

---

# Supplementary Material

Statistical Methods in Medical Research  
XX(X):1–46  
© The Author(s) 2016  
Reprints and permission:  
sagepub.co.uk/journalsPermissions.nav  
DOI: 10.1177/ToBeAssigned  
www.sagepub.com/



Francisco J. Rubio<sup>1</sup>, Laurent Remontet<sup>2</sup>, Nicholas P. Jewell<sup>3</sup>, and Aurélien Belot<sup>1</sup>

## 1 Comparison of the different hazard structures

We aim to compare in this section the different hazard shapes according to the assumed hazard structure. We focus on the Accelerated hazard model (AH) and the General Hazard (GH) model, and compare it with the well-know Proportional Hazard (PH) model. For these comparisons, we assume one binary covariate  $x$ . As a reminder (see main document), the GH model is defined as

$$h^{\text{GH}}(t; x) = h_0(t \exp(\beta_1 x)) \exp(\beta_2 x). \quad (\text{S1})$$

Both the AH and PH model are particular cases of the GH model, *i.e.* the GH model coincides with the AH model when  $\beta_2 = 0$ , and the GH model coincides with the PH model when  $\beta_1 = 0$ .

---

<sup>1</sup> Cancer Survival Group, Faculty of Epidemiology and Population Health,  
Department of Non-Communicable Disease Epidemiology,  
London School of Hygiene & Tropical Medicine,  
London, WC1E 7HT, United Kingdom.

<sup>2</sup>Hospices Civils de Lyon,  
Service de Biostatistique-Bioinformatique,  
Lyon, 69003, France.

<sup>3</sup> University of California, Berkeley,  
School of Public Health,  
Berkeley, CA 94720-7360.

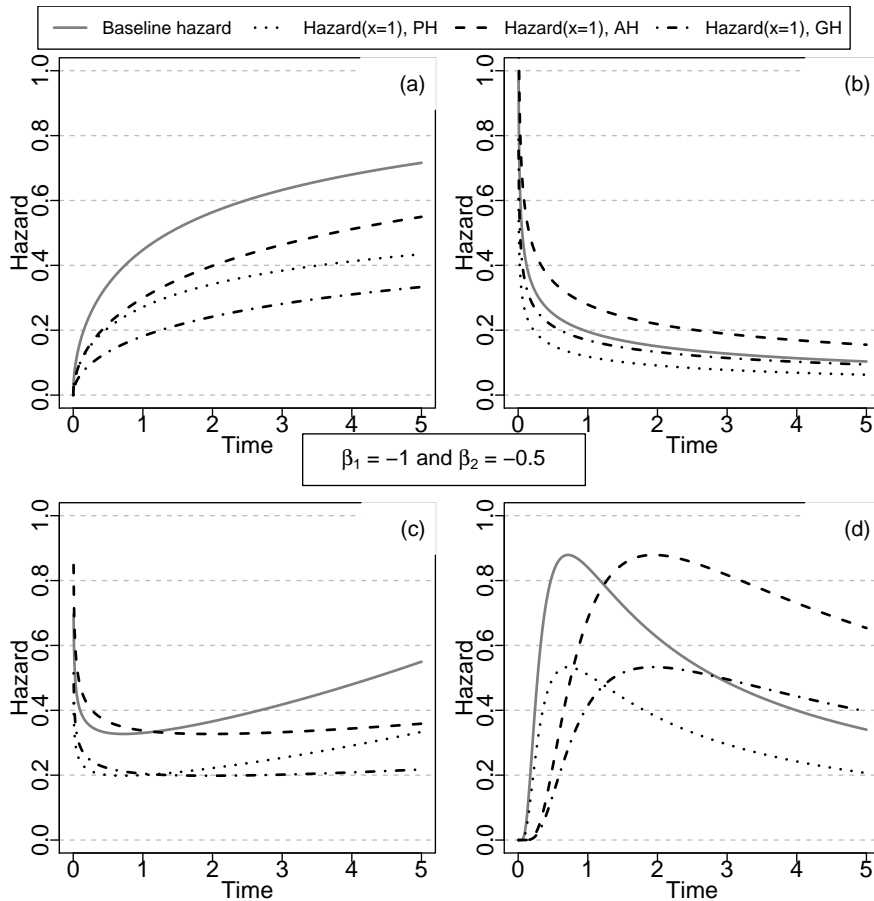
### Corresponding author:

Francisco J. Rubio, Cancer Survival Group, Faculty of Epidemiology and Population Health,  
Department of Non-Communicable Disease Epidemiology,  
London School of Hygiene & Tropical Medicine,  
London, WC1E 7HT, United Kingdom.  
Email: Francisco.Rubio@lshtm.ac.uk

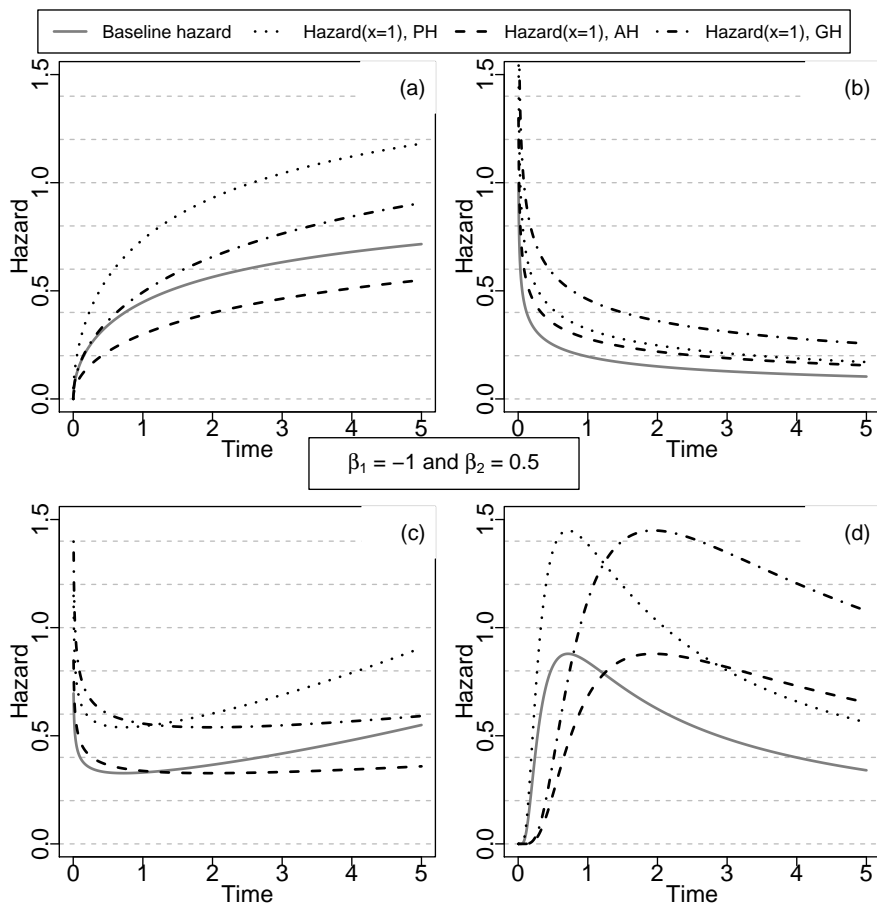
We defined 4 different scenarios, depending on the values assumed for  $\beta_1$  and  $\beta_2$ , and in each scenario (Figures S1-S4), we distinguish 4 different baseline hazard shapes (monotonic increasing -panels (a)-, monotonic decreasing -panels (b)-, bathtub -panels (c)-, and unimodal -panels (d)).

In each panel of each figure (Figures S1-S4), we plotted the baseline hazard (*i.e.* for group with  $x = 0$ ) -grey solid lines-, and the hazard for group with  $x = 1$  assuming the (i) PH model (thus  $\beta_1$  set to 0) -dotted black lines-, (ii) AH model (thus  $\beta_2$  set to 0) -dashed black lines-, (iii) GH model -dotdashed black lines-.

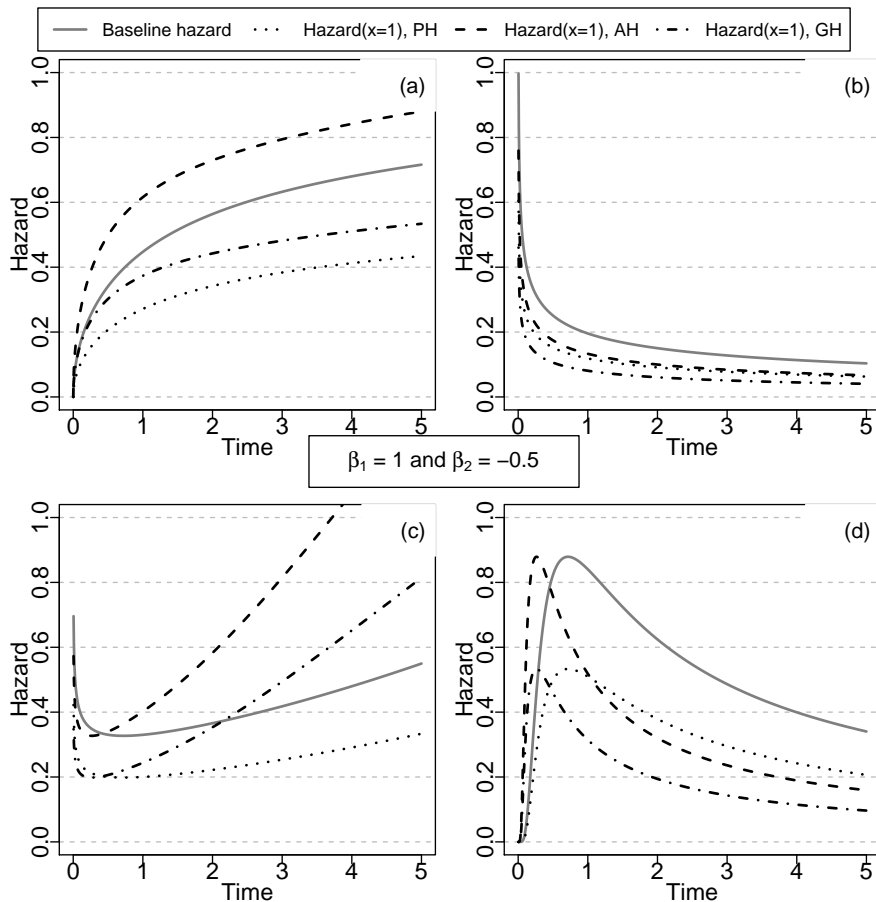
For full details on the interpretation of these figures, we refer the reader to the Section 2.3 of the main document.



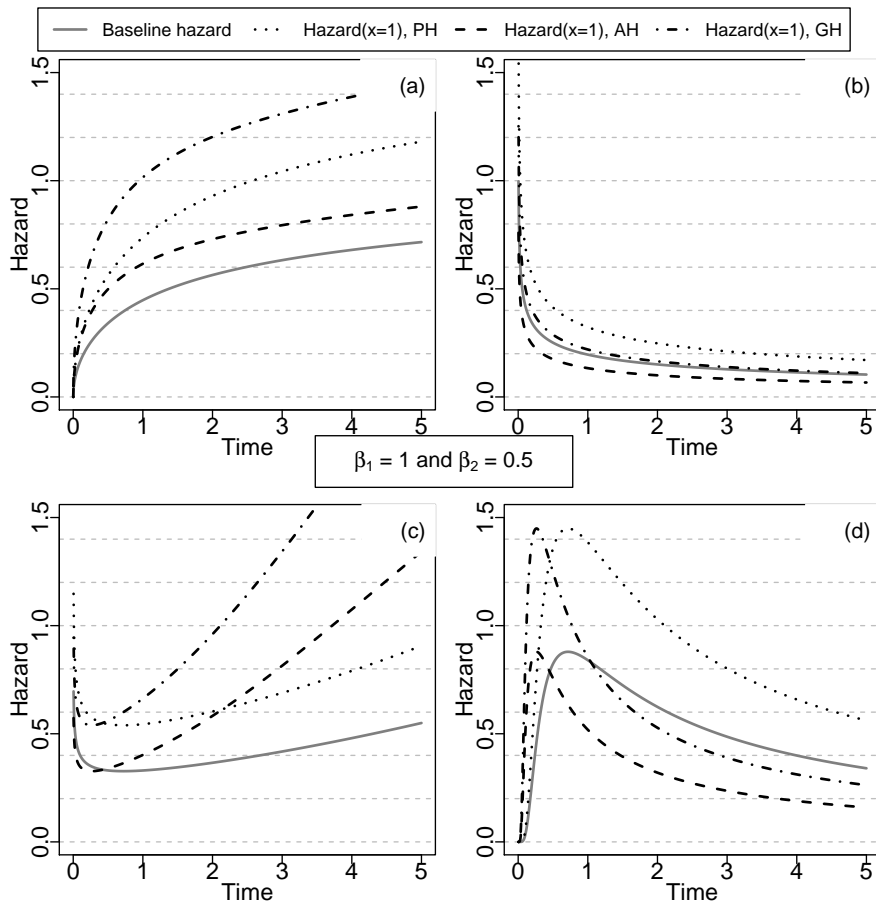
**Figure S1.** Baseline hazard ( $X = 0$ ) and hazards for group with covariate  $x = 1$  according to the hazard structure assumed (PH, AH, GH) - scenario with  $\beta_1$  and  $\beta_2$  are both negatives.



**Figure S2.** Baseline hazard ( $X = 0$ ) and hazards for group with covariate  $x = 1$  according to the hazard structure assumed (PH, AH, GH) - scenario with  $\beta_1$  negative and  $\beta_2$  positive.



**Figure S3.** Baseline hazard ( $X = 0$ ) and hazards for group with covariate  $x = 1$  according to the hazard structure assumed (PH, AH, GH) - scenario with  $\beta_1$  positive and  $\beta_2$  negative.



**Figure S4.** Baseline hazard ( $X = 0$ ) and hazards for group with covariate  $x = 1$  according to the hazard structure assumed (PH, AH, GH) - scenario with  $\beta_1$  and  $\beta_2$  are both positives.

## 2 Theoretical results

**Proposition 1.** *The Exponentiated Weibull distribution is identifiable.*

**Proof.** The strategy of the proof is by contradiction. Suppose that there are two sets of parameters  $(\sigma, \kappa, \alpha) \neq (\sigma', \kappa', \alpha')$  such that

$$F_{EW}(t; \sigma, \kappa, \alpha) = F_{EW}(t; \sigma', \kappa', \alpha'), \quad \text{for all } t > 0. \quad (\text{S2})$$

Consider the cases: (i)  $\alpha = \alpha'$ , (ii)  $\alpha > \alpha'$ , and (iii)  $\alpha < \alpha'$ .

- (i) Suppose that  $\alpha = \alpha'$ . Given that  $F_{EW}(t; \sigma, \kappa, \alpha) = F_W(t; \sigma, \kappa)^\alpha$ , where  $F_W(t; \sigma, \kappa)$  is the Weibull cumulative distribution function, it follows that (S2) implies that  $F_W(t; \sigma, \kappa) = F_W(t; \sigma', \kappa')$ . Since the Weibull distribution is identifiable, this implies that  $\sigma = \sigma'$  and  $\kappa = \kappa'$ , which contradicts the assumption  $(\sigma, \kappa, \alpha) \neq (\sigma', \kappa', \alpha')$ .
- (ii) Suppose that  $\alpha > \alpha'$ . Given that  $F_{EW}(t; \sigma, \kappa, \alpha) = F_W(t; \sigma, \kappa)^\alpha$ , it follows that  $F_W(t; \sigma, \kappa)^\alpha \leq F_W(t; \sigma, \kappa)^{\alpha'}$ . Consequently,  $F_W(t; \sigma, \kappa)^\alpha = F_W(t; \sigma', \kappa')^{\alpha'} \leq F_W(t; \sigma, \kappa)^{\alpha'}$ , and  $F_W(t; \sigma', \kappa') \leq F_W(t; \sigma, \kappa)$ . By using the expression of the Weibull distribution function, we obtain the condition  $\frac{t^{\kappa'}}{\sigma'^{\kappa'}} \leq \frac{t^\kappa}{\sigma^\kappa}$ , for all  $t > 0$ . Equivalently, we have  $t^{\kappa'} \leq \frac{\sigma'^{\kappa'}}{\sigma^\kappa} t^\kappa$ , for all  $t > 0$ . Given that these are polynomials of different power, this inequality cannot hold on both regions  $t \in (0, A)$  and  $t \in [A, \infty)$ , for some  $A = A(\sigma, \sigma', \kappa, \kappa') > 0$ , simultaneously for either  $\kappa > \kappa'$  or  $\kappa < \kappa'$ . Thus, we need  $\kappa = \kappa'$ , which in turns implies the need for  $\sigma' > \sigma$ .

Now, using (S2) and that  $\alpha > \alpha'$ ,  $\kappa = \kappa'$  and  $\sigma' > \sigma$ , it follows that:

$$F_{EW}(t; \sigma, \kappa, \alpha) = F_{EW}(t; \sigma', \kappa, \alpha'), \quad \text{for all } t > 0.$$

Consequently,

$$\alpha \log \left[ 1 - \exp \left\{ - \left( \frac{t}{\sigma} \right)^\kappa \right\} \right] = \alpha' \log \left[ 1 - \exp \left\{ - \left( \frac{t}{\sigma'} \right)^\kappa \right\} \right],$$

and

$$\frac{\alpha}{\alpha'} = \frac{\log \left[ 1 - \exp \left\{ - \left( \frac{t}{\sigma} \right)^\kappa \right\} \right]}{\log \left[ 1 - \exp \left\{ - \left( \frac{t}{\sigma'} \right)^\kappa \right\} \right]} > 1.$$

Now, taking the limit as  $t \rightarrow 0$ , we get (after sequentially applying L'Hôpital's rule)

$$\lim_{t \rightarrow 0} \frac{\log \left[ 1 - \exp \left\{ - \left( \frac{t}{\sigma} \right)^\kappa \right\} \right]}{\log \left[ 1 - \exp \left\{ - \left( \frac{t}{\sigma'} \right)^\kappa \right\} \right]} = \frac{\sigma^\kappa}{\sigma'^{\kappa}} \lim_{t \rightarrow 0} \frac{\exp \left\{ \left( \frac{t}{\sigma} \right)^\kappa \right\} - 1}{\exp \left\{ \left( \frac{t}{\sigma'} \right)^\kappa \right\} - 1} = \lim_{t \rightarrow 0} \frac{\exp \left\{ \left( \frac{t}{\sigma} \right)^\kappa \right\}}{\exp \left\{ \left( \frac{t}{\sigma'} \right)^\kappa \right\}} = 1,$$

which contradicts the assumption that  $\alpha > \alpha'$ .

- (iii) The case  $\alpha < \alpha'$  follows analogously as (ii).

### 3 True values of the parameters and corresponding net survival simulated in the different scenarios

	PH	AFT	HH	AH	GH	CH
$\sigma$	1.75	1.4	1.8	1.75	1.75	0.114
$\kappa$	0.5	0.5	0.5	0.5	0.6	0.366
$\alpha$	2.5	2.5	2.5	1.75	2.5	18.306
$\beta_{11}$	–	0.05	0.1	-0.1	0.1	–
$\beta_{12}$	–	0.2	–	-0.3	0.1	–
$\beta_{13}$	–	0.3	–	-0.5	0.1	0.5
$\beta_{21}$	0.035	0.05	0.05	–	0.05	0.05
$\beta_{22}$	0.2	0.2	0.2	–	0.2	0.3
$\beta_{23}$	0.3	0.3	0.5	–	0.25	-0.6
$r$	0.230	0.275	0.230	0.325	0.200	0.150

**Table S1.** True values of the parameters used in each scenario, where  $\beta_{.1}$  corresponds to variable age,  $\beta_{.2}$  corresponds to sex, and  $\beta_{.3}$  corresponds to W.

sex	comorb.	age	NS(1)	NS(5)
0	0	60	0.85	0.52
0	1	60	0.80	0.42
1	0	60	0.82	0.45
1	1	60	0.77	0.34
0	0	70	0.80	0.40
0	1	70	0.73	0.29
1	0	70	0.76	0.33
1	1	70	0.69	0.22
0	0	80	0.72	0.27
0	1	80	0.64	0.17
1	0	80	0.67	0.20
1	1	80	0.58	0.11

**Table S2.** Simulation design: PH with  $(\sigma, \kappa, \alpha) = (1.75, 0.5, 2.5)$  and  $\beta = (0.035, 0.2, 0.3)$

sex	comorb.	age	NS(1)	NS(5)
0	0	60	0.84	0.48
0	1	60	0.79	0.39
1	0	60	0.81	0.42
1	1	60	0.75	0.34
0	0	70	0.75	0.34
0	1	70	0.69	0.26
1	0	70	0.71	0.28
1	1	70	0.64	0.20
0	0	80	0.64	0.21
0	1	80	0.57	0.14
1	0	80	0.59	0.16
1	1	80	0.51	0.11

**Table S3.** Simulation design: AFT with  $(\sigma, \kappa, \alpha) = (1.4, 0.5, 2.5)$  and  $\beta = (0.05, 0.2, 0.3)$ .

sex	comorb.	age	NS(1)	NS(5)
0	0	60	0.87	0.53
0	1	60	0.80	0.35
1	0	60	0.85	0.46
1	1	60	0.76	0.27
0	0	70	0.80	0.41
0	1	70	0.69	0.22
1	0	70	0.76	0.33
1	1	70	0.64	0.16
0	0	80	0.72	0.32
0	1	80	0.58	0.15
1	0	80	0.67	0.25
1	1	80	0.51	0.10

**Table S4.** Simulation design: HH with  $(\sigma, \kappa, \alpha) = (1.8, 0.5, 2.5)$ ,  $\beta_1 = 0.1$ , and  $\beta_2 = (0.05, 0.2, 0.5)$ .



sex	comorb.	age	NS(1)	NS(5)
0	0	60	0.74	0.44
0	1	60	0.71	0.37
1	0	60	0.72	0.39
1	1	60	0.68	0.33
0	0	70	0.67	0.30
0	1	70	0.63	0.24
1	0	70	0.65	0.26
1	1	70	0.61	0.21
0	0	80	0.60	0.19
0	1	80	0.56	0.14
1	0	80	0.57	0.16
1	1	80	0.53	0.12

**Table S5.** Simulation design: AH with  $(\sigma, \kappa, \alpha) = (1.75, 0.5, 1.75)$  and  $\beta = (-0.1, -0.3, -0.5)$ .

sex	comorb.	age	NS(1)	NS(5)
0	0	60	0.90	0.51
0	1	60	0.87	0.42
1	0	60	0.88	0.44
1	1	60	0.85	0.35
0	0	70	0.81	0.34
0	1	70	0.76	0.25
1	0	70	0.77	0.27
1	1	70	0.72	0.19
0	0	80	0.69	0.22
0	1	80	0.63	0.15
1	0	80	0.64	0.16
1	1	80	0.56	0.10

**Table S6.** Simulation design: GH with  $(\sigma, \kappa, \alpha) = (1.75, 0.6, 2.5)$ ,  $\beta_1 = (0.1, 0.1, 0.1)$ , and  $\beta_2 = (0.05, 0.2, 0.25)$ .

sex	treatment	age	NS(1)	NS(5)
0	0	60	0.93	0.47
0	1	60	0.94	0.67
1	0	60	0.90	0.36
1	1	60	0.92	0.59
0	0	70	0.88	0.29
0	1	70	0.90	0.52
1	0	70	0.84	0.19
1	1	70	0.87	0.41
0	0	80	0.81	0.13
0	1	80	0.84	0.34
1	0	80	0.75	0.06
1	1	80	0.80	0.23

**Table S7.** Simulation design: CH with  $(\sigma, \kappa, \alpha) = (0.114, 0.366, 18.306)$ ,  $\beta_1 = (0, 0, 0.5)$ , and  $\beta_2 = (0.05, 0.3, -0.6)$ .

## 4 Performance of the MLEs

### 4.1 PH structure

$$h_E^{PH}(t; \mathbf{x}_j) = h_{EW}(t; 1.75, 0.5, 2.5) \times \exp(0.035 \times \text{age} + 0.2 \times \text{sex} + 0.3 \times W).$$

Model	Parameter	MMLE	mMLE	ESD	Mean Std Error	RMSE	Coverage
30% censoring							
GHEW	$\sigma$ (1.75)	1.652	1.508	1.055	1.078	1.059	0.947
	$\kappa$ (0.5)	0.489	0.477	0.129	0.138	0.129	0.947
	$\alpha$ (2.5)	3.221	2.705	1.818	1.675	1.954	0.958
	$\beta_{11}$ (0)	0.004	0.000	0.042	0.037	0.042	0.932
	$\beta_{12}$ (0)	-0.285	-0.006	2.117	4.756	2.136	0.972
	$\beta_{13}$ (0)	-0.422	-0.110	2.413	6.452	2.449	0.967
	$\beta_{21}$ (0.035)	0.035	0.035	0.007	0.007	0.007	0.931
	$\beta_{22}$ (0.2)	0.187	0.203	0.188	0.314	0.189	0.957
$\beta_{23}$ (0.3)	0.271	0.284	0.195	0.283	0.197	0.952	
PHEW	$\sigma$ (1.75)	1.717	1.571	1.085	1.078	1.085	0.954
	$\kappa$ (0.5)	0.495	0.483	0.126	0.126	0.126	0.969
	$\alpha$ (2.5)	3.082	2.682	1.565	1.487	1.669	0.977
	$\beta_1$ (0.035)	0.035	0.035	0.004	0.004	0.004	0.954
	$\beta_2$ (0.2)	0.198	0.198	0.088	0.090	0.088	0.954
	$\beta_3$ (0.3)	0.300	0.302	0.092	0.091	0.092	0.949
50% censoring							
GHEW	$\sigma$ (1.75)	1.486	1.321	1.065	1.130	1.097	0.957
	$\kappa$ (0.5)	0.465	0.459	0.138	0.157	0.142	0.948
	$\alpha$ (2.5)	3.835	2.901	3.131	2.653	3.403	0.969
	$\beta_{11}$ (0)	0.009	0.006	0.048	0.040	0.049	0.927
	$\beta_{12}$ (0)	-0.519	-0.103	2.732	9.247	2.780	0.984
	$\beta_{13}$ (0)	-0.419	-0.104	2.272	3.757	2.309	0.979
	$\beta_{21}$ (0.035)	0.035	0.035	0.007	0.007	0.007	0.935
	$\beta_{22}$ (0.2)	0.180	0.186	0.213	0.498	0.214	0.960
$\beta_{23}$ (0.3)	0.279	0.285	0.193	0.257	0.194	0.954	
PHEW	$\sigma$ (1.75)	1.583	1.406	1.118	1.163	1.130	0.970
	$\kappa$ (0.5)	0.476	0.464	0.137	0.145	0.139	0.959
	$\alpha$ (2.5)	3.529	2.856	2.594	2.210	2.790	0.987
	$\beta_1$ (0.035)	0.035	0.035	0.005	0.005	0.005	0.950
	$\beta_2$ (0.2)	0.199	0.195	0.108	0.106	0.108	0.939
	$\beta_3$ (0.3)	0.300	0.299	0.107	0.106	0.107	0.947

**Table S8.** Simulation from PH structure with  $(\sigma, \kappa, \alpha) = (1.75, 0.5, 2.5)$ ,  $\beta = (0.035, 0.2, 0.3)$ , and  $n = 1000$ . Mean of the MLEs (MMLE), median of the MLEs (mMLE), empirical standard deviation (ESD), mean standard error, root-mean-square error (RMSE), and coverage proportions (Coverage).

Model	Parameter	MMLE	mMLE	ESD	Mean Std Error	RMSE	Coverage
30% censoring							
GHEW	$\sigma$ (1.75)	1.784	1.749	0.597	0.561	0.597	0.944
	$\kappa$ (0.5)	0.505	0.504	0.068	0.064	0.068	0.943
	$\alpha$ (2.5)	2.580	2.489	0.544	0.523	0.550	0.946
	$\beta_{11}$ (0)	0.000	-0.000	0.017	0.015	0.017	0.949
	$\beta_{12}$ (0)	0.017	0.029	0.370	0.362	0.370	0.962
	$\beta_{13}$ (0)	0.007	0.017	0.388	0.367	0.387	0.953
	$\beta_{21}$ (0.035)	0.035	0.035	0.003	0.003	0.003	0.937
	$\beta_{22}$ (0.2)	0.205	0.208	0.070	0.069	0.070	0.951
	$\beta_{23}$ (0.3)	0.302	0.306	0.073	0.070	0.073	0.932
PHEW	$\sigma$ (1.75)	1.788	1.756	0.553	0.531	0.554	0.943
	$\kappa$ (0.5)	0.503	0.503	0.058	0.056	0.059	0.950
	$\alpha$ (2.5)	2.568	2.492	0.513	0.492	0.517	0.954
	$\beta_1$ (0.035)	0.035	0.035	0.002	0.002	0.002	0.936
	$\beta_2$ (0.2)	0.202	0.201	0.041	0.040	0.041	0.952
	$\beta_3$ (0.3)	0.301	0.302	0.041	0.041	0.041	0.955
50% censoring							
GHEW	$\sigma$ (1.75)	1.750	1.696	0.632	0.629	0.631	0.956
	$\kappa$ (0.5)	0.501	0.495	0.078	0.076	0.077	0.949
	$\alpha$ (2.5)	2.634	2.537	0.655	0.623	0.668	0.962
	$\beta_{11}$ (0)	0.000	-0.000	0.019	0.018	0.019	0.939
	$\beta_{12}$ (0)	-0.001	-0.004	0.457	0.435	0.457	0.960
	$\beta_{13}$ (0)	-0.038	-0.033	0.452	0.433	0.453	0.957
	$\beta_{21}$ (0.035)	0.035	0.035	0.003	0.003	0.003	0.929
	$\beta_{22}$ (0.2)	0.201	0.205	0.073	0.072	0.073	0.953
	$\beta_{23}$ (0.3)	0.296	0.296	0.075	0.073	0.076	0.941
PHEW	$\sigma$ (1.75)	1.774	1.751	0.580	0.597	0.580	0.958
	$\kappa$ (0.5)	0.502	0.501	0.065	0.066	0.065	0.961
	$\alpha$ (2.5)	2.595	2.512	0.577	0.573	0.585	0.959
	$\beta_1$ (0.035)	0.035	0.035	0.002	0.002	0.002	0.938
	$\beta_2$ (0.2)	0.201	0.201	0.046	0.047	0.046	0.957
	$\beta_3$ (0.3)	0.302	0.300	0.050	0.047	0.050	0.946

**Table S9.** Simulation from PH structure with  $(\sigma, \kappa, \alpha) = (1.75, 0.5, 2.5)$ ,  $\beta = (0.035, 0.2, 0.3)$ , and  $n = 5000$ . Mean of the MLEs (MMLE), median of the MLEs (mMLE), empirical standard deviation (ESD), mean standard error, root-mean-square error (RMSE), and coverage proportions (Coverage).

## 4.2 AFT structure

$$\begin{aligned}
 h_E^{\text{AFT}}(t; \mathbf{x}_j) &= h_{EW}(t \times \exp(0.05 \times \text{age} + 0.2 \times \text{sex} + 0.3 \times \mathbf{W}); 1.4, 0.5, 2.5) \\
 &\times \exp(0.05 \times \text{age} + 0.2 \times \text{sex} + 0.3 \times \mathbf{W}).
 \end{aligned}$$

Model	Parameter	MMLE	mMLE	ESD	Mean Std Error	RMSE	Coverage
30% censoring							
GHEW	$\sigma$ (1.75)	1.756	1.730	0.409	0.399	0.409	0.943
	$\kappa$ (0.5)	0.501	0.498	0.046	0.045	0.046	0.947
	$\alpha$ (2.5)	2.547	2.515	0.371	0.360	0.374	0.945
	$\beta_{11}$ (0)	0.000	0.000	0.010	0.011	0.010	0.960
	$\beta_{12}$ (0)	-0.001	0.005	0.246	0.248	0.246	0.962
	$\beta_{13}$ (0)	0.005	0.002	0.257	0.249	0.257	0.938
	$\beta_{21}$ (0.035)	0.035	0.035	0.002	0.002	0.002	0.942
	$\beta_{22}$ (0.2)	0.201	0.200	0.048	0.048	0.047	0.951
$\beta_{23}$ (0.3)	0.301	0.301	0.049	0.049	0.049	0.948	
PHEW	$\sigma$ (1.75)	1.765	1.738	0.391	0.376	0.391	0.937
	$\kappa$ (0.5)	0.501	0.498	0.041	0.039	0.041	0.944
	$\alpha$ (2.5)	2.537	2.503	0.349	0.339	0.351	0.938
	$\beta_1$ (0.035)	0.035	0.035	0.001	0.001	0.001	0.938
	$\beta_2$ (0.2)	0.201	0.202	0.028	0.028	0.028	0.952
	$\beta_3$ (0.3)	0.300	0.300	0.029	0.029	0.029	0.942
50% censoring							
GHEW	$\sigma$ (1.75)	1.775	1.763	0.454	0.453	0.455	0.949
	$\kappa$ (0.5)	0.504	0.503	0.054	0.054	0.054	0.957
	$\alpha$ (2.5)	2.540	2.479	0.420	0.415	0.421	0.958
	$\beta_{11}$ (0)	-0.000	0.000	0.014	0.013	0.014	0.941
	$\beta_{12}$ (0)	-0.006	-0.004	0.287	0.292	0.287	0.959
	$\beta_{13}$ (0)	0.004	0.012	0.292	0.294	0.292	0.969
	$\beta_{21}$ (0.035)	0.035	0.035	0.002	0.002	0.002	0.940
	$\beta_{22}$ (0.2)	0.201	0.202	0.050	0.050	0.050	0.950
$\beta_{23}$ (0.3)	0.300	0.301	0.051	0.051	0.051	0.956	
PHEW	$\sigma$ (1.75)	1.785	1.776	0.420	0.428	0.421	0.956
	$\kappa$ (0.5)	0.504	0.504	0.046	0.047	0.046	0.962
	$\alpha$ (2.5)	2.527	2.472	0.387	0.387	0.388	0.961
	$\beta_1$ (0.035)	0.035	0.035	0.001	0.001	0.001	0.954
	$\beta_2$ (0.2)	0.201	0.202	0.033	0.033	0.033	0.949
	$\beta_3$ (0.3)	0.300	0.301	0.034	0.033	0.034	0.948

**Table S10.** Simulation from PH structure with  $(\sigma, \kappa, \alpha) = (1.75, 0.5, 2.5)$ ,  $\beta = (0.035, 0.2, 0.3)$ , and  $n = 10000$ . Mean of the MLEs (MMLE), median of the MLEs (mMLE), empirical standard deviation (ESD), mean standard error, root-mean-square error (RMSE), and coverage proportions (Coverage).

### 4.3 HH structure

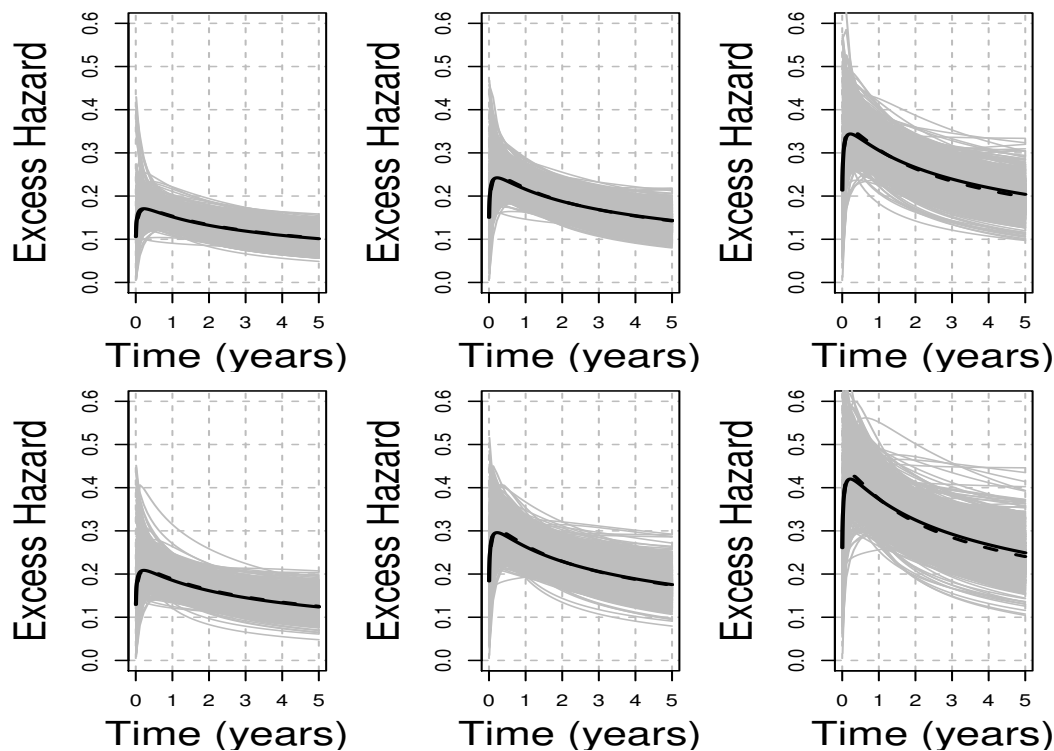
$$h_E^{HH}(t; \mathbf{x}_j) = h_{EW}(t \times \exp(0.1 \times \text{age}); 1.8, 0.5, 2.5) \times \exp(0.05 \times \text{age} + 0.2 \times \text{sex} + 0.5 \times W).$$

Model	30% censoring	50% censoring
$n = 1000$		
PHEW	65.4	59.8
AHEW	0.8	1.3
AFTEW	22.3	28.7
GHEW	11.5	10.2
$n = 5000$		
PHEW	83.3	78.4
AHEW	0	0
AFTEW	6.8	10.5
GHEW	9.9	11.4
$n = 10000$		
PHEW	88.2	85.2
AHEW	0	0
AFTEW	1.5	5.7
GHEW	10.3	9.1

**Table S11.** Simulation from PH structure with  $(\sigma, \kappa, \alpha) = (1.75, 0.5, 2.5)$ , and  $\beta = (0.035, 0.2, 0.3)$ . Percentage of models selected with AIC.

#### 4.4 AH structure

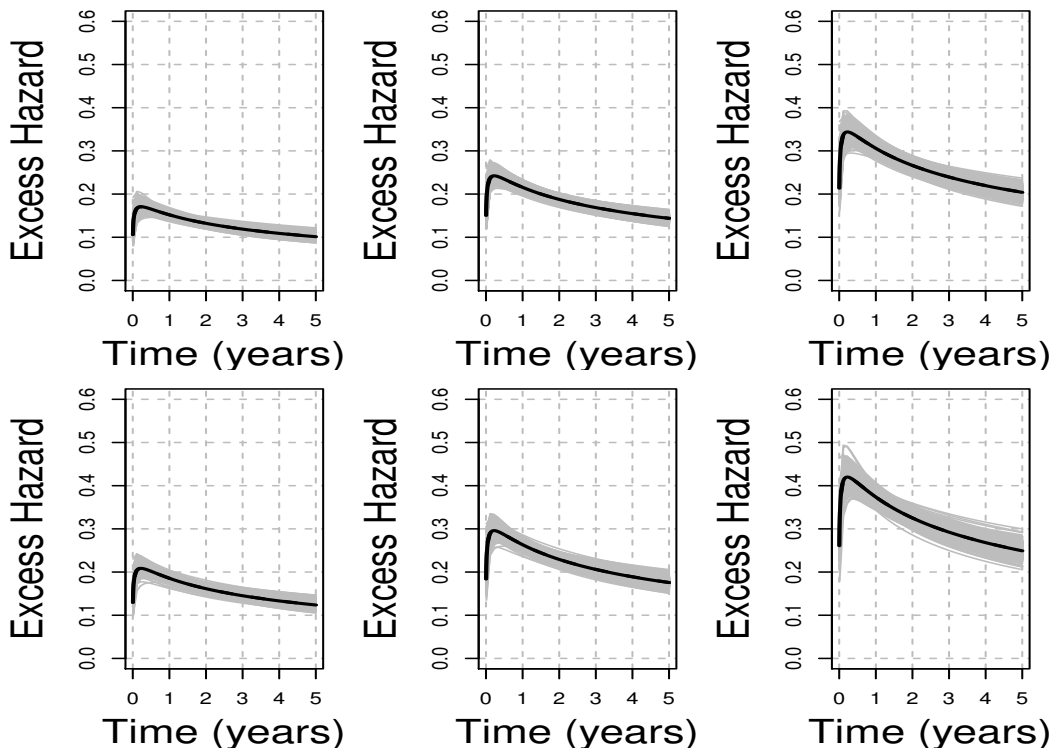
$$h_E^{AH}(t; \mathbf{x}_j) = h_{EW}(t \times \exp(-0.1 \times \text{age} - 0.3 \times \text{sex} - 0.5 \times \mathbf{W}); 1.75, 0.5, 1.75).$$



**Figure S5.** Best hazards in terms of AIC for  $n = 1000$  and 50% censoring PH: true generating hazard (continuous line), mean of best fitted hazards (dashed lines), and individual fitted hazards (gray lines). Panels from left to right correspond to covariate values (age, sex, comorbidity) = (60, 0, 0), (70, 0, 0), (80, 0, 0), (60, 1, 0), (70, 1, 0), (80, 1, 0), respectively.

#### 4.5 GH structure

$$h_E^{\text{GH}}(t; \mathbf{x}_j) = h_{EW}(t \times \exp(0.1 \times \text{age} + 0.1 \times \text{sex} + 0.1 \times \mathbf{W}); 1.75, 0.6, 2.5) \times \exp(0.05 \times \text{age} + 0.2 \times \text{sex} + 0.25 \times \mathbf{W}).$$



**Figure S6.** Best hazards in terms of AIC for  $n = 10000$  and 30% censoring PH: true generating hazard (continuous line), mean of best fitted hazards (dashed lines), and individual fitted hazards (gray lines). Panels from left to right correspond to covariate values (age, sex, comorbidity) = (60, 0, 0), (70, 0, 0), (80, 0, 0), (60, 1, 0), (70, 1, 0), (80, 1, 0), respectively.

#### 4.6 CH structure

$$h_E^{\text{CH}}(t; \mathbf{x}_j) = h_{EW}(t \times \exp(0.5 \times W); 0.114, 0.366, 18.306) \times \exp(0.05 \times \text{age} + 0.3 \times \text{sex} - 0.6 \times W)$$



Model	Parameter	MMLE	mMLE	ESD	Mean Std Error	RMSE	Coverage
30% censoring							
GHEW	$\sigma$ (1.4)	1.306	1.206	0.758	0.797	0.763	0.969
	$\kappa$ (0.5)	0.486	0.484	0.111	0.118	0.112	0.940
	$\alpha$ (2.5)	3.127	2.715	1.623	1.426	1.739	0.965
	$\beta_{11}$ (0.05)	0.049	0.049	0.039	0.034	0.039	0.926
	$\beta_{12}$ (0.2)	0.063	0.212	1.498	2.529	1.503	0.978
	$\beta_{13}$ (0.3)	0.161	0.205	1.279	1.398	1.286	0.973
	$\beta_{21}$ (0.05)	0.049	0.050	0.007	0.006	0.007	0.953
	$\beta_{22}$ (0.2)	0.186	0.198	0.206	0.278	0.206	0.953
	$\beta_{23}$ (0.3)	0.285	0.289	0.172	0.182	0.172	0.958
AFTEW	$\sigma$ (1.4)	1.395	1.324	0.746	0.741	0.745	0.963
	$\kappa$ (0.5)	0.496	0.489	0.095	0.095	0.095	0.956
	$\alpha$ (2.5)	2.901	2.598	1.326	1.101	1.384	0.964
	$\beta_1$ (0.05)	0.050	0.050	0.004	0.004	0.004	0.957
	$\beta_2$ (0.2)	0.197	0.199	0.108	0.107	0.108	0.943
	$\beta_3$ (0.3)	0.303	0.304	0.105	0.107	0.105	0.954
50% censoring							
GHEW	$\sigma$ (1.4)	1.232	1.147	0.791	0.868	0.808	0.970
	$\kappa$ (0.5)	0.475	0.473	0.126	0.140	0.128	0.948
	$\alpha$ (2.5)	3.461	2.796	2.322	2.014	2.512	0.975
	$\beta_{11}$ (0.05)	0.050	0.049	0.047	0.040	0.047	0.924
	$\beta_{12}$ (0.2)	-0.009	0.106	1.699	2.281	1.711	0.973
	$\beta_{13}$ (0.3)	0.076	0.227	1.559	2.005	1.575	0.964
	$\beta_{21}$ (0.05)	0.049	0.049	0.008	0.007	0.008	0.953
	$\beta_{22}$ (0.2)	0.180	0.186	0.207	0.241	0.208	0.952
	$\beta_{23}$ (0.3)	0.276	0.284	0.195	0.204	0.197	0.950
AFTEW	(1.4)	1.350	1.241	0.816	0.833	0.817	0.967
	$\kappa$ (0.5)	0.489	0.479	0.110	0.114	0.111	0.963
	$\alpha$ (2.5)	3.098	2.704	1.617	1.437	1.723	0.973
	$\beta_1$ (0.05)	0.050	0.050	0.005	0.005	0.005	0.948
	$\beta_2$ (0.2)	0.197	0.196	0.124	0.123	0.124	0.942
	$\beta_3$ (0.3)	0.299	0.297	0.125	0.123	0.125	0.947

**Table S12.** Simulation from AFT structure with  $(\sigma, \kappa, \alpha) = (1.4, 0.5, 2.5)$ ,  $\beta = (0.05, 0.2, 0.3)$ , and  $n = 1000$ . Mean of the MLEs (MMLE), median of the MLEs (mMLE), empirical standard deviation (ESD), mean standard error, root-mean-square error (RMSE), and coverage proportions (Coverage).

Model	Parameter	MMLE	mMLE	ESD	Mean Std Error	RMSE	Coverage
30% censoring							
GHEW	$\sigma$ (1.4)	1.419	1.386	0.409	0.405	0.410	0.943
	$\kappa$ (0.5)	0.503	0.500	0.056	0.055	0.056	0.940
	$\alpha$ (2.5)	2.563	2.519	0.472	0.461	0.476	0.950
	$\beta_{11}$ (0.05)	0.051	0.050	0.015	0.015	0.015	0.936
	$\beta_{12}$ (0.2)	0.212	0.227	0.324	0.311	0.324	0.949
	$\beta_{13}$ (0.3)	0.304	0.297	0.327	0.315	0.327	0.957
	$\beta_{21}$ (0.05)	0.050	0.050	0.003	0.003	0.003	0.955
	$\beta_{22}$ (0.2)	0.203	0.202	0.075	0.072	0.075	0.943
$\beta_{23}$ (0.3)	0.301	0.301	0.074	0.073	0.074	0.943	
AFTEW	$\sigma$ (1.4)	1.428	1.403	0.348	0.351	0.349	0.951
	$\kappa$ (0.5)	0.503	0.500	0.042	0.042	0.042	0.955
	$\alpha$ (2.5)	2.534	2.506	0.390	0.390	0.392	0.956
	$\beta_1$ (0.05)	0.050	0.050	0.002	0.002	0.002	0.952
	$\beta_2$ (0.2)	0.201	0.199	0.049	0.048	0.049	0.944
	$\beta_3$ (0.3)	0.302	0.303	0.048	0.048	0.048	0.952
50% censoring							
GHEW	$\sigma$ (1.4)	1.407	1.375	0.472	0.464	0.472	0.946
	$\kappa$ (0.5)	0.502	0.500	0.069	0.067	0.069	0.937
	$\alpha$ (2.5)	2.604	2.504	0.583	0.559	0.592	0.946
	$\beta_{11}$ (0.05)	0.051	0.051	0.019	0.018	0.019	0.945
	$\beta_{12}$ (0.2)	0.206	0.203	0.378	0.368	0.378	0.946
	$\beta_{13}$ (0.3)	0.298	0.276	0.382	0.373	0.382	0.958
	$\beta_{21}$ (0.05)	0.050	0.050	0.003	0.003	0.003	0.944
	$\beta_{22}$ (0.2)	0.200	0.200	0.076	0.077	0.076	0.950
$\beta_{23}$ (0.3)	0.300	0.300	0.079	0.077	0.079	0.950	
AFTEW	$\sigma$ (1.4)	1.423	1.411	0.400	0.404	0.401	0.953
	$\kappa$ (0.5)	0.503	0.502	0.051	0.051	0.051	0.951
	$\alpha$ (2.5)	2.558	2.497	0.471	0.462	0.474	0.951
	$\beta_1$ (0.05)	0.050	0.050	0.002	0.002	0.002	0.939
	$\beta_2$ (0.2)	0.199	0.202	0.053	0.054	0.053	0.957
	$\beta_3$ (0.3)	0.301	0.301	0.057	0.055	0.057	0.939

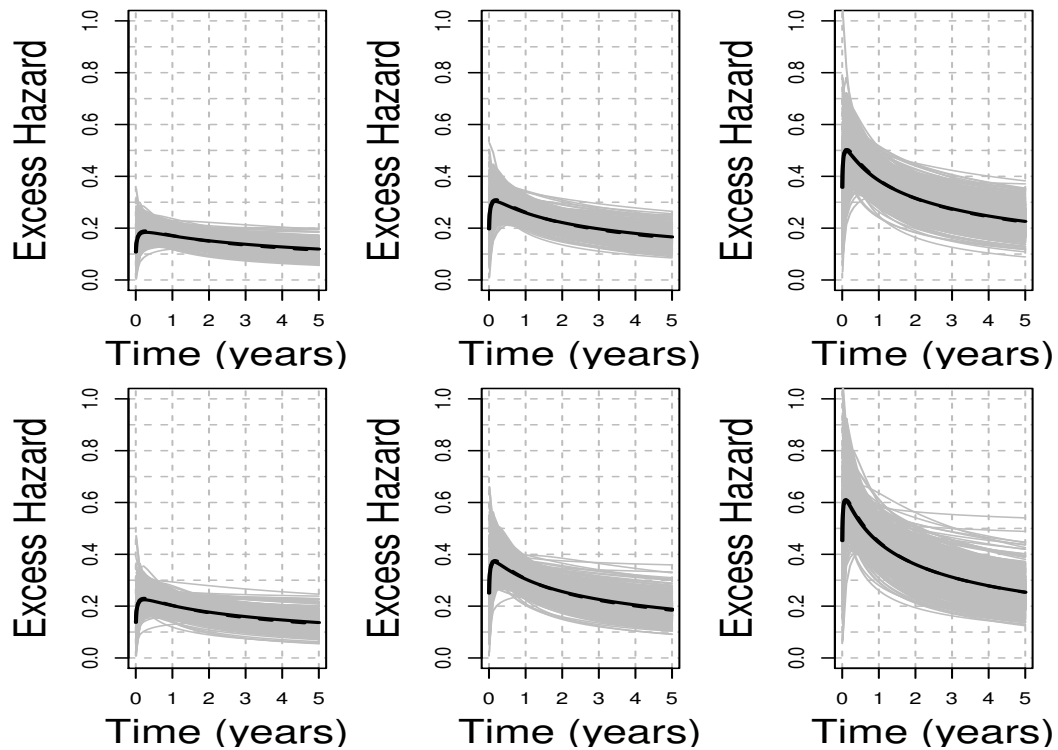
**Table S13.** Simulation from AFT structure with  $(\sigma, \kappa, \alpha) = (1.4, 0.5, 2.5)$ ,  $\beta = (0.05, 0.2, 0.3)$ , and  $n = 5000$ . Mean of the MLEs (MMLE), median of the MLEs (mMLE), empirical standard deviation (ESD), mean standard error, root-mean-square error (RMSE), and coverage proportions (Coverage).

Model	Parameter	MMLE	mMLE	ESD	Mean Std Error	RMSE	Coverage
30% censoring							
GHEW	$\sigma$ (1.4)	1.407	1.401	0.295	0.289	0.295	0.940
	$\kappa$ (0.5)	0.501	0.500	0.040	0.039	0.040	0.945
	$\alpha$ (2.5)	2.536	2.499	0.327	0.320	0.329	0.939
	$\beta_{11}$ (0.05)	0.050	0.050	0.010	0.010	0.010	0.954
	$\beta_{12}$ (0.2)	0.201	0.203	0.215	0.216	0.215	0.959
	$\beta_{13}$ (0.3)	0.305	0.297	0.222	0.218	0.222	0.948
	$\beta_{21}$ (0.05)	0.050	0.050	0.002	0.002	0.002	0.941
	$\beta_{22}$ (0.2)	0.200	0.201	0.051	0.051	0.051	0.958
$\beta_{23}$ (0.3)	0.302	0.300	0.051	0.051	0.051	0.942	
AFTEW	$\sigma$ (1.4)	1.414	1.399	0.255	0.248	0.256	0.944
	$\kappa$ (0.5)	0.501	0.500	0.031	0.030	0.031	0.946
	$\alpha$ (2.5)	2.520	2.504	0.278	0.272	0.279	0.940
	$\beta_1$ (0.05)	0.050	0.050	0.001	0.001	0.001	0.936
	$\beta_2$ (0.2)	0.200	0.201	0.034	0.034	0.034	0.950
	$\beta_3$ (0.3)	0.302	0.304	0.034	0.034	0.034	0.949
50% censoring							
GHEW	$\sigma$ (1.4)	1.419	1.405	0.337	0.335	0.337	0.940
	$\kappa$ (0.5)	0.504	0.501	0.048	0.048	0.049	0.941
	$\alpha$ (2.5)	2.532	2.494	0.381	0.378	0.382	0.939
	$\beta_{11}$ (0.05)	0.051	0.050	0.013	0.013	0.014	0.936
	$\beta_{12}$ (0.2)	0.204	0.203	0.258	0.255	0.258	0.944
	$\beta_{13}$ (0.3)	0.311	0.309	0.265	0.259	0.265	0.957
	$\beta_{21}$ (0.05)	0.050	0.050	0.002	0.002	0.002	0.948
	$\beta_{22}$ (0.2)	0.200	0.204	0.056	0.054	0.056	0.944
$\beta_{23}$ (0.3)	0.301	0.302	0.055	0.054	0.055	0.948	
AFTEW	$\sigma$ (1.4)	1.420	1.400	0.286	0.288	0.286	0.951
	$\kappa$ (0.5)	0.502	0.501	0.036	0.036	0.036	0.960
	$\alpha$ (2.5)	2.520	2.503	0.320	0.318	0.320	0.951
	$\beta_1$ (0.05)	0.050	0.050	0.002	0.002	0.002	0.948
	$\beta_2$ (0.2)	0.200	0.200	0.040	0.039	0.040	0.938
	$\beta_3$ (0.3)	0.300	0.300	0.038	0.039	0.038	0.950

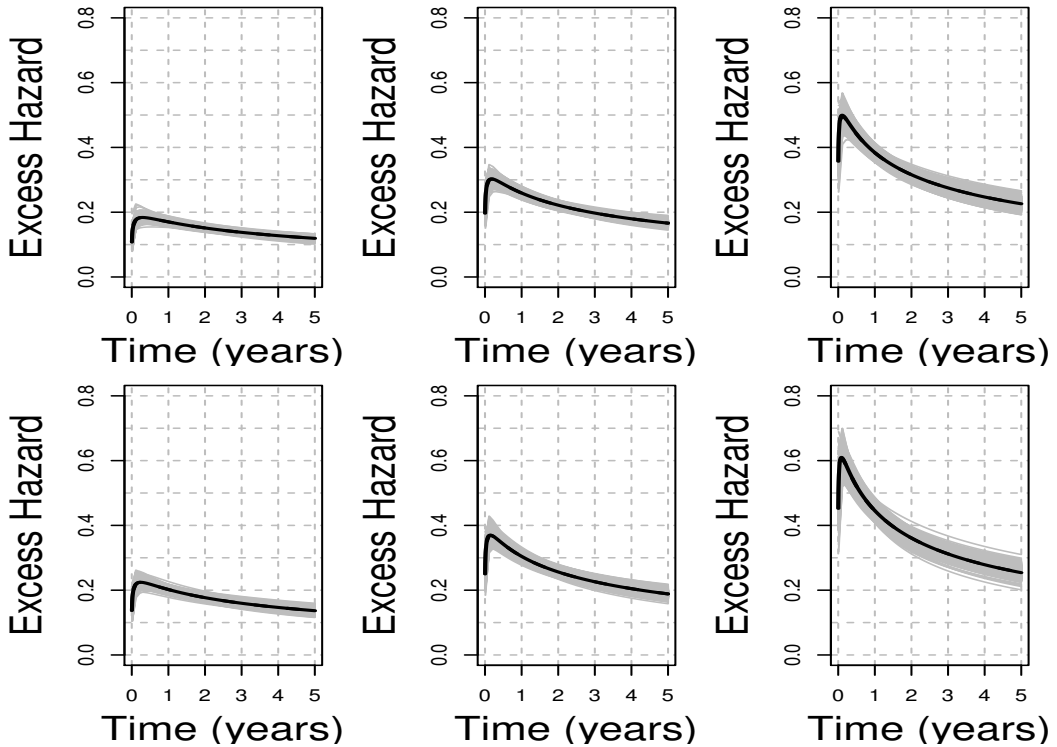
**Table S14.** Simulation from AFT structure with  $(\sigma, \kappa, \alpha) = (1.4, 0.5, 2.5)$ ,  $\beta = (0.05, 0.2, 0.3)$ , and  $n = 10000$ . Mean of the MLEs (MMLE), median of the MLEs (mMLE), empirical standard deviation (ESD), mean standard error, root-mean-square error (RMSE), and coverage proportions (Coverage).

Model	30% censoring	50% censoring
$n = 1000$		
PHEW	19.8	22.3
AHEW	0.1	0
AFTEW	71.8	65.9
GHEW	8.3	11.8
$n = 5000$		
PHEW	1.9	5.1
AHEW	0	0
AFTEW	86.9	84.7
GHEW	11.2	10.2
$n = 10000$		
PHEW	0.2	0.9
AHEW	0	0
AFTEW	89.2	87.9
GHEW	10.6	11.2

**Table S15.** Simulation from AFT structure with  $(\sigma, \kappa, \alpha) = (1.4, 0.5, 2.5)$ , and  $\beta = (0.05, 0.2, 0.3)$ . Percentage of models selected with AIC.



**Figure S7.** Best hazards in terms of AIC for  $n = 1000$  and 50% censoring AFT: true generating hazard (continuous line), mean of best fitted hazards (dashed lines), and individual fitted hazards (gray lines). Panels from left to right correspond to covariate values (age, sex, comorbidity) = (60, 0, 0), (70, 0, 0), (80, 0, 0), (60, 1, 0), (70, 1, 0), (80, 1, 0), respectively.



**Figure S8.** Best hazards in terms of AIC for  $n = 10000$  and 30% censoring AFT: true generating hazard (continuous line), mean of best fitted hazards (dashed lines), and individual fitted hazards (gray lines). Panels from left to right correspond to covariate values (age, sex, comorbidity) = (60, 0, 0), (70, 0, 0), (80, 0, 0), (60, 1, 0), (70, 1, 0), (80, 1, 0), respectively.

Model	Parameter	MMLE	mMLE	ESD	Mean Std Error	RMSE	Coverage
30% censoring							
GHEW	$\sigma$ (1.8)	1.647	1.577	0.940	1.026	0.952	0.976
	$\kappa$ (0.5)	0.483	0.477	0.114	0.125	0.115	0.960
	$\alpha$ (2.5)	3.184	2.719	1.779	1.562	1.905	0.982
	$\beta_{11}$ (0.1)	0.099	0.097	0.042	0.038	0.042	0.926
	$\beta_{12}$ (0)	-0.066	-0.038	0.859	0.911	0.861	0.973
	$\beta_{13}$ (0)	-0.081	-0.050	1.000	0.843	1.003	0.966
	$\beta_{21}$ (0.05)	0.049	0.049	0.006	0.006	0.006	0.965
	$\beta_{22}$ (0.2)	0.195	0.199	0.144	0.154	0.144	0.958
$\beta_{23}$ (0.5)	0.495	0.500	0.149	0.149	0.149	0.962	
HHEW	$\sigma$ (1.8)	1.688	1.611	0.931	1.004	0.937	0.972
	$\kappa$ (0.5)	0.487	0.481	0.109	0.116	0.109	0.954
	$\alpha$ (2.5)	3.082	2.680	1.587	1.419	1.689	0.976
	$\beta_{11}$ (0.1)	0.101	0.096	0.039	0.038	0.039	0.937
	$\beta_{21}$ (0.05)	0.050	0.049	0.005	0.006	0.005	0.966
	$\beta_{22}$ (0.2)	0.201	0.200	0.091	0.090	0.091	0.946
	$\beta_{23}$ (0.5)	0.506	0.510	0.088	0.091	0.088	0.956
	50% censoring						
GHEW	$\sigma$ (1.8)	1.575	1.467	1.011	1.118	1.035	0.963
	$\kappa$ (0.5)	0.474	0.466	0.129	0.147	0.131	0.967
	$\alpha$ (2.5)	3.464	2.805	2.346	2.131	2.535	0.980
	$\beta_{11}$ (0.1)	0.098	0.096	0.049	0.046	0.049	0.927
	$\beta_{12}$ (0)	-0.144	-0.039	1.425	1.419	1.432	0.965
	$\beta_{13}$ (0)	-0.102	-0.037	1.049	1.134	1.053	0.962
	$\beta_{21}$ (0.05)	0.049	0.049	0.007	0.007	0.007	0.953
	$\beta_{22}$ (0.2)	0.194	0.196	0.168	0.173	0.168	0.958
$\beta_{23}$ (0.5)	0.488	0.494	0.167	0.161	0.167	0.951	
HHEW	$\sigma$ (1.8)	1.630	1.517	1.002	1.109	1.016	0.969
	$\kappa$ (0.5)	0.479	0.471	0.121	0.135	0.123	0.967
	$\alpha$ (2.5)	3.336	2.778	2.235	1.931	2.385	0.982
	$\beta_{11}$ (0.1)	0.098	0.097	0.044	0.044	0.044	0.939
	$\beta_{21}$ (0.05)	0.049	0.049	0.007	0.007	0.007	0.958
	$\beta_{22}$ (0.2)	0.200	0.201	0.107	0.105	0.107	0.943
	$\beta_{23}$ (0.5)	0.499	0.497	0.107	0.106	0.107	0.945

**Table S16.** Simulation from HH structure with  $(\sigma, \kappa, \alpha) = (1.8, 0.5, 2.5)$ ,  $\beta_{11} = 0.1$ ,  $\beta_2 = (0.05, 0.2, 0.5)$ , and  $n = 1000$ . Mean of the MLEs (MMLE), median of the MLEs (mMLE), empirical standard deviation (ESD), mean standard error, root-mean-square error (RMSE), and coverage proportions (Coverage).

Model	Parameter	MMLE	mMLE	ESD	Mean Std Error	RMSE	Coverage
30% censoring							
GHEW	$\sigma$ (1.8)	1.812	1.810	0.510	0.519	0.510	0.959
	$\kappa$ (0.5)	0.502	0.501	0.057	0.058	0.057	0.953
	$\alpha$ (2.5)	2.571	2.506	0.472	0.476	0.477	0.957
	$\beta_{11}$ (0.1)	0.101	0.100	0.017	0.016	0.017	0.952
	$\beta_{12}$ (0)	0.001	0.004	0.290	0.284	0.290	0.957
	$\beta_{13}$ (0)	-0.003	-0.009	0.299	0.289	0.298	0.950
	$\beta_{21}$ (0.05)	0.050	0.050	0.002	0.002	0.002	0.945
	$\beta_{22}$ (0.2)	0.202	0.204	0.060	0.060	0.060	0.959
$\beta_{23}$ (0.5)	0.500	0.501	0.063	0.062	0.063	0.949	
HHEW	$\sigma$ (1.8)	1.822	1.831	0.491	0.500	0.491	0.957
	$\kappa$ (0.5)	0.503	0.502	0.052	0.053	0.053	0.960
	$\alpha$ (2.5)	2.558	2.486	0.455	0.457	0.458	0.960
	$\beta_{11}$ (0.1)	0.101	0.100	0.016	0.016	0.016	0.951
	$\beta_{21}$ (0.05)	0.050	0.050	0.002	0.002	0.002	0.945
	$\beta_{22}$ (0.2)	0.202	0.202	0.039	0.040	0.039	0.959
	$\beta_{23}$ (0.5)	0.502	0.500	0.041	0.041	0.041	0.952
	50% censoring						
GHEW	$\sigma$ (1.8)	1.813	1.800	0.584	0.589	0.584	0.949
	$\kappa$ (0.5)	0.503	0.500	0.070	0.070	0.070	0.947
	$\alpha$ (2.5)	2.599	2.489	0.601	0.566	0.609	0.955
	$\beta_{11}$ (0.1)	0.102	0.101	0.020	0.020	0.020	0.950
	$\beta_{12}$ (0)	-0.018	-0.001	0.342	0.338	0.342	0.964
	$\beta_{13}$ (0)	-0.000	0.002	0.348	0.345	0.348	0.960
	$\beta_{21}$ (0.05)	0.050	0.050	0.003	0.003	0.003	0.953
	$\beta_{22}$ (0.2)	0.198	0.200	0.064	0.066	0.064	0.969
$\beta_{23}$ (0.5)	0.501	0.502	0.070	0.068	0.070	0.946	
HHEW	$\sigma$ (1.8)	1.824	1.817	0.553	0.568	0.553	0.953
	$\kappa$ (0.5)	0.504	0.502	0.063	0.064	0.063	0.951
	$\alpha$ (2.5)	2.579	2.479	0.557	0.538	0.562	0.958
	$\beta_{11}$ (0.1)	0.102	0.101	0.020	0.020	0.020	0.956
	$\beta_{21}$ (0.05)	0.050	0.050	0.003	0.003	0.003	0.953
	$\beta_{22}$ (0.2)	0.200	0.200	0.045	0.046	0.045	0.956
	$\beta_{23}$ (0.5)	0.502	0.500	0.048	0.047	0.048	0.949

**Table S17.** Simulation from HH structure with  $(\sigma, \kappa, \alpha) = (1.8, 0.5, 2.5)$ ,  $\beta_{11} = 0.1$ ,  $\beta_2 = (0.05, 0.2, 0.5)$ , and  $n = 5000$ . Mean of the MLEs (MMLE), median of the MLEs (mMLE), empirical standard deviation (ESD), mean standard error, root-mean-square error (RMSE), and coverage proportions (Coverage).

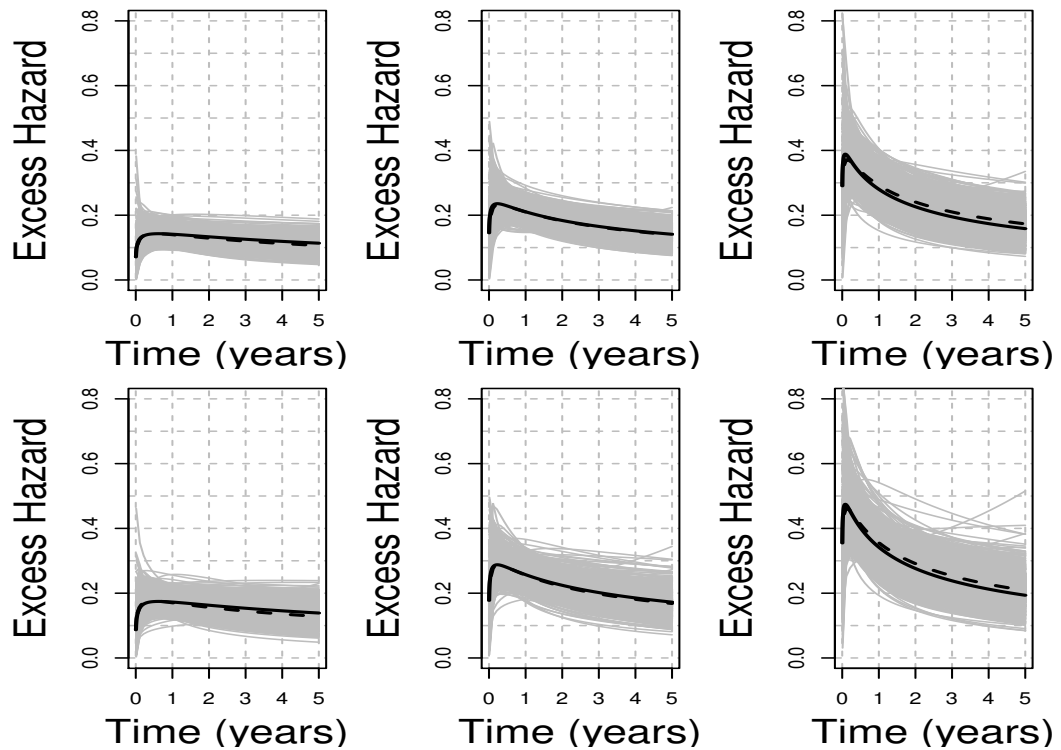


Model	Parameter	MMLE	mMLE	ESD	Mean Std Error	RMSE	Coverage
30% censoring							
GHEW	$\sigma$ (1.8)	1.804	1.793	0.375	0.371	0.375	0.953
	$\kappa$ (0.5)	0.501	0.498	0.041	0.041	0.041	0.945
	$\alpha$ (2.5)	2.541	2.510	0.336	0.329	0.338	0.948
	$\beta_{11}$ (0.1)	0.100	0.100	0.011	0.011	0.011	0.950
	$\beta_{12}$ (0)	0.005	0.007	0.201	0.199	0.201	0.946
	$\beta_{13}$ (0)	0.001	0.003	0.211	0.202	0.211	0.935
	$\beta_{21}$ (0.05)	0.050	0.050	0.002	0.002	0.002	0.945
	$\beta_{22}$ (0.2)	0.201	0.200	0.043	0.042	0.043	0.943
$\beta_{23}$ (0.5)	0.502	0.502	0.045	0.044	0.045	0.935	
HHEW	$\sigma$ (1.8)	1.804	1.788	0.361	0.356	0.361	0.949
	$\kappa$ (0.5)	0.500	0.500	0.038	0.038	0.038	0.947
	$\alpha$ (2.5)	2.539	2.504	0.326	0.318	0.328	0.947
	$\beta_{11}$ (0.1)	0.100	0.100	0.011	0.011	0.011	0.953
	$\beta_{21}$ (0.05)	0.050	0.050	0.002	0.002	0.002	0.949
	$\beta_{22}$ (0.2)	0.200	0.200	0.029	0.028	0.029	0.950
	$\beta_{23}$ (0.5)	0.502	0.501	0.029	0.029	0.029	0.953
	50% censoring						
GHEW	$\sigma$ (1.8)	1.819	1.803	0.429	0.424	0.429	0.940
	$\kappa$ (0.5)	0.503	0.500	0.050	0.049	0.050	0.944
	$\alpha$ (2.5)	2.537	2.500	0.390	0.382	0.392	0.955
	$\beta_{11}$ (0.1)	0.102	0.101	0.014	0.014	0.014	0.955
	$\beta_{12}$ (0)	0.000	0.002	0.242	0.235	0.242	0.938
	$\beta_{13}$ (0)	-0.000	0.003	0.243	0.239	0.243	0.952
	$\beta_{21}$ (0.05)	0.050	0.050	0.002	0.002	0.002	0.955
	$\beta_{22}$ (0.2)	0.202	0.202	0.049	0.046	0.049	0.941
$\beta_{23}$ (0.5)	0.500	0.500	0.049	0.048	0.049	0.946	
HHEW	$\sigma$ (1.8)	1.822	1.809	0.405	0.406	0.405	0.955
	$\kappa$ (0.5)	0.503	0.501	0.045	0.045	0.045	0.957
	$\alpha$ (2.5)	2.530	2.488	0.368	0.366	0.369	0.962
	$\beta_{11}$ (0.1)	0.102	0.101	0.014	0.014	0.014	0.957
	$\beta_{21}$ (0.05)	0.050	0.050	0.002	0.002	0.002	0.957
	$\beta_{22}$ (0.2)	0.202	0.201	0.034	0.033	0.034	0.941
	$\beta_{23}$ (0.5)	0.500	0.501	0.034	0.033	0.034	0.946

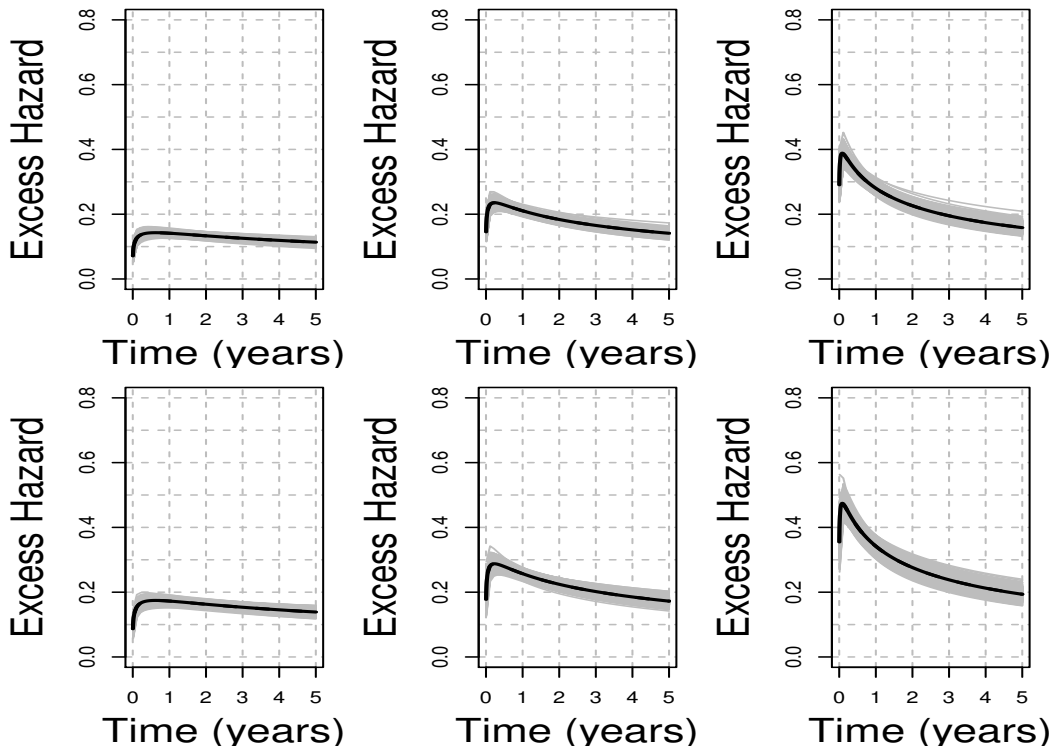
**Table S18.** Simulation from HH structure with  $(\sigma, \kappa, \alpha) = (1.8, 0.5, 2.5)$ ,  $\beta_{11} = 0.1$ ,  $\beta_2 = (0.05, 0.2, 0.5)$ , and  $n = 10000$ . Mean of the MLEs (MMLE), median of the MLEs (mMLE), empirical standard deviation (ESD), mean standard error, root-mean-square error (RMSE), and coverage proportions (Coverage).

Model	30% censoring	50% censoring
$n = 1000$		
PHEW	3.1	5.5
AHEW	0	0.1
AFTEW	23.5	30.0
GHEW	11.7	11.5
HHEW	61.7	52.9
$n = 5000$		
PHEW	0	0
AHEW	0	0
AFTEW	0.4	2.4
GHEW	14.4	13.2
HHEW	85.2	84.4
$n = 10000$		
PHEW	0	0
AHEW	0	0
AFTEW	0	0.1
GHEW	15.5	14.4
HHEW	84.5	85.5

**Table S19.** Simulation from HH structure with  $(\sigma, \kappa, \alpha) = (1.8, 0.5, 2.5)$ ,  $\beta_{11} = 0.1$ ,  $\beta_2 = (0.05, 0.2, 0.5)$ . Percentage of models selected with AIC.



**Figure S9.** Best hazards in terms of AIC for  $n = 1000$  and 50% censoring HH: true generating hazard (continuous line), mean of best fitted hazards (dashed lines), and individual fitted hazards (gray lines). Panels from left to right correspond to covariate values (age, sex, comorbidity) = (60, 0, 0), (70, 0, 0), (80, 0, 0), (60, 1, 0), (70, 1, 0), (80, 1, 0), respectively.



**Figure S10.** Best hazards in terms of AIC for  $n = 10000$  and 30% censoring HH: true generating hazard (continuous line), mean of best fitted hazards (dashed lines), and individual fitted hazards (gray lines). Panels from left to right correspond to covariate values (age, sex, comorbidity) = (60, 0, 0), (70, 0, 0), (80, 0, 0), (60, 1, 0), (70, 1, 0), (80, 1, 0), respectively.

Model	Parameter	MMLE	mMLE	ESD	Mean Std Error	RMSE	Coverage
30% censoring							
GHEW	$\sigma$ (1.75)	1.206	1.154	0.573	0.718	0.790	0.994
	$\kappa$ (0.5)	0.429	0.431	0.070	0.227	0.100	0.942
	$\alpha$ (1.75)	2.391	2.182	0.805	0.356	1.029	0.960
	$\beta_{11}$ (-0.1)	-0.075	-0.079	0.042	0.044	0.049	0.886
	$\beta_{12}$ (-0.3)	-0.613	-0.371	1.908	2.052	1.933	0.973
	$\beta_{13}$ (-0.5)	-0.863	-0.554	2.050	2.642	2.080	0.988
	$\beta_{21}$ (0)	0.007	0.006	0.015	0.016	0.017	0.911
	$\beta_{22}$ (0)	-0.082	-0.039	0.498	0.532	0.505	0.967
$\beta_{23}$ (0)	-0.097	-0.025	0.529	0.677	0.538	0.982	
AHEW	$\sigma$ (1.75)	1.420	1.404	0.657	0.749	0.735	0.989
	$\kappa$ (0.5)	0.457	0.457	0.077	0.091	0.088	0.951
	$\alpha$ (1.75)	2.165	2.002	0.712	0.710	0.824	0.975
	$\beta_1$ (-0.1)	-0.095	-0.095	0.013	0.016	0.014	0.914
	$\beta_2$ (-0.3)	-0.301	-0.288	0.281	0.271	0.281	0.951
	$\beta_3$ (-0.5)	-0.508	-0.508	0.282	0.274	0.282	0.958
50% censoring							
GHEW	$\sigma$ (1.75)	1.100	1.065	0.612	0.952	0.893	0.999
	$\kappa$ (0.5)	0.407	0.412	0.082	0.291	0.124	0.926
	$\alpha$ (1.75)	2.670	2.323	1.207	0.446	1.517	0.974
	$\beta_{11}$ (-0.1)	-0.065	-0.071	0.055	0.053	0.065	0.880
	$\beta_{12}$ (-0.3)	-0.779	-0.476	2.429	3.987	2.475	0.983
	$\beta_{13}$ (-0.5)	-1.126	-0.664	2.849	4.149	2.915	0.981
	$\beta_{21}$ (0)	0.010	0.008	0.018	0.017	0.020	0.893
	$\beta_{22}$ (0)	-0.116	-0.050	0.587	0.895	0.598	0.973
$\beta_{23}$ (0)	-0.147	-0.046	0.679	0.945	0.695	0.968	
AHEW	$\sigma$ (1.75)	1.372	1.320	0.745	0.870	0.836	0.996
	$\kappa$ (0.5)	0.444	0.443	0.094	0.114	0.109	0.948
	$\alpha$ (1.75)	2.337	2.088	1.014	0.985	1.171	0.975
	$\beta_1$ (-0.1)	-0.094	-0.093	0.015	0.019	0.016	0.923
	$\beta_2$ (-0.3)	-0.305	-0.298	0.331	0.338	0.331	0.960
	$\beta_3$ (-0.5)	-0.524	-0.511	0.344	0.343	0.344	0.961

**Table S20.** Simulation from AH structure with  $(\sigma, \kappa, \alpha) = (1.75, 0.2, 1.75)$ ,  $\beta = (-0.1, -0.3, -0.5)$ , and  $n = 1000$ . Mean of the MLEs (MMLE), median of the MLEs (mMLE), empirical standard deviation (ESD), mean standard error, root-mean-square error (RMSE), and coverage proportions (Coverage).

Model	Parameter	MMLE	mMLE	ESD	Mean Std Error	RMSE	Coverage
30% censoring							
GHEW	$\sigma$ (1.75)	1.660	1.654	0.355	0.247	0.367	0.977
	$\kappa$ (0.5)	0.489	0.489	0.042	0.099	0.043	0.974
	$\alpha$ (1.75)	1.842	1.807	0.242	0.146	0.259	0.977
	$\beta_{11}$ (-0.1)	-0.096	-0.096	0.019	0.024	0.020	0.956
	$\beta_{12}$ (-0.3)	-0.317	-0.299	0.454	0.473	0.454	0.956
	$\beta_{13}$ (-0.5)	-0.525	-0.515	0.475	0.477	0.475	0.966
	$\beta_{21}$ (0)	0.001	0.001	0.007	0.008	0.007	0.964
	$\beta_{22}$ (0)	-0.004	-0.004	0.142	0.145	0.142	0.962
$\beta_{23}$ (0)	-0.006	-0.005	0.149	0.147	0.149	0.965	
AHEW	$\sigma$ (1.75)	1.719	1.704	0.352	0.363	0.353	0.959
	$\kappa$ (0.5)	0.496	0.494	0.040	0.042	0.040	0.964
	$\alpha$ (1.75)	1.798	1.778	0.220	0.228	0.226	0.961
	$\beta_1$ (-0.1)	-0.100	-0.100	0.007	0.008	0.007	0.949
	$\beta_2$ (-0.3)	-0.303	-0.301	0.126	0.125	0.126	0.953
	$\beta_3$ (-0.5)	-0.505	-0.505	0.127	0.126	0.127	0.952
50% censoring							
GHEW	$\sigma$ (1.75)	1.594	1.600	0.390	0.299	0.420	0.983
	$\kappa$ (0.5)	0.479	0.483	0.049	0.125	0.053	0.956
	$\alpha$ (1.75)	1.901	1.839	0.308	0.177	0.343	0.972
	$\beta_{11}$ (-0.1)	-0.094	-0.095	0.022	0.027	0.023	0.945
	$\beta_{12}$ (-0.3)	-0.393	-0.366	0.563	0.586	0.570	0.971
	$\beta_{13}$ (-0.5)	-0.571	-0.546	0.612	0.590	0.616	0.955
	$\beta_{21}$ (0)	0.002	0.001	0.007	0.009	0.008	0.951
	$\beta_{22}$ (0)	-0.027	-0.027	0.161	0.169	0.164	0.970
$\beta_{23}$ (0)	-0.019	-0.015	0.178	0.171	0.179	0.943	
AHEW	$\sigma$ (1.75)	1.678	1.651	0.394	0.417	0.400	0.971
	$\kappa$ (0.5)	0.491	0.490	0.049	0.052	0.050	0.958
	$\alpha$ (1.75)	1.835	1.810	0.277	0.283	0.290	0.964
	$\beta_1$ (-0.1)	-0.099	-0.099	0.009	0.009	0.009	0.951
	$\beta_2$ (-0.3)	-0.299	-0.307	0.155	0.156	0.154	0.955
	$\beta_3$ (-0.5)	-0.505	-0.501	0.164	0.157	0.164	0.951

**Table S21.** Simulation from AH structure with  $(\sigma, \kappa, \alpha) = (1.75, 0.2, 1.75)$ ,  $\beta = (-0.1, -0.3, -0.5)$ , and  $n = 5000$ . Mean of the MLEs (MMLE), median of the MLEs (mMLE), empirical standard deviation (ESD), mean standard error, root-mean-square error (RMSE), and coverage proportions (Coverage).

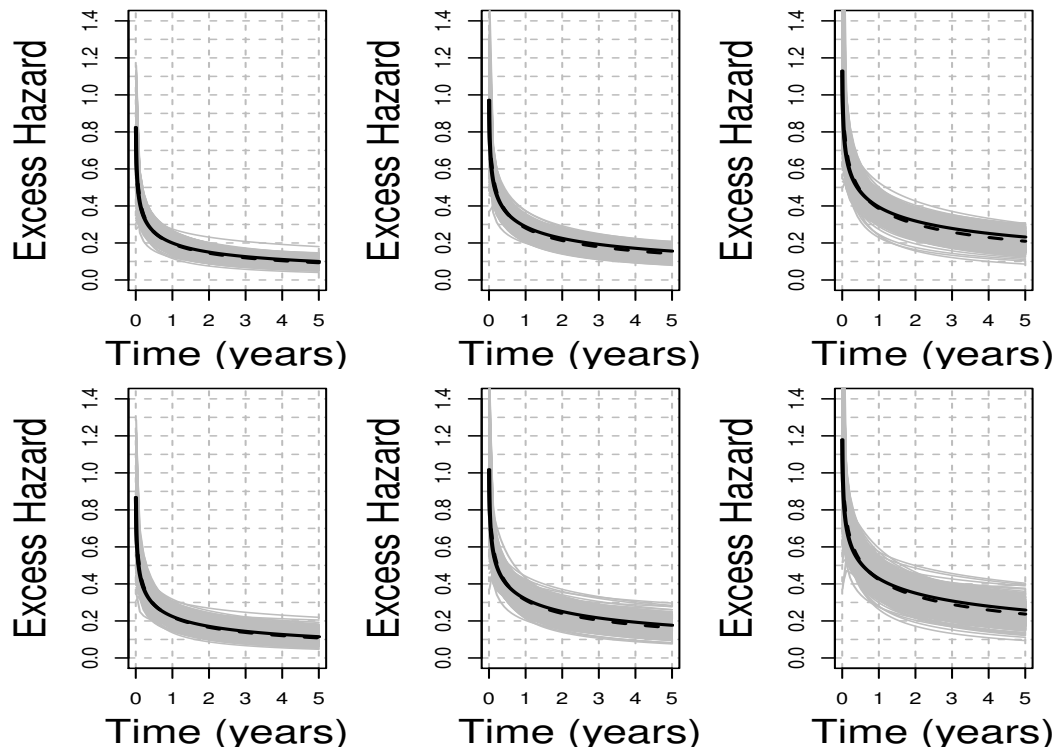
Model	Parameter	MMLE	mMLE	ESD	Mean Std Error	RMSE	Coverage
30% censoring							
GHEW	$\sigma$ (1.75)	1.743	1.737	0.211	0.167	0.211	0.984
	$\kappa$ (0.5)	0.499	0.499	0.025	0.070	0.025	0.984
	$\alpha$ (1.75)	1.767	1.763	0.130	0.101	0.131	0.985
	$\beta_{11}$ (-0.1)	-0.099	-0.099	0.013	0.017	0.013	0.978
	$\beta_{12}$ (-0.3)	-0.283	-0.276	0.288	0.335	0.288	0.981
	$\beta_{13}$ (-0.5)	-0.496	-0.487	0.310	0.339	0.310	0.963
	$\beta_{21}$ (0)	0.000	0.000	0.005	0.006	0.005	0.969
	$\beta_{22}$ (0)	0.006	0.008	0.088	0.103	0.089	0.974
$\beta_{23}$ (0)	0.003	0.006	0.094	0.104	0.094	0.962	
AHEW	$\sigma$ (1.75)	1.750	1.755	0.256	0.259	0.256	0.949
	$\kappa$ (0.5)	0.500	0.500	0.030	0.030	0.030	0.944
	$\alpha$ (1.75)	1.766	1.747	0.157	0.156	0.158	0.946
	$\beta_1$ (-0.1)	-0.100	-0.100	0.005	0.005	0.005	0.946
	$\beta_2$ (-0.3)	-0.302	-0.302	0.092	0.088	0.092	0.941
	$\beta_3$ (-0.5)	-0.505	-0.502	0.093	0.089	0.093	0.947
50% censoring							
GHEW	$\sigma$ (1.75)	1.728	1.734	0.226	0.196	0.227	0.987
	$\kappa$ (0.5)	0.498	0.499	0.028	0.088	0.028	0.988
	$\alpha$ (1.75)	1.775	1.761	0.149	0.123	0.151	0.990
	$\beta_{11}$ (-0.1)	-0.100	-0.101	0.016	0.020	0.015	0.981
	$\beta_{12}$ (-0.3)	-0.297	-0.291	0.361	0.413	0.361	0.979
	$\beta_{13}$ (-0.5)	-0.506	-0.496	0.355	0.417	0.355	0.982
	$\beta_{21}$ (0)	0.000	-0.000	0.005	0.006	0.005	0.976
	$\beta_{22}$ (0)	0.003	0.005	0.105	0.119	0.105	0.976
$\beta_{23}$ (0)	-0.001	0.002	0.102	0.120	0.102	0.984	
AHEW	$\sigma$ (1.75)	1.739	1.728	0.289	0.297	0.289	0.953
	$\kappa$ (0.5)	0.499	0.499	0.036	0.037	0.036	0.955
	$\alpha$ (1.75)	1.775	1.759	0.182	0.188	0.183	0.951
	$\beta_1$ (-0.1)	-0.100	-0.100	0.006	0.006	0.006	0.965
	$\beta_2$ (-0.3)	-0.305	-0.307	0.113	0.111	0.114	0.943
	$\beta_3$ (-0.5)	-0.500	-0.503	0.112	0.112	0.112	0.943

**Table S22.** Simulation from AH structure with  $(\sigma, \kappa, \alpha) = (1.75, 0.2, 1.75)$ ,  $\beta = (-0.1, -0.3, -0.5)$ , and  $n = 10000$ . Mean of the MLEs (MMLE), median of the MLEs (mMLE), empirical standard deviation (ESD), mean standard error, root-mean-square error (RMSE), and coverage proportions (Coverage).

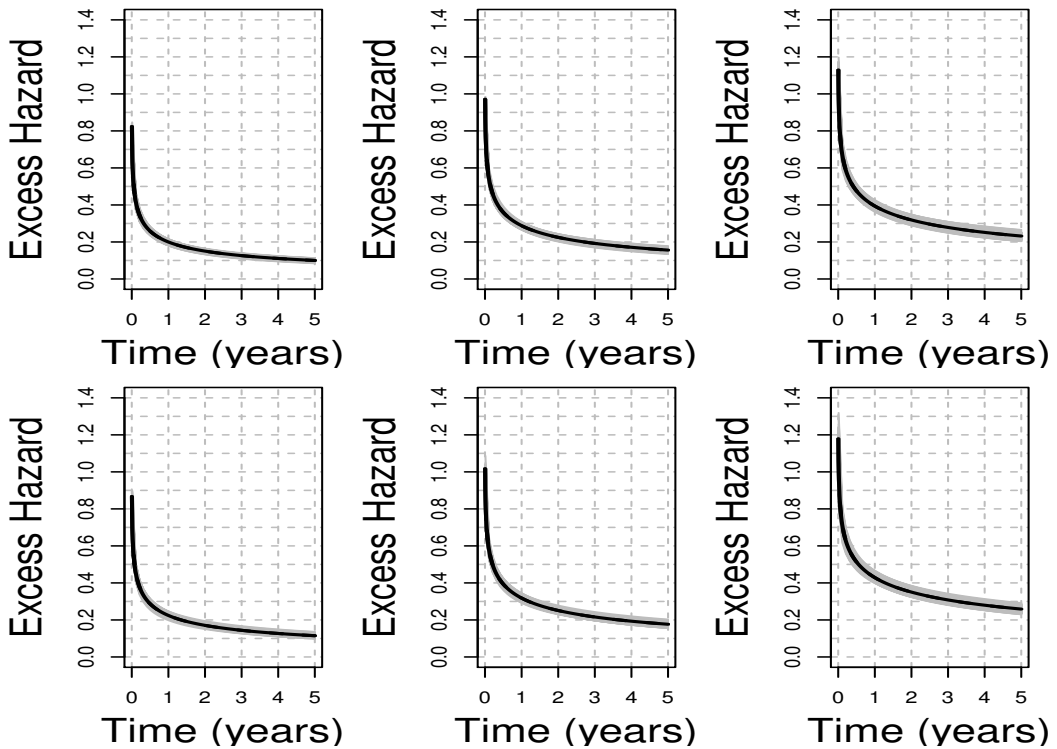
Model	30% censoring	50% censoring
$n = 1000$		
PHEW	13.1	16.6
AHEW	78.6	73.0
AFTEW	0.4	2.8
GHEW	7.9	7.6
$n = 5000$		
PHEW	0.1	1.3
AHEW	91.7	87.5
AFTEW	0	0
GHEW	8.2	11.2
$n = 10000$		
PHEW	0	0
AHEW	94.1	94.8
AFTEW	0	0
GHEW	5.9	5.2

**Table S23.** Simulation from AH structure with  $(\sigma, \kappa, \alpha) = (1.75, 0.2, 1.75)$ , and  $\beta = (-0.1, -0.3, -0.5)$ . Percentage of models selected with AIC.





**Figure S11.** Best hazards in terms of AIC for  $n = 1000$  and 50% censoring AH: true generating hazard (continuous line), mean of best fitted hazards (dashed lines), and individual fitted hazards (gray lines). Panels from left to right correspond to covariate values (age, sex, comorbidity) = (60, 0, 0), (70, 0, 0), (80, 0, 0), (60, 1, 0), (70, 1, 0), (80, 1, 0), respectively.



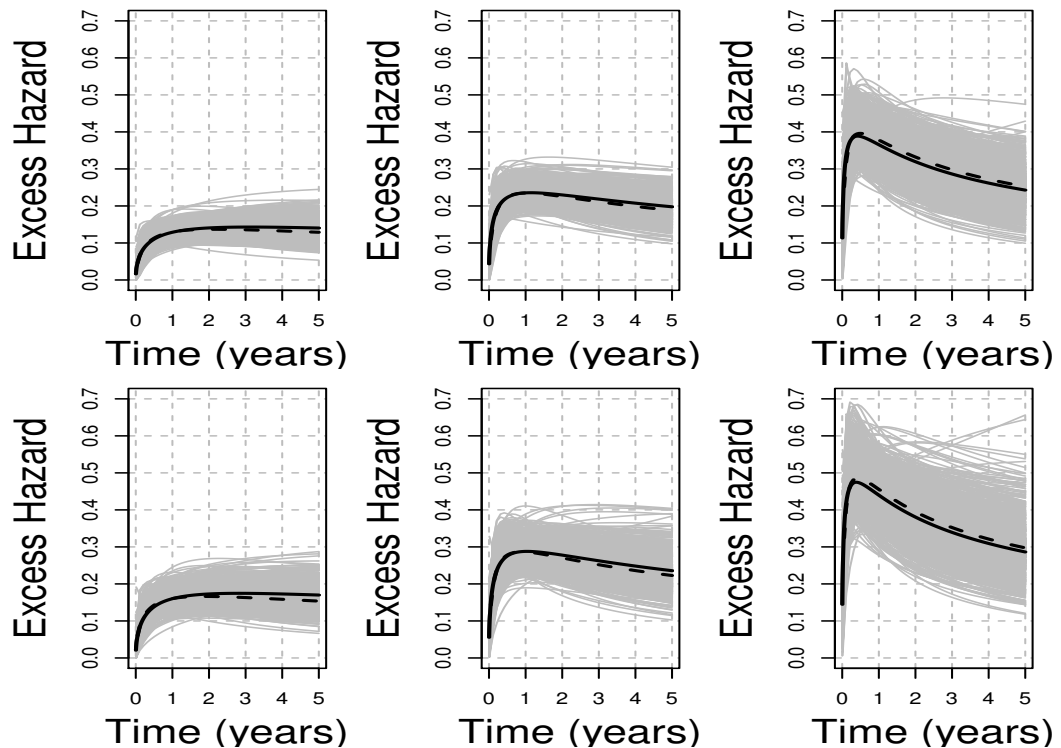
**Figure S12.** Best hazards in terms of AIC for  $n = 10000$  and 30% censoring AH: true generating hazard (continuous line), mean of best fitted hazards (dashed lines), and individual fitted hazards (gray lines). Panels from left to right correspond to covariate values (age, sex, comorbidity)=(60, 0, 0), (70, 0, 0), (80, 0, 0), (60, 1, 0), (70, 1, 0), (80, 1, 0), respectively.

Model	Parameter	MMLE	mMLE	ESD	Mean Std Error	RMSE	Coverage
30% censoring							
GHEW	$\sigma$ (1.75)	1.580	1.550	0.787	0.834	0.805	0.981
	$\kappa$ (0.6)	0.576	0.576	0.128	0.138	0.130	0.958
	$\alpha$ (2.5)	3.181	2.733	1.671	1.501	1.803	0.976
	$\beta_{11}$ (0.1)	0.099	0.098	0.033	0.033	0.033	0.941
	$\beta_{12}$ (0.1)	0.028	0.051	0.547	0.546	0.552	0.970
	$\beta_{13}$ (0.1)	0.032	0.054	0.720	0.679	0.723	0.969
	$\beta_{21}$ (0.05)	0.050	0.050	0.007	0.007	0.007	0.955
	$\beta_{22}$ (0.2)	0.202	0.204	0.093	0.098	0.093	0.959
$\beta_{23}$ (0.25)	0.251	0.252	0.094	0.103	0.094	0.961	
50% censoring							
GHEW	$\sigma$ (1.75)	1.560	1.527	0.841	0.928	0.862	0.975
	$\kappa$ (0.6)	0.573	0.580	0.143	0.163	0.146	0.956
	$\alpha$ (2.5)	3.467	2.687	2.621	2.094	2.793	0.980
	$\beta_{11}$ (0.1)	0.099	0.097	0.038	0.039	0.038	0.932
	$\beta_{12}$ (0.1)	0.044	0.082	0.809	0.719	0.810	0.977
	$\beta_{13}$ (0.1)	0.039	0.058	0.663	0.647	0.666	0.977
	$\beta_{21}$ (0.05)	0.050	0.050	0.008	0.008	0.008	0.963
	$\beta_{22}$ (0.2)	0.208	0.206	0.133	0.128	0.133	0.958
$\beta_{23}$ (0.25)	0.247	0.245	0.111	0.114	0.111	0.961	

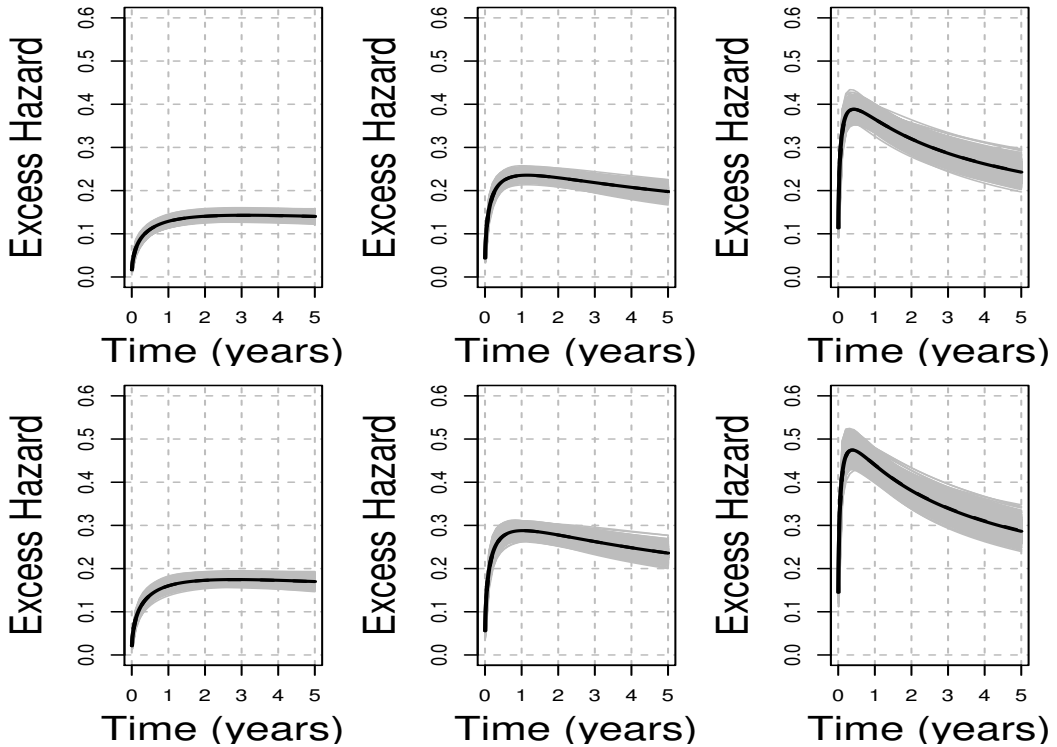
**Table S24.** Simulation from GH structure with  $(\sigma, \kappa, \alpha) = (1.75, 0.6, 2.5)$ ,  $\beta_1 = (0.1, 0.1, 0.1)$ ,  $\beta_2 = (0.05, 0.2, 0.25)$ , and  $n = 1000$ . Mean of the MLEs (MMLE), median of the MLEs (mMLE), empirical standard deviation (ESD), mean standard error, root-mean-square error (RMSE), and coverage proportions (Coverage).

Model	Parameter	MMLE	mMLE	ESD	Mean Std Error	RMSE	Coverage
30% censoring							
GHEW	$\sigma$ (1.75)	1.747	1.745	0.293	0.292	0.293	0.948
	$\kappa$ (0.6)	0.600	0.599	0.045	0.045	0.045	0.951
	$\alpha$ (2.5)	2.541	2.516	0.321	0.316	0.323	0.946
	$\beta_{11}$ (0.1)	0.100	0.100	0.010	0.009	0.010	0.947
	$\beta_{12}$ (0.1)	0.100	0.096	0.163	0.159	0.163	0.948
	$\beta_{13}$ (0.1)	0.100	0.109	0.163	0.160	0.163	0.946
	$\beta_{21}$ (0.05)	0.050	0.050	0.002	0.002	0.002	0.946
	$\beta_{22}$ (0.2)	0.201	0.201	0.029	0.029	0.029	0.949
$\beta_{23}$ (0.25)	0.252	0.250	0.029	0.029	0.029	0.952	
50% censoring							
GHEW	$\sigma$ (1.75)	1.753	1.746	0.334	0.332	0.334	0.946
	$\kappa$ (0.6)	0.602	0.600	0.054	0.054	0.054	0.951
	$\alpha$ (2.5)	2.540	2.509	0.366	0.364	0.368	0.953
	$\beta_{11}$ (0.1)	0.101	0.100	0.011	0.011	0.011	0.954
	$\beta_{12}$ (0.1)	0.101	0.107	0.185	0.186	0.185	0.950
	$\beta_{13}$ (0.1)	0.106	0.096	0.190	0.187	0.190	0.953
	$\beta_{21}$ (0.05)	0.050	0.050	0.003	0.003	0.003	0.954
	$\beta_{22}$ (0.2)	0.200	0.200	0.034	0.033	0.034	0.940
$\beta_{23}$ (0.25)	0.250	0.250	0.034	0.033	0.033	0.948	

**Table S25.** Simulation from GH structure with  $(\sigma, \kappa, \alpha) = (1.75, 0.6, 2.5)$ ,  $\beta_1 = (0.1, 0.1, 0.1)$ ,  $\beta_2 = (0.05, 0.2, 0.25)$ , and  $n = 10000$ . Mean of the MLEs (MMLE), median of the MLEs (mMLE), empirical standard deviation (ESD), mean standard error, root-mean-square error (RMSE), and coverage proportions (Coverage).



**Figure S13.** Best hazards in terms of AIC for  $n = 1000$  and 50% censoring GH: true generating hazard (continuous line), mean of best fitted hazards (dashed lines), and individual fitted hazards (gray lines). Panels from left to right correspond to covariate values (age, sex, comorbidity) = (60, 0, 0), (70, 0, 0), (80, 0, 0), (60, 1, 0), (70, 1, 0), (80, 1, 0), respectively.



**Figure S14.** Best hazards in terms of AIC for  $n = 10000$  and 30% censoring GH: true generating hazard (continuous line), mean of best fitted hazards (dashed lines), and individual fitted hazards (gray lines). Panels from left to right correspond to covariate values (age, sex, comorbidity) = (60, 0, 0), (70, 0, 0), (80, 0, 0), (60, 1, 0), (70, 1, 0), (80, 1, 0), respectively.

Model	Parameter	MMLE	mMLE	ESD	Mean Std Error	RMSE	Coverage	
30% censoring								
GHEW	$\sigma$ (0.113)	0.239	0.121	0.303	0.278	0.328	0.914	
	$\kappa$ (0.366)	0.387	0.368	0.117	0.131	0.119	0.957	
	$\alpha$ (18.305)	29.053	17.832	28.879	49.726	30.800	0.941	
	$\beta_{11}$ (0)	-0.000	0.000	0.012	0.011	0.012	0.945	
	$\beta_{12}$ (0)	-0.009	-0.008	0.230	0.223	0.230	0.946	
	$\beta_{13}$ (0.5)	0.518	0.495	0.255	0.260	0.256	0.959	
	$\beta_{21}$ (0.05)	0.050	0.050	0.005	0.005	0.005	0.949	
	$\beta_{22}$ (0.3)	0.302	0.302	0.099	0.104	0.099	0.958	
	$\beta_{23}$ (-0.6)	-0.610	-0.608	0.109	0.113	0.110	0.957	
CHEW	$\sigma$ (0.113)	0.249	0.134	0.304	0.281	0.333	0.905	
	$\kappa$ (0.366)	0.392	0.376	0.115	0.126	0.118	0.958	
	$\alpha$ (18.305)	28.059	16.906	30.104	44.256	31.630	0.933	
	$\beta_{11}$ (0.5)	0.517	0.501	0.249	0.258	0.250	0.964	
	$\beta_{21}$ (0.05)	0.050	0.050	0.005	0.005	0.005	0.940	
	$\beta_{22}$ (0.3)	0.300	0.300	0.093	0.096	0.093	0.950	
		$\beta_{23}$ (-0.6)	-0.610	-0.608	0.109	0.112	0.110	0.958
	50% censoring							
GHEW	$\sigma$ (0.113)	0.254	0.117	0.349	0.317	0.376	0.923	
	$\kappa$ (0.366)	0.391	0.368	0.130	0.155	0.132	0.967	
	$\alpha$ (18.305)	30.104	18.644	29.471	65.195	31.732	0.947	
	$\beta_{11}$ (0)	0.000	0.001	0.013	0.012	0.013	0.940	
	$\beta_{12}$ (0)	0.001	0.011	0.250	0.249	0.250	0.957	
	$\beta_{13}$ (0.5)	0.525	0.512	0.287	0.291	0.288	0.965	
	$\beta_{21}$ (0.05)	0.050	0.050	0.006	0.006	0.006	0.951	
	$\beta_{22}$ (0.3)	0.298	0.295	0.125	0.126	0.125	0.955	
	$\beta_{23}$ (-0.6)	-0.612	-0.605	0.131	0.136	0.132	0.965	
CHEW	$\sigma$ (0.113)	0.261	0.134	0.335	0.323	0.366	0.922	
	$\kappa$ (0.366)	0.395	0.377	0.123	0.147	0.126	0.976	
	$\alpha$ (18.305)	28.300	16.939	30.330	54.544	31.920	0.943	
	$\beta_{11}$ (0.5)	0.525	0.516	0.272	0.286	0.273	0.974	
	$\beta_{21}$ (0.05)	0.050	0.050	0.005	0.005	0.005	0.937	
	$\beta_{22}$ (0.3)	0.297	0.294	0.114	0.111	0.114	0.946	
		$\beta_{23}$ (-0.6)	-0.611	-0.606	0.129	0.134	0.130	0.966

**Table S26.** Simulation from CH structure with  $(\sigma, \kappa, \alpha) = (0.113, 0.366, 18.305)$ ,  $\beta_{13} = 0.5$ ,  $\beta_2 = (0.05, 0.3, -0.6)$ , and  $n = 1000$ . Mean of the MLEs (MMLE), median of the MLEs (mMLE), empirical standard deviation (ESD), mean standard error, root-mean-square error (RMSE), and coverage proportions (Coverage).

Model	Parameter	MMLE	mMLE	ESD	Mean Std Error	RMSE	Coverage
30% censoring							
GHEW	$\sigma$ (0.113)	0.138	0.114	0.111	0.101	0.114	0.948
	$\kappa$ (0.366)	0.368	0.366	0.055	0.055	0.055	0.960
	$\alpha$ (18.305)	21.509	18.537	11.053	10.913	11.502	0.958
	$\beta_{11}$ (0)	-0.000	-0.000	0.005	0.005	0.005	0.944
	$\beta_{12}$ (0)	0.005	0.004	0.092	0.094	0.092	0.956
	$\beta_{13}$ (0.5)	0.506	0.508	0.109	0.108	0.109	0.948
	$\beta_{21}$ (0.05)	0.050	0.050	0.002	0.002	0.002	0.951
	$\beta_{22}$ (0.3)	0.301	0.300	0.046	0.045	0.046	0.955
$\beta_{23}$ (-0.6)	-0.600	-0.600	0.049	0.048	0.049	0.952	
CHEW	$\sigma$ (0.113)	0.139	0.112	0.110	0.100	0.112	0.943
	$\kappa$ (0.366)	0.368	0.366	0.054	0.054	0.054	0.952
	$\alpha$ (18.305)	21.353	18.546	11.091	10.612	11.497	0.958
	$\beta_{11}$ (0.5)	0.506	0.506	0.109	0.108	0.109	0.952
	$\beta_{21}$ (0.05)	0.050	0.050	0.002	0.002	0.002	0.947
	$\beta_{22}$ (0.3)	0.301	0.300	0.044	0.043	0.044	0.956
	$\beta_{23}$ (-0.6)	-0.600	-0.600	0.049	0.048	0.049	0.950
	50% censoring						
GHEW	$\sigma$ (0.113)	0.143	0.105	0.125	0.116	0.128	0.936
	$\kappa$ (0.366)	0.367	0.363	0.063	0.064	0.063	0.964
	$\alpha$ (18.305)	22.816	19.184	13.824	13.818	14.535	0.950
	$\beta_{11}$ (0)	-0.000	-0.000	0.005	0.005	0.005	0.941
	$\beta_{12}$ (0)	0.000	-0.000	0.103	0.105	0.103	0.955
	$\beta_{13}$ (0.5)	0.497	0.488	0.117	0.119	0.117	0.961
	$\beta_{21}$ (0.05)	0.050	0.050	0.003	0.002	0.003	0.945
	$\beta_{22}$ (0.3)	0.299	0.300	0.052	0.054	0.052	0.960
$\beta_{23}$ (-0.6)	-0.600	-0.597	0.058	0.057	0.058	0.953	
CHEW	$\sigma$ (0.113)	0.145	0.111	0.124	0.115	0.127	0.934
	$\kappa$ (0.366)	0.368	0.365	0.061	0.062	0.061	0.961
	$\alpha$ (18.305)	22.503	18.793	14.161	13.463	14.763	0.945
	$\beta_{11}$ (0.5)	0.497	0.490	0.117	0.119	0.117	0.958
	$\beta_{21}$ (0.05)	0.050	0.050	0.002	0.002	0.002	0.935
	$\beta_{22}$ (0.3)	0.299	0.300	0.047	0.050	0.047	0.958
	$\beta_{23}$ (-0.6)	-0.600	-0.598	0.058	0.057	0.058	0.952

**Table S27.** Simulation from CH structure with  $(\sigma, \kappa, \alpha) = (0.113, 0.366, 18.305)$ ,  $\beta_{13} = 0.5$ ,  $\beta_2 = (0.05, 0.3, -0.6)$ , and  $n = 5000$ . Mean of the MLEs (MMLE), median of the MLEs (mMLE), empirical standard deviation (ESD), mean standard error, root-mean-square error (RMSE), and coverage proportions (Coverage).

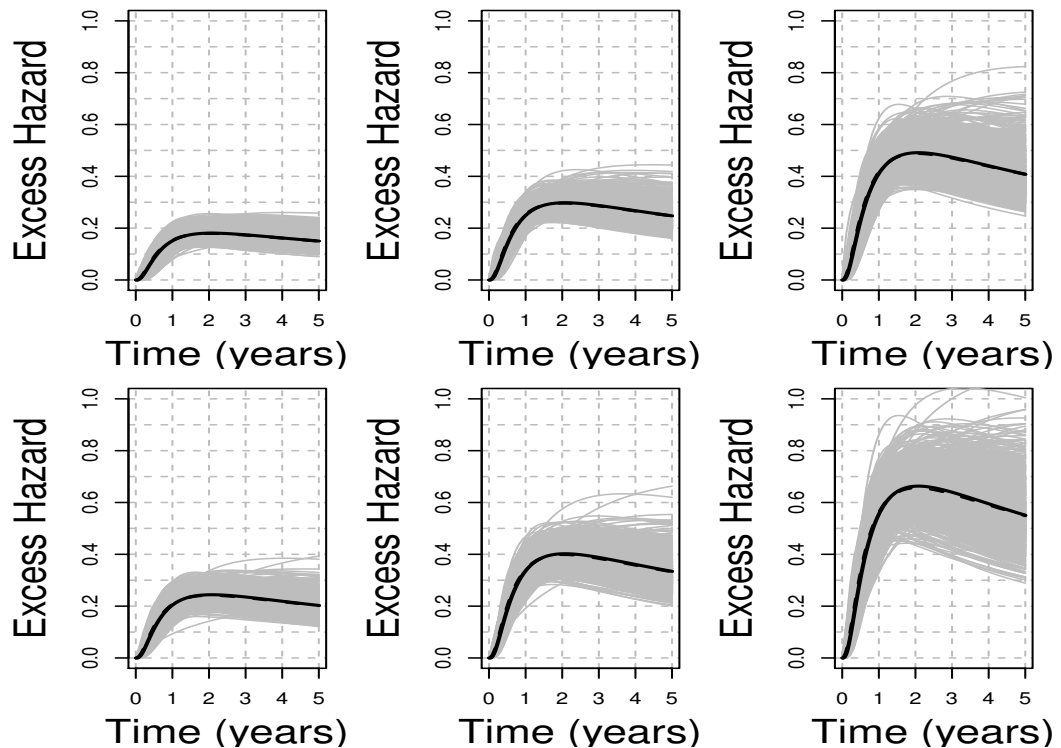


Model	Parameter	MMLE	mMLE	ESD	Mean Std Error	RMSE	Coverage
30% censoring							
GHEW	$\sigma$ (0.113)	0.133	0.120	0.078	0.072	0.080	0.943
	$\kappa$ (0.366)	0.370	0.370	0.040	0.039	0.040	0.949
	$\alpha$ (18.305)	19.414	17.848	6.931	6.484	7.016	0.945
	$\beta_{11}$ (0)	-0.000	0.000	0.003	0.003	0.003	0.959
	$\beta_{12}$ (0)	0.001	0.002	0.070	0.067	0.070	0.939
	$\beta_{13}$ (0.5)	0.504	0.501	0.076	0.076	0.076	0.953
	$\beta_{21}$ (0.05)	0.050	0.050	0.001	0.001	0.001	0.945
	$\beta_{22}$ (0.3)	0.299	0.300	0.032	0.032	0.032	0.952
	$\beta_{23}$ (-0.6)	-0.601	-0.602	0.036	0.034	0.036	0.944
CHEW	$\sigma$ (0.113)	0.133	0.121	0.077	0.071	0.079	0.941
	$\kappa$ (0.366)	0.370	0.370	0.039	0.038	0.039	0.949
	$\alpha$ (18.305)	19.255	17.706	6.902	6.311	6.963	0.947
	$\beta_{11}$ (0.5)	0.505	0.503	0.076	0.076	0.076	0.957
	$\beta_{21}$ (0.05)	0.050	0.050	0.001	0.001	0.001	0.949
	$\beta_{22}$ (0.3)	0.300	0.301	0.031	0.030	0.031	0.949
	$\beta_{23}$ (-0.6)	-0.601	-0.602	0.036	0.034	0.036	0.942
50% censoring							
GHEW	$\sigma$ (0.113)	0.133	0.115	0.085	0.081	0.087	0.940
	$\kappa$ (0.366)	0.369	0.367	0.045	0.045	0.045	0.952
	$\alpha$ (18.305)	19.961	18.307	7.911	7.618	8.078	0.953
	$\beta_{11}$ (0)	0.000	0.000	0.004	0.004	0.004	0.942
	$\beta_{12}$ (0)	-0.001	0.000	0.072	0.074	0.072	0.954
	$\beta_{13}$ (0.5)	0.502	0.499	0.082	0.084	0.082	0.950
	$\beta_{21}$ (0.05)	0.050	0.050	0.002	0.002	0.002	0.952
	$\beta_{22}$ (0.3)	0.300	0.301	0.037	0.038	0.037	0.954
	$\beta_{23}$ (-0.6)	-0.602	-0.601	0.040	0.040	0.040	0.951
CHEW	$\sigma$ (0.113)	0.134	0.115	0.083	0.080	0.085	0.950
	$\kappa$ (0.366)	0.370	0.367	0.043	0.044	0.043	0.957
	$\alpha$ (18.305)	19.714	18.163	7.620	7.401	7.745	0.956
	$\beta_{11}$ (0.5)	0.502	0.499	0.082	0.084	0.082	0.951
	$\beta_{21}$ (0.05)	0.050	0.050	0.002	0.002	0.002	0.950
	$\beta_{22}$ (0.3)	0.300	0.301	0.036	0.035	0.035	0.947
	$\beta_{23}$ (-0.6)	-0.602	-0.602	0.040	0.040	0.040	0.953

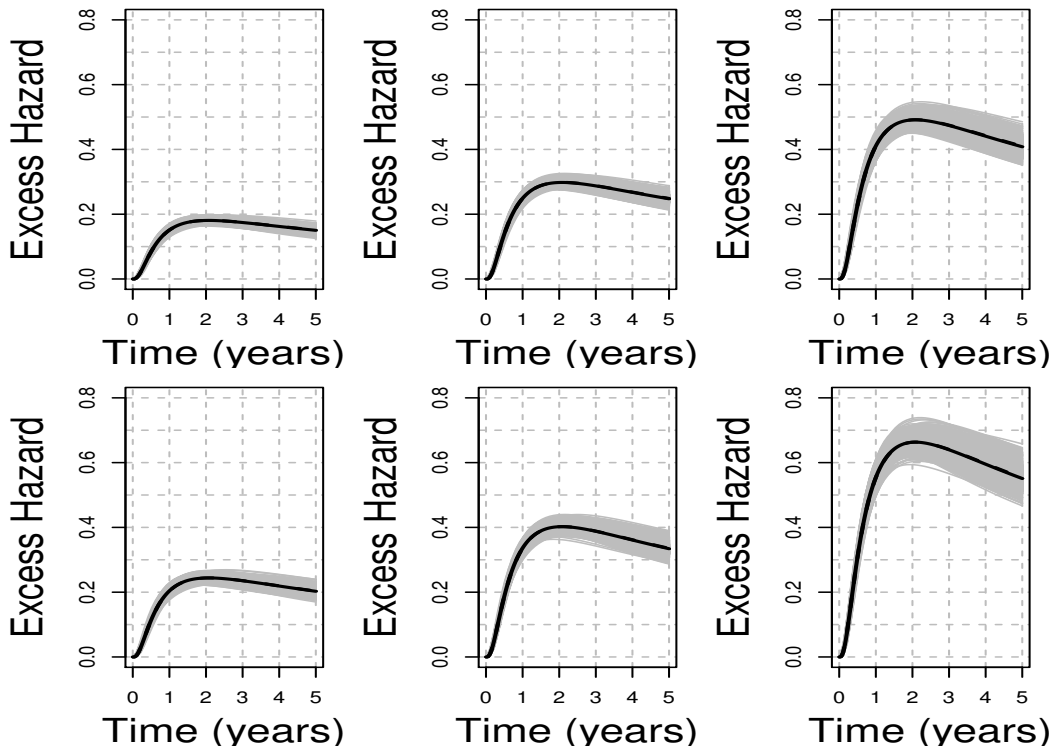
**Table S28.** Simulation from CH structure with  $(\sigma, \kappa, \alpha) = (0.113, 0.366, 18.305)$ ,  $\beta_{13} = 0.5$ ,  $\beta_2 = (0.05, 0.3, -0.6)$ , and  $n = 10000$ . Mean of the MLEs (MMLE), median of the MLEs (mMLE), empirical standard deviation (ESD), mean standard error, root-mean-square error (RMSE), and coverage proportions (Coverage).

Model	30% censoring	50% censoring
$n = 1000$		
PHEW	19.8	25.4
AHEW	0	0
AFTEW	0.1	0.9
GHEW	13.4	11.9
CHEW	66.7	61.8
$n = 5000$		
PHEW	0	0.1
AHEW	0	0
AFTEW	0	0
GHEW	13.4	14.4
CHEW	86.6	85.5
$n = 10000$		
PHEW	0	0
AHEW	0	0
AFTEW	0	0
GHEW	14.3	13.1
CHEW	85.7	86.9

**Table S29.** Simulation from CH structure with  $(\sigma, \kappa, \alpha) = (0.113, 0.366, 18.305)$ ,  $\beta_{13} = 0.5$ ,  $\beta_2 = (0.05, 0.3, -0.6)$ . Percentage of models selected with AIC.



**Figure S15.** Best hazards in terms of AIC for  $n = 1000$  and 50% censoring CH: true generating hazard (continuous line), mean of best fitted hazards (dashed lines), and individual fitted hazards (gray lines). Panels from left to right correspond to covariate values (age, sex, comorbidity) = (60, 0, 0), (70, 0, 0), (80, 0, 0), (60, 1, 0), (70, 1, 0), (80, 1, 0), respectively.



**Figure S16.** Best hazards in terms of AIC for  $n = 10000$  and 30% censoring CH: true generating hazard (continuous line), mean of best fitted hazards (dashed lines), and individual fitted hazards (gray lines). Panels from left to right correspond to covariate values (age, sex, comorbidity) = (60, 0, 0), (70, 0, 0), (80, 0, 0), (60, 1, 0), (70, 1, 0), (80, 1, 0), respectively.

## 5 Analysis of initial points

### 5.1 Performance of initial points

We now analyse the impact of using different initial points in the optimisation of the likelihood of the GH model, which corresponds to the most challenging structure. We compare the R commands ‘nlminb’ and ‘optim’ for the maximisation of the likelihood function, both with a maximum limit of 10,000 iterations. The command ‘optim’ implements, by default, the Nelder-Mead algorithm, while ‘nlminb’ implements a quasi-Newton algorithm. For this purpose, we simulated  $N = 1000$  data sets of size  $n = 1000$  from the GH model with parameters  $(\sigma, \kappa, \alpha) = (1.75, 0.6, 2.5)$ ,  $\beta_1 = (0.1, 0.1, 0.1)$  and  $\beta_2 = (0.05, 0.2, 0.25)$ , with administrative censoring at  $T_C = 5$  years, which induced approximately 30% censoring (simulation scenario (v)). We then checked the the performance of:

- Methods 1a and 1b. Initialising the optimisation algorithms (‘nlminb’ and ‘optim’) at  $(\sigma, \kappa, \alpha) = (1, 1, 1)$ ,  $\beta_1 = (0, \dots, 0)$ , and  $\beta_2 = (0, \dots, 0)$ . Method 1a refers to the use of ‘nlminb’, while Method 1b refers to the use of ‘optim’.
- Methods 2a and 2b. Initialising the optimisation algorithms at  $(\sigma, \kappa, \alpha) = (\widehat{\sigma}, \widehat{\kappa}, 1)$ ,  $\beta_1 = (0, \dots, 0)$ , and  $\beta_2 = \widehat{\beta}_2$ , where the notation “ $\widehat{\cdot}$ ” indicates the MLE obtained with the model with Weibull excess baseline hazard. That is, we are initialising at the estimated PH model with Weibull baseline hazard (an already known model). Method 2a refers to the use of ‘nlminb’, while Method 2b refers to the use of ‘optim’. The MLE for the Weibull excess hazard model is, in turn, obtained using the initial points  $(\sigma, \kappa) = (1, 1)$ ,  $\beta_1 = (0, \dots, 0)$ , and  $\beta_2 = (0, \dots, 0)$ , and the best outcome (highest maximum) from either ‘optim’ or ‘nlminb’.

Table S30 shows the percentages when each of the methods is preferred. This study shows that initialising the optimisation algorithm as in Methods 1a and 1b tends to provide the best results. Although Methods 2a and 2b were not superior in general, in some cases they provided a better performance than Methods 1a and 1b. For the best methods, the convergence codes were 0 (successful convergence) in 89.5% of the cases, 1 (iteration limit *maxit* had been reached) in 8.4% of the cases, and 10 (degeneracy of the Nelder-Mead simplex) in 2.1% of the cases. Moreover, Table S31 shows that Methods 1a and 1b converged (Code 0) to a solution in 92.3% of the cases, 6.7% of the cases required more than 10,000 iterations to converge (Code 1), and only 1% of the cases produced degeneracy of the Nelder-Mead simplex algorithm (Code 10), while Methods 2a and 2b tended to required more iterations to converge and produced Code 10 in a higher proportion. **In practice, we recommend checking different optimisation methods and different initial points, while allowing each method for a large enough number of iterations; and this study provides some guidelines to achieve an accurate optimisation.** In particular, one could also play with different initial values for  $\alpha$ , say  $\alpha = 0.5, 1, 2$ , as we illustrate in the next section. Alternatively (to Methods 2a and 2b) or in addition, the GH model can be initialised at the AFT model with Weibull baseline hazard.

Method 1a	Method 1b	Method 2a	Method 2b
57.5	27.1	5.5	9.9

**Table S30.** Percentages of preferred methods.

Convergence Code	0	1	10
Methods 1a	89.7	10.3	0
Methods 1b	99.3	0	0.7
Methods 2a	54.5	45.6	0
Methods 2b	80.8	0	19.2

**Table S31.** Percentages of convergence codes for Methods 1a and 1b compared to Methods 2a and 2b.

## 5.2 Re-analysis of the GH model with different initial points

We now re-analyse the GH simulation scenario described in Table S1, with  $n = 1000$  and censoring rates 30% and 50%, in order to assess the performance of using different “automatic” initial points instead of the true values of the parameters employed in the simulation study. We consider the following three initial points:

1.  $\sigma = \kappa = \alpha = 1, \beta_1 = \beta_2 = 0.$
2.  $\sigma = \kappa = 1, \alpha = 2, \beta_1 = \beta_2 = 0.$
3.  $\sigma = \kappa = 1, \alpha = 0.5, \beta_1 = \beta_2 = 0.$

coupled with the R command ‘nlminb()’. Thus, we initialise the optimiser at these three initial points, allowing for a maximum of 10,000 iterations, and we select the estimates associated to the maximum value of the likelihood function. We emphasise that, in the simulation study, we have used the true values of the parameters as initial points as a way of making the estimation process more efficient, since we were only interested in the properties of the maximum likelihood estimators rather than those of the optimisation process. Table S32 shows the properties of the MLEs obtained using these three initial points. We can notice that the numbers in this table are virtually the same as those in Table S24, as expected. The model selection also has virtually the same performance using these three initial values compared to that obtained using the true values of the parameters as initial values. Figures S17 and S18 show the shapes of the best (in terms of AIC) fitted baseline excess hazards using the proposed initial points. These figures closely resemble the results obtained in Figure S13, which is reassuring. This study shows that the use of the three proposed “automatic” initial points has a very good performance in terms of producing accurate estimators, and therefore we recommend their use in practice as an initial guess.

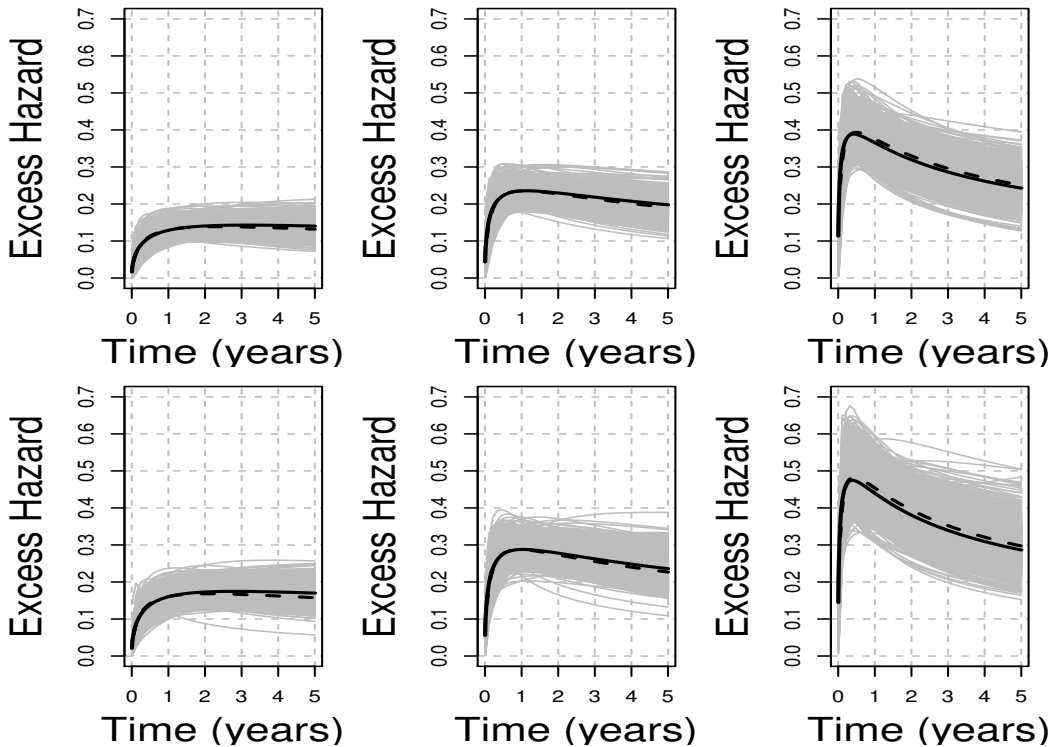
## References

Model	Parameter	MMLE	mMLE	ESD	Mean Std Error	RMSE	Coverage
30% censoring							
GHEW	$\sigma$ (1.75)	1.581	1.549	0.790	0.834	0.808	0.980
	$\kappa$ (0.6)	0.576	0.575	0.128	0.138	0.131	0.957
	$\alpha$ (2.5)	3.181	2.733	1.672	1.500	1.804	0.975
	$\beta_{11}$ (0.1)	0.099	0.098	0.033	0.033	0.033	0.941
	$\beta_{12}$ (0.1)	0.010	0.048	0.730	0.775	0.735	0.970
	$\beta_{13}$ (0.1)	0.028	0.054	0.738	0.682	0.741	0.970
	$\beta_{21}$ (0.05)	0.050	0.050	0.007	0.007	0.007	0.957
	$\beta_{22}$ (0.2)	0.204	0.205	0.106	0.129	0.106	0.959
$\beta_{23}$ (0.25)	0.251	0.251	0.095	0.105	0.095	0.961	
50% censoring							
GHEW	$\sigma$ (1.75)	1.559	1.521	0.851	0.919	0.872	0.970
	$\kappa$ (0.6)	0.572	0.578	0.145	0.162	0.147	0.954
	$\alpha$ (2.5)	3.503	2.712	2.743	2.202	2.919	0.975
	$\beta_{11}$ (0.1)	0.098	0.096	0.039	0.039	0.039	0.927
	$\beta_{12}$ (0.1)	0.027	0.078	0.898	0.835	0.901	0.978
	$\beta_{13}$ (0.1)	0.025	0.060	0.847	1.388	0.850	0.976
	$\beta_{21}$ (0.05)	0.050	0.050	0.008	0.008	0.008	0.963
	$\beta_{22}$ (0.2)	0.210	0.206	0.147	0.149	0.147	0.959
$\beta_{23}$ (0.25)	0.247	0.244	0.117	0.176	0.117	0.962	

**Table S32.** Simulation from GH structure with  $(\sigma, \kappa, \alpha) = (1.75, 0.6, 2.5)$ ,  $\beta_1 = (0.1, 0.1, 0.1)$ ,  $\beta_2 = (0.05, 0.2, 0.25)$ , and  $n = 1000$ . Mean of the MLEs (MMLE), median of the MLEs (mMLE), empirical standard deviation (ESD), mean standard error, root-mean-square error (RMSE), and coverage proportions (Coverage). The estimates are obtained using 6 different automatic initial points.

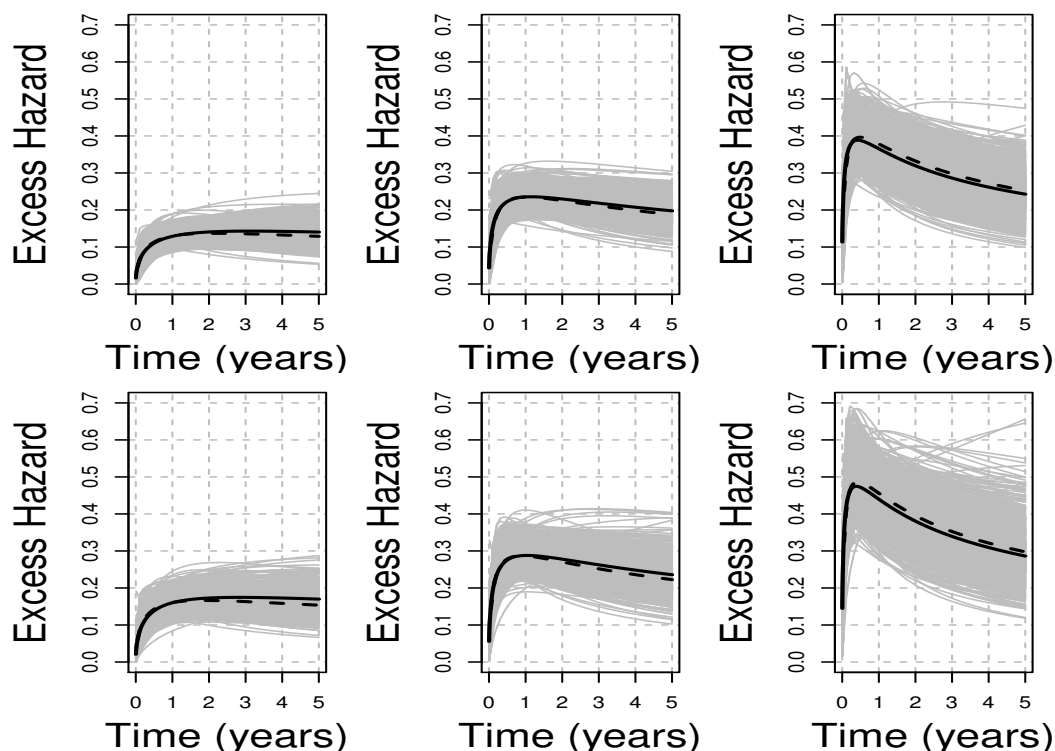
Model	30% censoring	50% censoring
$n = 1000$		
PHEW	1.6	2.4
AHEW	0	0
AFTEW	56.5	65.1
GHEW	41.9	32.5

**Table S33.** Simulation from GH structure with  $(\sigma, \kappa, \alpha) = (1.75, 0.6, 2.5)$ ,  $\beta_1 = (0.1, 0.1, 0.1)$ , and  $\beta_2 = (0.05, 0.2, 0.25)$ . Percentage of models selected with AIC.



**Figure S17.** Best hazards in terms of AIC for  $n = 1000$  and 30% censoring GH: true generating hazard (continuous line), mean of best fitted hazards (dashed lines), and individual fitted hazards (gray lines). Panels from left to right correspond to covariate values (age, sex, comorbidity) = (60, 0, 0), (70, 0, 0), (80, 0, 0), (60, 1, 0), (70, 1, 0), (80, 1, 0), respectively. The estimate values are obtained using different automatic initial points.





**Figure S18.** Best hazards in terms of AIC for  $n = 1000$  and 50% censoring GH: true generating hazard (continuous line), mean of best fitted hazards (dashed lines), and individual fitted hazards (gray lines). Panels from left to right correspond to covariate values (age, sex, comorbidity) = (60, 0, 0), (70, 0, 0), (80, 0, 0), (60, 1, 0), (70, 1, 0), (80, 1, 0), respectively. The estimate values are obtained using different automatic initial points.