Statistics in Medicine

WILEY

Statistics in Medicine

RESEARCH ARTICLE OPEN ACCESS

Modeling Chronic Disease Mortality by Methods From Accelerated Life Testing

Marina Zamsheva[1,2] | Alexander Kluttig[1,2] | Andreas Wienke[1,2] | Oliver Kuss[3,4]

[1]Institute of Medical Epidemiology, Biostatistics, and Informatics, Medical Faculty of the Martin-Luther-University Halle-Wittenberg, Halle, Saale, Germany | [2]Interdisciplinary Center for Health Sciences, Medical Faculty of the Martin-Luther-University Halle-Wittenberg, Halle, Saale, Germany | [3]German Diabetes Center, Leibniz Institute for Diabetes Research at Heinrich-Heine-University Düsseldorf, Institute for Biometrics and Epidemiology, Düsseldorf, Germany | [4]Center for Health and Society, Medical Faculty and University Hospital Düsseldorf, Heinrich-Heine-University Düsseldorf, Düsseldorf, Germany

Correspondence: Marina Zamsheva (marina.zamsheva@uk-halle.de)

Received: 8 September 2023 | **Revised:** 12 September 2024 | **Accepted:** 16 September 2024

Funding: The authors received no specific funding for this work.

Keywords: Gompertz distribution | lifetime analysis | maximum likelihood | tampered random variable | Type 2 diabetes

ABSTRACT

We propose a parametric model for describing chronic disease mortality from cohort data and illustrate its use for Type 2 diabetes. The model uses ideas from accelerated life testing in reliability theory and conceptualizes the occurrence of a chronic disease as putting the observational unit to an enhanced stress level, which is supposed to shorten its lifetime. It further addresses the issue of semi-competing risk, that is, the asymmetry of death and diagnosis of disease, where the disease can be diagnosed before death, but not after. With respect to the cohort structure of the data, late entry into the cohort is taken into account and prevalent as well as incident cases inform the analysis. We finally give an extension of the model that allows age at disease diagnosis to be observed not exactly, but only partially within an interval. Model parameters can be straightforwardly estimated by Maximum Likelihood, using the assumption of a Gompertz distribution we show in a small simulation study that this works well. Data of the Cardiovascular Disease, Living and Ageing in Halle (CARLA) study, a population-based cohort in the city of Halle (Saale) in the eastern part of Germany, are used for illustration.

1 | Introduction

Chronic diseases, for example, cardiovascular disease, cancer, or diabetes have overruled infectious diseases as leading causes of death since long times. Actually and following the World Health Organization, chronic (noncommunicable) diseases together accounted for 74% of global deaths in 2019 [1]. Detailed knowledge of chronic disease mortality and its dynamics is thus essential for policy and decision makers to organize health systems and health care in the 21st century.

From a statistical viewpoint, chronic disease mortality comes with some challenges. In essence, two time-to-event processes are involved, one leading to onset of disease, the other one leading to death. Both process are interrelated, in general have similar risk factors and are correlated within observational units. However, there is also a substantial asymmetry between the two processes: One can contract a chronic disease and will die with it, but not vice versa, that is, after death there is no possibility to be newly diagnosed with the disease. This asymmetry has been termed "semi-competing risk" in statistics, with death being denoted as the "terminal" or "absorbing" event. Several statistical methods have been used to model this semi-competing risk, most prominently illness-death models (see, e.g., Haneuse and Lee [2] for an overview), but also shared frailty models [3] or marginal regression models with corrections for dependent censoring [4].

Here, we propose an entirely new approach to semi-competing risk data, which borrows ideas from accelerated life testing (ALT) for technical items, for example, light bulbs or cable insulation (Nelson [5] and Pascual, Meeker, and Escobar [6]). As these items are in general highly reliable today, assessing their lifetime under normal conditions necessitates prohibitively long observation times. One solution is to put the devices to higher-than-normal environmental conditions or stresses (e.g., temperature, voltage, pressure) to achieve more failures in shorter times (thus making the experiments more efficient), and then extrapolating back to normal conditions.

With respect to chronic diseases in humans, we conceptualize the diagnosis of a chronic disease here as inducing stress for a human being, which is expected to shorten its residual lifetime, or, equivalently, accelerating its time to death. In ALT, this situation is termed a "step-stress" situation because the stress (=chronic disease) is induced at a specific point in time. Moreover, the term "partial" is used, because observations are observed under normal (=before disease diagnosis) as well as under accelerated (=after disease diagnosis) conditions. A step-stress partial ALT model (SSPALT) thus consists of a life distribution under normal conditions and an acceleration factor which governs lifetime under stress conditions [7]. There are several models for assessing step-stress (see, e.g., Kateri and Kamps [8, 9] for a review), and we focus here on the tampered random variable model (TRV) of Goel [10].

The methodical work presented here was motivated by the clinical example of Type 2 diabetes. The available data originated from a cohort study, the CARLA cohort, [11–13] which imposes several additional challenges for the model. To be more specific, we want to properly account for entry into the cohort during lifetime (rather than at birth) and include the information of prevalent (=diagnosis before cohort entry) as well as incident (=new diagnosis while on cohort follow-up) diabetes cases.

In Section 2, we introduce the model, its likelihood function, and methods for parameter estimation. Section 3 describes the design and results of a small simulation study, inspired by the CARLA data. In Section 4, we give the results for the analysis of the CARLA data, and Section 5 re-iterates strengths and limitations of the model and gives an outlook to future work.

2 | The Model

2.1 | Notation and Basic Assumptions

We consider the TRV model as defined by Goel [10]. The model was developed for step-stress experiments in reliability analysis and describes the total lifetime of the (in general, technical) items under observation. The total lifetime of an item tested under standard conditions, that is, before stress is introduced, is denoted by the random variable T^* with cumulative distribution function (CDF) $F(t^*;\theta)$, with θ a set of parameters to be estimated. A real number y from the same time scale as T^* is considered either fixed before observing or chosen by some random procedure independently of T^* and is called tampering point of the observation. The TRV T that describes the total lifetime of an item (depending on whether it experiences a stress condition during its entire lifetime or not) is then given by

$$T = \begin{cases} T^*, & T^* \le y \\ y + \frac{T^*-y}{\gamma}, & T^* > y \end{cases} \tag{1}$$

with $\gamma > 0$ ("the tampering coefficient" or "acceleration factor") an unknown parameter that describes the change in the remaining life time after inducing stress at y. The tampering coefficient quantifies external effects that change the residual life time after y. For $0 < \gamma < 1$ this remaining life time is prolonged (thus decelerating the time to failure), for $\gamma > 1$ the remaining life time is shortened (thus accelerating the time to failure). Note that $\frac{1}{\gamma}$ corresponds to the tampering coefficient α as it was defined by Goel [10].

The survival function of the TRV model (1) is

$$S_T(t) = \begin{cases} S(t), & t \le y \\ S(y + \gamma(t-y)), & t > y \end{cases} \tag{2}$$

We use model (1) for the analysis of chronic disease mortality in humans. In particular, we are interested to describe the change in lifetimes for individuals after the diagnosis of Type 2 diabetes. To this task, let the random variable T^* denote the lifetime, that is, the time from age at birth (0) to age at death, of an individual without diabetes during its whole lifetime. In terms of ALT terminology, this corresponds to the normal stress condition. If an individual gets a Type 2 diabetes diagnosis at the age of y, then its lifetime is tampered, so that $T = y + \gamma^{-1}(T^* - y)$ is observed and the residual lifetime after y is accelerated by γ. It is well known that the effect of diabetes is to shorten the residual lifetime of the diagnosed individual [14] and so we expect γ to be larger than 1.

We further note that we have to consider the age at diabetes diagnosis y to be fixed, but individual-specific (y_i), although we omit the index i for the tampered variables T_i and the corresponded individual-specific y_i for better readability. This is due to the fact that we observe highly unique ages (i.e., stress-level changes) of diabetes diagnoses, whereas in standard ALT testing the time point of stress-level change is typically identical for all items under study. In principle, this assumption invalidates the proofs for consistency and asymptotic normality of maximum likelihood estimators as given in Goel [10] and the assumptions 5.A from Goel [10] needed to be proven for the Gompertz distribution.

We assume that T^* follows a Gompertz distribution with parameters $\alpha > 0$ and $\beta > 0$ for $t \ge 0$, hazard function $\lambda(t;\alpha,\beta) = \lambda(t) = \alpha \cdot e^{\beta t}$, and survival function $S(t;\alpha,\beta) = S(t) = \exp\left(-\alpha/\beta \cdot \left(e^{\beta t} - 1\right)\right)$. The Gompertz distribution was recently shown to describe nearly perfectly diabetes mortality in a sample of 6.5 million people above 30 years of age in Germany, [15] similar fits were also seen in other countries, see, for example, Carstensen, Rønn, and Jørgensen [16] for Denmark. In addition, Gompertz distributions gave better fits [15] than Weibull and logistic distributions, which have also been proposed as biologically plausible for age-at-death distributions. In essence, the Gompertz distribution assumes that the logarithm

of the hazard is a linear function of age, $\log(\lambda(t)) = \log(\alpha) + \beta \cdot t$. The Gompertz parameter α gives the value of the hazard function at age 0, or equivalently, $\log(\alpha)$ corresponds to the intercept of the linear log-mortality function of time. The Gompertz parameter β describes the aging process of the population and quantifies the increase of mortality with age. As the linear relation on the log-hazard scale translates into a multiplicative relation on the original hazard scale, the value of β can be interpreted as the multiplicative annual increase in the mortality hazard. Typical values for β in Western societies are around 0.1, which means that the annual hazard of dying for an individual grows by 10% each year.

Returning to the TRV model, in the Gompertz case the survival function for the total lifetime of an individual with diabetes diagnosis ($t > y$) is

$$\begin{aligned} S_T(t) &= \mathrm{S}(y + \gamma(t-y)) \\ &= \exp\left(-\alpha/\beta \cdot \left(\mathrm{e}^{\beta(y+\gamma(t-y))} - 1\right)\right) \\ &= \exp\left(\alpha/\beta - \alpha/\beta \cdot e^{\beta y} \cdot e^{\beta\gamma(t-y)}\right) \\ &= \exp\left(-\alpha/\beta \cdot \mathrm{e}^{\beta y} + \alpha/\beta - \alpha/\beta \cdot \mathrm{e}^{\beta y} \cdot \mathrm{e}^{\beta\gamma(t-y)} + \alpha/\beta \cdot \mathrm{e}^{\beta y}\right) \\ &= \exp\left(-\alpha/\beta \cdot \left(\mathrm{e}^{\beta y} - 1\right) - \alpha/\beta \cdot \mathrm{e}^{\beta y}\left(\mathrm{e}^{\beta\gamma(t-y)} - 1\right)\right) \\ &= \underbrace{\exp\left(-\alpha/\beta \cdot \left(\mathrm{e}^{\beta y} - 1\right)\right)}_{=\mathrm{S}(y)} \cdot \underbrace{\exp\left(-\alpha/\beta \cdot \mathrm{e}^{\beta y}\left(\mathrm{e}^{\beta\gamma(t-y)} - 1\right)\right)}_{\text{Survival function of a Gompertz distribution for } t-y \text{ with } \alpha'=\alpha\gamma\mathrm{e}^{\beta y},\ \beta'=\beta\gamma} \\ &= \mathrm{S}(y;\alpha,\beta) \cdot \mathrm{S}(t-y;\alpha',\beta') \end{aligned} \tag{3}$$

The survival function $S(y + \gamma(t-y))$ factorizes in two factors before and after diabetes diagnosis. The first factor $S(y)$, the probability of surviving until y, is computed from the original Gompertz distribution with parameters α and β, independent of γ. The second factor is actually the survival function of a Gompertz distribution for $t-y$, the residual lifetime after diagnosis with new parameters $\alpha' = \alpha\gamma e^{\beta y}$ and $\beta' = \beta\gamma$. In Figure 1, we show the different survival functions from the TRV model. The hazard function for $t > y$ is given as $\lambda_T(t) = \gamma\alpha \ \exp(\beta(y + \gamma(t-y))) = \gamma\lambda(y + \gamma(t-y); \alpha, \beta)$. Just aside, the factorization in (3) generalizes the "setting-the-clock-back-to-zero"-property (SCBZ) of the Gompertz distribution as given by Rao [17] to the TRV case.

The expected residual lifetime for an individual of age t without a diabetes diagnosis during the whole life is

$$\begin{aligned} e_t^* &= E(T^*|T^* > t) - t = \frac{1}{S(t;\alpha,\beta)} \int_t^\infty S(x;\alpha,\beta)dx \\ &= \int_0^\infty \frac{S(x+t;\alpha,\beta)}{S(t;\alpha,\beta)} dx = \int_0^\infty \exp\left[-\frac{\alpha \mathrm{e}^{\beta t}}{\beta}\left(\mathrm{e}^{\beta x} - 1\right)\right] dx \end{aligned} \tag{4}$$

It should be noted that (4) takes an individual perspective and describes the remaining lifetime of an individual that will be undiagnosed with diabetes for the rest of its life. From an epidemiological perspective, that is, when describing the remaining lifetime of a population, then some of the individuals within this group would actually be diagnosed with diabetes and the average remaining lifetime of this population would be overestimated when averaging across the individual lifetimes.

Rao [17] gives the expected residual lifetime for the step-stress accelerated life test, according to the TRV model. The expected residual lifetime after a diabetes diagnosis at age y is given, using (1) and (4), as follows:

$$\begin{aligned} e_y^{TRV} &= E(T - y|T > y) \overset{(1)}{=} \frac{1}{\gamma} E(T^* - y|T > y) \\ &= \frac{1}{\gamma} e_y^* \overset{(4)}{=} \frac{1}{\gamma} \int_0^\infty \exp\left[-\frac{\alpha \mathrm{e}^{\beta y}}{\beta}\left(\mathrm{e}^{\beta x} - 1\right)\right] dx \end{aligned} \tag{5}$$

The relation $e_y^{TRV} = \frac{1}{\gamma} e_y^*$ in Equation (5) explicates the term "accelerating factor" for γ. The expected residual lifetime after a diabetes diagnosis at age y is only $\frac{1}{\gamma}$ of the time without a diabetes diagnosis. In other words, the time until death is accelerated by the factor γ after a diagnosis of diabetes.

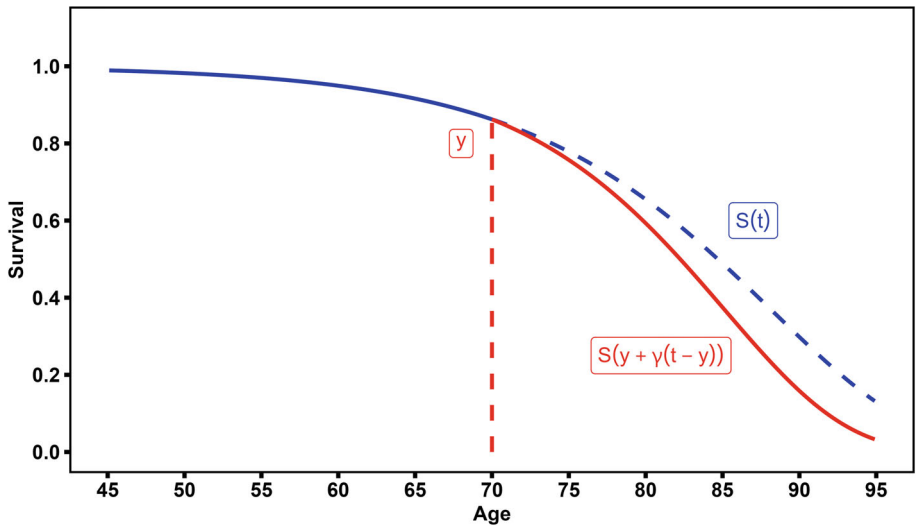

FIGURE 1 | Survival functions of the TRV model for lifetimes without diabetes diagnosis during the whole life (blue broken line) and lifetimes with diabetes diagnosis at age $y = 70$.

2.2 | Likelihood Function and Parameter Estimation

As our data originate from a population-based cohort study, we have to account for two specific challenges when defining the likelihood function of the model. First, cohort entry is observed during lifetime (rather than from birth) which requires methods for left-truncated data. Second, we have to distinguish prevalent cases which had their diabetes diagnoses before cohort entry from incident cases which are diagnosed with diabetes while on cohort follow-up.

Overall, six different cases (I–VI) have to be distinguished when defining the likelihood function and we give an overview of them in Figure 2 and Table 1. We define three indicator functions to describe the respective cases in terms of death, diabetes diagnosis, and being a prevalent or an incident disease case. For n individuals, we consider a sample of n independent random variables $T_1, T_2, \ldots, T_n$, that describe the total lifetime of n individuals according to the model (1) with observed values $t_1, t_2, \ldots, t_n, i = 1, 2, \ldots, n$, and set $\delta_i \in \{0, 1\}$ for censoring versus death, $d_i \in \{0, 1\}$ for diabetes no versus yes, and $\zeta_i \in \{0, 1\}$ for being an incident ($y_i \geq t_{0i}$) versus a prevalent ($y_i < t_{0i}$) case, where t_{0i} equals the age at cohort entry.

Collecting terms for all six cases finally yields the logarithm of the likelihood function

$$\begin{aligned} LL(\alpha, \beta, \gamma) \propto & \sum_i (1 - d_i) \cdot \delta_i \cdot \log\left[S(t_i) \cdot \lambda(t_i)/S(t_{0i})\right] \\ & + \sum_i (1 - d_i) \cdot (1 - \delta_i) \cdot \log\left[S(t_i)/S(t_{0i})\right] \\ & + \sum_i d_i \cdot (1 - \zeta_i) \cdot \delta_i \cdot \log[S(y_i + \gamma(t_i - y_i)) \\ & \cdot \gamma\lambda(y_i + \gamma(t_i - y_i))/S(t_{0i})] \\ & + \sum_i d_i \cdot (1 - \zeta_i) \cdot (1 - \delta_i) \cdot \log\left[S(y_i + \gamma(t_i - y_i))/S(t_{0i})\right] \\ & + \sum_i d_i \cdot \zeta_i \cdot \delta_i \cdot \log[S(y_i + \gamma(t_i - y_i)) \\ & \cdot \gamma\lambda(y_i + \gamma(t_i - y_i))/S(y_i + \gamma(t_{0i} - y_i))] \\ & + \sum_i d_i \cdot \zeta_i \cdot (1 - \delta_i) \cdot \log\left[S(y_i + \gamma(t_i - y_i))/S(y_i + \gamma(t_{0i} - y_i))\right] \end{aligned} \tag{6}$$

Under the assumption of a Gompertz distribution for the total lifetime the unknown parameters α, β and γ of model (1) can be straightforwardly estimated by maximum likelihood methods. In principle, each software that allows maximizing a non-linear function with several parameters could be used. We used the NLMIXED procedure in SAS 9.4 (SAS Institute Inc., Cary, NC, USA), the respective code is given in the online supplement.

2.3 | An Extension for Age at Diagnosis Only Partially Observed

The likelihood function (6) is appropriate only if age at diabetes diagnosis is known exactly. However, we noticed a number of observations in our example cohort, where age at diagnosis was only reported as having occurred within an interval or at a respective full age, given as an integer value. We thus give here an extension of the likelihood function (6) if only partial information about age at diagnosis is available.

In terms of notation, we assume that the age of diabetes onset is observed within an interval $[y_l, y_r]$. The random variable Y

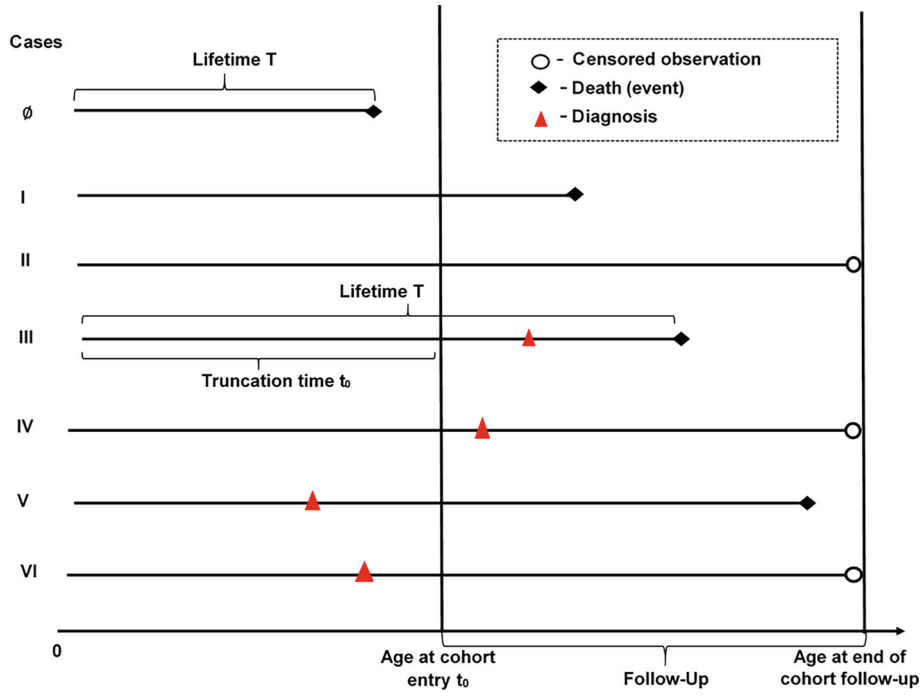

FIGURE 2 | Illustration of the six different possibilities ("cases") for observed life courses with respect to death, diabetes diagnosis and prevalence/incidence of diabetes with respect to cohort entry. Please note, that for simplicity it is assumed that all individuals enter the cohort in the same age, which is not the case in our example data set.

TABLE 1 | Contributions to the likelihood function ("cases") with respect to death, diabetes diagnosis, and prevalence/incidence of diabetes with respect to cohort entry. $S(t)$ and $\lambda(t)$ are the survival and the hazard functions, defined in Section 2.1.

Case	Description	d_i	ζ_i	δ_i	Contribution to likelihood function
∅	Death before cohort entry	—	—	—	—
I	Death during cohort follow-up without diabetes diagnosis before cohort entry or during cohort follow-up	0	—	1	$\frac{S(t)\cdot\lambda(t)}{S(t_0)}$
II	Neither death nor diabetes diagnosis before cohort entry or during cohort follow-up	0	—	0	$\frac{S(t)}{S(t_0)}$
III	Incident diagnosis with death during cohort follow-up	1	0	1	$\frac{S(y+\gamma(t-y))\cdot\gamma\lambda(y+\gamma(t-y))}{S(t_0)}$
IV	Incident diagnosis without death during cohort follow-up	1	0	0	$\frac{S(y+\gamma(t-y))}{S(t_0)}$
V	Prevalent diagnosis with death during cohort follow-up	1	1	1	$\frac{S(y+\gamma(t-y))\cdot\gamma\lambda(y+\gamma(t-y))}{S(y+\gamma(t_0-y))}$
VI	Prevalent diagnosis without death during cohort follow-up	1	1	0	$\frac{S(y+\gamma(t-y))}{S(y+\gamma(t_0-y))}$

denotes the age at diagnosis onset and follows some prespecified distribution with density function g and cumulative density function G. The density function for age at diagnosis on an interval $[y_l, y_r]$ is given by

$$g^*(y) = \begin{cases} \frac{g(y)}{G(y_r)-G(y_l)}, & \text{for } y \in [y_l, y_r] \\ 0, & \text{otherwise} \end{cases}$$

with

$$\int_{y_l}^{y_r} g^*(v)\mathrm{d}v = 1, \qquad g^*(y) \geq 0, \quad \forall y \in [y_l, y_r]$$

With respect to a random variable Y the conditional survival and density functions of the lifetime due to the TRV model are $S_{T|Y}(t|y) = S(y+\gamma(t-y))$ and $f_{T|Y}(t|y) = S(y+\gamma(t-y))\cdot\gamma\lambda(y+\gamma(t-y))$, respectively. From the general definition of conditional probability, the marginal density function for T, where $T^* > y_l$, can then be written as

$$\begin{aligned} f_T(t) &= \int_{y_l}^{y_r} f_{T,Y}(t,v)dv = \int_{y_l}^{y_r} f_{T|Y}(t|v)\cdot g^*(v)dv \\ &= \gamma\int_{y_l}^{y_r} S(v_i+\gamma(t_i-v_i))\cdot\lambda(v_i+\gamma(t_i-v_i))\cdot g^*(v)dv \end{aligned} \tag{7}$$

and the marginal survival function as

$$S_T(t) = \int_{y_l}^{y_r} S_{T|Y}(t|v)\cdot g^*(v)\mathrm{d}v = \int_{y_l}^{y_r} S(v_i+\gamma(t_i-v_i))\cdot g^*(v)\mathrm{d}v \tag{8}$$

Hence the entries of the likelihood function for the individuals with established diagnosis (corresponding to cases III–VI) change according to the marginal density and the marginal survival functions. The log-likelihood function for the situation when age at diagnosis is only partially observed is

$$\begin{aligned} LL(\alpha,\beta,\gamma) \propto &\sum_i (1-d_i)\cdot\delta_i\cdot\log\left[S(t_i)\cdot\lambda(t_i)/S\left(t_{0_i}\right)\right] \\ &+\sum_i (1-d_i)\cdot(1-\delta_i)\cdot\log\left[S(t_i)/S(t_{0i})\right] \\ &+\sum_i d_i\cdot\left(1-\zeta_i\right)\cdot\delta_i\cdot\log\left[f_T(t_i)/S(t_{0i})\right] \\ &+\sum_i d_i\cdot\left(1-\zeta_i\right)\cdot(1-\delta_i)\cdot\log\left[S_T(t_i)/S(t_{0i})\right] \\ &+\sum_i d_i\cdot\zeta_i\cdot\delta_i\cdot\log\left[f_T(t_i)/S_T(t_{0i})\right] \\ &+\sum_i d_i\cdot\zeta_i\cdot(1-\delta_i)\cdot\log\left[S_T(t_i)/S_T(t_{0i})\right] \end{aligned} \tag{9}$$

For parameter estimation, maximizing (9) involves numerical integration with respect to the marginal density and the marginal survival function.

3 | Simulation Study

In this section, we report on a small simulation study to investigate the statistical properties of the TRV model with respect to the epidemiology of chronic diseases. The simulation settings, that is, the true values for the simulation parameters, were motivated by the CARLA study [11–13] or more generally and by diabetes mortality in Germany. In particular and with a view toward challenges of cohort data, we compare different strategies to handle partially observed diagnosis age and include censoring as well as truncation in the simulation procedure. The simulation study was performed using SAS (SAS Institute Inc., Cary, NC, USA, Version 9.4) and reported according to Morris, White, and Crowther [18].

3.1 | Data-Generating Process

For each individual, the year of birth was generated from a uniform distribution between 1920 and 1958. Based on the year of birth, the age at cohort entry t_{0i} in 2003 was assigned to each individual. To simulate age at death without diabetes diagnosis T^*, we used a Gompertz distribution with parameters $\alpha = 0.00001$ and $\beta = 0.105$, approximately averaging those values across sexes in Germany.

Age at diabetes diagnosis was simulated using a Weibull distribution with shape parameter k and scale parameter l. The density and the cumulative density functions are

$$f(y) = \frac{k}{l}\left(\frac{y}{l}\right)^{k-1} \exp\left(-\left(\frac{y}{l}\right)^{k}\right),$$
$$F(y) = 1 - \exp\left(-\left(\frac{y}{l}\right)^{k}\right), \text{for } y \geq 0 \tag{10}$$

Values of $k = 5$ and $l = 95$ were used for the simulation of age at diagnosis, resulting in about 20% diabetes diagnoses as observed in the CARLA cohort. Individuals with diabetes diagnosis during lifetime ($T^* > y$) were switched to an increased stress level in the model and their total life span was calculated according to the model (1) as

$$T = y + \frac{T^* - y}{\gamma}$$

We varied the tampering coefficient γ using the values 1, 1.2, and 1.6, with $\gamma = 1$ denoting the null effect, that is, no influence of diabetes on residual lifetime. The value of $\gamma = 1.2$ is motivated by the respective value in the CARLA cohort, and $\gamma = 1.6$ would describe a more severe chronic disease, for example, an aggressive tumor, leading to death more rapidly.

With respect to the observed age at diabetes diagnosis, we distinguished an exact setting (where the age of diagnosis was observed exactly) and a partially observed setting. In the latter, age at diagnosis was observed only in an interval, and the respective interval limits y_l and y_r were simulated by two uniform distributions

$$y_l \in \text{Unif}(y - 8, y), \quad y_r \in \text{Unif}(y, \max(y + 8, T)) \tag{11}$$

resulting in a median (minimal, maximal) length of 6.5 (0.09, 15.4) years for the intervals. Random censoring for the time to death was taken into account also by using Gompertz distribution with parameters 0.000012 and 0.125. This resulted in about 17% of deaths and a corresponding 83% proportion of right censored data, again mirroring the numbers from the CARLA cohort. For each simulation setting, we generated $N = 1000$ data sets with sample of sizes of 3000 or 1000. Due to truncation of observations with age of death before age of cohort entry (corresponding to case $\emptyset$ in Table 1), the simulated data sets were actually smaller, roughly including 60% of the intended sample size.

3.2 | Estimands

The parameters of interest were those from model (1), which are the parameters α and β of the Gompertz distribution for the disease-free lifetime and the tampering coefficient γ which affects the lifetime after a diabetes diagnosis. We focus in particular on the tampering coefficient γ, which quantifies the impact of diabetes on the residual lifetime of an individual.

3.3 | Performance Measures

The performance of the estimation procedure was summarized in terms of mean bias over 1000 simulation sets, the mean squared error (MSE) and 95% coverage.

3.4 | Methods

Parameters were estimated by maximum likelihood and we used Newton–Raphson optimization with ridging in the NLMIXED procedure in SAS 9.4 (SAS Institute Inc., Cary, NC, USA, Version 9.4). In terms of the likelihood function, we have to distinguish between the two different simulation settings with respect to age at diagnosis. In the "exact" setting where age at diagnosis is given exactly the log-likelihood function (6) was used. In the "partially observed" (henceforth abbreviated with "PO") setting we observe only an interval $[y_{li}, y_{ri}]$ in which diabetes was diagnosed, but not the exact age of diagnosis y_i, that is, $y_i \in [y_{li}, y_{ri}]$. For parameter estimation, we considered three different approaches PO1 to PO3 to deal with the missing exact age at diagnosis.

PO1: Age at diagnosis $\bar{y}_i$ was fixed as the middle of the known interval $[y_{li}, y_{ri}]$.
$$\bar{y}_i = \frac{y_{li} + y_{ri}}{2}$$

PO2: Age at diabetes diagnosis $\tilde{y}_i$ was fixed by generating a random number from a uniformly distributed random variable on the interval $[y_{li}, y_{ri}]$.

PO3: Age at diabetes diagnosis $\tilde{y}_i$ was assumed to follow a Weibull distribution with known parameters $k = 5$ and $l = 67$ according to the distribution of the random variable Y from Section 2.3. The values of the Weibull parameters were chosen after finding the Weibull distribution as a plausible fit to nationwide data from all statutorily insured persons in Germany [19, 20].

For the approaches PO1 and PO2 where the age at diagnosis had been fixed, the log-likelihood function (6) was used for parameter estimation. For approach PO3, we used the log-likelihood function (9), numerical integration was needed for estimating the marginal distribution in PO3.

3.5 | Results

The results of the simulation study are presented below in Table 2. Overall, there were no problems with numerical robustness of the estimation procedure, we found sensible results for all 1000 simulated data sets and no indication of nonconvergence of algorithms. The additional results of the simulation study for larger ($n = 7000$) and smaller ($n = 100$) sample sizes are presented in the Appendix S1.

3.5.1 | Bias

In terms of bias, we find only small and negligible deviations from the true values for all three parameters (α, β, γ) of the model. In addition, also the approaches PO1–PO3 perform quite similar. Comparing the bias between the "exact" setting and the PO approaches, the bias for α and β parameters is also similar. However, γ tends to be underestimated by the PO approaches as compared to the "exact" setting. This results in a negative bias for γ for all PO approaches and both sample sizes.

TABLE 2 | Results of the simulation study with respect to bias, MSE, and 95% coverage.

		$\alpha = 0.00001$, $\beta = 0.105$, $\gamma = 1.0$					
		$n = 1791$[a]			$n = 598$[a]		
Setting	**Parameter**	**Bias**	**MSE**	**Cov. (%)**	**Bias**	**MSE**	**Cov.** (%)
Exact	α	1.2e − 6	1.6e − 12	90.0	4.6e − 6	2.1e − 11	85.8
	β	5.3e − 4	3.5e − 7	95.8	1.1e − 3	1.4e − 6	95.2
	γ	−7.2e − 4	3.3e − 6	94.8	−4.1e − 3	2.6e − 5	96.0
PO1	α	1.2e − 6	1.4e − 12	90.0	4.5e − 6	2.1e − 11	85.8
	β	6.1e − 4	4.4e − 7	95.8	1.2e − 3	1.7e − 6	95.1
	γ	−3.3e − 3	1.3e − 5	94.9	−7.1e − 3	6.0e − 5	95.6
PO2	α	1.1e − 6	1.4e − 12	89.7	4.5e − 6	2.1e − 11	85.9
	β	6.1e − 4	4.4e − 7	95.9	1.2e − 3	1.7e − 6	95.1
	γ	−3.3e − 3	1.3e − 5	94.8	−7.0e − 3	5.8e − 5	95.9
PO3	α	1.1e − 6	1.3e − 12	89.9	4.4e − 6	2.0e − 11	85.7
	β	6.9e − 4	5.4e − 7	95.8	1.3e − 3	1.8e − 6	95.1
	γ	−5.5e − 3	3.3e − 5	94.8	−9.4e − 3	9.7e − 5	96.0
		$\alpha = 0.00001$, $\beta = 0.105$, $\gamma = 1.2$					
		$n = 1785$[a]			$n = 595$[a]		
Setting	**Parameter**	**Bias**	**MSE**	**Cov. (%)**	**Bias**	**MSE**	**Cov. (%)**
Exact	α	1.0e − 6	1.1e − 12	90.2	3.8e − 6	1.5e − 11	85.1
	β	6.2e − 4	4.5e − 7	95.6	1.5e − 3	2.3e − 6	94.9
	γ	6.6e − 4	2.8e − 6	94.8	2.8e − 3	1.6e − 5	95.1
PO1	α	9.6e − 7	9.5e − 13	90.3	3.7e − 6	1.4e − 11	85.1
	β	7.1e − 4	5.6e − 7	95.6	1.5e − 3	2.5e − 6	95.1
	γ	−2.1e − 3	6.7e − 6	94.4	−2.8e − 4	8.0e − 6	94.8
PO2	α	9.2e − 7	8.8e − 13	89.8	3.7e − 6	1.4e − 11	85.0
	β	7.8e − 4	6.6e − 7	95.7	1.6e − 3	2.8e − 6	95.0
	γ	−5.1e − 3	2.9e − 5	94.2	−3.4e − 3	1.9e − 5	94.7
PO3	α	9.7e − 7	9.7e − 13	90.2	3.7e − 6	1.4e − 11	85.1
	β	6.9e − 4	5.4e − 7	95.6	1.5e − 3	2.5e − 6	95.1
	γ	−2.5e − 3	8.8e − 6	94.4	−7.8e − 4	8.7e − 6	95.1
		$\alpha = 0.00001$, $\beta = 0.105$, $\gamma = 1.6$					
		$n = 1766$[a]			$n = 589$[a]		
Setting	**Parameter**	**Bias**	**MSE**	**Cov. (%)**	**Bias**	**MSE**	**Cov. (%)**
Exact	α	9.2e − 7	8.8e − 13	91.1	3.2e − 6	1.1e − 11	86.1
	β	4.3e − 4	2.4e − 7	95.5	1.0e − 3	1.2e − 6	95.1
	γ	2.5e − 3	9.3e − 6	94.6	7.2e − 3	6.3e − 5	95.0
PO1	α	8.6e − 7	7.6e − 13	91.0	3.1e − 6	9.9e − 12	85.7
	β	4.4e − 4	2.5e − 7	95.7	1.0e − 3	1.2e − 6	95.3
	γ	−3.7e − 3	1.7e − 5	94.3	1.1e − 3	1.1e − 5	94.7
PO2	α	1.0e − 6	1.0e − 12	91.1	3.0e − 6	1.1e − 11	86.4
	β	3.8e − 4	2.0e − 7	95.4	9.4e − 4	1.0e − 6	95.1
	γ	−1.6e − 2	2.7e − 4	91.9	−1.2e − 2	1.4e − 4	93.4
PO3	α	8.7e − 7	7.9e − 13	91.1	3.2e − 6	1.0e − 11	85.7
	β	3.9e − 4	2.0e − 7	95.5	9.8e − 4	1.1e − 6	95.5
	γ	2.1e − 3	7.6e − 6	94.5	6.5e − 3	5.3e − 5	95.1

[a]Median sample sizes of 1000 data sets, due to truncation.

3.5.2 | Mean Squared Error

Overall, MSE values are very close to zero, indicating the accuracy of the estimation methods. Similar to the bias outcome, there is no large difference in MSE between the "exact" setting and PO approaches in the parameters α and β. For the parameter γ, MSE values are larger with the PO approaches if compared to the "exact" setting. When comparing MSE values for different values of γ, it is noticeable that the larger values of γ, the more bias and MSE values differ between the various approaches. For $\gamma = 1.0$, all approaches give very similar results with respect to performance measures. However, for $\gamma = 1.6$, the values of bias and MSE differ between the approaches and, especially for PO2, we observe compromised estimates. Moreover, both bias and MSE increase with the smaller sample size for all three different values of γ.

3.5.3 | Coverage

In terms of coverage, we find rather different results for the three parameters of the model. The smallest coverage is observed for the parameter α and ranges between 85.0% and 91.1%. This is probably a consequence of the small true value of the parameter α, which complicates estimation. It appears that with increasing γ, the proportion that confidence interval contains the true α value improves, however, only for larger sample size. In terms of the parameter β, all observed values for coverage are around 95%, which, allowing some random simulation error over the 1000 simulated data sets, indicates correct coverage of the estimands. With respect to the parameter γ, we also find satisfying values for coverage over all sample sizes with values ranging in general between 94.2% and 96%. Only in the PO2 approach with a large value of γ $(= 1.6)$ in the data, we find coverage values of less than 92%.

4 | Analysis of the CARLA Cohort

In this section, we use the TRV model to analyze data of the Cardiovascular Disease, Living and Ageing in Halle (CARLA) study, a population-based cohort in the city of Halle (Saale) in the eastern part of Germany. The primary aim of the study was to investigate risk factors for cardiovascular disease based on comprehensive cardiological phenotyping of study participants and was extended to study factors related to healthy aging [11, 12]. We use data from all $n = 1779$ (54% men and 46% women) study participants, who were aged $45 - 83$ during the baseline examination which took place between 2002 and 2005. The participants were drawn from the population registry of the city of Halle (Saale), Germany in 2002. The detailed recruitment procedure has been described by Greiser et al. [12] Diabetes status was recorded during baseline and two follow-up investigations (follow-up 1 between 2007 and 2010, follow-up 2 in 2013) by self-reported, physician-confirmed diabetes diagnosis or self-reported antidiabetic medication within the preceding 7 days. This means that some study participants were already diagnosed with diabetes before being included in the study (prevalent case), while others were diagnosed with diabetes during the follow-up period of the study (incident case). Therefore, the study entry is not at the same time as diabetes diagnosis (incident and prevalent cases). Furthermore, vital status was ascertained before the beginning of two follow-ups and additionally in the years 2012 and 2019 by contacting the residents' registration office [13]. To avoid misclassification of participants who were diagnosed with diabetes after 2013, we only use information from follow-up 2 and especially ignore the information of the vital status from 2019.

Overall, we recorded 406 Type 2 diabetes diagnoses, of which 276 were prevalent at baseline, and 130 were incident during the cohort follow-up. With respect to the age at diagnosis, 375 diagnosis ages were observed on small intervals with $y_{li} - y_{ri} \leq 1$. Ten individuals reported having diabetes at baseline, but without age at diagnosis or year of diagnosis. In these cases, we set $y_{li} = 18$ and $y_{ri} = t_{0i}$, resulting in a median (minimal, maximal) length of intervals of 49.5 $(39.7, 62.6)$ years. The remaining 21 diabetes diagnoses were observed on intervals $[y_{li}, y_{ri}]$ with a median (minimal, maximal) length of 4.1 $(3.7, 5.4)$. Finally, 305 deaths and 1474 censored cases (unreachable, declined participation or no event was occurred) were observed.

In Table 3, we give the results of fitting the TRV model to the CARLA data, where we distinguish between sex-specific and joint γ parameters for men and women. Note that we report the parameter $log(\alpha)$ instead of α to enhance readability. Indeed and to avoid numerical problems we fitted all models in the simulation also using $log(\alpha)$. In any case, γ is estimated to be larger than 1.0, pointing, as expected, to a reduction of the residual lifetime after diabetes diagnosis. According to the sex-constant γ, the residual lifetime after diagnosis is shortened on average by a factor of 1.19 $[95\%CI : 1.10, 1.27]$ independent of sex. With respect to the sex-specific analyses, estimates for γ are rather similar for sexes, amounting to 1.17 $[95\%CI : 1.04, 1.29]$ for men and 1.20 $[95\%CI : 1.09, 1.31]$ for women. Correspondingly, information criteria prefer the smaller model with a sex-constant γ.

Converting the γ factor into percentages provides the alternative interpretation of the acceleration factor of time. The loss of residual lifetime for an individual after diabetes diagnosis is on average 16% $[9\%, 21\%]$ $(= (1 - 1/\gamma) * 100)$ in comparison to an individual without diabetes diagnosis.

As the CARLA data contain partially observed ages at diabetes diagnosis we used all three given strategies (PO1 – PO3) to deal with this. As expected from the results of the simulation study, estimates for all model parameters were very similar (data not shown). Interestingly and referring to AIC/BIC values (PO1: 2481.2/2514.1; PO2: 2481.4/2514.3), the PO3 (2435.5/2462.9) approach was chosen as the favorite model.

Using the results for the sex-constant γ from Table 3, the expected residual lifetime of an individual who survived y years without diabetes diagnosis and an individual with a diagnosis at age of y can be compared according to (4) and (5). Figure 3 gives the difference in the residual life expectancy according to the TRV model and stratified for sex. As expected from clinical epidemiological knowledge, the residual lifetime of women is larger than that of men both for the groups with and without diabetes. Moreover, the solid lines in Figure 3 show the loss of lifetime due to the influence of diabetes for both sexes according to the TRV model. For instance, the expected residual lifetime

TABLE 3 | Results of fitting the TRV to the CARLA data using the PO3 approach for handling partially observed ages at diabetes diagnosis.

Parameter	Estimate [95%CI] for men	Estimate [95%CI] for women	Fit statistics
$\log(\alpha)$	-10.8 $[-12.1, -9.5]$	-16.1 $[-18.4, -13.8]$	
β	0.098 [0.083, 0.116]	0.161 [0.135, 0.192]	AIC: 2435.5
γ	1.19 [1.10, 1.27]		BIC: 2462.9
$\log(\alpha)$	-10.8 $[-12.1, -9.5]$	-16.0 $[-18.4, -13.7]$	
β	0.099 [0.083, 0.117]	0.160 [0.133, 0.192]	AIC: 2437.2
γ	1.17 [1.04, 1.29]	1.20 [1.09, 1.31]	BIC: 2470.1

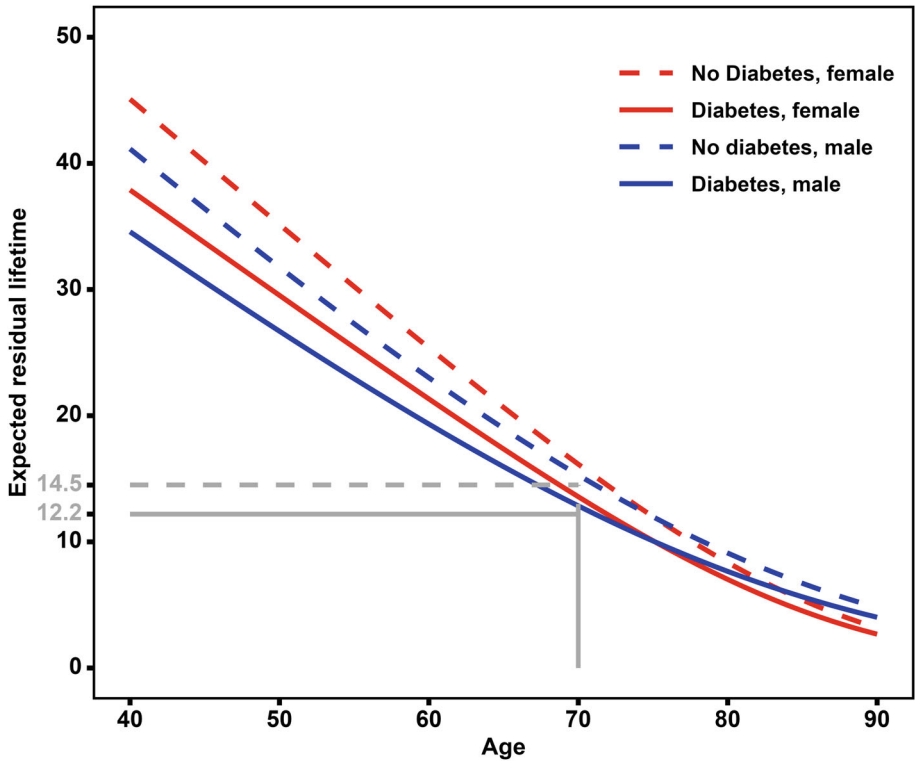

FIGURE 3 | Expected residual lifetime with and without diabetes diagnosis according to TRV model using the estimated parameters from the CARLA cohort. For instance, the expected residual lifetime of a man diagnosed with Type 2 diabetes at age of $y = 70$ is 12.2 years. In comparison, the expected residual lifetime of a man without diabetes diagnosis at the same age is 14.5 years.

of a man diagnosed with Type 2 diabetes at age of $y = 70$ is 12.2 years. In comparison, the expected residual lifetime of a man without diabetes diagnosis at the same age is 14.5 years. Dividing those two expected residual lifetimes yields, as expected, the acceleration factor γ ($14.5/12.2 \approx 1.19$).

It is straightforward to include covariates into the TRV model. We prefer an AFT interpretation on the original age scale also for the covariates. We therefore have to reparametrize the Gompertz TRV model with new parameters η and ρ and the transformation $(\eta, \rho) \to \left(\frac{\alpha}{\beta}, \frac{1}{\beta}\right)$. This yields the survival and the hazard functions of an AFT model as

$$S(t) = \exp\left(-\eta\left(e^{t/\rho} - 1\right)\right), \quad \lambda(t) = \frac{\eta}{\rho}\exp(t/\rho)$$

where the parameter $1/\rho$ models the acceleration of time and can be parametrized by covariates via an exponential function of a linear predictor $\frac{1}{\rho} = exp(\phi_0 + \phi_1 X_1 + \cdots + \phi_m X_m)$ with m covariates X_i, $j = 1, \cdots, m$ and the corresponding coefficients $\phi_1, \cdots, \phi_m$ and intercept ϕ_0. To facilitate comparison with the sex-stratified results as given above we give the results of a model with the single covariate sex in the non-stratified CARLA data set in Table 4. Interestingly, the covariate model replicates the value (1.19) of the γ parameter from the sex-stratified model. With respect to sex itself, we observe $S_m(t) = S_w\left(e^{0.028}t\right) = S_w(1.028 \cdot t)$, with S_m, S_w survival functions for men and women, respectively. That is, the time toward death is accelerated by a factor of 1.028 for men as compared to women. In terms of the model fit, the covariate model is judged inferior as compared to the sex-stratified model, AIC as well as BIC values are larger. Just aside and from a technical viewpoint, this AFT parametrization of the Gompertz distribution in the TRV model can be interpreted as a general AFT model with several time-constant covariates and one time-dependent

TABLE 4 | Results of fitting the TRV model with covariate sex to the CARLA data.

Parameter	Estimate [95%CI]	Fit statistics
$\log(\eta)$	-10.2 $[-11.5, -8.9]$	
ϕ_0	-2.17 $[-2.30, -2.04]$	AIC: 2449.7
ϕ_1 (man)	0.028 $[0.003, 0.053]$	BIC: 2471.6
γ	1.19 $[1.10, 1.28]$	

covariate and we show the respective equivalence in the Appendix S1.

5 | Discussion

We proposed a parametric model for describing chronic disease mortality from cohort data and illustrated its use for Type 2 diabetes. The innovative aspect of our proposal is that we use a model from ALT in reliability theory and conceptualize the occurrence of a chronic disease as putting the observational unit to an enhanced stress level, which is supposed to shorten its lifetime. With respect to the cohort structure of our motivating data set, our model addresses various challenges and solves several problems for such data. In particular, it properly accounts for the semi-competing risk character of the data, where diabetes can be diagnosed before death, but not vice versa. Using parametric distributions for age at diagnosis as well as for age of death allows a number of additional and easy-to-communicate insights on the time scale, for example, in terms of residual life expectancy or years of life lost due to disease. Especially by using the Gompertz distribution for age of death there is a minimal loss of information as compared with a semi- or even non-parametric modeling [15]. In addition, the parameter γ that describes the dynamic of disease can be conveniently interpreted as an acceleration factor of time. By using methods for left-truncated data, late entry into the cohort can also be taken into account. With respect to the age at diagnoses, prevalent as well as incident cases of disease are allowed. Finally, we presented an extension of the model which allows age at disease diagnosis to be observed not exactly, but only partially within an interval. To assess the statistical properties of the model and the estimation procedure a small simulation study was conducted. This shows that the model works well and the estimation algorithm was numerically stable with no convergence problems. Model parameters can be straightforwardly estimated by maximum likelihood and can be realized with every software tool that allows defining and maximizing a non-linear function.

It is fair to point to some limitations of our approach. First, the TRV is a parametric model that comes with various assumptions. While we have shown in previous work [15] that the Gompertz distribution is plausible for overall as well as for diabetes mortality in Germany we are not aware of empirical evidence for the TRV structure of the model, especially of the accelerating character of γ. There are alternative models in ALT, which might give better fits or model the process of the chronic disease toward death more plausible. To be more specific, possible alternatives to the TRV model are the tampered failure rate (TFR) model of Bhattacharyya and Soejoeti [21] or the cumulative exposure (CE) model by Nelson [5]. However, in the case of the Gompertz distribution, the TRV and the TFR model coincide, [17] and there is no additional insight from the TFR model. For more general cases, for example, when allowing different distributions for different stress levels, the hazard-based approach of the TFR model has some advantages also in comparison to the CE model [8].

Second, in the TRV model as used here, we assume a constant γ, that is a constant effect of diabetes stress on mortality. In ALT, there are also models allowing progressive stress (see, e.g. Yin and Sheng [22] for an early application with a Weibull distribution as assumed for the constant stress situation and the progressive stress being linearly related to time) and these might be further fruitful extensions for the diabetes case, but also for other diseases. In particular, a phenomenon known as compression of mortality can be observed in case of Type 2 diabetes [15, 16]. This means that the association between diabetes and mortality becomes smaller with increasing age. Thus, a model with regressive (instead of progressive) stress would be of great interest for future work, especially in the diabetes area.

Third, we have to assume that the probability of cohort entry is equal for people with and without the disease diagnosis. It is unclear if this assumption is true for the case of Type 2 diabetes. On one hand, people with diabetes might be more concerned about their health status and thus more willing by participate in a cohort study. On the other hand, they might be too stressed from managing the disease, making study participation less likely.

Fourth, the TRV does not allow to explicitly model the distribution of age at disease diagnosis. Instead, it is only concerned with modeling age at death, assuming age at diagnosis to be fixed. This has the limitation, if compared to illness-death models, that no inference on age at disease onset can be made. However, this also comes with an advantage of the TRV model, results with respect to disease mortality are valid independent of the distributional form of age at disease diagnosis.

As a further limitation, at least if the model is compared to the illness-death model, we consider the fact that the TRV model does not allow the age of disease diagnosis to be censored. In particular, it is not possible that information on mortality can be used for the cohort if disease status is not yet known. This was a problem here with the analysis of the CARLA data when we could not use the information from the mortality follow-up in 2019 for modeling disease mortality because information on disease status was only available from 2013.

Finally, there are models in ALT which would allow age-of-death not observed exactly, but only within weekly or monthly intervals, that is, as numbers of people that died within a specific interval. These situations are termed "interval monitoring" in the ALT literature, see, for example, Bobotas and Kateri, [23] for an example. In epidemiological research, these models might be useful when confidentiality issues play a role, for example, when published strata from registries are too sparsely occupied and would allow re-identification of single individuals.

There are several possibilities for future work or extensions of the model. First, we restricted here to pure modeling of diabetes mortality for men and women. Of course, it would

be straigtforward to include further covariates in the model, for example, BMI or smoking behavior, for additional insights. Similarly, random effects might be used for modeling correlated observations. Second, we did not yet think too deeply about assessing the model fit, that is, comparing observed deaths to those predicted from the TRV model.

Third, we know from external sources that the Gompertz distribution is especially well suited for modeling diabetes mortality [15] and we used it exclusively here. There might be other chronic diseases where other distributions, for example, the Weibull, might improve the fit of the TRV model. Fourth, with respect to other chronic diseases, for example, hypertension or hypercholesterolemia, it might be illuminating to include them as additional stress factors and treat them in the sense of step-stress modeling.

Finally, considering the multiple step-stress model in epidemiological context is an interesting area for further research. To extend the TRV model to more stress levels in the context of diabetes, it would be possible to consider prediabetes as an additional stress level between normal stress (without diabetes) and the stress level after a diabetes diagnosis, giving a TRV model with three stress levels.

To conclude, we think that the TRV model is a valuable extension to the epidemiologists' toolbox for modeling chronic disease mortality from cohort data. We are further confident that there are numbers of other methods in ALT or reliability theory in general that can be fruitfully used in epidemiology or clinical medicine. Interestingly, a reviewer (obviously from the field of reliability theory) proposed that this exchange should be bi-directional, because reliability theory and engineering applications can and should also learn from epidemiology.

Acknowledgment

Open Access funding enabled and organized by Projekt DEAL.

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

References

1. World Health Organization WHO, "The Top 10 Causes of Death" (2020), https://www.who.int/news-room/fact-sheets/detail/the-top-10-causes-of-death.

2. S. Haneuse and K. H. Lee, "Semi-Competing Risks Data Analysis," *Circulation. Cardiovascular Quality and Outcomes* 9, no. 3 (2016): 322–331, https://doi.org/10.1161/CIRCOUTCOMES.115.001841.

3. I. D. Ha, L. Xiang, M. Peng, J. H. Jeong, and Y. Lee, "Frailty Modelling Approaches for Semi-Competing Risks Data," *Lifetime Data Analysis* 26, no. 1 (2020): 109–133, https://doi.org/10.1007/s10985-019-09464-2.

4. A. A. Ding, G. Shi, W. Wang, and J. J. Hsieh, "Marginal Regression Analysis for Semi-Competing Risks Data Under Dependent Censoring," *Scandinavian Journal of Statistics* 36, no. 3 (2009): 481–500, https://doi.org/10.1111/j.1467-9469.2008.00635.x.

5. W. Nelson, "Accelerated Life Testing—Step-Stress Models and Data Analyses," *IEEE Transactions on Reliability* R-29, no. 2 (1980): 103–108, https://doi.org/10.1109/TR.1980.5220742.

6. F. Pascual, W. Meeker, and L. Escobar, "Accelerated Life Test Models and Data Analysis," in *Springer Handbook of Engineering Statistics*, ed. H. Pham (Springer: London, UK, 2006), 397–426.

7. M. H. Degroot and P. K. Goel, "Bayesian Estimation and Optimal Designs in Partially Accelerated Life Testing," *Naval Research Logistics Quarterly* 26, no. 2 (1979): 223–235, https://doi.org/10.1002/nav.3800260204.

8. M. Kateri and U. Kamps, "Inference in Step-Stress Models Based on Failure Rates," *Statistical Papers* 56 (2015): 639–660, https://doi.org/10.1007/s00362-014-0601-y.

9. M. Kateri and U. Kamps, "Hazard Rate Modeling of Step-Stress Experiments," *Annual Review of Statistics and Its Application* 4, no. 1 (2017): 147–168, https://doi.org/10.1146/annurev-statistics-060116-054214.

10. P. K. Goel, "Consistency and Asymptotic Normality of Maximum Likelyhood Estimators," *Scandinavian Actuarial Journal* 2 (1975): 109–118, https://doi.org/10.1080/03461238.1975.10405085.

11. K. H. Greiser, A. Kluttig, B. Schumann, et al., "Cardiovascular Disease, Risk Factors and Heart Rate Variability in the Elderly General Population: Design and Objectives of the CARdiovascular Disease, Living and Ageing in Halle (CARLA) Study," *BMC Cardiovascular Disorders* 5, no. 1 (2005): 33, https://doi.org/10.1186/1471-2261-5-33.

12. K. H. Greiser, A. Kluttig, B. Schumann, et al., "Cardiovascular Diseases, Risk Factors and Short-Term Heart Rate Variability in an Elderly General Population: The Carla Study 2002–2006," *European Journal of Epidemiology* 24, no. 3 (2009): 123–142, https://doi.org/10.1007/s10654-009-9317-z.

13. L. Hassan, L. Efremov, A. Großkopf, et al., "Cardiovascular Risk Factors, Living and Ageing in Halle: The CARLA Study," *European Journal of Epidemiology* 37, no. 1 (2022): 103–116, https://doi.org/10.1007/s10654-021-00824-7.

14. D. J. Magliano, L. Chen, B. Carstensen, et al., "Trends in All-Cause Mortality Among People With Diagnosed Diabetes in High-Income Settings: A Multicountry Analysis of Aggregate Data," *Lancet Diabetes & Endocrinology* 10, no. 2 (2022): 112–119, https://doi.org/10.1016/S2213-8587(21)00327-2.

15. O. Kuss, J. Baumert, C. Schmidt, and T. Tönnies, "Mortality of Type 2 Diabetes in Germany: Additional Insights From Gompertz Models," *Acta Diabetologica* 61 (2024): 765–771, https://doi.org/10.1007/s00592-024-02237-w.

16. B. Carstensen, P. Rønn, and M. Jørgensen, "Prevalence, Incidence and Mortality of Type 1 and Type 2 Diabetes in Denmark 1996–2016," *BMJ Open Diabetes Research & Care* 8 (2020): e001071, https://doi.org/10.1136/bmjdrc-2019-001071.

17. B. R. Rao, "Equivalence of the Tampered Random Variable and the Tampered Failure Rate Models in Accelerated Lifetesting for a Class of Life Distributions Having the 'Setting the Clock Back to Zero Property'," *Communications in Statistics - Theory and Methods* 21, no. 3 (1992): 647–664, https://doi.org/10.1080/03610929208830805.

18. T. P. Morris, I. R. White, and M. J. Crowther, "Using Simulation Studies to Evaluate Statistical Methods," *Statistics in Medicine* 38, no. 11 (2019): 2074–2102, https://doi.org/10.1002/sim.8086.

19. E. Jacobs, W. Rathmann, T. Tönnies, et al., "Age at Diagnosis of Type 2 Diabetes in Germany: A Nationwide Analysis Based on Claims Data From 69 Million People," *Diabetic Medicine* 37 (2020): 1723–1727, https://doi.org/10.1111/dme.14100.

20. T. Tönnies, A. Hoyer, R. Brinks, O. Kuss, R. Hering, and M. Schulz, "Spatio-Temporal Trends in the Incidence of Type 2 Diabetes in Germany," *Dtsch Arztebl International* 120, no. 11 (2023): 173–179, https://doi.org/10.3238/arztebl.m2022.0405.

21. G. K. Bhattacharyya and Z. Soejoeti, "A Tampered Failure Rate Model for Step-Stress Accelerated Life Test," *Communications in Statistics - Theory and Methods* 18, no. 5 (1989): 1627–1643, https://doi.org/10.1080/03610928908829990.

22. X. K. Yin and B. Z. Sheng, "Some Aspects of Accelerated Life Testing by Progressive Stress," *IEEE Transactions on Reliability* R-36 (1987): 150–155, https://doi.org/10.1109/TR.1987.5222320.

23. P. Bobotas and M. Kateri, "The Step-Stress Tampered Failure Rate Model Under Interval Monitoring," *Statistical Methodology* 27 (2015): 100–122, https://doi.org/10.1016/j.stamet.2015.06.002.

Supporting Information

Additional supporting information can be found online in the Supporting Information section.