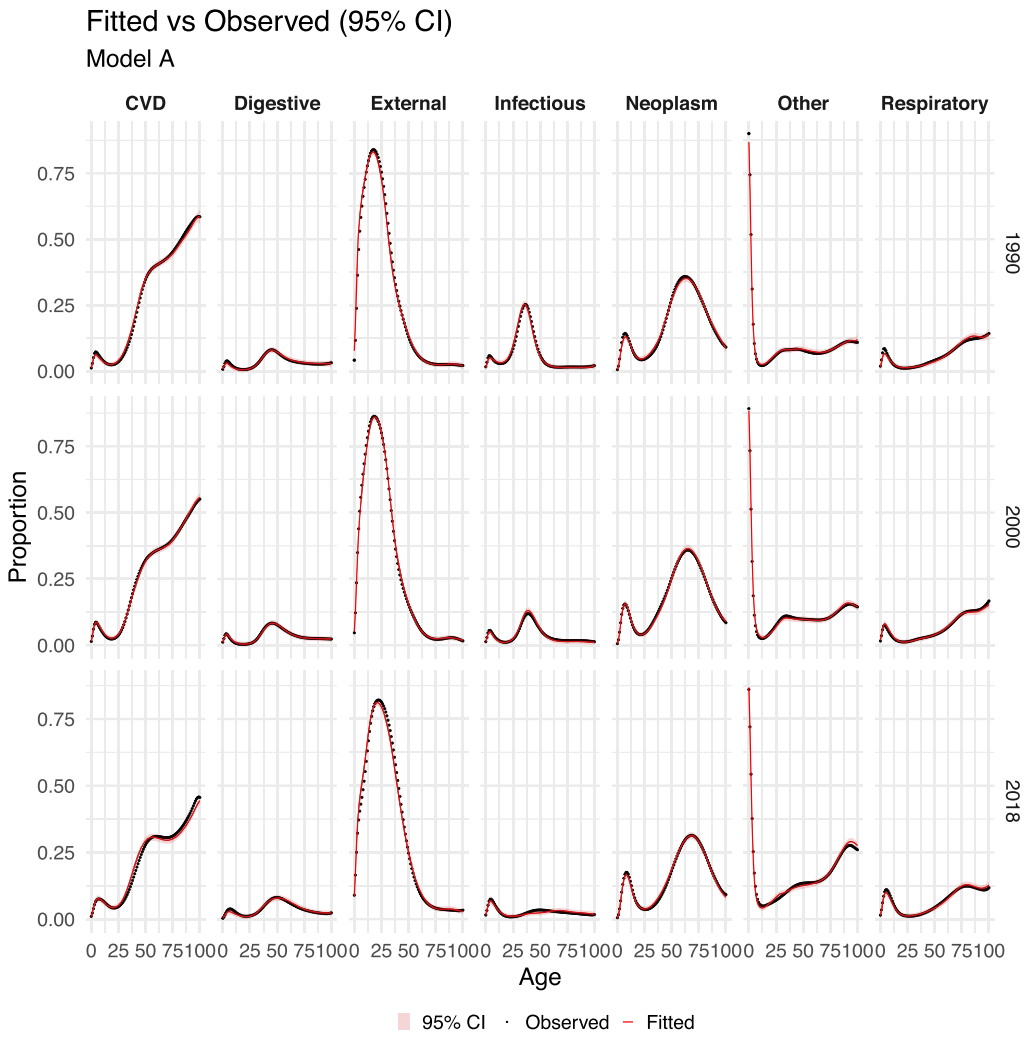


Fig. 1. ModelA. Comparison of observed and predicted (model-fitted) proportions of cause-specific mortality for years 1990, 2000, 2018 and age groups. Dots represent observed proportions, solid lines indicate model estimates, and shaded areas show 95 % credible intervals.

on APC analysis. A widely used solution is the constrained generalized linear model (CGLM) suggested by Mason et al. (1973). The idea is to introduce at least one additional identification coefficient constraint. In general, the constraints are equalities. For example, the effects of the first two age groups, periods, or cohorts may be constrained to be equal based on theoretical or external information (Mason et al., 1973). Clayton and Schifflers (1987) restrict the first-period effect differences to be, on average, equal to zero. Bayesian methods based on ICAR priors include zero-sum constraints and penalties on the linear trends in age, period, and cohort effects (see Berzuini and Clayton, 1994; Besag et al., 1995; Knorr-Held and Rainer, 2001). An approach of this kind suffers two main problems. First, the assumptions underlying the imposed constraints require strong prior information and cannot, in general, be empirically validated. Second, the results of the analyses rely strongly on the imposed constraints: different choices of identifying constraints can lead to different estimates for the age, period, or cohort effects (Rodgers, 1982; Holford, 1983). A large literature in demography, epidemiology, and statistics has discussed the sensitivity of the APC model to the assumptions specified (see e.g. Mason and Wolfinger, 2001; O'Brien, 2011).

In this paper, we explore the approach to age, period, and cohort analysis first suggested by Yang and Land (2006), in which one or more of ages, periods, and cohorts are treated as random effects. Yang and Land (2006) referred to such models as hierarchical APC (HAPC) models. The same approach is found in the mixed model suggested by O'Brien et al. (2008), and is used in several fields; see e.g. Masters et al. (2012) for total mortality. O'Brien (2017) discusses the sense in which a generalized linear mixed model offers a solution to the identification problem intrinsic to APC models, and labels this identification as statistical model identification. The idea is that the mixed model incorporates constraints on the effects without needing to specify them as in the CGLM. In the case of one random effect, O'Brien (2017) shows that the implicit constraint consists of forcing to zero the estimated linear trend in the component that has been modelled as random. This is proved also in Luo and Hodges (2020). The type of constraint introduced in the case of two random effects is less clear. Luo and Hodges (2020) claim, based on simulation studies, that the constraints implicit in the two-random-effects model depend on the size of the nonlinear components of the random effects. The APC mixed model then suffers the same drawbacks as the constraint generalized model. Moreover, the estimated trend in the fixed effects might change according to the effect(s) selected as random. However, we think that, in a multivariate framework such as ours, in which several causes of death are jointly analysed, a mixed-model approach offers the clear computational advantage of avoiding the need to specify constraints for each and every cause.

3. Methodological specification

Our proposal, then, is to turn model (3) into a generalized linear mixed Dirichlet regression model by treating period, cohort, or both as

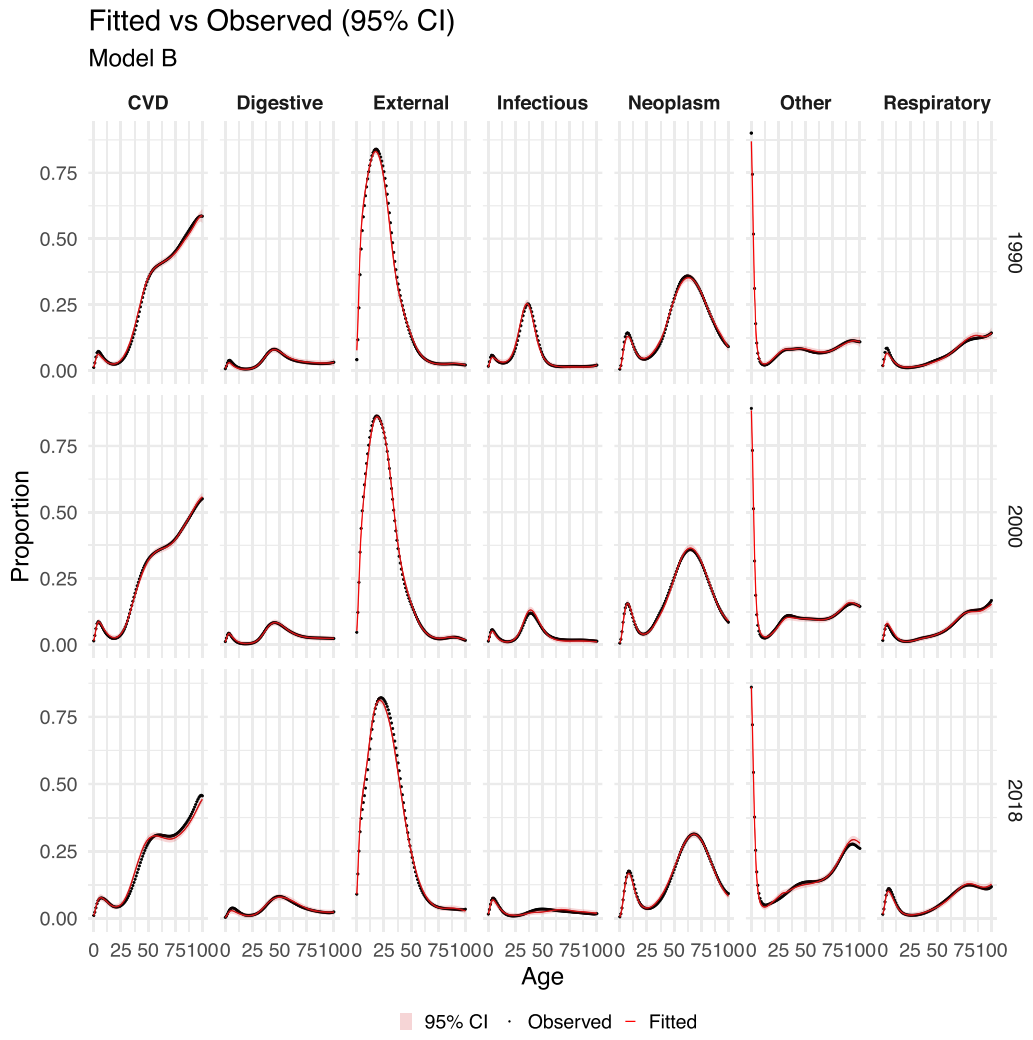


Fig. 2. ModelB. Comparison of observed and predicted (model-fitted) proportions of cause-specific mortality for years 1990, 2000, 2018 and age groups. Dots represent observed proportions, solid lines indicate model estimates, and shaded areas show 95 % credible intervals.

random effects. We keep age as a fixed effect since, out of the three time dimensions, it is agreed to capture the biological component of the different causes of death.

Let $\mathbf{m}_{a,t} = (m_{a,t,1}, \dots, m_{a,t,D})^\top \in \mathbb{R}^D$ denote the vector of cause-specific mortality proportions for age $a = 1, \dots, A$ and period $t = 1, \dots, T$, where $\sum_{d=1}^{D} m_{a,t,d} = 1$. We assume

$$\mathbf{m}_{a,t} \sim \text{Dirichlet}(\boldsymbol{\theta}_{a,t} \cdot \tau_{a,t}), \tag{4}$$

where $\boldsymbol{\theta}_{a,t} = (\theta_{a,t,1}, \dots, \theta_{a,t,D})^\top \in \mathbb{R}^D$ is the mean vector of the Dirichlet distribution with $\sum_d \theta_{a,t,d} = 1$, and $\tau_{a,t}$ is the concentration (precision) parameter.

To model $\boldsymbol{\theta}_{a,t}$, we use a multinomial logit link function with the Dth cause of death as the reference:

$$\log\left(\frac{\theta_{a,t,d}}{\theta_{a,t,D}}\right) = \eta_{a,t,d}, \quad \text{for } d = 1, \dots, D-1. \tag{5}$$

We write the entire model in compact matrix form as:

$$\boldsymbol{\eta} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\boldsymbol{\gamma} \tag{6}$$

where:

- $\boldsymbol{\eta} \in \mathbb{R}^{n(D-1)}$, with $n = A \times T$ observations (age-period combinations), and each observation has $D-1$ logits;
- $\mathbf{X} \in \mathbb{R}^{n(D-1)\times p}$ is the fixed effects design matrix;
- $\boldsymbol{\beta} \in \mathbb{R}^p$ is the vector of fixed effects coefficients;
- $\mathbf{Z} \in \mathbb{R}^{n(D-1)\times q}$ is the random effects design matrix;
- $\boldsymbol{\gamma} \in \mathbb{R}^q$ is the vector of random effects coefficients.

We construct the design matrices $\mathbf{X}$ and $\mathbf{Z}$ as follows. Denote by:

- $\mathbf{X}^{(a)} \in \mathbb{R}^{n\times A}$ the matrix of age indicators (fixed effects),
- $\mathbf{X}^{(t)} \in \mathbb{R}^{n\times T}$ the matrix of period indicators,
- $\mathbf{X}^{(k)} \in \mathbb{R}^{n\times(A+T-1)}$ the matrix of cohort indicators (with $k = t - a + A$).

Then, the full fixed effect matrix is defined by:

$$\mathbf{X} = \left[\mathbf{X}^{(a)} \otimes \mathbf{I}_{D-1} | \mathbf{X}^{(t)} \otimes \mathbf{I}_{D-1} | \mathbf{X}^{(k)} \otimes \mathbf{I}_{D-1}\right], \tag{7}$$

with corresponding coefficient vector:

$$\boldsymbol{\beta} = \begin{bmatrix} \boldsymbol{\beta}^{(a)} \\ \boldsymbol{\beta}^{(t)} \\ \boldsymbol{\beta}^{(k)} \end{bmatrix}.$$

Similarly, the random effects matrix is:

$$\mathbf{Z} = \left[\mathbf{Z}^{(a)} \otimes \mathbf{I}_{D-1} | \mathbf{Z}^{(t)} \otimes \mathbf{I}_{D-1} | \mathbf{Z}^{(k)} \otimes \mathbf{I}_{D-1}\right], \tag{8}$$

with random effects vector:

$$\boldsymbol{\gamma} = \begin{bmatrix} \boldsymbol{\gamma}^{(a)} \\ \boldsymbol{\gamma}^{(t)} \\ \boldsymbol{\gamma}^{(k)} \end{bmatrix}, \quad \boldsymbol{\gamma} \sim \mathcal{N}(\mathbf{0}, \mathbf{G}),$$

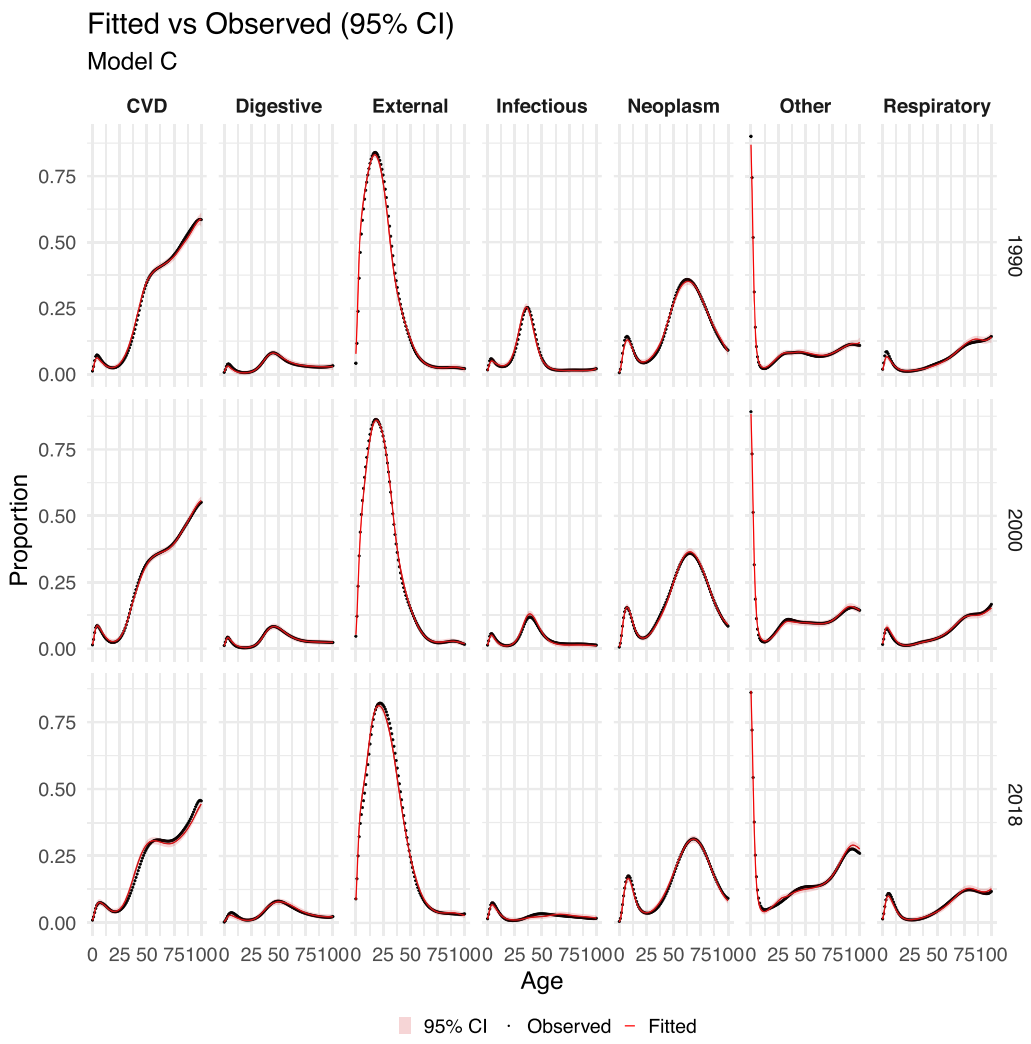


Fig. 3. ModelC. Comparison of observed and predicted (model-fitted) proportions of cause-specific mortality for years 1990, 2000, 2018 and age groups. Dots represent observed proportions, solid lines indicate model estimates, and shaded areas show 95 % credible intervals.

where $\mathbf{G} = \text{diag}(\sigma_a^2\mathbf{I}, \sigma_t^2\mathbf{I}, \sigma_k^2\mathbf{I})$ depending on model choice. For clarity of exposition and to facilitate comparison across models, we adopt a simplified representation of the linear predictor that reflects the classical age-period-cohort decomposition. Specifically, rather than writing the full matrix form $\boldsymbol{\eta} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\boldsymbol{\gamma}$ for each cause-specific logit, we express the model in its factor-based form:

$$\log\left(\frac{\theta_{a,t,d}}{\theta_{a,t,D}}\right) = \beta_{a,d} + \beta_{t,d} + \gamma_{k,d},$$

where age effects $\beta_{a,d}$ are treated as fixed, while period and cohort components can be specified either as fixed effects $\beta_{t,d}, \beta_{k,d}$ or random effects $\gamma_{t,d}, \gamma_{k,d}$, depending on the model configuration summarized in Table 1, where ModelA treats only the cohort as a random effect, ModelB treats only the period as a random effect, and, in ModelC, both the period and cohort are random effects.

This notation is aligned with the demographic convention of decomposing temporal variation into age, period, and cohort contributions, while allowing flexible modeling of each component.

3.1. Bayesian hierarchical prior specification

We adopt a fully Bayesian hierarchical approach to estimate model parameters. Let $\boldsymbol{\beta} \in \mathbb{R}^p$ denote the vector of fixed effect coefficients, and $\boldsymbol{\gamma} \in \mathbb{R}^q$ the vector of random effects, partitioned as:

Table 1
Model specifications: random effects γ are used for period and/or cohort.

Model	Equation
ModelA	$\log\left(\frac{\theta_{a,t,d}}{\theta_{a,t,D}}\right) = \beta_{a,d} + \beta_{t,d} + \gamma_{k,d}$
ModelB	$\log\left(\frac{\theta_{a,t,d}}{\theta_{a,t,D}}\right) = \beta_{a,d} + \gamma_{t,d} + \beta_{k,d}$
ModelC	$\log\left(\frac{\theta_{a,t,d}}{\theta_{a,t,D}}\right) = \beta_{a,d} + \gamma_{t,d} + \gamma_{k,d}$

Our formulation assumes weakly informative priors defined as follows:

$$\boldsymbol{\beta} \sim \mathcal{N}(\mathbf{0}, \mathbf{I}_p), \quad \text{fixed effects coefficients,} \tag{9}$$

$$\boldsymbol{\gamma} \sim \mathcal{N}(\mathbf{0}, \sigma_\gamma^2\mathbf{I}_q), \quad \text{random effects coefficients,} \tag{10}$$

$$\sigma_\gamma \sim t_3^+(0, 2.5), \quad \text{random effects SD (half-}t\text{),} \tag{11}$$

$$\tau_{a,t} \sim \text{Gamma}(0.01, 0.01), \quad \text{precision parameter.} \tag{12}$$

Here, $\mathcal{N}(\mathbf{0}, \mathbf{I})$ denotes a standard multivariate normal prior with independent components and unit variance. The prior for $\boldsymbol{\gamma}$ induces shrinkage on the random effects, while uncertainty in the amount of shrinkage is accounted for by placing a weakly informative half-Student-t prior on the standard deviation parameter σ_γ. This is denoted as:

$$\sigma_\gamma \sim t_3^+(0, 2.5),$$

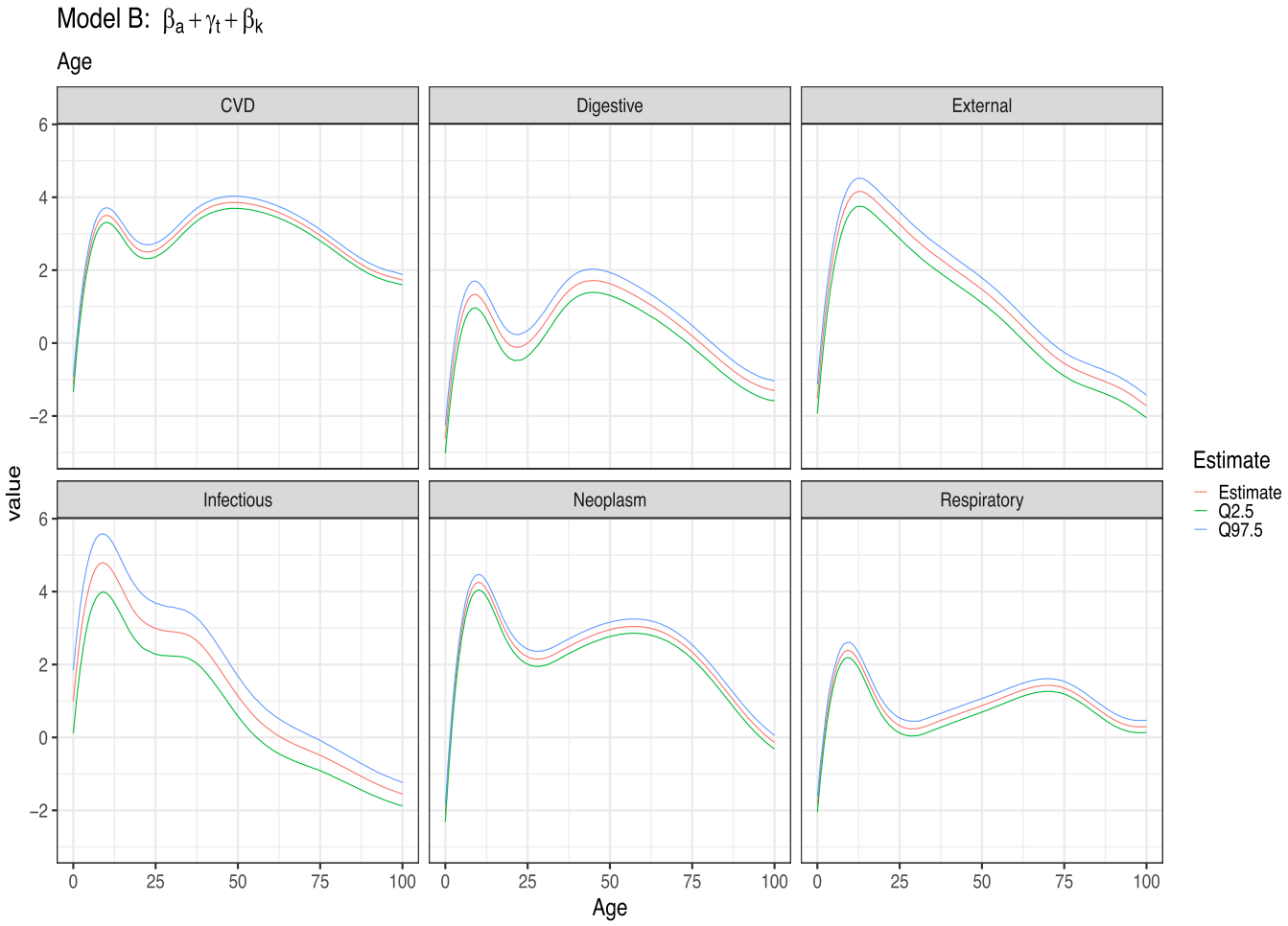


Fig. 4. Estimated age effects for each cause of death based on ModelB (Age as fixed effect, Period as random effect, Cohort as fixed effect). The solid red line displays the posterior mean, and the green and blue lines indicate the pointwise 95 % credible intervals. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

i.e., a Student-t distribution with 3 degrees of freedom, centered at 0, scaled by 2.5, and truncated to the positive reals.

The Dirichlet precision parameter $\tau_{a,t}$ is modeled independently for each observation using a diffuse gamma prior, to allow substantial flexibility in dispersion:

$$\tau_{a,t} \sim \text{Gamma}(0.01, 0.01),$$

which corresponds to a mean of 1 and a variance of 100, offering a flat prior over the log-precision scale.

All priors are specified independently across components, and apply identically across all logits $d = 1, \dots, D-1$ in the multinomial logit transformation:

$$\log\left(\frac{\theta_{a,t,d}}{\theta_{a,t,D}}\right) = \eta_{a,t,d}.$$

These choices ensure regularization of high-dimensional parameter spaces, while retaining enough flexibility to capture temporal structure in the age-period-cohort decomposition. Posterior inference is conducted via Hamiltonian Monte Carlo (HMC) using the Stan platform (Team, 2015).

4. Data and posterior estimation

4.1. Data

The model outlined in the previous section is applied to data extracted from the period life tables by sex and single year of age (0–100+) from the Human Mortality Database (000) (HMD) from 1990 to 2018. We extracted death counts by cause of death from Human Cause of Death Database (000) (HCD) for males in the USA, thus computing the proportion of deaths by cause, age, and cohort in a given year. The HCD Database provides high-quality data on cause-specific mortality. It is coded using the International Classification of Diseases (ICD), and provides three aggregation levels: full list, intermediate list, and short list. Each classification has been developed using the same criteria for all countries, ensuring homogeneity and comparability. By using these data, we obtain a universal and standardized methodology to redistribute deaths between 104 disease categories in five-year age groups, avoiding issues regarding the ICD revisions and ensuring cross-country comparability despite differing coding practices. Starting from the Human Cause of Death Database (000) classification of causes of death, we chose to group major causes of death up to age 100+.

For our purpose of capturing conditions that might have affected the evaluation of longevity in the USA, we group granular information into seven major categories: neoplasms, cardiovascular diseases (CVDs), digestive, respiratory diseases, infectious diseases, external diseases, and 'other' to group all remaining causes. For ICD codes and details on the classification system, see Table 2.

Since cause-of-death data are available in 5-year age groups, we perform estimation based on the penalized composite link model (PCLM) (Rizzi et al., 2015) to ungroup data into single years of age. The method is based on the composite link model (Thompson and Baker, 1981), with

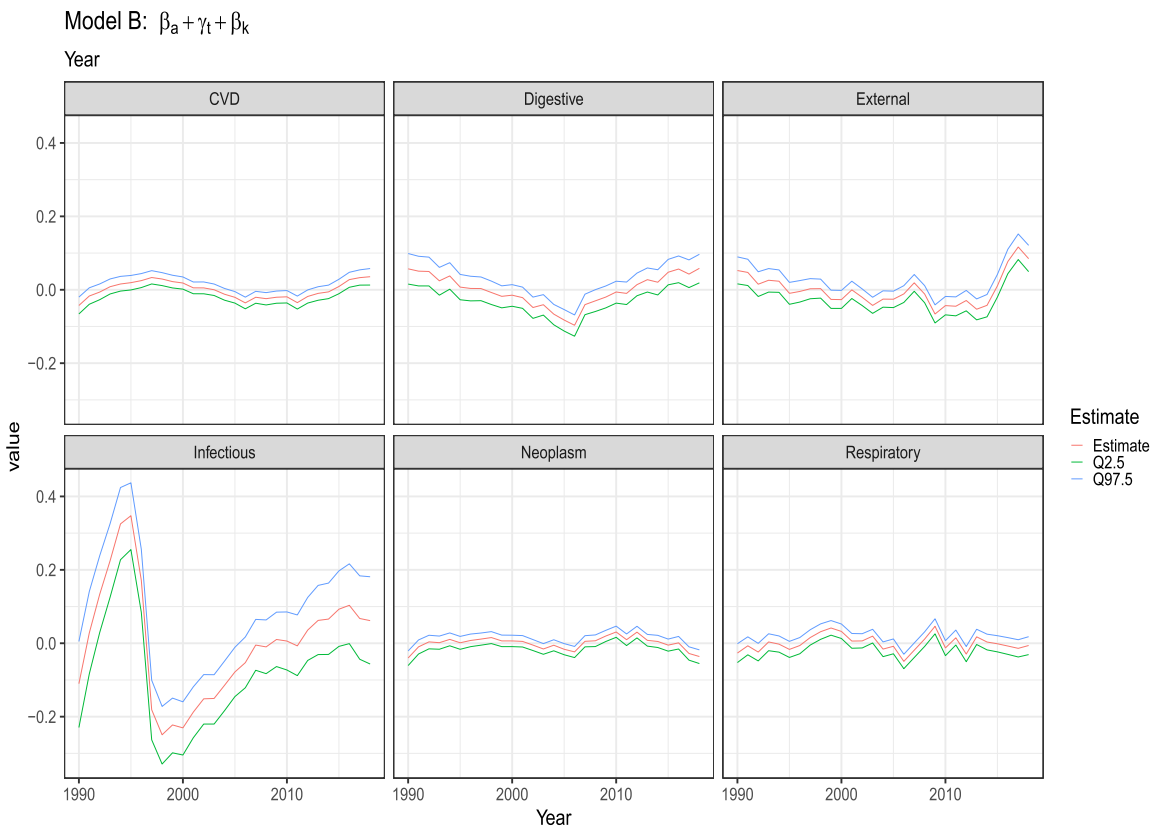


Fig. 5. Estimated period effects for each cause of death based on ModelB (Age as fixed effect, Period as random effect, Cohort as fixed effect). The solid red line displays the posterior mean, and the green and blue lines indicate the pointwise 95 % credible intervals. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

a penalty added to ensure the smoothness of the target distribution (Eilers, 2007). Fig. A1 presents the age patterns of the male cause-specific proportions from 1990 to 2018.

In order to link the data visualization with the model output, we provide in figure Fig. B.1 a representation of the dataset with the log-odds transformation in which the six causes are related to the reference cause (Other).

4.2. Posterior estimation and model comparison

Samples from the posterior distributions of the parameters were drawn by using Hamiltonian Monte Carlo sampling and specifically using the `stan` software package Team (2015). Stan and its interface in the R programming language (R Core Team, 2017) allows the construction of a Hamiltonian Monte Carlo (HMC) sampling "no U-turns sampler" from a simple user specification of the Bayesian model to be estimated. Hamiltonian Monte Carlo simulates movement through the parameter space by analogy to a physical system where the potential energy is equal to negative log-posterior (Brooks et al., 2011), and it is a special case of the more general Metropolis-Hastings algorithm for Markov Chain Monte Carlo sampling. Details on HMC procedure are reported in the Appendix A

Three parallel chains were constructed and used to assess convergence to the posterior distribution, each with 3000 samples, and the first half of each chain was used as a warm-up period. For the majority of model and country combinations, the split $\hat{r}$ diagnostics are below the suggested 1.05 threshold, and examination of the traceplots indicates convergence to the target distribution (Gelman and Rubin, 1992; Team, 2015).

We assessed the models' predictive accuracy through the leave-one-out information criterion (LOOIC) as discussed by Vehtari et al. (2017). LOOIC is a measure of how well we might expect a model to perform in predicting a data point without including it in the data that are used to fit the model. It can be used for model comparison. Consider some data $y_1, \dots, y_n$, modelled as independent given some parameters θ; thus, $p(y \mid \theta) = \prod_{i=1}^{n} p(y_i \mid \theta)$. This formulation also encompasses latent variable models with $p(y_i \mid f_i, \theta)$, where f_i are latent variables. Also suppose we have a prior distribution $p(\theta)$, thus yielding a posterior distribution $p(\theta \mid y)$ and a posterior predictive distribution $p(\tilde{y} \mid y) = \int p(\tilde{y}_i \mid \theta) p(\theta \mid y) d\theta$. To maintain comparability with the given dataset and to facilitate interpretation of scale differences in the effective number of parameters, we define a measure of predictive accuracy for the n data points taken one at a time: the expected log pointwise predictive density for a new dataset (elpd), given by

$$\text{elpd} = \sum_{i=1}^{n} \int p_t(\tilde{y}_i) \log p(\tilde{y}_i \mid y) d\tilde{y}_i,$$

where $p_t(\tilde{y}_i)$ is the distribution representing the true data-generating process for $\tilde{y}_i$. These distributions are unknown, and we will use cross-validation. The Bayesian LOOIC estimate of goodness predictive fit is $-2\widehat{\text{elpd}}_{\text{loo}}$, where, $\widehat{\text{elpd}}_{\text{loo}} = \sum_{i=1}^{n} \log p(y_i \mid y_{-i})$, and $p(y_i \mid y_{-i}) = \int p(y_i \mid \theta) p(\theta \mid y_{-i}) d\theta$. The model with the best predictive accuracy is the one that shows the largest elpd and, therefore, the smallest LOOIC estimate. is the leave-one-out predictive density given the data without the ith data point. LOOIC is also widely used for demographic applications, for example, by Hilton et al. (2019).

5. Results

The results of the fit of the three models in Table 1 are provided in Appendix A through graphical visualizations that display, for each cause of death, the estimated age, period, and cohort patterns of the

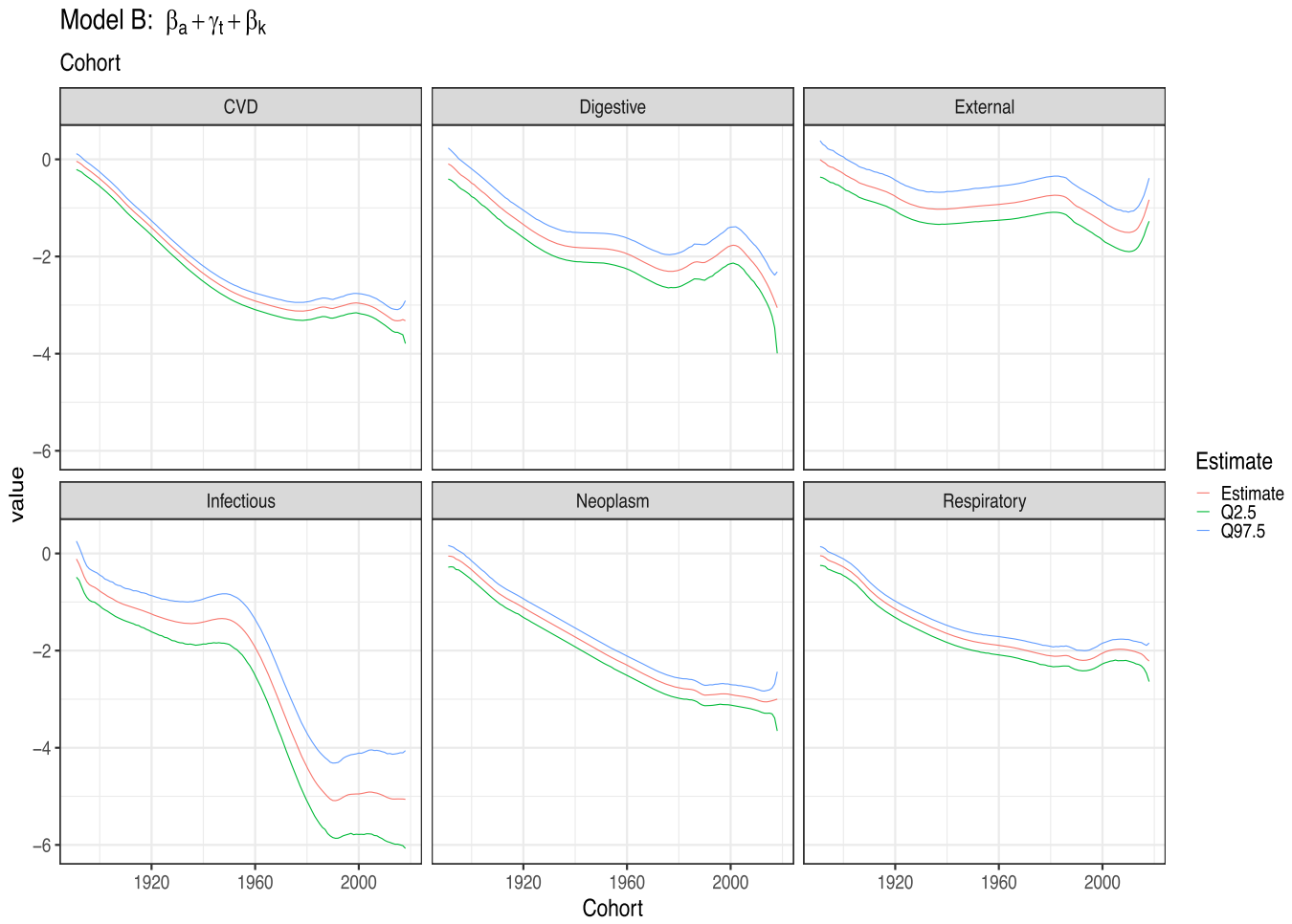


Fig. 6. Estimated cohort effects for each cause of death based on ModelB (Age as fixed effect, Period as random effect, Cohort as fixed effect). The solid red line displays the posterior mean, and the green and blue lines indicate the pointwise 95 % credible intervals. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

corresponding log odds of mortality, along with the confidence intervals. Table 3 provides the LOOIC and elpd estimates (4.2) and values of accuracy measures found using the `loo` function in R.

ModelB shows better accuracy than the other two models. Indeed, as we can see from Table 3, ModelB achieves the greatest elpd and consequently the lowest LOOIC value.

Let y_i denote the observed proportion of deaths due to a given cause for the ith data point, and $\hat{y}_i$ denote the corresponding predicted value from the model. For a given cause and a given model, let n be the number of data points (e.g., age-period combinations).

The Mean Absolute Error (MAE) is defined as:

$$\text{MAE} = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i|$$

The Root Mean Squared Error (RMSE) is defined as:

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}$$

Both metrics are computed for each cause of death and each model under comparison. The MAE quantifies the average magnitude of the absolute errors, providing a measure of overall accuracy that is less sensitive to outliers. The RMSE, being based on squared errors, penalizes larger deviations more heavily, thus highlighting model performance with respect to extreme prediction errors.

The performance of the three models—ModelA (random cohort effect), ModelB (random period effect), and ModelC (random period and cohort effects)—was evaluated using the MAE and RMSE metrics for each major cause of death. Table 4 reports these values, together with their column-wise averages.

A close inspection of the results reveals that the three model specifications perform similarly across most causes, with only minor differences in MAE and RMSE values. Nevertheless, ModelB systematically achieves the lowest average MAE and RMSE across all causes, albeit by a small margin. This suggests that treating the period effect as random allows the model to more accurately capture the inherent temporal fluctuations and improve generalization.

For most causes of death, including the leading categories such as cardiovascular diseases (CVD) and neoplasms, the prediction errors remain consistently low across all models, reflecting the overall adequacy of the proposed approach. Notably, higher RMSE and MAE values are observed for “External” causes of death, which is expected given the larger volatility and unpredictability associated with this category (e.g., accidental deaths, homicides), as also discussed in the literature.

The reported metrics also highlight that the model specification has limited impact on the accuracy for causes where the age effect is dominant and the temporal structure is stable (e.g., neoplasms, CVD). However, for causes with more marked period or cohort effects, the choice of random effect structure can influence predictive performance.

In summary, the comparative analysis based on MAE and RMSE indicates that all three models provide accurate and robust predictions of cause-specific mortality proportions, with ModelB offering the best overall trade-off between model complexity and accuracy. These

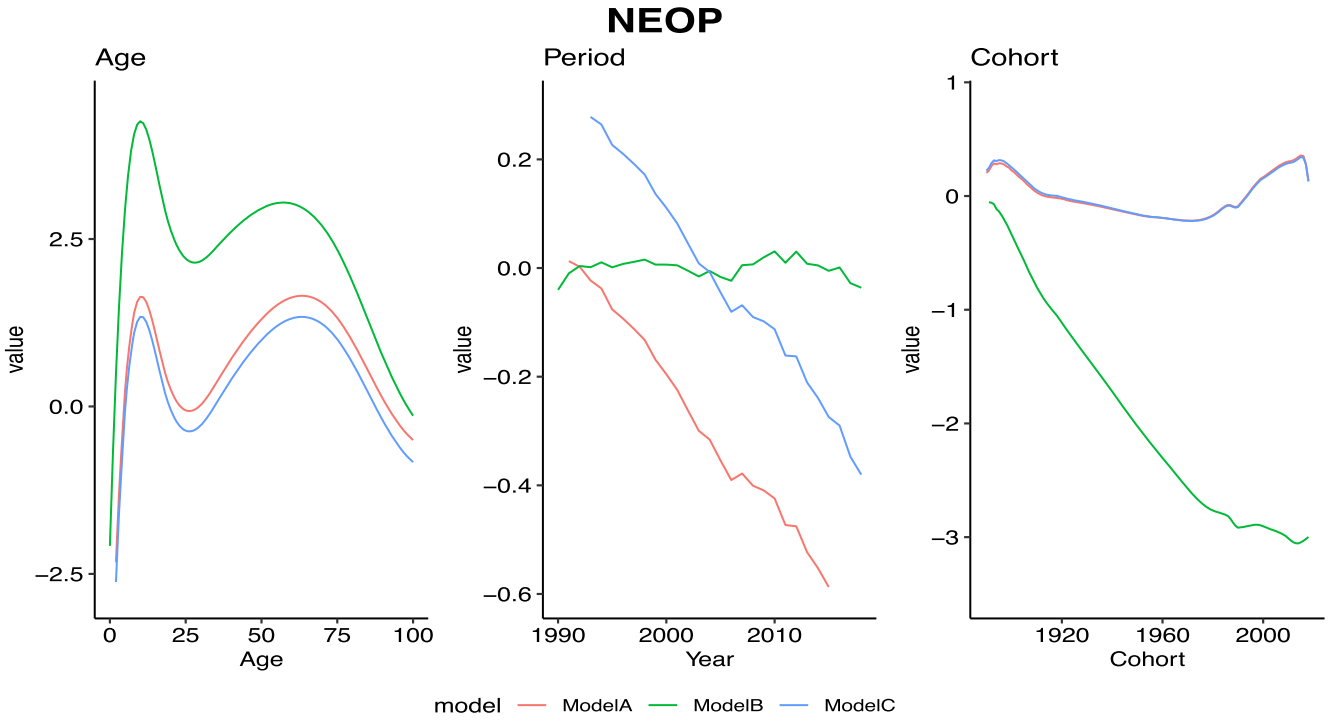


Fig. 7. Estimated age effects for each cause of death based on Model A, B, C.

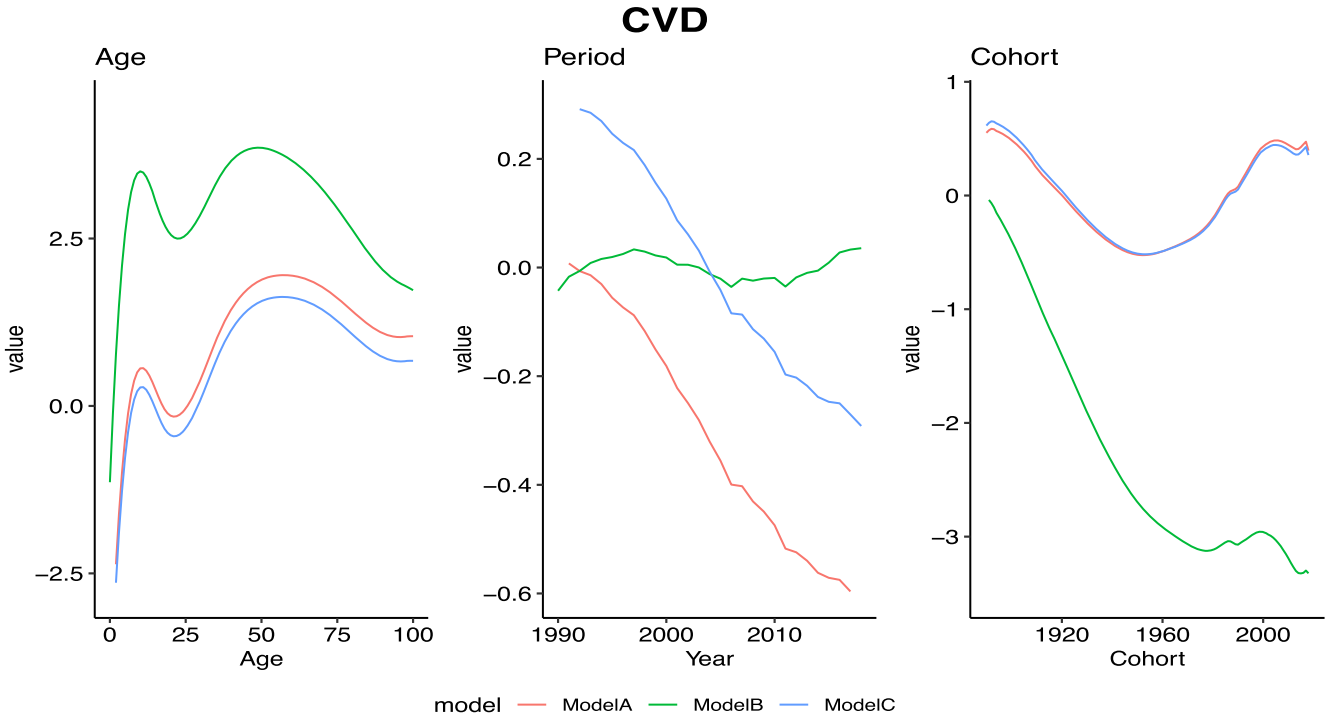


Fig. 8. Estimated period effects for each cause of death based on Model A, B, C.

findings are coherent with the results obtained using information criteria (e.g., LOOIC) and confirm the suitability of the mixed-effects Dirichlet regression framework for the joint analysis of compositional mortality data.

To further support our comparative analysis, we provide a graphical assessment of the fitted values for all causes of death, across the three model specifications, at three selected calendar years. These visualizations enable a direct and intuitive evaluation of the models' ability to reproduce the observed cause-specific mortality patterns over time and across age groups.

Specifically, in Figs. 1–3, for each model (ModelA, ModelB, and ModelC), we plot the fitted cause-specific death proportions as a function of age, separately for each major cause of death, for the years 1990, 2000, and 2018. By juxtaposing the fitted values across models and years, we

Table 2
ICD codes and classification.

Title	ICD 10	Cause
Certain infectious diseases	A00–B99	Infectious
Neoplasms	C00–D48	Neoplasms
Heart diseases	I00–I52	CVD
Cerebrovascular diseases	G45, I60–I69	CVD
Other and unspecified disorders of the circulatory system	I70–I99	CVD
Acute respiratory diseases	J00–J22, U04	Respiratory
Other respiratory diseases	J30–J98	Respiratory
Diseases of the digestive system	K00–K93	Digestive
Diseases of the skin and subcutaneous tissue, musculoskeletal system, and	L00–M99	Other
Diseases of the genitourinary system and complications of pregnancy	N00–O99	Other
Certain conditions originating in the perinatal period and congenital	P00–Q99, R95	Other
Diseases of the blood and blood-forming organs	D50–D89	Other
Endocrine, nutritional, and metabolic diseases	E00–E90	Other
Mental and behavioural disorders	F00–F99	Other
Diseases of the nervous system and the sense organs	G00–G44, G47–H95	Other
External causes	V01–Y98	External

Table 3
Model accuracy.

Model	LOOIC	elpd_{loo}
ModelB	−138949.2	69474.6
ModelC	−138897.8	69448.9
ModelA	−138896.2	69448.1

Table 4
Comparison of RMSE and MAE values for Models A, B, and C across different causes of death.

Cause of Death	ModelA		ModelB		ModelC	
	RMSE	MAE	RMSE	MAE	RMSE	MAE
CVD	0.00712	0.00529	0.00710	0.00525	0.00712	0.00529
Digestive	0.00204	0.00144	0.00203	0.00141	0.00204	0.00143
External	0.01260	0.00752	0.01260	0.00749	0.01260	0.00752
Infectious	0.00553	0.00379	0.00554	0.00379	0.00553	0.00379
Neoplasms	0.00446	0.00348	0.00446	0.00348	0.00446	0.00349
Other	0.00628	0.00402	0.00620	0.00402	0.00627	0.00403
Respiratory	0.00367	0.00252	0.00365	0.00250	0.00366	0.00252
Average	**0.00596**	**0.00429**	**0.00594**	**0.00428**	**0.00596**	**0.00429**

can visually inspect both the temporal consistency of the fit and the ability of each model to capture the age structure of mortality for each cause.

This graphical comparison complements the quantitative results based on MAE and RMSE, and highlights any systematic differences between the models that may not be immediately apparent from aggregate error measures alone. Such an approach is particularly useful for detecting model deficiencies in specific age ranges, causes, or periods, and for confirming the robustness of the selected model across the entire observed time span.

5.1. Model interpretation

We can now move to interpreting the results of the fit of ModelB for all causes, referring to Figs. 4, 5, and 6, which display for each cause the estimated age, period, and cohort patterns of the log-odds ratio of mortality associated with this cause to all other causes, along with the prediction intervals.

The model effectively captures the age trends for each cause (see Fig. 1 in the supplementary material). The major causes of death, neoplasms and CVD, exhibit comparable characteristics. The model accurately captures the rise in occurrence with age, followed by a brief slowdown during the young-adult phase. This is followed by a consistent proportion during the productive and mature years, with a relative decline in older age groups.

In the context of external mortality, the model's estimates exhibit an upward trajectory in the mortality proportion within the younger age groups, followed by a persistent and gradual decline. According to the literature, this phenomenon can be explained by the prevalence of elevated homicide rates, correlated with the heightened availability of firearms in the USA. Additionally, motor vehicle fatalities can be interpreted as an outcome of risky behaviours like widespread alcohol consumption and consequent drunk driving. Finally, infectious, digestive, and respiratory diseases share a similar pattern in the 0–40 age range, after which the proportion of infectious mortality stalls and becomes constant.

We point out that all the estimates are coherent with the log-odds representation provided in the supplementary material, thus obtaining $k - 1$ classes due to the reference category (Other).

As discussed in Section 2.2, when treating the period as a random effect, we implicitly zero the linear trend in the period. Indeed, as we can see from Figs. 5 and 6, it is not possible to discern a linear trend based on the period. The effect of the time dimension is captured by the cohort, and the contribution of the period reduces to idiosyncratic fluctuations.

The cohort effect (Fig. 6) is a novelty in this study, as it has been quite neglected in the study of causes of death. In some cases, therefore, we do not have a suitable benchmark and the results are difficult to interpret. In other cases, however, our expectations can be confirmed, such as the peak shown by the proportion of deaths from infectious diseases and its decline from 1960 onwards. We may speculate that this is related, for example, to transitory pandemics or to some cohorts having greater exposure to specific risks, such as in the case of HIV. Other causes, such as CVD and neoplasms, show a U-shape, which can be explained by behavioural and lifestyle changes between different cohorts, or in other cases by the emergence of a widespread diagnosis at the population level that was not performed for the previous cohorts.

We point out that the cohort effect is linked to variation in the exposure to certain risk factors or illnesses that, in turn, are linked to the period in which a person was born or through the development of habits during childhood or early adulthood that affect later health (Holford, 1991). Smoking, for example, is a prime example of a habit that subjects start in youth – most commonly in early adolescence – and continue for most of their lives. Although, this link between the year of birth and cause of death can appear strong and straightforward for some causes, it is not necessarily so for others. We can speculate, for instance, that CVD mortality has increased risk associated with several co-morbidities such as high cholesterol, obesity, and diabetes that are, in turn, linked to poor diet and exercise as preventable risk factors. Thus, the cohort effects for each cause of death will often be the result of a potentially complex combination of single-risk-factor cohort effects. We also need to be mindful that a single-risk-factor cohort effect differs from the underlying lifestyle factor in a number of ways. However, our work is conducted on a population level and does not use microdata with individual features, preventing our speculation from being corroborated by an investigation of causal relationships.

As discussed, a statistical comparison of the three models has identified ModelB as the model that most accurately describes the APC dynamics of the US male cause of death proportions from 1980 to 2018. However, as outlined in Section 2.2, any APC model entails constraints that cannot be statistically verified, as stressed by Bell (2020), who furthermore recalls how the same level of goodness-of-fit to the data can be achieved by models with different constraints. The selection of an APC model cannot rely solely on statistical reasoning.

Given the inherently complex interplay between the three temporal components, statistically based criteria must be complemented by a thorough analysis of potential solutions grounded in a deep understanding of the phenomenon under scrutiny. We agree with Bell (2020) in stressing that, in the initial stages, it is crucial to define the assumptions explicitly and then to question whether the constraints imposed by the model align with the existing knowledge of the phenomenon. Subsequently, we think that the fitting results of various models need to be compared, taking into account their ability to capture the dynamics that an understanding of the phenomenon leads us to expect; see also Smith (2020). Such comparative evaluation helps to ensure that the model's choices are not only statistically sound but also conceptually valid, enhancing its ability to reflect the underlying dynamics accurately. This type of evaluation is particularly complex when undertaking a joint analysis of different causes of death. For this reason, we begin by restricting our attention to two causes – neoplasms and CVDs – that have been identified as the primary contributors to mortality.

5.2. The cases of neoplasms and cardiovascular diseases

Figs. 7 and 8 display estimates of the age, period, and cohort dynamics for the three models under consideration, depicting the effects on these specific causes. For the age effect, the dynamics estimated under the three models are quite similar and in line with our previous expectations. All models capture the increasing proportion over age and the short deceleration around young adulthood. Subsequently, we can observe a constant proportion during the working and adult ages and a decline at older ages. The three models differ with respect to the estimation of the two other time effects. When only the period is treated as random (ModelB), its linear trend is set to zero, and the period effect reduces to positive and negative idiosyncratic fluctuations. The effect of the time dimension is entirely captured by the cohort. For both causes of death, we observe a decrease over time of the mortality odds for each of the two considered causes relative to other causes.

Treating the period as a random effect allows for a more nuanced exploration of how cohorts contribute to observed trends, while the temporal variations are accounted for as inherent cohort differences. This modelling choice is particularly useful when trying to disentangle the effects of age, period, and cohort in complex datasets, helping to reveal a deeper understanding of the underlying dynamics at play. This scenario is supported by demographic insights about increased life expectancy due to reductions in mortality for CVD and cancer. The reverse is observed when the cohort is treated as a random effect (ModelA), or when both period and cohort are treated as random effects (ModelC). In both cases, the period effect shows a clear decreasing trend, while the cohort effect shows fluctuations around zero, with a U-shaped pattern.

Ultimately, this comparison of the results for the three models supports the selection of ModelB as providing the best description of the APC dynamics. In particular, treating the period as a random effect allows for a better exploration of how cohorts contribute to the observed trends, which represents a further contribution of this work since the cohort effect is quite often neglected in cause-of-death studies. Moreover, even if it is beyond the scope of the current work, as stressed by Smith (2020) for fertility, we expect that a forecasting structure with all historical trends allocated to the cohort will work better.

6. Conclusions

This paper extends and improves on the APC framework for cause-specific death analysis by incorporating, for the first time, the simultaneous modelling of causes of death.

Adopting the Dirichlet distribution allows for explicit modelling of the proportions of cause-specific mortality. Thus, by jointly modelling cause-specific variation in the APC, we ensure coherence with respect to total mortality.

Using data on US mortality from the HCD Database, we describe the statistical details of our method, offering three different model specifications based on the mixed-effect scheme. Our investigation provides two main insights into the related literature: i) we introduce an original approach to studying APC dynamics, and ii) we study the different specifications with the double lens of the model accuracy evaluation and by providing an assessment of the demographic significance of our approach, with a specific focus on cancer and CVD.

Our model comparison based on LOOIC indicate the ModelB as the best model compared to ModelA and ModelC. Nevertheless, while both ModelA and ModelC yield nearly identical predictive performance and parameter estimates, there are important practical considerations in choosing between them. ModelA introduces a single random effect (cohort), while ModelC includes two random effects (period and cohort), making it structurally more complex. This increased complexity in ModelC translates to a higher number of estimated parameters and a more challenging computational landscape, typically resulting in longer computation times and potentially greater convergence issues in a Bayesian MCMC framework. When the improvement in predictive performance is negligible—as is the case here—ModelA may be preferable for practical applications due to its parsimony and computational efficiency. On the other hand, ModelC allows for more flexibility in capturing potential unobserved heterogeneity in both period and cohort dimensions, which could be advantageous in settings with richer or more volatile data. Ultimately, the choice between these models should balance predictive accuracy, model interpretability, and computational considerations.

The proposed framework provides valuable information that could be used to improve the efficiency of healthcare and public actions. In particular, healthcare planning can benefit from our approach as it provides a useful snapshot of the magnitude of a specific cause of death relative to the total population. The overall survival rates of the US population could be improved by better tailoring investments aimed at tackling CVDs. By introducing the cohort effect, we can also speculate that causes which, in the recent past, were considered as natural public health targets because of their magnitude might be less important today. Medical innovations can take many years to become available to the whole population, and thus some cohorts can benefit from these innovations while some others may not.

We have stressed the novelties and related implications of our proposal. Our model provides significant insights despite its low complexity, as we work in a GLM-like framework using a Dirichlet probabilistic assumption. Our proposal also offers relevant cues for discussion in the population studies domain. In particular, we consider that it is important to include the cohort effect in joint modelling of causes of death. As pointed out by Kjærgaard et al. (2019), when taking the CODA approach for cause-of-death analysis, "it is not straightforward to implement cohort effects because of the unique relationship between age, time and cohort". In the Kjærgaard approach, none of the models considered include cohort effects to account for specific survival in some cohorts. The same authors also conclude that it is very likely that cohort components could improve the fit and forecast in some populations. In light of this, we do not deem it appropriate to compare our model with other models such as Lee–Carter; that would be beyond the scope of this work.

Further studies could consider a natural extension that includes a forecast component. Nevertheless, the structure of the proposed model makes the forecasting phase challenging. Studying the dependence

structure of causes-of-death proportions and forecasting, over period and cohort components, deserves a separate investigation.

CRediT authorship contribution statement

Rebecca Graziani: Writing – review & editing, Writing – original draft, Methodology, Conceptualization; **Andrea Nigri:** Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Methodology, Formal analysis, Data curation, Conceptualization.

Data availability

Data will be made available on request.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A.

In the implementation of Hamiltonian Monte Carlo as employed in RStan, the momentum variables r are independently sampled from a reference distribution. Together with the position parameters $\boldsymbol{\vartheta}$, the momentum defines a joint density $p(\boldsymbol{\vartheta}, r)$, from which the Hamiltonian function $H(\boldsymbol{\vartheta}, r)$ is derived. This function governs the evolution of the dynamic system and is given by:

$$\begin{aligned} H(\boldsymbol{\vartheta}, r) &= -\log p(\boldsymbol{\vartheta}, r) \\ &= -\log p(r \mid \boldsymbol{\vartheta}) - \log p(\boldsymbol{\vartheta}) \\ &= K(r \mid \boldsymbol{\vartheta}) + V(\boldsymbol{\vartheta}), \end{aligned} \tag{13}$$

where $K(r \mid \boldsymbol{\vartheta})$ denotes the kinetic energy term corresponding to the auxiliary momentum distribution, and $V(\boldsymbol{\vartheta})$ denotes the potential energy term, associated with the negative log of the target distribution. For a detailed exposition of HMC principles and implementation, see Neal (2011).

The HMC algorithm proceeds iteratively via two primary steps. First, new momentum variables r are sampled. Then, the Hamiltonian dynamics are simulated to update both $\boldsymbol{\vartheta}$ and r over a fictitious time interval. The evolution of the system is governed by Hamilton's equations:

$$\begin{aligned} \frac{d\boldsymbol{\vartheta}}{dt} &= +\frac{\partial H}{\partial r} = \frac{\partial K}{\partial r}, \\ \frac{dr}{dt} &= -\frac{\partial H}{\partial \boldsymbol{\vartheta}} = -\frac{\partial K}{\partial \boldsymbol{\vartheta}} - \frac{\partial V}{\partial \boldsymbol{\vartheta}}, \end{aligned} \tag{14}$$

where $\partial V / \partial \boldsymbol{\vartheta}$ represents the gradient of the negative log-target density with respect to $\boldsymbol{\vartheta}$.

These equations are typically approximated using the leapfrog integrator, which balances numerical stability and computational efficiency. The leapfrog updates proceed as follows:

$$\begin{aligned} r(n+\epsilon/2) &= r(n) - \frac{\epsilon}{2}\frac{\partial V}{\partial \boldsymbol{\vartheta}}(\boldsymbol{\vartheta}(n)), \\ \boldsymbol{\vartheta}(n+\epsilon) &= \boldsymbol{\vartheta}(n) + \epsilon \frac{r(n+\epsilon/2)}{m}, \\ r(n+\epsilon) &= r(n+\epsilon/2) - \frac{\epsilon}{2}\frac{\partial V}{\partial \boldsymbol{\vartheta}}(\boldsymbol{\vartheta}(n+\epsilon)), \end{aligned} \tag{15}$$

in which ϵ denotes the integration step size, m is the mass parameter, and n indexes the discrete integration steps.

After simulating the trajectory for a predetermined number of steps L, a proposal $(\boldsymbol{\vartheta}^*, r^*)$ is obtained. This proposal is accepted with probability:

$$\alpha = \min\left[1, \exp\left(-H(\boldsymbol{\vartheta}^*, r^*) + H(\boldsymbol{\vartheta}, r)\right)\right], \tag{16}$$

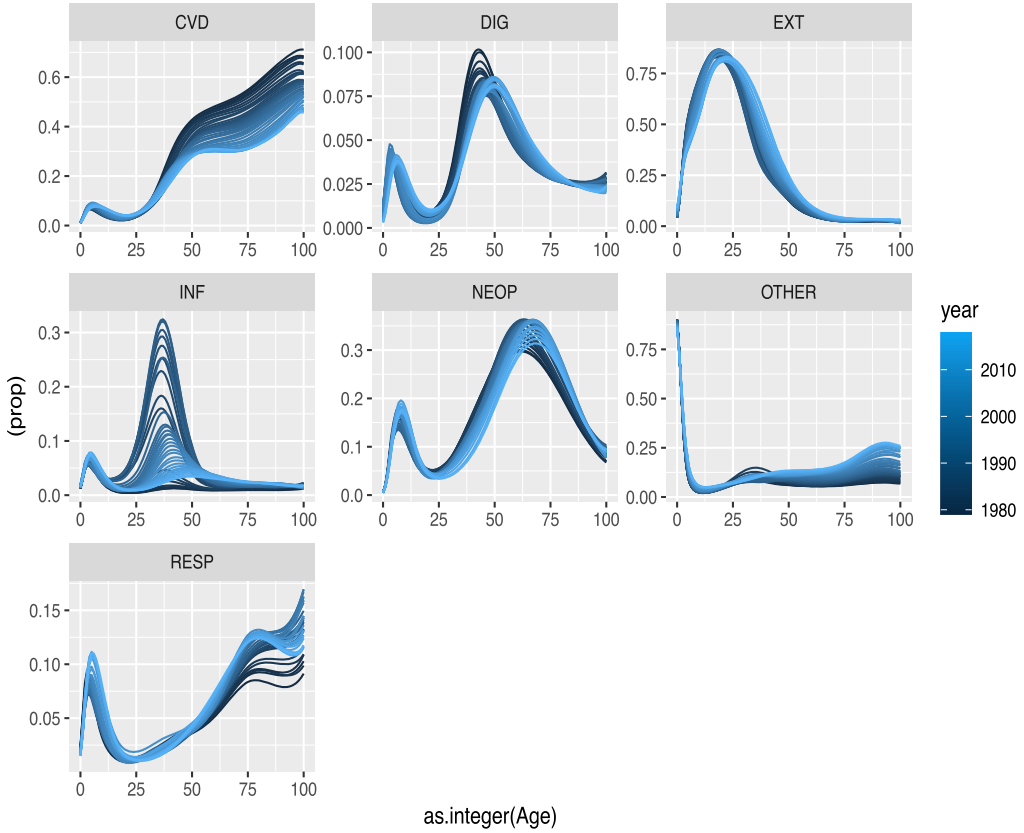


Fig. A1. Male population. Age and cause-specific proportions, by year from 1990 to 2018. Estimation using CLM and our ICD classification.

ensuring detailed balance and preserving the target distribution as the stationary distribution of the Markov chain.

The principal strength of HMC lies in its ability to generate proposals that follow informed trajectories through the parameter space, guided by the underlying geometry of the target density. This results in more efficient exploration, particularly in high-dimensional or highly correlated settings, thereby reducing autocorrelation and accelerating convergence Figs. A2–A7.

A.1. One random effect

A.1.1. Random effects: Cohort

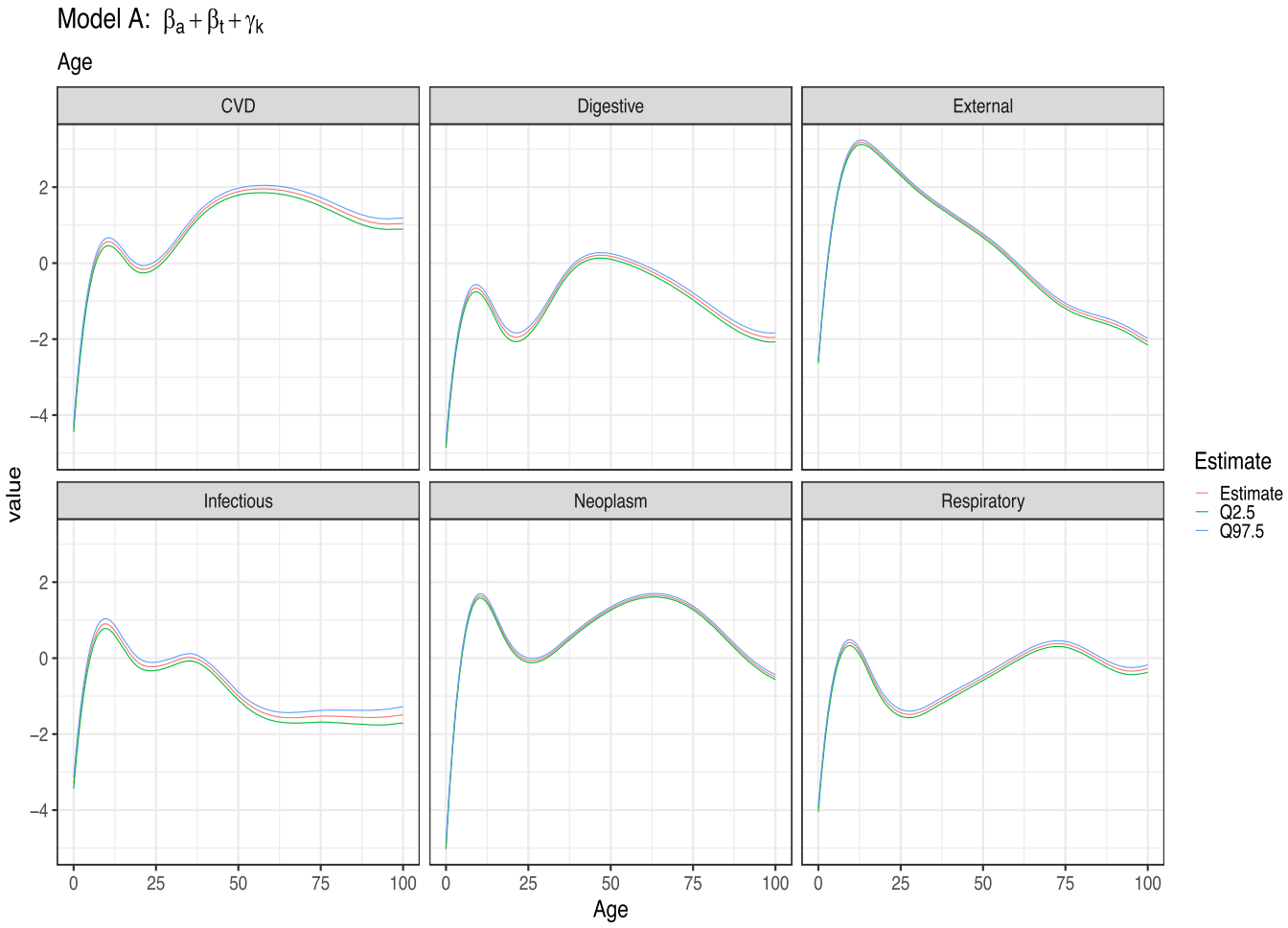


Fig. A2. Estimated age effects for each cause of death based on ModelA (age as fixed effect, cohort as random effect, period as fixed effect). The solid line shows the posterior mean, and the shaded area indicates the 95 % credible interval.

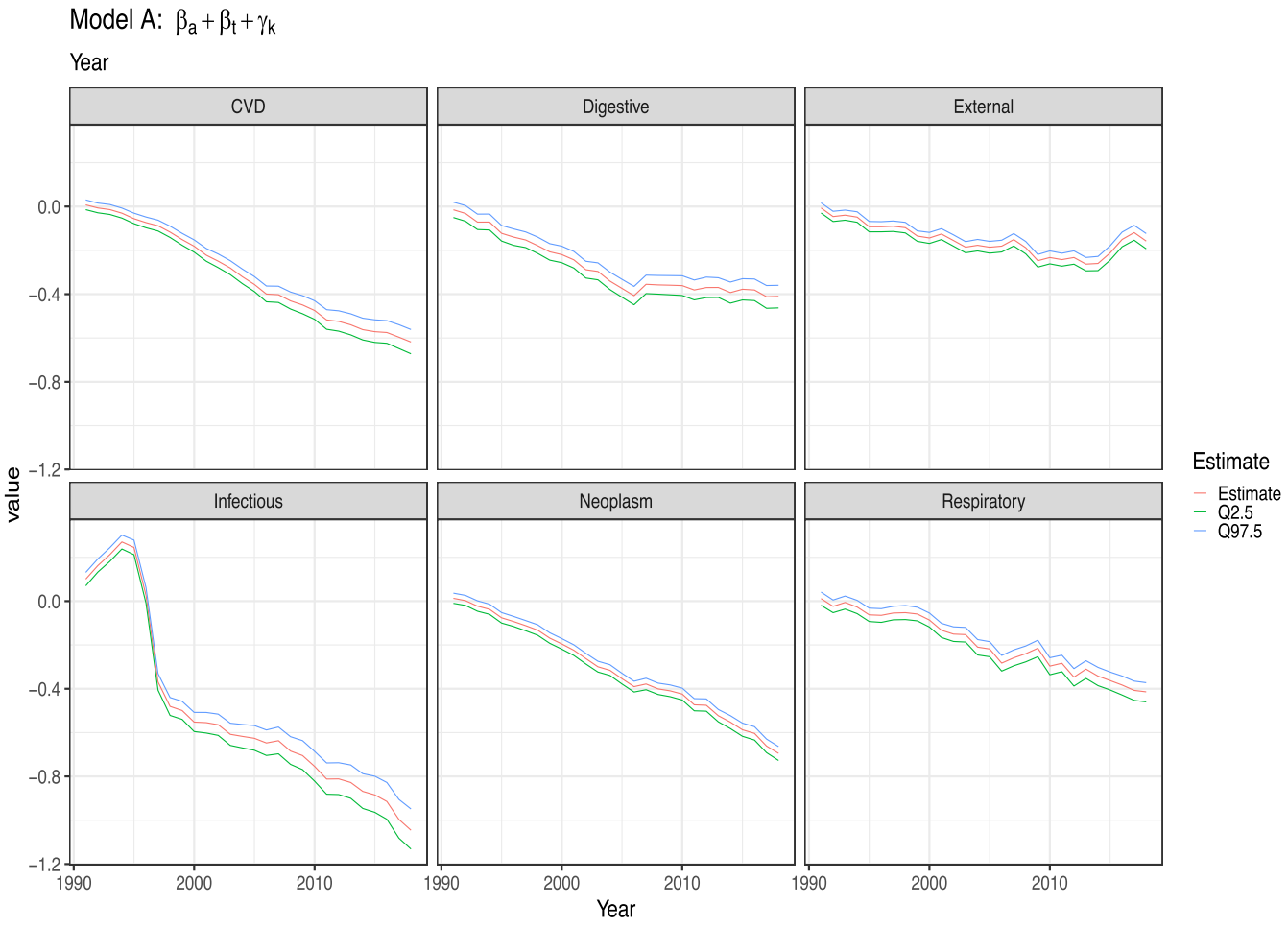


Fig. A3. Estimated year effects for each cause of death based on ModelA (age as fixed effect, cohort as random effect, period as fixed effect). The solid line shows the posterior mean, and the shaded area indicates the 95% credible interval. .

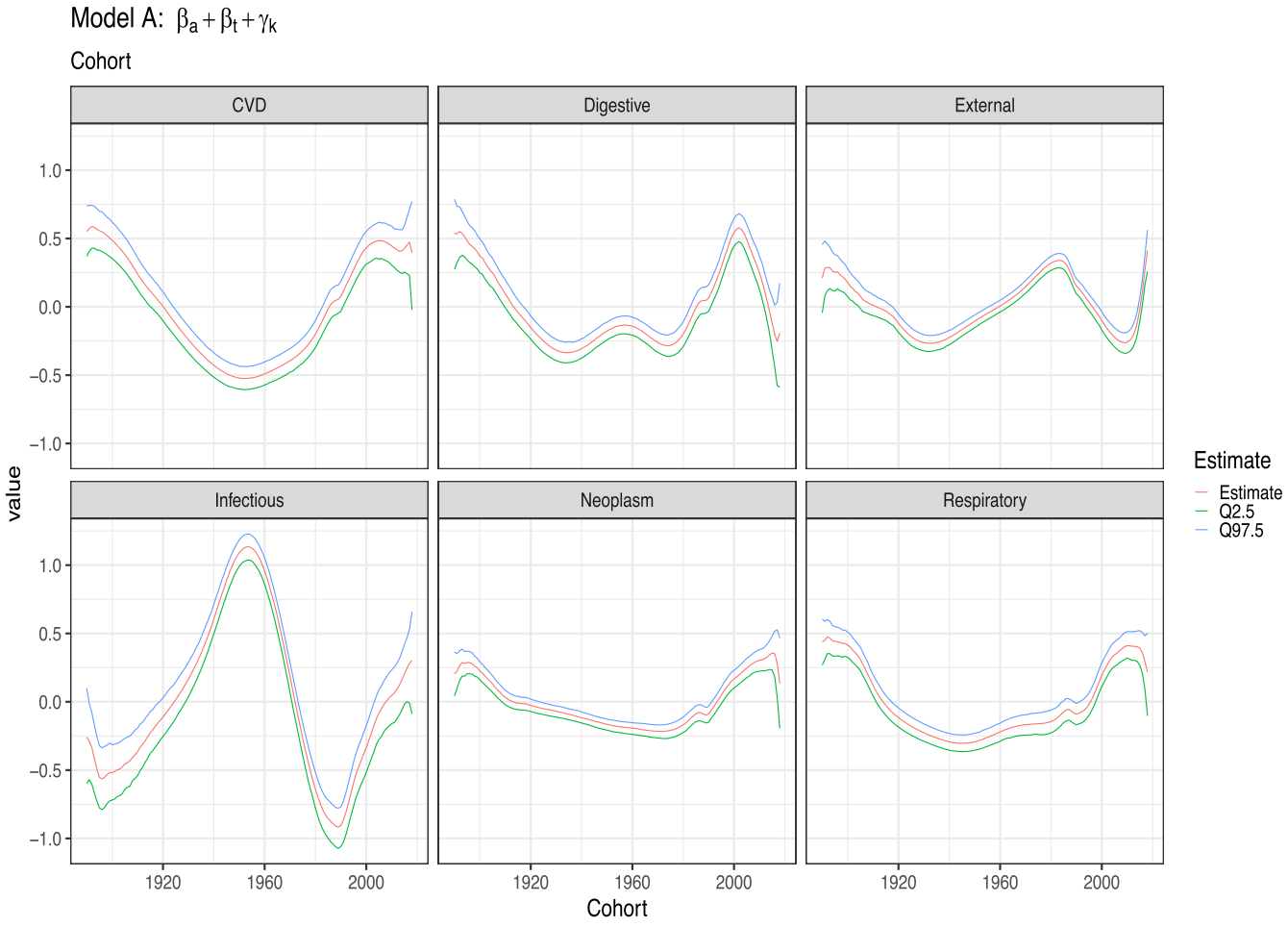


Fig. A4. Estimated cohort effects for each cause of death based on ModelA (age as fixed effect, cohort as random effect, period as fixed effect). The solid line shows the posterior mean, and the shaded area indicates the 95 % credible interval.

A.2. Two random effects

A.3. Random effects: Period, cohort

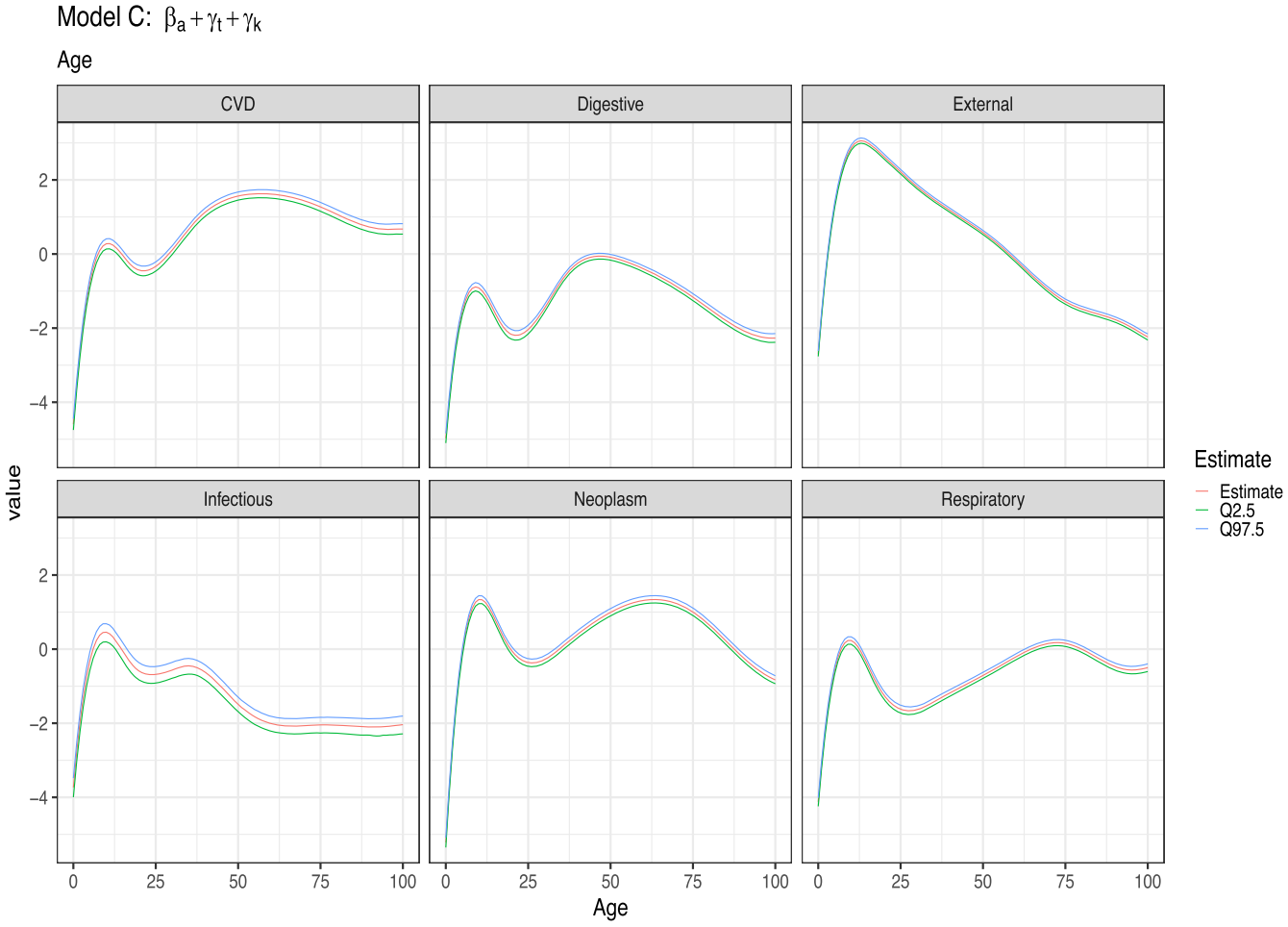


Fig. A5. Estimated age effects for each cause of death based on ModelC (age as fixed effect, cohort and period as random effects). The solid line shows the posterior mean, and the shaded area indicates the 95 % credible interval.

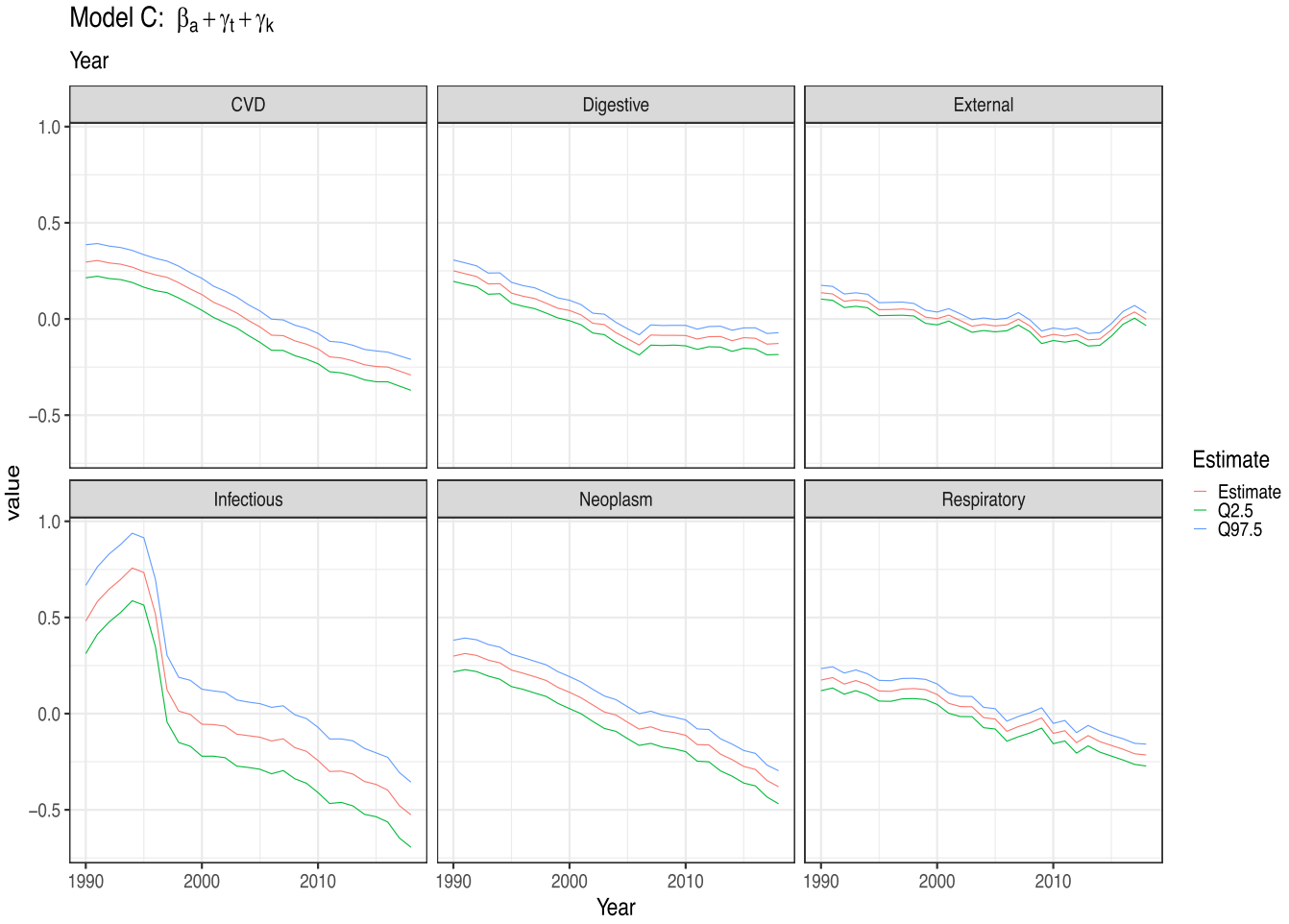


Fig. A6. Estimated period effects for each cause of death based on ModelC (age as fixed effect, cohort and period as random effects). The solid line shows the posterior mean, and the shaded area indicates the 95 % credible interval.

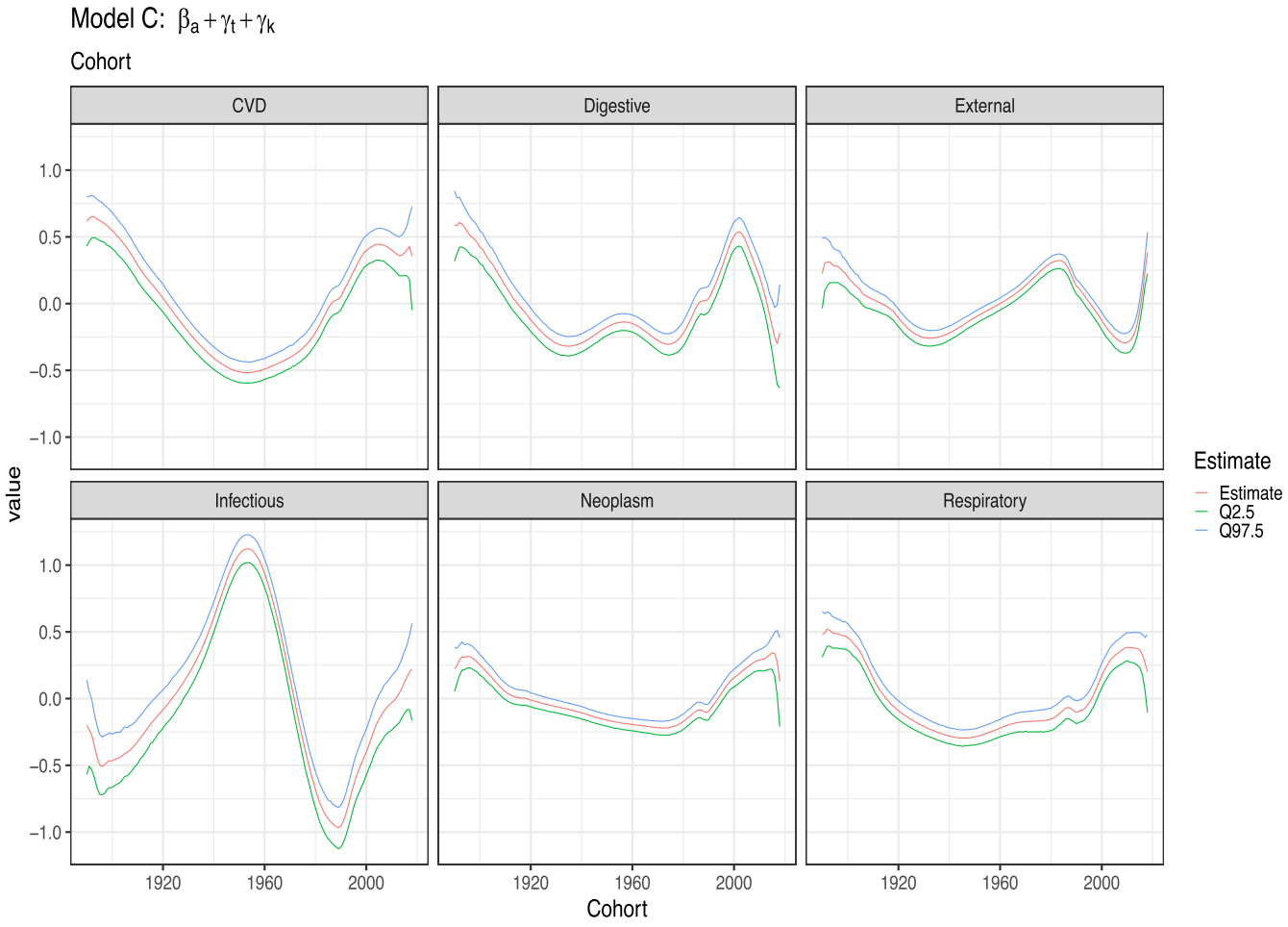


Fig. A7. Estimated cohort effects for each cause of death based on ModelC (age as fixed effect, cohort and period as random effects). The solid line shows the posterior mean, and the shaded area indicates the 95 % credible interval.

Appendix B. Log Odds

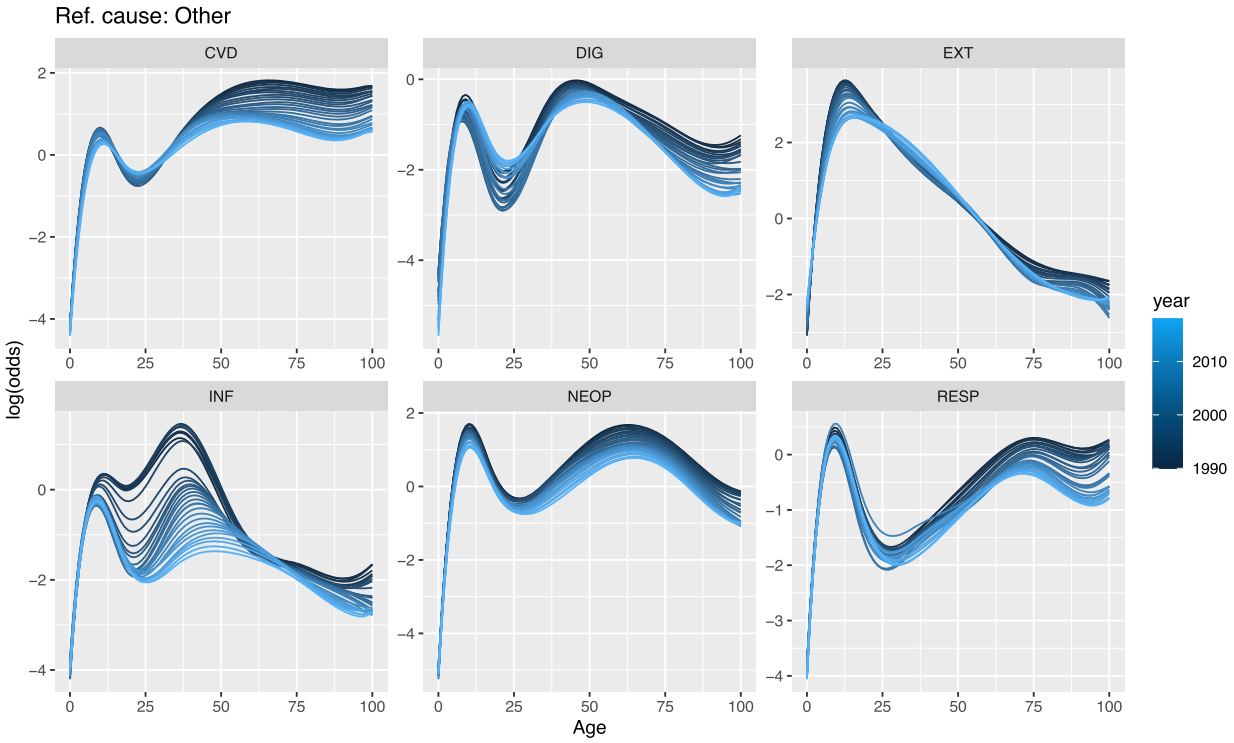


Fig. B.1. Log Odds, with the respect to Other cause.

References

Aitchison, J., 1986. The statistical analysis of compositional data, monographs on statistics and applied probability (chapman & hall, london. In: Reprinted in 2003 with Additional Material by the Blackburn Press, pp. 416–p.

Alai, D.H., Arnold, S., Sherris, M., 2015. Modelling cause-of-death mortality and the impact of cause-elimination. Ann. Actuarial Sci. 9 (1), 167–186.

Arnold, S., Sherris, M., 2013. Forecasting mortality trends allowing for cause-of-death mortality dependence. North Am. Actuarial J. 17 (4), 273–282.

Bell, A., 2020. Age, Period and Cohort Effects: Statistical Analysis and the Identification Problem. Routledge.

Bergeron-Boucher, M.-P., Canudas-Romo, V., Oeppen, J., Vaupel, J.W., 2017. Coherent forecasts of mortality with compositional data analysis. Demogr. Res. 37, 527–566.

Berzuini, C., Clayton, D., 1994. Bayesian analysis of survival on multiple time scales. Stat. Med. 13 (8), 823–838.

Besag, J., Green, P., Higdon, D., Mengersen, K., 1995. Bayesian computation and stochastic systems. Stat. sci. 3–41.

Billari, F., Graziani, R., 2024. Age-period-cohort analysis of US fertility: a realistic approach. Qual. Quant., 58, 3021–3040. https://doi.org/10.1007/s11135-023-01787-5

Booth, H., Tickle, L., 2008. Mortality modelling and forecasting: a review of methods. Ann. Actuarial Sci. 3 (1–2), 3–43.

Brooks, S., Gelman, A., Jones, G., Meng, X.-L., 2011. Handbook of Markov Chain Monte Carlo. CRC press.

Brunsdon, T.M., Smith, T., 1998. The time series analysis of compositional data. J. Off. Stat. 14 (3), 237.

Cairns, A. J.G., Blake, D., Dowd, K., Coughlan, G.D., Epstein, D., Ong, A., Balevich, I., 2009. A quantitative comparison of stochastic mortality models using data from england and wales and the united states. North Am. Actuarial J. 13 (1), 1–35.

Cardillo, G., Giordani, P., Levantesi, S., Nigri, A., Spelta, A., 2023. Mortality forecasting using the four-way CANDECOMP/PARAFAC decomposition. Scand. Actuar. J., 2023 (9), 916–932.

Chiang, C.L., 1968. Introduction to Stochastic Processes in Biostatistics, New York.

Clayton, D., Schifflers, E., 1987. Models for temporal variation in cancer rates. i: age–period and age–cohort models. Stat. Med. 6 (4), 449–467.

Currie, I.D., 2006. Smoothing and forecasting mortality rates with p-splines. Talk Institute Actuaries. http://www.ma.hw.ac.uk/iain/research/talks.html

Di Cesare, M., Murphy, M., 2009. Forecasting mortality, different approaches for different cause of deaths? The cases of lung cancer; influenza, pneumonia, and bronchitis; and motor vehicle accidents. British Actuarial J. 15 (S1), 185–211.

Dimitrova, D.S., Haberman, S., Kaishev, V.K., 2013. Dependent competing risks: cause elimination and its impact on survival. Insur. Math. Econ. 53 (2), 464–477.

Eilers, P. H.C., 2007. Ill-posed problems with counts, the composite link model and penalized likelihood. Stat. Modell. 7 (3), 239–254.

Foreman, K.J., Li, G., Best, N., Ezzati, M., 2017. Small area forecasts of cause-specific mortality: application of a bayesian hierarchical model to US vital registration data. J. R. Stat. Soc. Series C: Appl. Stat. 66 (1), 121–139.

Gelman, A., Rubin, D.B., 1992. Inference from iterative simulation using multiple sequences. Stat. Sci. 7 (4), 457–472.

Grunwald, G.K., Raftery, A.E., Guttorp, P., 1993. Time series of continuous proportions. J. R. Stat. Soci. Series B: Stat. Methodol. 55 (1), 103–116.

Gueorguieva, R., Rosenheck, R., Zelterman, D., 2008. Dirichlet component regression and its applications to psychiatric data. Comput. Stat. Data Anal. 52 (12), 5344–5355.

Hijazi, R.H., Jernigan, R.W., 2009. Modelling compositional data using dirichlet regression models. J. Appl. Probab. Stat. 4 (1), 77–91.

Hilton, J., Dodd, E., Forster, J.J., Smith, P. W.F., 2019. Projecting UK mortality by using bayesian generalized additive models. J. R. Stat. Soci. Series C: Appl. Stat. 68 (1), 29–49.

Hobcraft, J., Menken, J., Preston, S., 1985. Age, Period, and Cohort Effects in Demography: a Review. Springer.

Holford, T.R., 1983. The estimation of age, period and cohort effects for vital rates. Biometrics, 39 (2), 311–324.

Human, Cause of death database. https://www.causesofdeath.org/cgi-bin/main.php.

Human, Mortality database . https://www.mortality.org/.

Hunt, A., Blake, D., 2014. A general procedure for constructing mortality models. North Am. Actuarial J. 18 (1), 116–138.

Kaishev, V.K., Dimitrova, D.S., Haberman, S., 2007. Modelling the joint distribution of competing risks survival times using copula functions. Insur. Math. Econ. 41 (3), 339–361.

Kjærgaard, S., Ergemen, Y.E., Kallestrup-Lamb, M., Oeppen, J., Lindahl-Jacobsen, R., 2019. Forecasting causes of death by using compositional data analysis: the case of cancer deaths. J. R. Stat. Soci. Series C: Appl. Stat. 68 (5), 1351–1370.

Knorr-Held, L., Rainer, E., 2001. Projections of lung cancer mortality in west germany: a case study in bayesian prediction. Biostatistics 2 (1), 109–129.

Lanfiuti Baldi, G., Nigri, A., Shang, H.L., 2025. Age-period modeling of mortality gaps: the cases of cancer and circulatory diseases. J. Off. Stat. https://doi.org/10.1177/0282423X251350772

Luo, L., Hodges, J.S., 2020. Constraints in random effects age-period-cohort models. Sociol. Methodol. 50 (1), 276–317.

Manton, K.G., Poss, S.S., 1979. Effects of dependency among causes of death for cause elimination life table strategies. Demography 16 (2), 313–327.

Manton, K.G., Stallard, E., Vaupel, J.W., 1986. Alternative models for the heterogeneity of mortality risks among the aged. J. Am. Stat. Assoc. 81 (395), 635–644.

Mason, K.O., Mason, W.M., Winsborough, H.H., Poole, W.K., 1973. Some methodological issues in cohort analysis of archival data. Am. Sociol. Rev., 38, 242–258. https://doi.org/10.2307/2094398

Mason, W.M., Wolfinger, N.H., 2001. Cohort analysis.

Masters, R.K., Hummer, R.A., Powers, D.A., 2012. Educational differences in US adult mortality: a cohort perspective. Am. Sociol. Rev. 77 (4), 548–572.

Mathers, C.D., Loncar, D., 2006. Projections of global mortality and burden of disease from 2002 to 2030. PLoS Med. 3 (11), e442.

Van der Merwe, S., 2019. A method for bayesian regression modelling of composition data. South African Stat. J. 53 (1), 55–64.

Mills, T.C., 2010. Forecasting compositional time series. Quality Quantity 44 (4), 673–690.

Neal, R.M., 2011. MCMC using Hamiltonian Dynamics. In: Brooks, S., Gelman, A., Jones, G. & Meng X. (Eds.). Handbook of Markov Chain Monte Carlo. Chapman and Hall/CRC.

O'Brien, R.M., 2017. Mixed models, linear dependency, and identification in age-period-cohort models. Stat. Med. 36 (16), 2590–2600.

O'Brien, R.M., Hudson, K., Stockard, J., 2008. A mixed model estimation of age, period, and cohort effects. Sociol. Methods Res. 36 (3), 402–428.

Oeppen, J., et al., 2008. Coherent forecasting of multiple-decrement life tables: a test using japanese cause of death data, European Population Conference, Barcelona, Spain (2008). http://epc2008.princeton.edu/papers/

O'Brien, R.M., 2011. Constrained estimators and age-period-cohort models. Sociol. Methods Res. 40 (3), 419–452.

Peltonen, M., Asplund, K., 1996. Age-period-cohort effects on stroke mortality in sweden 1969–1993 and forecasts up to the year 2003. Stroke 27 (11), 1981–1985.

Plat, R., 2009. On stochastic mortality modeling. Insur. Math. Econ. 45 (3), 393–404.

Preston, S.H., 2013. Mortality Patterns in National Populations: with Special Reference to Recorded Causes of Death. Elsevier.

Quinn, M.J., d'Onofrio, A., Møller, B., Black, R., Martinez-Garcia, C., Møller, H., Rahu, M., Robertson, C., Schouten, L.J., La Vecchia, C., et al., 2003. Cancer mortality trends in the EU and acceding countries up to 2015. Ann. Oncol. 14 (7), 1148–1152.

Renshaw, A., Haberman, S., 2003. Lee–carter mortality forecasting: a parallel generalized linear modelling approach for england and wales mortality projections. J. R. Stat. Soci. Series C: Appl. Stat. 52 (1), 119–137.

Renshaw, A.E., Haberman, S., 2006. A cohort-based extension to the lee–carter model for mortality reduction factors. Insur. Math. Econ. 38 (3), 556–570.

Richards, S.J., 2009. Selected issues in modelling mortality by cause and in small populations. British Actuarial J, 15 (S1), 267–283.

Richards, S.J., Currie, I.D., Kleinow, T., Ritchie, G.P., 2019. A stochastic implementation of the APCI model for mortality projections. British Actuarial J, 24, e13.

Rizzi, S., Gampe, J., Eilers, P. H.C., 2015. Efficient estimation of smooth distributions from coarsely grouped data. Am. J. Epidemiol. 182 (2), 138–147.

Rodgers, W.L., 1982. Estimable functions of age, period, and cohort effects. Am. Sociol. Rev., 47, 774–787. https://doi.org/10.2307/2095213

Smith, H.L., 2020. Age-period-cohort analysis: what is it good for?

Smith, T.R., Wakefield, J., 2016. A review and comparison of age–period–cohort models for cancer incidence. Statist. Sci., 31 (4), 591–610. https://doi.org/10.1214/16-STS580

Stefanucci, M., Mazzuco, S., 2022. Analysing cause-specific mortality trends using compositional functional data analysis. J. R. Stat. Soci. Series A: Stat. Soci. 185 (1), 61–83.

Team, S.D., 2015. Stan modeling language users guide and reference manual.

Thompson, R., Baker, R.J., 1981. Composite link functions in generalized linear models. J. R. Stat. Soci. Series A: Stat. Soci.) 30 (2), 125–131.

Tsiatis, A., 1975. A nonidentifiability aspect of the problem of competing risks. Proc. Nat. Acad. Sci. 72 (1), 20–22.

Vehtari, A., Gelman, A., Gabry, J., 2017. Practical bayesian model evaluation using leave-one-out cross-validation and WAIC. Stat. Comput. 27, 1413–1432.

Wilmoth, J.R., 1995. Are mortality projections always more pessimistic when disaggregated by cause of death? Math. Popul. Stud. 5 (4), 293–319.

Yang, Y., Land, K.C., 2006. A mixed models approach to the age-period-cohort analysis of repeated cross-section surveys, with an application to data on trends in verbal test scores. Sociol. Methodol. 36 (1), 75–97.