

## Article

# An Accelerated Failure Time Cure Model with Shifted Gamma Frailty and Its Application to Epidemiological Research

Haro Aida <sup>1</sup>, Kenichi Hayashi <sup>2,\*</sup> , Ayano Takeuchi <sup>3</sup> , Daisuke Sugiyama <sup>4</sup> and Tomonori Okamura <sup>3</sup> <sup>1</sup> Graduate School of Science and Technology, Keio University, Yokohama 223-0061, Japan; haro\_aida@a5.keio.jp<sup>2</sup> Department of Mathematics, Keio University, Yokohama 223-0061, Japan<sup>3</sup> Department of Predictive Medicine and Public Health, Keio University School of Medicine, Tokyo 160-0016, Japan; ayanotakeuchi@keio.jp (A.T.); okamura@z6.keio.jp (T.O.)<sup>4</sup> Faculty of Nursing and Medical Care, Keio University, Fujisawa 252-0816, Japan; dsugiyama@keio.jp

\* Correspondence: hayashi@math.keio.ac.jp

**Abstract:** Survival analysis is a set of methods for statistical inference concerning the time until the occurrence of an event. One of the main objectives of survival analysis is to evaluate the effects of different covariates on event time. Although the proportional hazards model is widely used in survival analysis, it assumes that the ratio of the hazard functions is constant over time. This assumption is likely to be violated in practice, leading to erroneous inferences and inappropriate conclusions. The accelerated failure time model is an alternative to the proportional hazards model that does not require such a strong assumption. Moreover, it is sometimes plausible to consider the existence of cured patients or long-term survivors. The survival regression models in such contexts are referred to as cure models. In this study, we consider the accelerated failure time cure model with frailty for uncured patients. Frailty is a latent random variable representing patients' characteristics that cannot be described by observed covariates. This enables us to flexibly account for individual heterogeneities. Our proposed model assumes a shifted gamma distribution for frailty to represent uncured patients' heterogeneities. We construct an estimation algorithm for the proposed model, and evaluate its performance via numerical simulations. Furthermore, as an application of the proposed model, we use a real dataset, Specific Health Checkups, concerning the onset of hypertension. Results from a model comparison suggest that the proposed model is superior to existing alternatives.

**Keywords:** survival data analysis; accelerated failure time model; cure model; frailty model; epidemiological research



**Citation:** Aida, H.; Hayashi, K.; Takeuchi, A.; Sugiyama, D.; Okamura, T. An Accelerated Failure Time Cure Model with Shifted Gamma Frailty and Its Application to Epidemiological Research. *Healthcare* **2022**, *10*, 1383. <https://doi.org/10.3390/healthcare10081383>

Academic Editors: Maria Cristina Monti and Ottavia Eleonora Ferraro

Received: 26 June 2022

Accepted: 22 July 2022

Published: 25 July 2022

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Survival analysis statistically infers the time until the occurrence of an event of interest, and its main subject of interest is the inference of the survival function. Particularly, investigating how the characteristics of the individuals (covariates) affect survival time has been an epidemiologically important challenge. To tackle this, the proportional hazard model is most commonly used as the regression model for survival time [1]. The advantage of the proportional hazard model is that it allows intuitive interpretation of the results if the covariates are the characteristics obtained at the beginning of the observation of the individuals and are independent of time. In contrast, because the model requires the assumption that the ratio of the hazard functions is independent of time (proportional hazard assumption), its disadvantage is that it may give inappropriate analysis results if the data do not meet the assumption. Furthermore, it has been pointed out that this proportional hazard assumption is strong and often unrealistic [2].

In addition, some of the events of interest may not occur in all individuals (for example, lifestyle-related diseases and the onset of diabetes in AIDS patients). Thus, the proportional hazard model, which assumes that an event always occurs under a sufficiently long observation period, is not suitable for the problem setting mentioned above, and the

cure model has been proposed as a model that can be applied to such situations [3,4]. In particular, the mixture cure model divides individuals into latent groups with or without the occurrence of an event (uncured individuals in the case of a curable disease), expressing the survival function as their mixture distribution [5].

To solve these problems, the present study considers the accelerated failure time mixture cure model, which is a flexible model that directly expresses the relationship between the survival function and the covariates and does not necessarily require an assumption regarding the hazard ratios. In this study, we propose the mixture cure model based on the accelerated failure time (AFT) model with frailty, a latent variable that expresses the heterogeneity of the individuals [6], which enables the model to be more flexible. One of our ideas is to assume that the frailty for survival time of the uncured individuals follows a shifted gamma distribution, allowing clearer characterization of uncured individuals. In addition, we analyze the dataset on hypertension using the proposed model and compare its results with those of the existing models. One of the competitive models is a mixture cure model that assumes the proportional hazard model in the survival function for uncured patients. In this case, however, the proportional hazard assumption mentioned above is assumed in the survival function for uncured patients, leading to a model with a strong assumption. On the contrary, the assumption of the AFT model in the survival function for uncured patients allows us to consider a more flexible mixture cure model that does not require an assumption on hazard ratios [7,8].

Our major contribution of the present study is that introducing two aspects of heterogeneity and relaxation of model assumptions in parametric settings. Heterogeneity in this context allows existence of (i) uncured patients (experience an event in the long run) and cured patients in population and (ii) individual difference among uncured patients. Although similar models have been considered in the previous studies (for example, see [9]), our approach is novel in the sense that a shifted gamma frailty is found to improve model fitting and a generalized gamma distribution is considered for the uncured component for flexibility.

The overall organization of this paper is structured as follows: First, Section 2 explains survival time and censoring and defines the functions, such as the survival function, used in survival analysis. It also describes the proportional hazard and AFT models and introduces the survival time regression model considering the cure rate. Section 3 defines the AFT frailty model and the parameter estimation method for frailty. Section 4 describes a numerical example of the proposed model, in which we perform a simulation to evaluate the behavior of estimated values and summarize the results of data analysis on the onset of hypertension for discussion. Finally, Section 5 summarizes this study and discusses future challenges.

## 2. Regression Models for Survival Time Response

### 2.1. Literature Review

We develop a survival model that relaxes some assumptions in the proportional hazard model. Prior to model formulation, we briefly review some related previous studies without showing statistical models. When survival (time-to-event) outcome is of interest, statistical models are often presume that all patients will experience an event if they are observed sufficiently long time. However, this implicit assumption is justified depending on characteristics of the event or context of the study. The cure model considers a fraction of cured patients in a population, and is generally classified into non-mixture and mixture cure models [4]. Since the non-mixture cure model is not suitable for our application, the mixture cure model are considered in our study. While the proportional hazard model is most popular to model a censored survival outcome, it requires a strict “proportionality” assumption on the outcome distribution. Hence, some approach to overcome the restriction the accelerated failure time model becomes popular as an alternative [2,10,11]. To take into account heterogeneity of uncured patients, the concept of frailty is also introduced [6,12].

Frailty is a kind of random effect for each patient; we can interpret it as variability of outcome that does not come from covariates' effect.

The above reviews are three different generalizations of the proportional hazard model in survival analysis. They have developed independently or integrated partially until recent years [9,13–16]. In this study, a new model that had all of these elements is proposed with one additional twist: a shifted gamma frailty. We tabulate elements in our proposed model to compare related literature in Table 1.

**Table 1.** Selected related literature and elements of our proposed method. Elements of our proposed method and relationship among existing studies. The symbol  $\checkmark$  means “Considered”. PH and AFT means proportional hazard and accelerated failure time models, respectively.

Literature	Cured Patients	Uncured Model	Frailty
Sy and Taylor [5]	$\checkmark$	PH	–
Vaupel [6], Aalen [12]	–	PH	gamma
Pan [13]	–	AFT	gamma, log-normal
Chen et al. [15]	–	AFT	generalized gamma
Yu [9], Price and Manatunga [14]	$\checkmark$	PH	gamma
He [16]	$\checkmark$	AFT	generalized gamma
Present study	$\checkmark$	AFT	shifted gamma

## 2.2. Problem Formulation

This section provides an overview of the problem setting and the regression model for survival time. Let  $T$  be a continuous non-negative random variable representing the time of occurrence of an event of interest, and let  $C$  be a non-negative random variable that represents the censoring time. Let  $O = \min(T, C)$  and  $\Delta = I(T \leq C)$  be the actual observed survival time (right censoring), and  $\mathbf{Z}$  be a  $p$ -dimensional covariate vector. Based on these, let  $\{O, \Delta, \mathbf{Z}\}$  be the observed variables. Let the sample size be  $n$ , with the observed values of the  $i$ -individual following the probability distribution of  $(O, \Delta, \mathbf{Z}^\top)^\top$  independently of each other, and their realization is expressed as  $(t_i, \delta_i, \mathbf{z}_i^\top)^\top$ . Therefore, the sample is  $\{(t_i, \delta_i, \mathbf{z}_i), i = 1, \dots, n\}$ .

Let  $S(t|\mathbf{z}) = P(T > t | \mathbf{Z} = \mathbf{z})$ ,  $t > 0$  be the conditional survival function of  $T$  given  $\mathbf{Z} = \mathbf{z}$ . In addition,

$$h(t|\mathbf{z}) = \lim_{\epsilon \rightarrow 0^+} \frac{P(t \leq T < t + \epsilon | T \geq t, \mathbf{Z} = \mathbf{z})}{\epsilon}$$

is called the conditional hazard function of  $T$ . Furthermore, because of the expression  $h(t|\mathbf{z}) = -d \log S(t|\mathbf{z})/dt$ , the relationship between the survival function and the hazard function is as follows:

$$S(t|\mathbf{Z} = \mathbf{z}) = \exp \left[ - \int_0^t h(s|\mathbf{z}) ds \right]$$

The proportional hazard model, proposed by [1], assumes that the conditional hazard function is

$$h(t|\mathbf{Z}) = h_0(t) \exp(\mathbf{Z}^\top \boldsymbol{\beta}), \quad t > 0$$

where  $\boldsymbol{\beta} \in \mathbb{R}^p$  is the regression coefficient for the covariates, and  $h_0(t)$  is the baseline hazard function. The proportional hazard model possesses the property that the ratio of hazard

functions between different individuals does not depend on time (proportional hazard assumption). Therefore, for different  $\mathbf{Z}_1, \mathbf{Z}_2$ , the following is expressed:

$$\frac{h(t|\mathbf{Z}_1)}{h(t|\mathbf{Z}_2)} = \frac{\exp(\mathbf{Z}_1^\top \boldsymbol{\beta})}{\exp(\mathbf{Z}_2^\top \boldsymbol{\beta})}. \quad (1)$$

The most well-known regression model for survival analysis is the proportional hazard model, which is often used in actual medical research. The partial likelihood method for the semi-parametric model and the maximum likelihood method assuming a specific distribution for the baseline hazard function can be considered for the estimation of the regression coefficient  $\boldsymbol{\beta}$  [17]. The advantage of the proportional hazard model is that the effects of covariates on the hazard function can be easily interpreted. On the other hand, because this model has a strong assumption that the proportional hazard assumption (1) always holds, its disadvantage is that it cannot draw appropriate conclusions due to the biased estimation when the true probability distribution violates this assumption.

### 2.3. Mixture Cure Models

It is reasonable to consider that events, such as cancer recurrence or the onset of lifestyle-related diseases, do not occur (or are cured) in some cases. The mixture cure model [3], one of the cure models, takes this into consideration. Let  $D$  be a random variable that is 1 when it belongs to the event occurrence group and is 0 otherwise. The conditional survival function of  $T$  can then be expressed as

$$S(t|z) = P(D = 1|Z = z)P(T > t|D = 1, Z = z) + 1 \times P(D = 0|Z = z). \quad (2)$$

The right-hand side of Equation (2) contains the probability of belonging to the uncured group  $P(D = 1|Z = z)$ , in which the survival function is  $P(T > t|D = 1, Z = z)$ . Let  $p(\mathbf{Z})$ , be the model for the former and  $S_u(t|\mathbf{Z})$  be the model for the latter. The mixture cure model can then be expressed as

$$S(t|\mathbf{Z}) = p(\mathbf{Z})S_u(t|\mathbf{Z}) + (1 - p(\mathbf{Z})), \quad (3)$$

where  $\lim_{t \rightarrow \infty} S_u(t|\mathbf{Z}) = 0$ ,  $0 \leq p(\mathbf{Z}) \leq 1$  a.e. is assumed. Under the assumption described above,  $\lim_{t \rightarrow \infty} S(t|\mathbf{Z}) = 1 - p(\mathbf{Z}) \geq 0$ , in which  $1 - p(\mathbf{Z})$  can be interpreted as the cured probability.

In general, for the uncured probability  $p(\mathbf{Z})$ , the logistic regression model with the parameter  $\boldsymbol{\gamma} \in \mathbb{R}^{p+1}$  is considered to be

$$p(\mathbf{Z}; \boldsymbol{\gamma}) = \frac{\exp(\tilde{\mathbf{Z}}^\top \boldsymbol{\gamma})}{1 + \exp(\tilde{\mathbf{Z}}^\top \boldsymbol{\gamma})}, \quad (4)$$

where  $\tilde{\mathbf{Z}} = (1, \mathbf{Z}^\top)^\top$ . In addition, the proportional hazard and AFT models can be applied to the survival function for uncured patients  $S_u(t|\mathbf{Z})$  [5,7]. The observed data likelihood function is maximized in the estimation of model (3). Let  $f_u(t|\mathbf{Z})$  be the probability density function corresponding to  $S_u(t|\mathbf{Z})$ . For the observed data  $\{(t_i, \delta_i, \mathbf{z}_i) ; i = 1, \dots, n\}$ , the observation likelihood function  $L$  of model (3) is expressed as

$$\begin{aligned} L(\boldsymbol{\beta}, \boldsymbol{\gamma}, \boldsymbol{\psi}) &= \prod_{i=1}^n \left[ -\frac{\partial}{\partial t} S(t_i|\mathbf{z}_i) \right]^{\delta_i} [S(t_i|\mathbf{z}_i)]^{1-\delta_i} \\ &= \prod_{i=1}^n [p(\mathbf{x}_i) f_u(t|\mathbf{z}_i)]^{\delta_i} [p(\mathbf{x}_i) S_u(t|\mathbf{z}_i) + (1 - p(\mathbf{x}_i))]^{1-\delta_i}. \end{aligned} \quad (5)$$

In addition, the maximization of Equation (5) can be performed with the EM algorithm (expectation maximization algorithm) [18]. The non-mixture cure model, an approach to the cure model [4], originated from the biological mathematical model of cancer cells. Therefore, its application to the interpretation of data in epidemiological studies, which is the main objective of the present study, is difficult. In fact, the mixture cure model is often applied to clinical trials and observational studies [7,19].

#### 2.4. Accelerated Failure Time Models

The AFT model is a regression model that assumes

$$\log(T) = \mathbf{Z}^\top \boldsymbol{\beta} + \zeta \quad (6)$$

for  $T$  [20], with the assumption that  $\zeta$  is a random variable that represents the error and is independent of  $\mathbf{Z}$ . Here, let  $S_\varepsilon(t|\mathbf{Z})$  be the conditional survival function where  $\varepsilon = e^\zeta$ ; then,

$$S(t|\mathbf{Z}) = S_\varepsilon\left(te^{-\mathbf{Z}^\top \boldsymbol{\beta}}\right). \quad (7)$$

In addition, using the conditional probability density function  $f_\varepsilon(t)$  of  $\varepsilon$ , the conditional probability density function  $f(t|\mathbf{Z})$  of  $T$  is expressed as

$$f(t|\mathbf{Z}) = f_\varepsilon\left(te^{-\mathbf{Z}^\top \boldsymbol{\beta}}\right)e^{-\mathbf{Z}^\top \boldsymbol{\beta}}.$$

The advantage of the AFT model is that it models the relationship between  $T$  and  $\mathbf{Z}$ , and it is equivalent to a linear regression model without censoring. In addition, a model without proportional hazard assumption can also be expressed depending on the distribution that assumes for  $\zeta$  or  $\varepsilon$  (it becomes the parametric proportional hazard model in the case of exponential and Weibull distributions). As with the proportional hazard model, there are semiparametric and parametric methods for inferring the AFT model (e.g., [21]). In the present study, we consider the parametric method and make statistical inferences assuming a specific distribution, such as Weibull and logarithmic normal distributions, for  $\varepsilon$ .

In this chapter, we describe the survival time regression model that incorporates frailty and proposes a model using frailty. Frailty in survival analysis represents heterogeneity that cannot be expressed by covariates. Since Ref. [6] first proposed the frailty model, many have studied the regression models incorporating frailty and their parameter estimation methods (e.g., [13,22]). Furthermore, previous studies, including [14,23], reported cases in which the cure rate was considered. In this chapter, we propose the mixture cure model using the accelerated failure time frailty model that assumes a shifted gamma distribution of frailty in the survival function for uncured patients, and we describe its estimation method.

#### 2.5. Frailty Models

Ref. [6] proposed a model incorporating a latent variable, known as frailty, for situations that assume heterogeneity between individuals that cannot be expressed by covariates. Let  $Y$  be a non-negative random variable that represents frailty and assume that it is independent of  $(T, \mathbf{Z})$ . The frailty model for the proportional hazard model is thus the conditional hazard function  $h(t|\mathbf{Z})$  with

$$h(t|\mathbf{Z}) = Yh_0(t) \exp\left(\mathbf{Z}^\top \boldsymbol{\beta}\right). \quad (8)$$

Model (8) is a normal proportional hazard model when  $Y = 1$ . In this frailty model (8), the value of the hazard function increases when  $Y > 1$ , while it decreases when  $Y < 1$ . Therefore, it can express the changes in the magnitude of the event occurrence rate for the same  $\mathbf{Z}$  depending on the magnitude of  $Y$ .

Additionally, taking frailty into consideration based on Equations (6) and (7), the following can be considered for the AFT model:

$$\begin{cases} T = \varepsilon \exp(\mathbf{Z}^\top \boldsymbol{\beta}) \\ S_\varepsilon(t) = \{\tilde{S}(t)\}^Y, \end{cases} \quad (9)$$

where  $\tilde{S}(t)$  is the survival function. Model (9) is obtained by multiplying the hazard function by a variable representing the frailty from the relationship between the survival function and the hazard function. Based on these, when the covariates and frailty values are given, the survival function is expressed as

$$S(t|\mathbf{Z}, Y = y) = \tilde{S}\left(te^{-\mathbf{Z}^\top \boldsymbol{\beta}}\right)^y. \quad (10)$$

There are two methods of estimation for the AFT frailty model (8): a method assuming that  $\tilde{S}(t)$  follows a parametric probability distribution [24] and a semiparametric method that does not assume a specific distribution [15].

Since the variable  $Y$  is a latent variable, we marginalize the conditional survival function  $S(t|\mathbf{Z}, Y)$  with respect to  $Y$ . When we denote the distribution function of  $Y$  by  $F_Y$ , the conditional survival function  $S(t|\mathbf{Z})$  of  $T$  in the frailty model (8) can be calculated by

$$\begin{aligned} S(t|\mathbf{Z}) &= \int_0^\infty S(t|\mathbf{Z}, Y = y) dF_Y(y) \\ &= \int_0^\infty \exp\left[-yH_0(t) \exp(\mathbf{Z}^\top \boldsymbol{\beta})\right] dF_Y(y), \end{aligned} \quad (11)$$

where  $H_0(t) = \int_0^t h_0(s) ds$ . Similarly, the accelerated failure time frailty model (9) can be calculated as

$$\begin{aligned} S(t|\mathbf{Z}) &= \int_0^\infty S(t|\mathbf{Z}, Y = y) dF_Y(y) \\ &= \int_0^\infty \tilde{S}\left(te^{-\mathbf{Z}^\top \boldsymbol{\beta}}\right)^y dF_Y(y) \\ &= \int_0^\infty \exp\left[-y\tilde{H}\left(te^{-\mathbf{Z}^\top \boldsymbol{\beta}}\right)\right] dF_Y(y), \end{aligned} \quad (12)$$

where  $\tilde{H}(t)$  is a cumulative hazard function corresponding to  $\tilde{S}$ , and  $\tilde{S}(t) = \exp[-\tilde{H}(t)]$ .

An example of frailty  $Y$  is the gamma distribution  $\text{Gamma}(\zeta, \tau)$  with the scale parameter  $\zeta$  and the shape parameter  $\tau$ . In this frailty model, we assume a one-parameter gamma distribution with  $\zeta = \theta$ ,  $\tau = 1/\theta$  for identifiability. The probability density function of  $Y$  is then expressed as

$$f(y; \theta) = \frac{1}{\theta^{1/\theta} \Gamma(1/\theta)} y^{1/\theta-1} e^{-y/\theta}, \quad y > 0, \theta > 0 \quad (13)$$

and  $E[Y] = 1$  regardless of the value of  $\theta$  [25].

Assuming this gamma distribution is the distribution of  $Y$ , model (11) is expressed as

$$S(t|\mathbf{Z}) = \left(1 + \theta H_0(t) \exp(\mathbf{Z}^\top \boldsymbol{\beta})\right)^{-\frac{1}{\theta}} \quad (14)$$

and model (12) obtains

$$S(t|\mathbf{Z}) = \left(1 + \theta \tilde{H}\left(te^{-\mathbf{Z}^\top \boldsymbol{\beta}}\right)\right)^{-\frac{1}{\theta}}. \quad (15)$$



They are referred to as the gamma frailty model and the AFT gamma frailty model, respectively.

### 3. A Novel Accelerated Failure Time Frailty Mixture Cure Model

Our study uses a shifted gamma distribution for the random variable  $Y$ . In this section, we address our proposed model, the reason why the above distribution is assumed, and an estimation algorithm.

#### 3.1. Proposed Model

The mixture cure model, which is proposed in this section, is based on the survival function  $S(t|\mathbf{Z}) = p(\mathbf{Z})S_u(t|\mathbf{Z}) + (1 - p(\mathbf{Z}))$  in Equation (3). More specifically, we assume the AFT frailty model (12) in the survival function for the uncured group  $S_u(t|\mathbf{Z})$  and the logistic regression model (4) in the uncured probability  $p(\mathbf{Z})$ . In addition, this study considers the case in which  $\tilde{S}(t)$  is the survival function of a generalized gamma distribution, and we use a reparametrized representation of the generalized gamma distribution given in the original study by [26]. Therefore, using three parameters  $\mu \in \mathbb{R}, \sigma > 0, q \in \mathbb{R}$ , the probability density function is expressed as

$$f_{gg}(x; \mu, \sigma, q) = \begin{cases} |q| \frac{\{q^{-2}\}^{q-2}}{\sigma x \Gamma(q-2)} \exp \left[ q^{-2} \left( q \frac{\log x - \mu}{\sigma} - \exp \left( q \frac{\log x - \mu}{\sigma} \right) \right) \right] & (q \neq 0) \\ \frac{1}{\sqrt{2\pi}\sigma x} \exp \left\{ -\frac{(\log x - \mu)^2}{2\sigma^2} \right\} & (q = 0) \end{cases} \quad (16)$$

for  $x > 0$  [27]. The function (16) becomes the probability density function of a Weibull distribution when  $q = 1$ , a gamma distribution when  $q = \sigma$ , and a logarithmic normal distribution when  $q = 0$ . The generalized gamma distribution is the flexible random model that contains the probability distribution described above, which is often used in survival analysis. Hereafter, the survival function and cumulative hazard function of a generalized gamma distribution are expressed as  $S_{gg}(t; \mu, \sigma, q)$ ,  $H_{gg}(t; \mu, \sigma, q)$ , respectively.

Furthermore, we assume a shifted gamma distribution for the frailty  $Y$  [28]. If  $Y' \sim \text{Gamma}(\zeta, \tau)$ , then a random variable  $Y$  that follows a shifted gamma distribution with three parameters,  $\eta, \zeta$ , and  $\tau$ , is defined as

$$Y = \eta + Y'$$

Then,  $E[Y] = \eta + \zeta\tau$  and  $\text{Var}[Y] = \zeta^2\tau$ . In this study, the parameter  $\eta$  of a shifted gamma distribution is fixed at  $\eta = 1$ , and

$$Y = 1 + Y', Y' \sim \text{Gamma}(\theta, 1/\theta) \quad (17)$$

is assumed for the frailty. When the distribution function of  $Y$  is expressed as  $F_{sg}(y)$ , the marginalized survival function  $\check{S}(t|\mathbf{Z})$  is as follows:

$$\check{S}(t|\mathbf{Z}) = \int_1^\infty \left\{ \tilde{S} \left( te^{-\mathbf{Z}^\top \boldsymbol{\beta}} \right) \right\}^y dF_{sg}(y) = \tilde{S} \left( te^{-\mathbf{Z}^\top \boldsymbol{\beta}} \right) \left( 1 + \theta \tilde{H} \left( te^{-\mathbf{Z}^\top \boldsymbol{\beta}} \right) \right)^{-\frac{1}{\theta}}. \quad (18)$$

The uncured individuals can be characterized more clearly by assuming this shifted gamma distribution. The reason for this is that, assuming a gamma distribution  $\text{Gamma}(\theta, 1/\theta)$  for  $Y$ , it is highly probable that the frailty value becomes 1 or less [29]. That is,

$$\int_0^1 \frac{1}{\theta^{1/\theta} \Gamma(1/\theta)} y^{1/\theta-1} e^{-y/\theta} dy \geq \frac{1}{2}$$

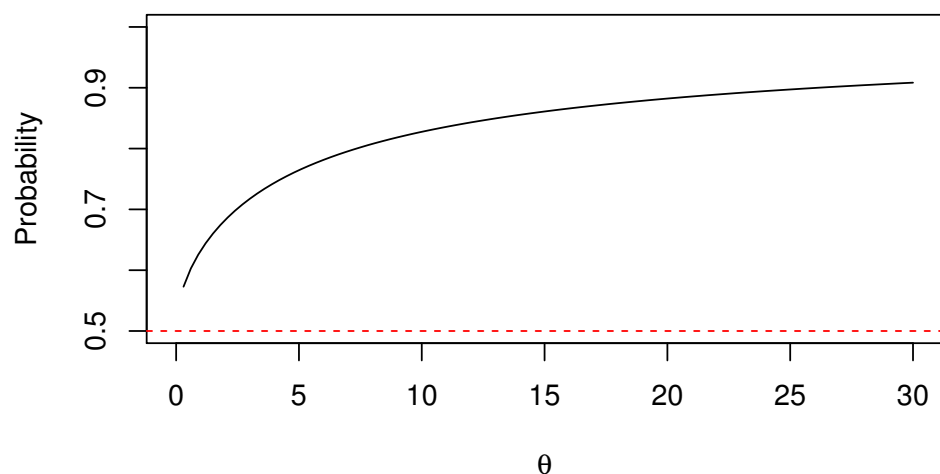
holds for almost all  $\theta$  (Figure 1). Therefore, the normal gamma frailty model considers the case in which events are less likely to occur, compared to the case in which frailty is not assumed. In fact, in the analysis of the real dataset discussed below, the estimated

value of  $\theta$  in this model was significantly close to 0, suggesting that the normal gamma frailty model [16] is inappropriate. When the shifted gamma distribution (17) is assumed for frailty,

$$\check{S}(t|\mathbf{Z}) = \tilde{S}\left(te^{-\mathbf{Z}^\top \boldsymbol{\beta}}\right) \left(1 + \theta \tilde{H}\left(te^{-\mathbf{Z}^\top \boldsymbol{\beta}}\right)\right)^{-\frac{1}{\theta}} \leq \tilde{S}\left(te^{-\mathbf{Z}^\top \boldsymbol{\beta}}\right)$$

holds for any  $t \in [0, \infty)$ . Therefore, the proposed model can characterize the survival function of the uncured group more clearly compared to the case in which frailty is not assumed. Based on these, the proposed model of the survival rate function is expressed more specifically as

$$S(t|\mathbf{Z}) = p(\mathbf{Z}; \gamma) S_{\text{gg}}\left(te^{-\mathbf{Z}^\top \boldsymbol{\beta}}; \mu, \sigma, q\right) \left(1 + \theta H_{\text{gg}}\left(te^{-\mathbf{Z}^\top \boldsymbol{\beta}}; \mu, \sigma, q\right)\right)^{-\frac{1}{\theta}} + (1 - p(\mathbf{Z}; \gamma)). \quad (19)$$



**Figure 1.** Graph of probability  $P(Y < 1)$  for  $Y \sim \text{Gamma}(\theta, 1/\theta)$ .

### 3.2. Estimation Method and Its Algorithm

In this section, we construct the parameter estimation method for the proposed model (19) based on the EM algorithm. Recall that  $\boldsymbol{\beta} \in \mathbb{R}^p$ ,  $\gamma \in \mathbb{R}^{p+1}$ , and let  $\boldsymbol{\kappa} = (\mu, \sigma, q)^\top$  and  $\theta$  be the parameters of the generalized gamma distribution and shifted gamma distribution, respectively.

Let  $D_i$  ( $i = 1, \dots, n$ ) be a random variable that is 1 when the  $i$ -th individual belongs to the event occurrence group and is 0, otherwise ( $i = 1, \dots, n$ ). Let the observation time of the  $i$ -th individual be  $O_i = \min(T_i, C_i)$  and the latent sample be  $\{(O_i, \Delta_i, \mathbf{Z}_i, D_i) ; i = 1, \dots, n\}$ . The likelihood function  $L(\boldsymbol{\beta}, \boldsymbol{\kappa}, \theta, \gamma)$  of the proposed model (19) is expressed as

$$L(\boldsymbol{\beta}, \boldsymbol{\kappa}, \theta, \gamma) = \prod_{i=1}^n \{p(\mathbf{Z}_i) f_u(O_i|\mathbf{Z}_i)\}^{\Delta_i D_i} \{p(\mathbf{Z}_i) S_u(O_i|\mathbf{Z}_i)\}^{(1-\Delta_i) D_i} \{1 - p(\mathbf{Z}_i)\}^{1-D_i}.$$

Note that indication of parameters on the right-hand side of the equation above is suppressed for simplicity of expression. Here, because

$$f_u(t|\mathbf{z}_i) = e^{-\mathbf{Z}_i^\top \boldsymbol{\beta}} f_{\text{gg}}(te^{-\mathbf{Z}_i^\top \boldsymbol{\beta}}; \boldsymbol{\kappa}) \left(1 + \theta H_{\text{gg}}(te^{-\mathbf{Z}_i^\top \boldsymbol{\beta}})\right)^{-\frac{1}{\theta}-1} \left(2 + \theta H_{\text{gg}}(te^{-\mathbf{Z}_i^\top \boldsymbol{\beta}})\right),$$

the log-likelihood function can be written as follows:



$$\begin{aligned}
\ell(\boldsymbol{\beta}, \boldsymbol{\kappa}, \theta, \gamma) = & \sum_{i=1}^n \left\{ \Delta_i D_i \left[ \log f_{\text{gg}}(O_i e^{-\mathbf{Z}_i^\top \boldsymbol{\beta}}; \boldsymbol{\kappa}) - \left( \frac{1}{\theta} + 1 \right) \log \left( 1 + \theta H_{\text{gg}}(O_i e^{-\mathbf{Z}_i^\top \boldsymbol{\beta}}; \boldsymbol{\kappa}) \right) \right] \right\} \\
& + \sum_{i=1}^n \left\{ \Delta_i D_i \left[ \log \left( 2 + \theta H_{\text{gg}}(O_i e^{-\mathbf{Z}_i^\top \boldsymbol{\beta}}; \boldsymbol{\kappa}) \right) - \mathbf{Z}_i^\top \boldsymbol{\beta} \right] \right\} \\
& + \sum_{i=1}^n \left\{ (1 - \Delta_i) D_i \left[ \log S_{\text{gg}}(O_i e^{-\mathbf{Z}_i^\top \boldsymbol{\beta}}; \boldsymbol{\kappa}) - \frac{1}{\theta} \log \left( 1 + \theta H_{\text{gg}}(O_i e^{-\mathbf{Z}_i^\top \boldsymbol{\beta}}; \boldsymbol{\kappa}) \right) \right] \right\} \\
& + \sum_{i=1}^n \left\{ D_i \tilde{\mathbf{Z}}_i^\top \boldsymbol{\gamma} - \log \left\{ 1 + \exp(\tilde{\mathbf{Z}}_i^\top \boldsymbol{\gamma}) \right\} \right\}. \tag{20}
\end{aligned}$$

where  $\tilde{\mathbf{z}}_i = (1, \mathbf{Z}_i^\top)^\top$ . Since the log-likelihood function (20) is divided into the terms related to  $\{\boldsymbol{\beta}, \boldsymbol{\kappa}, \theta\}$  and the terms related to  $\boldsymbol{\gamma}$ , their maximization can be performed separately. In the algorithm given below, let the updated value of the  $k$ -th parameter be  $\boldsymbol{\phi}^{(k)} = (\boldsymbol{\beta}^{(k)\top}, \boldsymbol{\kappa}^{(k)\top}, \theta^{(k)\top}, \boldsymbol{\gamma}^{(k)\top})^\top$  and the observed samples be  $\mathcal{D} = \{(O_i, \Delta_i, \mathbf{Z}_i); i = 1, \dots, n\}$ .

To simplify the following notation,  $E[\cdot | \mathcal{D}; \boldsymbol{\phi}^{(k)}]$  is written as  $\bar{E}[\cdot]$ . If the conditional expected value of the function (20) is expressed as  $\bar{\ell}(\boldsymbol{\beta}, \boldsymbol{\kappa}, \theta, \boldsymbol{\gamma})$ , the following is then expressed:

$$\begin{aligned}
\bar{\ell}(\boldsymbol{\beta}, \boldsymbol{\kappa}, \theta, \boldsymbol{\gamma}) = & E[\tilde{\ell}(\boldsymbol{\beta}, \boldsymbol{\kappa}, \theta, \boldsymbol{\gamma}) | \mathcal{D}; \boldsymbol{\phi}^{(k)}] \\
= & \sum_{i=1}^n \left\{ \Delta_i \bar{E}[D_i] \left[ \log f_{\text{gg}}(O_i e^{-\mathbf{Z}_i^\top \boldsymbol{\beta}}; \boldsymbol{\kappa}) - \left( \frac{1}{\theta} + 1 \right) \log \left( 1 + \theta H_{\text{gg}}(O_i e^{-\mathbf{Z}_i^\top \boldsymbol{\beta}}; \boldsymbol{\kappa}) \right) \right] \right\} \\
& + \sum_{i=1}^n \left\{ \Delta_i \bar{E}[D_i] \left[ \log \left( 2 + \theta H_{\text{gg}}(O_i e^{-\mathbf{Z}_i^\top \boldsymbol{\beta}}; \boldsymbol{\kappa}) \right) - \mathbf{Z}_i^\top \boldsymbol{\beta} \right] \right\} \\
& + \sum_{i=1}^n \left\{ (1 - \Delta_i) \bar{E}[D_i] \left[ \log S_{\text{gg}}(O_i e^{-\mathbf{Z}_i^\top \boldsymbol{\beta}}; \boldsymbol{\kappa}) - \frac{1}{\theta} \log \left( 1 + \theta H_{\text{gg}}(O_i e^{-\mathbf{Z}_i^\top \boldsymbol{\beta}}; \boldsymbol{\kappa}) \right) \right] \right\} \\
& + \sum_{i=1}^n \left\{ \bar{E}[D_i] \tilde{\mathbf{Z}}_i^\top \boldsymbol{\gamma} - \log \left\{ 1 + \exp(\tilde{\mathbf{Z}}_i^\top \boldsymbol{\gamma}) \right\} \right\} \tag{21}
\end{aligned}$$

In addition,  $\bar{E}[D_i]$  can be calculated as follows:

$$\begin{aligned}
\bar{E}[D_i] = & E[D_i | \mathcal{D}; \boldsymbol{\phi}^{(k)}] \\
= & 1 \times P(D_i = 1 | O_i, \Delta_i, \mathbf{Z}_i; \boldsymbol{\phi}^{(k)}) + 0 \times P(D_i = 0 | O_i, \Delta_i, \mathbf{Z}_i; \boldsymbol{\phi}^{(k)}) \\
= & \Delta_i P(D_i = 1 | \Delta_i = 1, O_i, \Delta_i, \mathbf{Z}_i; \boldsymbol{\phi}^{(k)}) \\
& + (1 - \Delta_i) P(D_i = 1 | \Delta_i = 0, O_i, \Delta_i, \mathbf{Z}_i; \boldsymbol{\phi}^{(k)}) \\
= & \Delta_i + (1 - \Delta_i) \frac{p_{(k)}(\mathbf{Z}_i) S_{\mathbf{u},(k)}(O_i | \mathbf{Z}_i)}{1 - p_{(k)}(\mathbf{Z}_i) + p_{(k)}(\mathbf{Z}_i) S_{\mathbf{u},(k)}(O_i | \mathbf{Z}_i)}, \tag{22}
\end{aligned}$$

where

$$S_{\mathbf{u},(k)}(t | \mathbf{z}) = S_{\text{gg}} \left( t e^{-\mathbf{z}^\top \boldsymbol{\beta}^{(k)}}; \boldsymbol{\kappa}^{(k)} \right) \left( 1 + \theta^{(k)} H_{\text{gg}} \left( t e^{-\mathbf{z}^\top \boldsymbol{\beta}^{(k)}}; \boldsymbol{\kappa}^{(k)} \right) \right)^{-\frac{1}{\theta^{(k)}}}$$

and  $p_{(k)}(\mathbf{z})$  is a function in which  $\boldsymbol{\gamma}$  in  $p(\mathbf{z}; \boldsymbol{\gamma})$  is replaced with  $\boldsymbol{\gamma}^{(k)}$ .

In the M step, the value of the parameter that maximizes Equation (21) is obtained based on the calculation result of the E step. As mentioned above, since Equation (21) is divided into the term related to  $\{\boldsymbol{\beta}, \boldsymbol{\kappa}, \theta\}$  and the term related to  $\boldsymbol{\gamma}$ , the maximization of the two parameter sets can be performed individually. Therefore, the function  $\bar{\ell}(\boldsymbol{\beta}, \boldsymbol{\kappa}, \theta, \boldsymbol{\gamma})$  is decomposed into

$$\begin{aligned}\bar{\ell}_1(\boldsymbol{\beta}, \boldsymbol{\kappa}, \theta) = & \sum_{i=1}^n \left\{ \Delta_i \bar{E}[D_i] \left[ \log f_{\text{gg}}(O_i e^{-\mathbf{Z}_i^\top \boldsymbol{\beta}}; \boldsymbol{\kappa}) - \left( \frac{1}{\theta} + 1 \right) \log \left( 1 + \theta H_{\text{gg}}(O_i e^{-\mathbf{Z}_i^\top \boldsymbol{\beta}}; \boldsymbol{\kappa}) \right) \right] \right\} \\ & + \sum_{i=1}^n \left\{ \Delta_i \bar{E}[D_i] \left[ \log \left( 2 + \theta H_{\text{gg}}(O_i e^{-\mathbf{Z}_i^\top \boldsymbol{\beta}}; \boldsymbol{\kappa}) \right) - \mathbf{Z}_i^\top \boldsymbol{\beta} \right] \right\} \\ & + \sum_{i=1}^n \left\{ (1 - \Delta_i) \bar{E}[D_i] \left[ \log S_{\text{gg}}(O_i e^{-\mathbf{Z}_i^\top \boldsymbol{\beta}}; \boldsymbol{\kappa}) - \frac{1}{\theta} \log \left( 1 + \theta H_{\text{gg}}(O_i e^{-\mathbf{Z}_i^\top \boldsymbol{\beta}}; \boldsymbol{\kappa}) \right) \right] \right\}\end{aligned}\quad (23)$$

$$\bar{\ell}_2(\gamma) = \sum_{i=1}^n \left\{ \bar{E}[D_i] \tilde{\mathbf{Z}}_i^\top \boldsymbol{\gamma} - \log \left\{ 1 + \exp(\tilde{\mathbf{Z}}_i^\top \boldsymbol{\gamma}) \right\} \right\} \quad (24)$$

and the updated value of  $(k+1)$ -th parameter can be obtained as follows:

$$\begin{aligned}\left\{ \boldsymbol{\beta}^{(k+1)}, \boldsymbol{\kappa}^{(k+1)}, \theta^{(k+1)} \right\} &= \underset{\boldsymbol{\beta}, \boldsymbol{\kappa}, \theta}{\operatorname{argmax}} \bar{\ell}_1(\boldsymbol{\beta}, \boldsymbol{\kappa}, \theta) \\ \boldsymbol{\gamma}^{(k+1)} &= \underset{\boldsymbol{\gamma}}{\operatorname{argmax}} \bar{\ell}_2(\boldsymbol{\gamma}).\end{aligned}$$

The above calculation is organized as an algorithm. Let the sample as the realization  $D$  be  $\{(t_i, \delta_i, \mathbf{z}_i) ; i = 1, \dots, n\}$ .

Step 1 Set the initial values  $\boldsymbol{\beta}^{(1)}, \boldsymbol{\kappa}^{(1)}, \theta^{(1)}, \boldsymbol{\gamma}^{(1)}$ .

Step 2 Calculate the sample version of Equation (22) for  $k = 1, 2, \dots$ . That is,

$$\hat{E}[D_i] = \delta_i + (1 - \delta_i) \frac{p(\mathbf{z}_i)^{(k)} S_{\mathbf{u},(k)}(t|\mathbf{z})}{p^{(k)}(\mathbf{z}_i) S_{\mathbf{u},(k)}(t|\mathbf{z}) + p(\mathbf{z}_i)^{(k)} S_{\mathbf{u},(k)}(t|\mathbf{z})}.$$

Step 3 Find the updated value for each parameter

$$(\boldsymbol{\beta}^{(k+1)}, \boldsymbol{\kappa}^{(k+1)}, \theta^{(k+1)}) = \underset{\boldsymbol{\beta}, \boldsymbol{\kappa}, \theta}{\operatorname{argmax}} \bar{\ell}_1(\boldsymbol{\beta}, \boldsymbol{\kappa}, \theta), \quad \boldsymbol{\gamma}^{(k+1)} = \underset{\boldsymbol{\gamma}}{\operatorname{argmax}} \bar{\ell}_2(\boldsymbol{\gamma}).$$

Step 4 If the convergence condition is satisfied, terminate the algorithm and set the estimated values to  $(\boldsymbol{\beta}^{(k+1)}, \boldsymbol{\kappa}^{(k+1)}, \theta^{(k+1)}, \boldsymbol{\gamma}^{(k+1)})$ . Otherwise, increase the value of  $k$  by 1 and return to step 2.

The estimators estimated by the EM algorithm are the value that maximizes the observation likelihood and are the maximum likelihood estimated values. Therefore, the estimators estimated by this algorithm have consistency and asymptotic normality under appropriate regularity conditions [30]. The log-likelihood function of the observation likelihood in the proposed model (19) is expressed as with  $\boldsymbol{\phi} = (\boldsymbol{\beta}^\top, \boldsymbol{\kappa}^\top, \theta, \boldsymbol{\gamma}^\top)^\top$

$$\begin{aligned}\tilde{\ell}(\boldsymbol{\beta}, \boldsymbol{\kappa}, \theta, \boldsymbol{\gamma}) &= \sum_{i=1}^n [\Delta_i \log f(O_i | \mathbf{Z}_i) + (1 - \Delta_i) \log S(O_i | \mathbf{Z}_i)] \\ &= \sum_{i=1}^n [\Delta_i \{ \log p(\mathbf{Z}_i) + \log f_{\mathbf{u}}(O_i | \mathbf{Z}_i) \} + (1 - \Delta_i) \log \{ p(\mathbf{Z}_i) S_{\mathbf{u}}(O_i | \mathbf{Z}_i) + (1 - p(\mathbf{Z}_i)) \}] \end{aligned}\quad (25)$$

Then, the information matrix of  $\boldsymbol{\phi} \in \mathbb{R}^{2p+4}$  is expressed as

$$I(\boldsymbol{\phi}) = -E \left[ \frac{\partial^2}{\partial \boldsymbol{\phi} \partial \boldsymbol{\phi}'} \ell(\boldsymbol{\phi}) \right]$$

and, from the asymptotic normality of the maximum likelihood estimator, the following holds:

$$\sqrt{n}(\hat{\boldsymbol{\phi}} - \boldsymbol{\phi}_0) \xrightarrow{d} N_{2p+4}(\mathbf{0}, I^{-1}(\boldsymbol{\phi}_0)) \text{ as } n \rightarrow \infty$$

where  $\hat{\boldsymbol{\phi}} = (\hat{\boldsymbol{\beta}}^\top, \hat{\boldsymbol{\kappa}}^\top, \hat{\theta}, \hat{\boldsymbol{\gamma}}^\top)^\top$  is the maximum likelihood estimator, and  $\boldsymbol{\phi}_0$  is the true value of  $\boldsymbol{\phi}$ .

#### 4. Numerical Examples

In this section, the behavior of estimated values of the parameters of the proposed model is confirmed by simulation. In addition, the dataset on the onset of hypertension will be analyzed using the proposed model and compared with that analyzed using existing models for discussion. The statistical software R was used for all analyses.

##### 4.1. Simulations

###### 4.1.1. Setting

The objective of this section is to confirm the behavior of estimated values of the model through a Monte Carlo simulation. First, generate the elements of the random variable vector  $\mathbf{Z} = (\mathbf{Z}_1, \mathbf{Z}_2, \mathbf{Z}_3)^\top$  representing the covariates follow

$$\mathbf{Z}_1 \sim \text{Bernoulli}(0.5), \mathbf{Z}_2 \sim \text{Unif}(0, 1), \mathbf{Z}_3 \sim \text{Bernoulli}(0.4)$$

independently of each other, where  $\text{Bernoulli}(r)$  is a Bernoulli distribution with the parameter  $r \in (0, 1)$  and  $\text{Unif}(a, b)$  is a uniform distribution on the interval  $(a, b)$ . Let  $n$  be the sample size and  $\mathbf{Z}_1, \dots, \mathbf{Z}_n$  be independent copies of  $\mathbf{Z}$ . Next, for  $\tilde{\mathbf{Z}}_i = (1, \mathbf{Z}_i^\top)^\top$  and  $\boldsymbol{\gamma} = (\gamma_1, \gamma_2, \gamma_3)^\top$ , let

$$p_i = \frac{\exp(\tilde{\mathbf{Z}}_i^\top \boldsymbol{\gamma})}{1 + \exp(\tilde{\mathbf{Z}}_i^\top \boldsymbol{\gamma})}$$

be the uncured probability of the  $i$ -th individual and generate  $D_i \sim \text{Bernoulli}(p_i)$  independently of each other. Assuming that the frailty  $Y_i$  of the  $i$ -th individual follows the shifted gamma distribution (17) with the parameter  $\theta$  independently of  $\mathbf{Z}_i$ , the event occurrence time  $T_i$  for  $\boldsymbol{\beta} = (\beta_1, \beta_2, \beta_3)^\top$  is generated based on

$$T_i | D_i, Y_i \sim \begin{cases} \infty & (D_i = 0) \\ S_{\text{gg}}\left(te^{-\mathbf{Z}_i^\top \boldsymbol{\beta}}; \boldsymbol{\kappa}\right)^{Y_i} & (D_i = 1) \end{cases}$$

In addition, it is assumed that the censoring time  $C_i$  ( $i = 1, \dots, n$ ) follows the uniform distribution  $\text{Unif}(0, 5)$  independently of any variables. Based on the variables generated above,  $O_i = \min(T_i, C_i)$  and  $\Delta_i = I(T_i \leq C_i)$  are defined, and the observed sample  $\{O_i, \Delta_i, \mathbf{Z}_i \mid i = 1, \dots, n\}$  is obtained to estimate the parameters of the proposed model. In this simulation, the following settings are used for the true values of the parameters:

- (i)  $\boldsymbol{\beta} = (-0.5, 0.5, -0.8)$ ,  $(\mu, \sigma, q) = (3, 1, 2)$ ,  $\theta = 0.5$ ,  $\boldsymbol{\gamma} = (0.1, 0.5, -1, 0.6)$
- (ii)  $\boldsymbol{\beta} = (-0.5, 0.5, -0.8)$ ,  $(\mu, \sigma, q) = (3, 1, 2)$ ,  $\theta = 3$ ,  $\boldsymbol{\gamma} = (0.1, 0.5, -1, 0.6)$

The difference between settings (i) and (ii) is the value of parameter  $\theta$  of the shifted gamma distribution. We will examine the effects and tendencies of the estimation results due to the difference in the parameter on frailty.

###### 4.1.2. Results

The mean parameter estimate values for settings (i) and (ii) are shown in Tables 2 and 3, respectively.

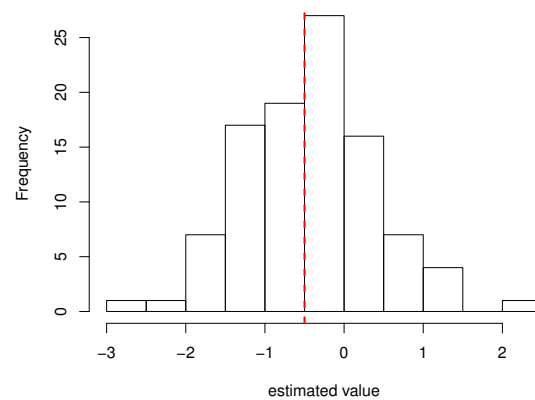
**Table 2.** Means and standard deviations (SD) of estimate values in setting (i).

Parameter	True Value	Mean (SD)					
		$n = 100$		$n = 500$		$n = 1000$	
$\beta_1$	−0.5	−0.355	(1.602)	−0.507	(0.948)	−0.529	(0.803)
$\beta_2$	0.5	0.278	(3.000)	0.796	(1.556)	0.814	(1.203)
$\beta_3$	−0.8	−0.402	(1.830)	−0.651	(0.925)	−0.656	(0.737)
$\mu$	3	2.091	(4.390)	2.889	(2.074)	2.844	(0.985)
$\sigma$	1	1.141	(0.852)	1.004	(0.641)	1.179	(0.743)
$q$	2	1.435	(3.412)	2.804	(1.835)	2.600	(2.012)
$\theta$	0.5	8.214	(20.992)	3.053	(9.309)	2.273	(5.812)
$\gamma_0$	0.1	−0.063	(1.526)	0.135	(1.110)	0.192	(0.927)
$\gamma_1$	0.5	0.540	(1.216)	0.576	(0.853)	0.553	(0.818)
$\gamma_2$	−1	−1.031	(2.294)	−0.878	(1.556)	−0.865	(1.308)
$\gamma_3$	0.6	1.045	(1.077)	0.785	(0.852)	0.829	(0.725)

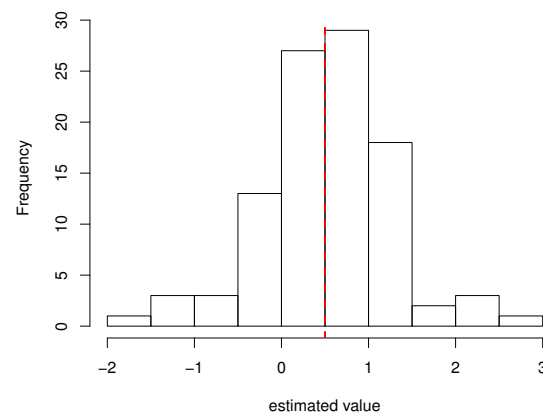
**Table 3.** Means and standard deviations (SD) of estimate values in setting (ii).

Parameter	True Value	Mean (SD)					
		$n = 100$		$n = 500$		$n = 1000$	
$\beta_1$	−0.5	−0.092	(1.796)	−0.316	(1.072)	−0.452	(0.823)
$\beta_2$	0.5	0.269	(3.277)	0.552	(1.875)	0.548	(1.285)
$\beta_3$	−0.8	−0.386	(1.619)	−0.777	(1.032)	−0.698	(0.904)
$\mu$	3	1.378	(2.993)	3.253	(4.296)	2.840	(1.163)
$\sigma$	1	0.910	(0.768)	1.206	(0.878)	1.077	(0.718)
$q$	2	7.276	(49.95)	2.706	(2.197)	2.928	(2.184)
$\theta$	3	23.076	(72.067)	8.874	(21.858)	8.258	(17.101)
$\gamma_0$	0.1	−0.093	(1.333)	0.185	(1.335)	0.109	(0.962)
$\gamma_1$	0.5	0.699	(1.191)	0.704	(0.926)	0.534	(0.748)
$\gamma_2$	−1	−1.159	(2.140)	−1.133	(1.587)	−0.971	(1.141)
$\gamma_3$	0.6	0.933	(1.269)	0.608	(0.771)	0.692	(0.844)

For almost all of the estimated values in settings (i) and (ii), the mean approaches the true value, and the standard deviation decreases as the sample size increases. In addition, Figures 2 and 3 show the histograms of estimated values of the main subjects of interest in setting (ii) ( $n = 1000$ ). They show that the estimated values of the regression coefficients  $\beta_1$  and  $\gamma_1$  are symmetrically distributed around the true value. A similar result was obtained for  $\sigma$ , a component other than  $\beta$  and  $\gamma$ . Although most of the estimated values of  $\theta$  are close to their true values, some were far from their true values. In addition, the results suggest that the magnitude of  $\theta$  affects the parameter  $q$  of the generalized gamma distribution and that the estimation of  $\theta$  affects the estimation of  $q$ . The initial values of all estimations were fixed to the same value in this simulation, but the estimated values were converged to those closer to the true value when the initial values were modified. Therefore, the use of multiple initial values for parameter estimation of the proposed model would be a measure for obtaining an appropriate estimate.



**Figure 2.** Histogram of the estimated values of  $\beta_1$  in setting (ii). The red dotted line represents the true value.



**Figure 3.** Histogram of the estimated values of  $\gamma_1$  in setting (ii). The red dotted line represents the true value.

#### 4.2. Real Data Example

In this section, we analyze the real dataset on the onset of hypertension using survival time regression models and compare the results.

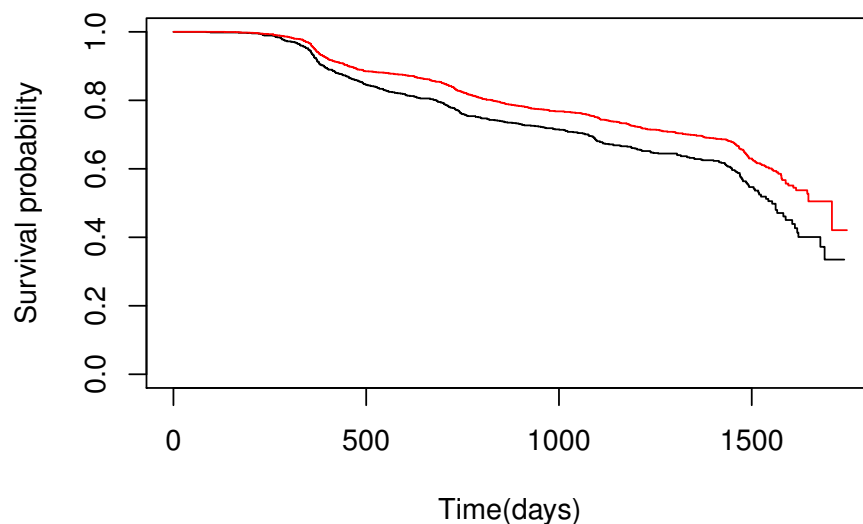
##### 4.2.1. Dataset and Previous Study

The subject of analysis is the dataset on the Specific Health Checkups and Specific Health Guidance, the program started in April 2008 under the National Health Insurance System to maintain lifestyle-related diseases aged between 40 and 74 years. We analyzed the data from Specific Health Checkups conducted in Habikino City, Osaka Prefecture, Japan. They consist of 8325 residents who participated in Specific Health Checkups in 2008 and were followed until the end of March 2013. Of these, 4993 were females and 3332 were males. In this program, lifestyle information for the participants was collected using a self-report questionnaire about the lifestyle and body measurement values, such as waist circumference, age and body weight, and triglyceride level.

We analyzed the dataset of size 3326 (2202 females and 1124 males) obtained by the same exclusion criteria of [31]. They studied the relationship between the onset of hypertension and lifestyle habits, collected via a standard questionnaire, using the proportional hazard model. The present study used the 14 variables (such as age, waist circumference, eating speed, and amount of drinking) used in this previous study as covariates and considered a maximum of 18 linear models as parameters as it also included categorical variables (refer to the paper for details on the variables).

In the present analysis, the onset of hypertension was regarded as an event of interest, and the regression model was applied for each gender. First, the survival functions for males and females were estimated using the Kaplan–Meier estimator (Figure 4). It showed

the five-year survival rates of 0.33 and 0.42 for males and females, respectively, indicating that they were not low. In addition, we performed a hypothesis testing of the proportional hazard assumption based on the Schoenfeld residuals for this dataset [32]. The result indicated that the null hypothesis was rejected for the variable “eating snacks after dinner at least 3 times a week” in the datasets of both males and females at a significance level of 5% ( $p$ -values for the datasets of males and females were 0.0155 and 0.0203, respectively). Therefore, it was suggested that the survival function for the onset of hypertension does not have the proportional hazard assumption. This result indicates that assuming a model without the proportional hazard assumption is appropriate, considering the probability that an event does not occur.



**Figure 4.** Kaplan–Meier estimator of the survival function for the onset of hypertension for each gender. The black and red lines represent the estimators for males and females, respectively.

The following five types of models are applied in this section:

1. Proportional hazard model;
2. AFT model;
3. Mixture cure model with the proportional hazard model;
4. Mixture cure model with the AFT model;
5. Proposed model.

The survival distributions assumed for the model are an exponential distribution and a Weibull distribution for the proportional hazard model. A logarithmic normal distribution and a generalized gamma distribution were used for the AFT model. It should be noted that, based on the definitions of these two models, the model used when an exponential or Weibull distribution is assumed for the baseline hazard function of the proportional hazard model is the same as that used when an exponential or Weibull distribution is assumed as the error variable of the AFT model. A logistic regression model is assumed for all uncured probabilities in the mixture cure model. The distribution mentioned above is used for the survival distribution assumed for the survival function for uncured patients.

In some cases during the analysis using the mixture cure model, the absolute value of the estimated value of the regression coefficient  $\gamma$  in the logistic regression model was extremely large. Therefore, we adopted the method of handling the problem of complete separation by [33], which maximizes the objective function

$$\ell^*(\gamma) = \ell_{\text{logis}}(\gamma) + \frac{1}{2} \log |\hat{I}(\gamma)| \quad (26)$$

where a penalty function is added to the log-likelihood function  $\ell_{\text{logis}}(\gamma)$  of the logistic regression model [34]. For the sample covariates  $z_1, \dots, z_n$ ,

$$\hat{I}(\gamma) = \sum_{i=1}^n z_i z_i^{\top} \frac{e^{z_i^{\top} \gamma}}{(1 + e^{z_i^{\top} \gamma})^2} \quad (27)$$

is an estimated value of the information matrix of  $\ell_{\text{logis}}(\gamma)$ . This method was applied to  $\bar{\ell}_2$  for the proposed model (19). In addition, the Akaike information criterion (AIC) was used as a criterion for comparing the goodness of fit of the model [35].

#### 4.2.2. Results

The regression model used in the analysis and AIC values are shown in Tables 4 and 5.

**Table 4.** AIC for the regression model of the onset of hypertension in males. The asterisk (\*) represents that variable selection is performed for the uncured probability  $p(\mathbf{Z})$ .

Model	Distribution	Number of Parameters	AIC
Proportional hazard (PH)	Exponential	19	7120.407
	Weibull	20	7030.615
AFT model	Log-normal	20	7002.352
	Generalized gamma	21	6999.232
Mixture cure + PH	Exponential	38	7103.040
	Weibull	39	7047.378
Mixture cure + AFT	Log-normal	39	7005.633
	Generalized gamma	40	7004.674
	Generalized gamma *	29	6992.934
Mixture cure + AFT frailty	Generalized gamma	41	7003.011
	Shifted gamma		
	Generalized gamma Shifted gamma *	30	<b>6987.012</b>

**Table 5.** AIC for the regression model of the onset of hypertension in females. The asterisk (\*) represents that variable selection is performed for the uncured probability  $p(\mathbf{Z})$ .

Model	Distribution	Number of Parameters	AIC
Proportional hazard (PH)	Exponential	19	11,804.01
	Weibull	20	11,644.16
AFT	Log-normal	20	11,596.45
	Generalized gamma	21	11,586.00
Mixture cure + PH	Exponential	38	11,798.48
	Weibull	39	11,669.49
Mixture cure + AFT	Log-normal	39	11,609.63
	Generalized gamma *	40	11,600.23
	Generalized gamma *	27	11,579.07
Mixture cure + AFT frailty	Generalized gamma	41	11,596.26
	Shifted gamma *		
	Generalized gamma Shifted gamma *	26	<b>11,575.05</b>

As shown in Tables 4 and 5, the lowest AIC is observed in the proposed model when the variable selection is performed for variables in the logistic regression model in both



males and females. The estimated values, 95% confidence intervals, and  $p$ -values for these estimation results are summarized in Tables 6–9.

**Table 6.** Estimation result of regression coefficient when variable selection is performed by applying the proposed model in males. CI is the confidence interval. The asterisk (\*) and dagger (†) indicate that  $p$ -value is less than 0.10 and 0.05, respectively.

Covariate	Inference of $\beta$ (Regression Coefficients for the Uncured Group)		
	Estimates	95% CI	$p$ -Value
age	−0.0071	(−0.0188, 0.0046)	0.232
waist	−0.0077	(−0.0183, 0.0029)	0.155
exe1h_day	−0.0392	(−0.2006, 0.1221)	0.634
exe30_2_week	0.0470	(−0.0112, 0.0517)	0.562
sleep_good	0.0138	(−0.1732, 0.2009)	0.885
walk_speed	−0.1112	(−0.2617, 0.0393)	0.148
eat_speed_n	0.1386	(−0.1019, 0.3791)	0.259
eat_speed_f	0.0672	(−0.1617, 0.2961)	0.565
eat_b_sleep	0.0025	(−0.1920, 0.1970)	0.980
snacking	−0.3154	(−0.6195, −0.0112)	0.042 †
breakfast	−0.1784	(−0.4329, 0.0760)	0.169
weight_move	−0.0227	(−0.1956, 0.1502)	0.797
plus10kg	0.0162	(−0.1896, 0.2219)	0.878
smoking	−0.0673	(−0.2229, 0.0884)	0.397
drink_amount2	−0.1636	(−0.4831, 0.1558)	0.315
drink_amount3	−0.2134	(−0.4313, 0.0045)	0.054 *
drink_amount4	−0.1120	(−0.3288, 0.1048)	0.311
drink_amount5	−0.1824	(−0.4299, 0.0650)	0.148

Covariate	Inference of $\gamma$ (Regression Coefficients for the Cured Group)		
	Point Estimates	95% CI	$p$ -Value
Intercept	−4.9869	(−7.7813, −2.1924)	0.0005 †
age	0.0882	(0.0423, 0.1340)	0.0002 †
eat_speed_n	0.8394	(−0.2355, 1.9144)	0.1259
eat_speed_f	0.8427	(−0.0571, 1.7424)	0.0664 *
kanshyoku	−1.4826	(−2.4026, −0.5626)	0.0016 †
plus10kg	0.9488	(0.0316, 1.8666)	0.0426 †
drink_amount2	−0.9510	(−1.9441, 0.0420)	0.0605 *
drink_amount4	1.1910	(0.0918, 2.2901)	0.0337 †

**Table 7.** Results of parameter estimation when variable selection is performed by applying the proposed model in males.

Parameter	Distribution			
	Generalized Gamma			Shifted Gamma
	$\mu$	$\sigma$	$q$	$\theta$
Estimates	8.4404	0.8767	−0.3141	13.1350
Standard error	0.6150	0.0866	0.2701	7.0537

As shown in Tables 4 and 5, the proposed model with the variable selection of the regression coefficient in the uncured probability had the lowest AIC values. However, other models had lower AIC values than the proposed model when variable selection was not performed. This is likely due to the instability resulting from the 18-parameter logistic regression model in the mixture cure model.

**Table 8.** Estimation result of regression coefficient when variable selection is performed by applying the proposed model in females. CI is the confidence interval. The asterisk (\*) and dagger (†) indicate that  $p$ -value is less than 0.10 and 0.05, respectively.

Covariate	Inference of $\beta$ (Regression Coefficients for the Uncured Group)		
	Estimates	95% CI	$p$ -Value
age	−0.0103	(−0.0222, 0.0016)	0.089 *
waist	−0.0034	(−0.0100, 0.0033)	0.321
exel1h_day	−0.0493	(−0.1717, 0.0732)	0.430
exe30_2_week	−0.0230	(−0.1456, 0.0995)	0.712
sleep_good	0.0855	(−0.1126, 0.1297)	0.890
walk_speed	0.0401	(−0.0726, 0.1528)	0.486
eat_speed_n	−0.0651	(−0.2045, 0.0742)	0.360
eat_speed_f	0.0520	(−0.1051, 0.2092)	0.516
eat_b_sleep	−0.0342	(−0.2395, 0.1711)	0.744
snacking	0.0442	(−0.1075, 0.1958)	0.568
breakfast	−0.0781	(−0.3625, 0.2063)	0.591
weight_move	−0.0210	(−0.1589, 0.1169)	0.765
plus10kg	0.0301	(−0.1391, 0.1993)	0.727
smoking	−0.1611	(−0.0638, 0.3859)	0.160
drink_amount2	−0.0098	(−0.1464, 0.1268)	0.888
drink_amount3	−0.0500	(−0.2688, 0.1688)	0.654
drink_amount4	−0.1566	(−0.5165, 0.2032)	0.394
drink_amount5	−0.2162	(−0.8744, 0.4419)	0.520

Covariate	Inference of $\gamma$ (Regression Coefficients for the Cured Group)		
	Point Estimates	95% CI	$p$ -Value
Intercept	−6.0049	(−8.3249, −3.6849)	<0.001 †
age	0.1045	(0.0645, 0.1445)	<0.001 †
eat_speed_f	0.5600	(−0.1536, 1.2736)	0.124
plus10kg	0.6271	(−0.1873, 1.4415)	0.131

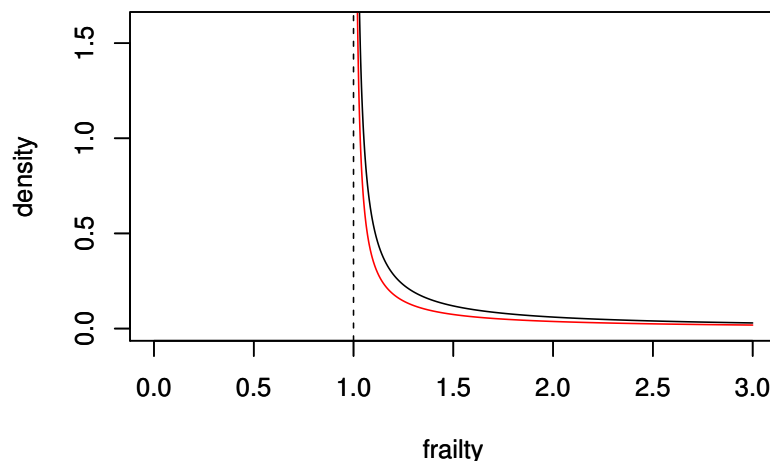
**Table 9.** Results of parameter estimation when variable selection is performed by applying the proposed model in females.

Parameter	Distribution			
	Generalized Gamma			Shifted Gamma
	$\mu$	$\sigma$	$q$	$\theta$
Estimates	8.1034	0.9653	−0.8065	22.9620
Standard error	0.4688	0.0794	0.2772	9.9425

For males, seven variables were selected for the uncured probability after variable selection: “age”, “eating speed: normal”, “eating speed: fast”, “eating snacks after dinner at least 3 times a week”, “increase in body weight by 10 kg or more compared to that at the age of 20 years”, “amount of drinking: occasionally”, and “amount of drinking: less than 1–2 go daily”. Of these, “age”, “eating snacks after dinner at least 3 times a week”, “increase in body weight by 10 kg or more compared to that at the age of 20 years”, and “amount of drinking: less than 1–2 go daily” were judged to have a significant regression coefficient at the 5% significance level. Among the variables included in the survival function for uncured patients, “eating snacks after dinner at least 3 times a week” was judged to have a significant regression coefficient at the 5% significance level. For females, three variables were selected for the logistic regression model of the uncured probability after variable selection: “age”, “eating speed: fast”, and “increase in body weight by 10 kg or more compared to that at the age of 20 years”. Of these, only “age” was judged to be significant at the 5% significance level. There were no variables in the survival function for uncured

patients judged to be significant at the 5% significance level. This result suggests that “age” is the only variable that affects the onset of hypertension in females.

Figure 5 shows the probability density function of the estimated shifted gamma distribution. Estimation was also performed when the accelerated failure time gamma frailty model was assumed for the mixture cure model, showing that the estimated values of  $\theta$  were  $9.42 \times 10^{-6}$  and  $6.93 \times 10^{-5}$  for males and females, respectively, which were significantly close to 0. The AIC values were 7006.649 and 11,603.85 for males and females, respectively. These were higher than those observed without the assumption of frailty, suggesting that this model was not appropriate.



**Figure 5.** Probability density function of the estimated shifted gamma distribution. The black and red lines show the results for males and females, respectively.

The number of variables selected for the logistic regression model differed between males and females, and some variables were judged to be significant only in males, suggesting that the covariates that affect the onset of hypertension differ between males and females. In the distribution of frailty, females had larger estimated values of  $\theta$  and variance than males. This indicates that females have greater differences in the onset of hypertension among individuals.

## 5. Discussion

In this study, we proposed a model that assumes the AFT frailty model in the survival function for uncured patients in the mixture cure model. The characteristic of the proposed method is that the variable representing frailty is assumed to follow a shifted gamma distribution, which characterizes the uncured individuals more clearly. In addition, we constructed a parameter estimation method for the proposed model using the EM algorithm and confirmed that the estimation could be performed appropriately using a simulation. Furthermore, we analyzed the dataset on the onset of hypertension in epidemiological research. Compared with existing models, the proposed model showed the lowest AIC values for both males and females.

Ref. [31] reported the results of the analysis using a semiparametric proportional hazard model. Since a parametric model was used in the analysis of the present study, the results of these two studies cannot be compared directly using AIC. Therefore, comparison criteria between semiparametric and parametric models need to be considered in the future. Previously, comparison criteria between semiparametric and parametric models in the proportional hazard model were studied by [36,37], who used approaches based on the focused information criterion (FIC).

The model proposed in this study uses the parametric AFT model, but the method using the semiparametric AFT model can be considered as an alternative. The semiparametric AFT frailty model was previously studied by [38]. Because this model can estimate

regression parameters without assuming a specific distribution in the error term, a more flexible model can be considered. On the other hand, its estimation method is more complicated than that of the parametric model. One future challenge is to extend the model proposed in this study to the semiparametric model and construct an estimation algorithm.

In addition, while this study constructed the estimation algorithm assuming only right censoring, there are other types of censoring, one of which is interval censoring. Interval censoring refers to the cases in which only the occurrence of an event between two observation periods is recorded [39]; it is a more generalized type of censoring and includes right censoring. The construction of an estimation algorithm for the datasets that contain interval censoring is also needed in the future. While models considered in the study are linear static models, dynamic prediction for survival data is an important issue [40]. In addition, recent development for big data such as machine learning techniques would be incorporated [41].

**Author Contributions:** The authors carried out this work and drafted the manuscript collaboratively. In particular, H.A. and K.H. focused statistical modeling and construction of the estimation algorithm. A.T., D.S. and T.O. evaluated the results of analyses. All authors have read and agreed to the published version of the manuscript.

**Funding:** Kenichi Hayashi is supported by the Japan Society for the Promotion of Science (JSPS) KAKENHI (Grant-in-Aid for Scientific Research) under Grant No. 18K11197.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Interested readers can contact the corresponding author.

**Acknowledgments:** The authors are grateful to the Insurance and Pension Division of Habikino city, Osaka Prefecture for permission to use the dataset in the work of Tsutani et al. (2017). The authors would also like to acknowledge the associate editor and anonymous reviewers for their comments and suggestions that served to materially improve the manuscript.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Cox, D.R. Regression Models and Life-Tables. *J. R. Stat. Soc. Ser. B Stat. Methodol.* **1972**, *34*, 87–220. [\[CrossRef\]](#)
2. Orbe, J.; Ferreira, E.; Núñez-Antón, V. Comparing proportional hazards and accelerated failure time models for survival analysis. *Stat. Med.* **2002**, *21*, 3493–3510. [\[CrossRef\]](#) [\[PubMed\]](#)
3. Berkson, J.; Gage, R. Survival Curve for Cancer Patients Following Treatment. *J. Am. Stat. Assoc.* **1952**, *47*, 501–515. [\[CrossRef\]](#)
4. Chen, M.; Ibrahim, J.; Sinha, D. A New Bayesian Model for Survival Data with a Surviving Fraction. *J. Am. Stat. Assoc.* **1999**, *94*, 909–919. [\[CrossRef\]](#)
5. Sy, J.P.; Taylor, J. Estimation in a Cox Proportional Hazards Cure Model. *Biometrics* **2000**, *56*, 227–236. [\[CrossRef\]](#)
6. Vaupel, J.W.; Manton, K.G.; Stallard, E. The Impact of Heterogeneity in Individual Frailty on the Dynamics of Mortality. *Demography* **1979**, *16*, 439–454. [\[CrossRef\]](#)
7. Yamaguchi, K. Accelerated Failure-Time Regression Models with a Regression Model of Surviving Fraction: An Application to the Analysis of “Permanent Employment” in Japan. *J. Am. Stat. Assoc.* **1992**, *87*, 284–292.
8. Li, C.; Taylor, J. A semi-parametric accelerated failure time cure model. *Stat. Med.* **2002**, *21*, 3235–3247. [\[CrossRef\]](#)
9. Yu, B. A frailty mixture cure model with application to hospital readmission data. *Biom. J.* **2008**, *50*, 386–394. [\[CrossRef\]](#)
10. Hutton, J.L.; Monaghan, P.F. Choice of Parametric Accelerated Life and Proportional Hazards Models for Survival Data: Asymptotic Results. *Lifetime Data Anal.* **2002**, *8*, 375–393. [\[CrossRef\]](#)
11. Patel, K.; Kay, R.; Rowell, L. Comparing proportional hazards and accelerated failure time models: An application in influenza. *Pharmaceut. Statist.* **2006**, *5*, 213–224. [\[CrossRef\]](#) [\[PubMed\]](#)
12. Aalen, O.O. Heterogeneity in survival analysis. *Statist. Med.* **1988**, *7*, 1121–1137. [\[CrossRef\]](#) [\[PubMed\]](#)
13. Pan, W. Using Frailties in the Accelerated Failure Time Model. *Lifetime Data Anal.* **2001**, *7*, 55–64. [\[CrossRef\]](#) [\[PubMed\]](#)
14. Price, D.; Manatunga, A. Modelling survival data with a cured fraction using frailty models. *Stat. Med.* **2001**, *20*, 1515–1527. [\[CrossRef\]](#)
15. Chen, P.; Zhang, J.; Zhang, R. Estimation of the accelerated failure time frailty model under generalized gamma frailty. *Comput. Stat. Data Anal.* **2013**, *62*, 171–180. [\[CrossRef\]](#)
16. He, M. Some Flexible Families of Mixture Cure Frailty Models and Associated Inference. Ph.D. Thesis, McMaster University, Hamilton, ON, Canada, 2021. Available online: <http://hdl.handle.net/11375/26258> (accessed on 20 July 2022).

17. O'Quinley, J. *Proportional Hazards Regression*; Springer: New York, NY, USA, 2008; 542p.
18. Dempster, A.P.; Laird, N.M.; Rubin, D.B. Maximum Likelihood from Incomplete Data via the EM algorithm. *J. R. Stat. Soc. Ser. B Stat. Methodol.* **1977**, *39*, 1–38.
19. Scolas, S.; Ghouch, A.; Legrabd, C.; Oulhaj, A. Variable selection in a flexible parametric mixture cure model with interval-censored data. *Stat. Med.* **2016**, *35*, 1210–1225. [[CrossRef](#)]
20. Kalbfleisch, J.; Prentice, R. *The Statistical Analysis of Failure Time Data*, 2nd ed.; Wiley Series in Probability and Statistics; John Wiley & Sons, Inc.: Hoboken, NJ, USA, 2002; 462p.
21. Wei, L.J. The accelerated failure time model: A useful alternative to the cox regression model in survival analysis. *Stat. Med.* **1992**, *11*, 1871–1879. [[CrossRef](#)]
22. Aalen, O.O. Two Examples of Modelling Heterogeneity in Survival Analysis. *Scand. Stat. Theory Appl.* **1987**, *14*, 19–25.
23. Peng, Y.; Zhang, J. Estimation method of the semiparametric mixture cure gamma frailty model. *Stat. Med.* **2008**, *27*, 5177–5194. [[CrossRef](#)]
24. Lambert, P.; Collett, D.; Kimber, A.; Johnson, R. Parametric accelerated failure time models with random effects and an application to kidney transplant survival. *Stat. Med.* **2004**, *23*, 3177–3192. [[CrossRef](#)] [[PubMed](#)]
25. Elbers, C.; Ridder, G. True and spurious duration dependence: The identifiability of the proportional hazards model. *Rev. Econ. Stud.* **1982**, *49*, 403–409. [[CrossRef](#)]
26. Stacy, E.W. A generalization of the gamma distribution. *Ann. Math. Statist.* **1962**, *33*, 1187–1192. [[CrossRef](#)]
27. Prentice, R.L. A log gamma model and its maximum likelihood estimation. *Biometrika* **1974**, *61*, 539–544. [[CrossRef](#)]
28. Kim, S.; Lee, J.Y.; Sung, D.K. A Shifted Gamma Distribution Model for Long-Range Dependent Internet Traffic. *IEEE Commun. Lett.* **2003**, *7*, 124–126.
29. You, X. Approximation of the median of the gamma distribution. *J. Number Theory* **2017**, *174*, 487–493. [[CrossRef](#)]
30. Casella, G.; Berger, R.L. *Statistical Inference*, 2nd ed.; Thomson Learning: Pacific Grove, CA, USA, 2002; 660p.
31. Tsutani, H.; Funamoto, M.; Sugiyama, D.; Kuwabara, K.; Miyamatsu, N.; Watanabe, K.; Okamura, T. Association between lifestyle factors assessed by standard question items of specific health checkup and the incidence of metabolic syndrome and hypertension in community dwellers: A five-year cohort study of National Health Insurance beneficiaries in Habikino City. *Nihon Koshu Eisei Zasshi* **2017**, *64*, 258–269.
32. Grambsch, P.M.; Therneau, T.M. Proportional Hazards Tests and Diagnostics Based on Weighted Residuals. *Biometrika* **1994**, *81*, 515–526. [[CrossRef](#)]
33. Heinze, G.; Schemper, M. A solution to the problem of separation in logistic regression. *Stat. Med.* **2002**, *21*, 2409–2419. [[CrossRef](#)]
34. Firth, D. Bias reduction of Maximum Likelihood Estimates. *Biometrika* **1993**, *80*, 27–38. [[CrossRef](#)]
35. Akaike, H. Information theory and an extension of the maximum likelihood principle. In Proceedings of the 2nd International Symposium on Information Theory, Tsahkadsor, Armenia, 2–8 September 1971; Akademiai Kiado: Budapest, Hungary, 1973; pp. 267–281.
36. Hjort, N.L.; Claeskens, G. Focused Information Criteria and Model Averaging for the Cox Hazard Regression Model. *J. Am. Stat. Assoc.* **2006**, *101*, 1449–1464. [[CrossRef](#)]
37. Jullum, M.; Hjort, N.L. What price semiparametric Cox regression? *Lifetime Data Anal.* **2019**, *25*, 405–438. [[CrossRef](#)] [[PubMed](#)]
38. Xu, L.; Zhang, J. An EM-like algorithm for the semiparametric accelerated failure time gamma frailty model. *Comput. Stat. Data Anal.* **2010**, *54*, 1467–1474. [[CrossRef](#)]
39. Sun, J. *The Statistical Analysis of Interval-Censored Failure Time Data*; Springer: New York, NY, USA, 2006; 304p.
40. Karthik, S.; Bhadoria, R.S.; Lee, J.G.; Sivaraman, A.K.; Samanta, S.; Balasundaram, A.; Chaurasia, B.K.; Ashokkumar, S. Prognostic Kalman Filter Based Bayesian Learning Model for Data Accuracy Prediction. *Comput. Mater. Contin.* **2022**, *72*, 243–259. [[CrossRef](#)]
41. Singh, L.K.; Garg, H.; Khanna, M.; Bhadoria, R.S. An enhanced deep image model for glaucoma diagnosis using feature-based detection in retinal fundus. *Med. Biol. Eng. Comput.* **2021**, *59*, 333–353. [[CrossRef](#)]