DAV/DGVFM
Jahrestagung
2026

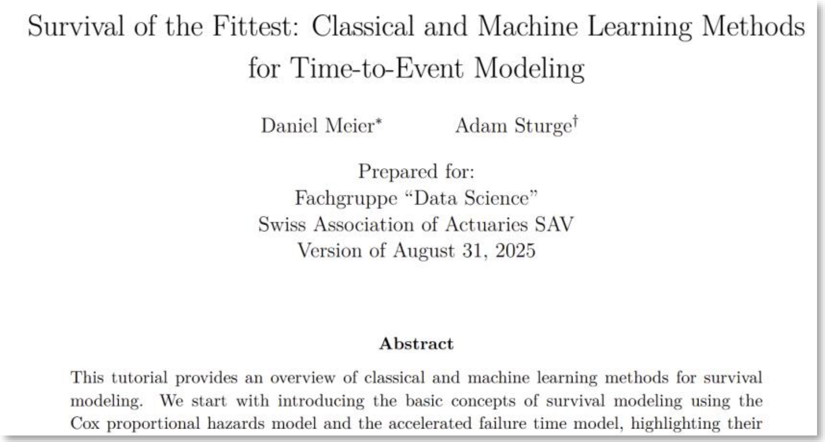
Daniel Meier, Swiss Re

Klassische und moderne Survival Modelle

Anwendung der Modelle zum Einfluss von GLP-1 Abnehmspritzen auf Lebensversicherer

30. April 2026

Two articles



Survival of the Fittest: Classical and Machine Learning Methods for Time-to-Event Modeling

Daniel Meier* Adam Sturge†

Prepared for:
Fachgruppe "Data Science"
Swiss Association of Actuaries SAV
Version of August 31, 2025

Abstract

This tutorial provides an overview of classical and machine learning methods for survival modeling. We start with introducing the basic concepts of survival modeling using the Cox proportional hazards model and the accelerated failure time model, highlighting their

Swiss Re Institute

September 2025

The future of metabolic health and weight loss drugs

Projecting mortality reductions in the US and UK populations

02 Executive summary
03 Metabolic ill-health and mortality
07 Metabolic ill-health trends: the US and UK
08 GLP-1 drugs and mortality reductions: our modelling approach
11 The US and UK 2045 mortality reduction projections
14 Risks associated with GLP-1 drugs
16 Implications for L&H insurance
19 Swiss Re's vision of metabolic health

Case study 16 on survival models available at actuarialdatascience.org

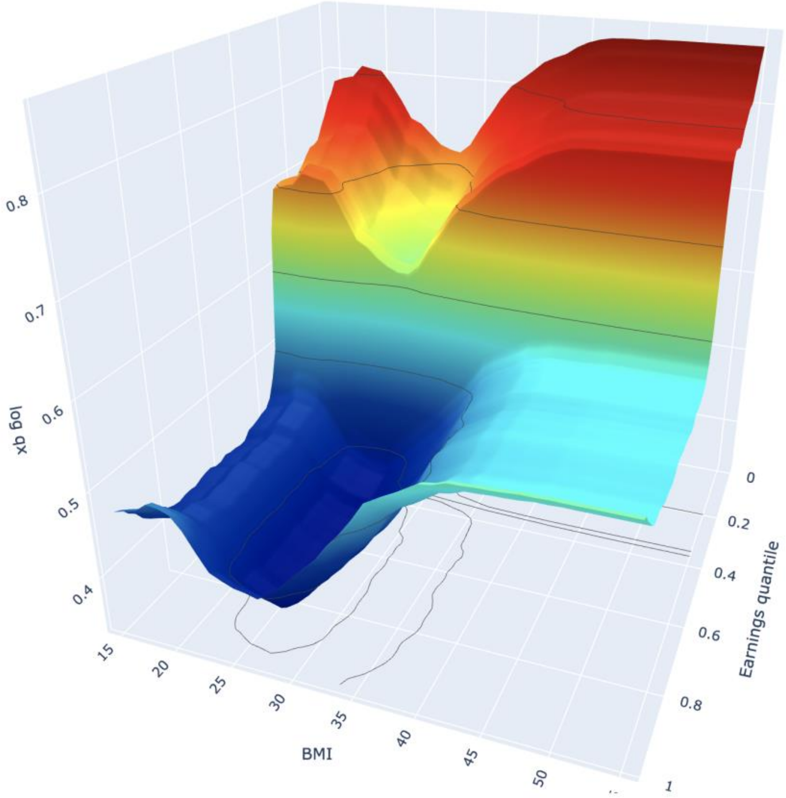
What is the impact of GLP-1 weight loss drugs on age-standardized mortality in the US in 20 years?

DAV
Deutsche Aktuarvereinigung e.V.

DGVFM
Deutsche Gesellschaft für Versicherungs- und Finanzmathematik e.V.

Where are survival models applied?

- Life & Health Underwriting
- Scenario testing, e.g., weight loss drugs
- US cancer registry SEER: Underwriting
- Pensioner mortality tables
- Unemployment times
- Public health
- Any other use case where **time-to-event** is important, e.g., credit default, lapse, engineering, etc.



Linear regression

$$y(\boldsymbol{x}) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_M x_M$$

```
                            OLS Regression Results
==============================================================================
Dep. Variable:                      y   R-squared:                       0.738
Model:                            OLS   Adj. R-squared:                  0.734
Method:                 Least Squares   F-statistic:                     184.4
Date:                Fri, 15 Aug 2025   Prob (F-statistic):           8.17e-57
Time:                        14:50:57   Log-Likelihood:                -277.39
No. Observations:                 200   AIC:                             562.8
Df Residuals:                     196   BIC:                             576.0
Df Model:                           3
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
const          2.1888      0.293      7.458      0.000       1.610       2.768
x1             0.4899      0.034     14.387      0.000       0.423       0.557
x2            -0.3280      0.025    -13.149      0.000      -0.377      -0.279
x3             1.1735      0.068     17.172      0.000       1.039       1.308
==============================================================================
Omnibus:                        0.265   Durbin-Watson:                   2.082
Prob(Omnibus):                  0.876   Jarque-Bera (JB):                0.402
Skew:                           0.064   Prob(JB):                        0.818
Kurtosis:                       2.822   Cond. No.                         48.1
==============================================================================
```

Logistic regression

$$p(\boldsymbol{x}) = \text{logistic}(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_M x_M)$$

$$\text{logistic}(x) = \frac{1}{1 + \exp(-x)}$$

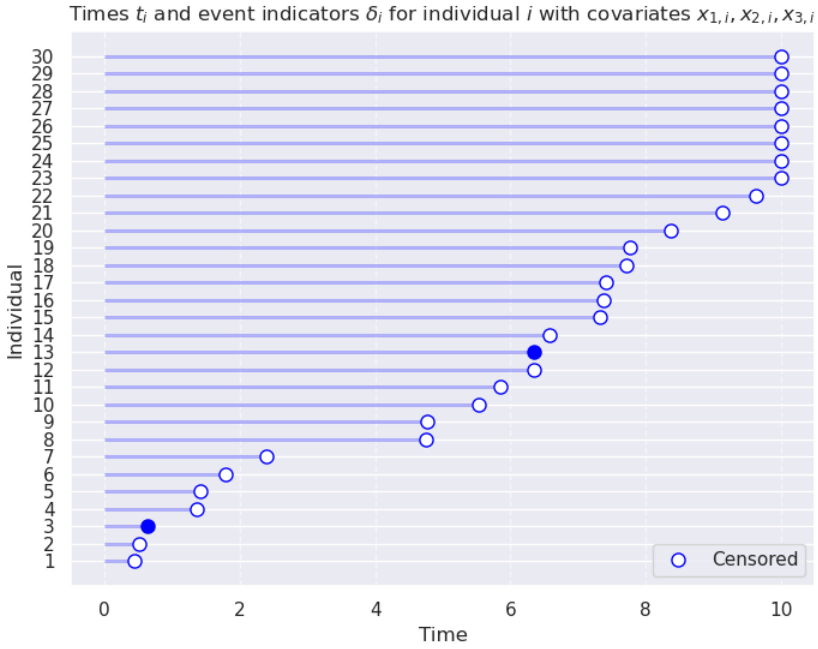
```
                           Logit Regression Results
==============================================================================
Dep. Variable:                      y   No. Observations:                  200
Model:                          Logit   Df Residuals:                      196
Method:                           MLE   Df Model:                            3
Date:                Fri, 15 Aug 2025   Pseudo R-squ.:                  0.4304
Time:                        14:48:34   Log-Likelihood:                -66.445
converged:                       True   LL-Null:                       -116.65
Covariance Type:            nonrobust   LLR p-value:                 1.266e-21
==============================================================================
                 coef    std err          z      P>|z|      [0.025      0.975]
------------------------------------------------------------------------------
const          0.9304      0.915      1.017      0.309      -0.863       2.723
x1             0.8313      0.155      5.354      0.000       0.527       1.136
x2            -0.6647      0.112     -5.961      0.000      -0.883      -0.446
x3             1.1915      0.270      4.413      0.000       0.662       1.721
==============================================================================
```

Cox regression (the most common survival model)

$$h(t|\boldsymbol{x}) = h_0(t)\exp(\beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_M x_M)$$

- **Data** consists of individuals i with
  - features $x_{1,i}, x_{2,i}, \ldots$
  - time t_i
  - event indicator δ_i, where
    - $\delta_i = 0$ denotes (right-)censoring
    - $\delta_i = 1$ denotes, e.g., mortality

- What is the distribution (CDF F, PDF f) of **survival time** T?

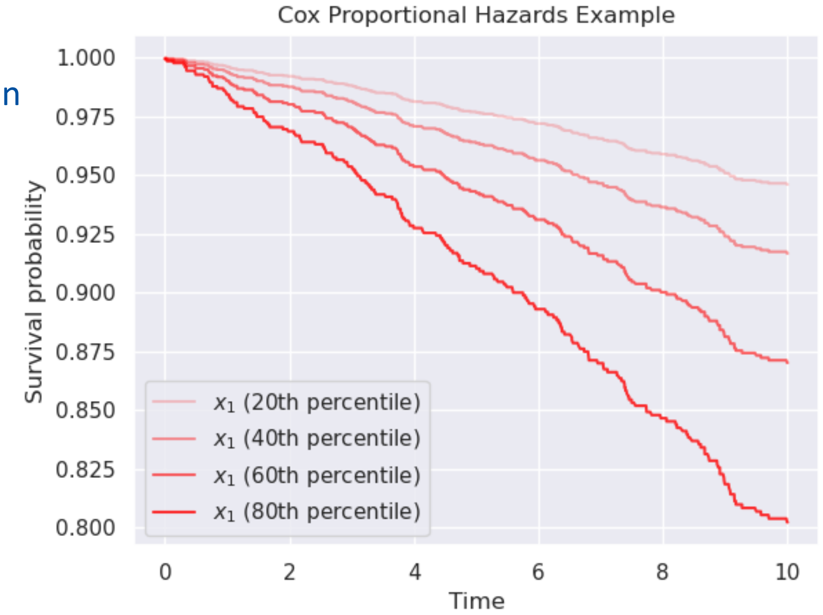


Cox regression (the most common survival model)

$$h(t|\boldsymbol{x}) = h_0(t)\exp(\beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_M x_M)$$

- **Hazard rates** $h(t|\boldsymbol{x})$, correspond to force of mortality $\mu_{\boldsymbol{x}}(t)$ in continuous time and $q_{\boldsymbol{x},t}$ or $m_{\boldsymbol{x},t}$ in discrete time
- **Proportional hazards**: $h(t|\boldsymbol{x}_i)/h(t|\boldsymbol{x}_j)$ const.
- **Survival probability function** $S(t|\boldsymbol{x})$, corresponds to ${}_t p_{\boldsymbol{x}}$

- $S(t|\boldsymbol{x}) = 1 - F(t|\boldsymbol{x})$
- $h(t|\boldsymbol{x}) = -\frac{\partial}{\partial t}\log S(t|\boldsymbol{x}) = \frac{f(t|\boldsymbol{x})}{S(t|\boldsymbol{x})}$

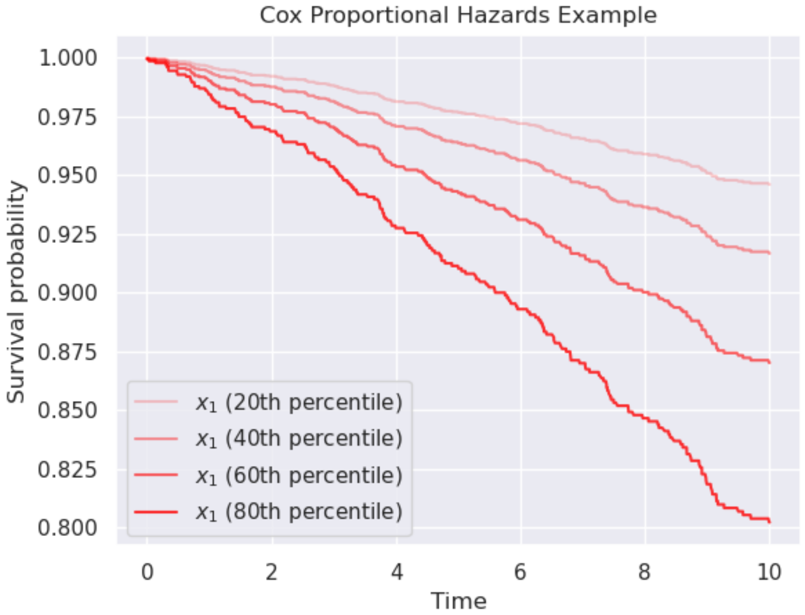


Cox regression (the most common survival model)

$$h(t|\boldsymbol{x}) = h_0(t)\exp(\beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_M x_M)$$

- **Baseline hazard rates** $h_0(t)$ via
  - Kaplan-Meier: $S(t) = \prod_{t_i \le t}\left(1 - \frac{d_i}{n_i}\right)$
  - Nelson-Aalen: $H(t) = \sum_{t_i \le t}\frac{d_i}{n_i}$
- **Coefficients** $\beta_1, \beta_2, \ldots$ via partial likelihood function maximization (Breslow method for tied events)

$$\mathcal{L} = \prod_{i:\delta_i=1}\prod_{j:t_j=t_i}\frac{\exp(\beta_1 x_{1,j} + \cdots)}{\sum_{k:t_k \ge t_j}\exp(\beta_1 x_{1,k} + \cdots)}$$



DAV Deutsche Aktuarvereinigung e.V.
DGVFM Deutsche Gesellschaft für Versicherungs- und Finanzmathematik e.V.

Cox regression (the most common survival model)

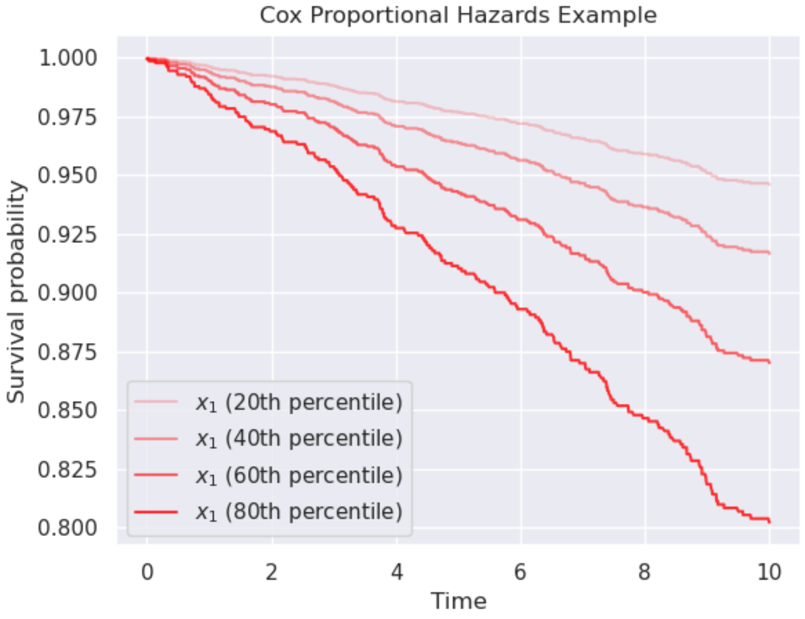
$$h(t|\boldsymbol{x}) = h_0(t)\exp(\beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_M x_M)$$

model	lifelines.CoxPHFitter
duration col	'time'
event col	'event'
baseline estimation	breslow
number of observations	2000
number of events observed	178
partial log-likelihood	-1248.33
time fit was run	2025-08-18 08:31:08 UTC

	coef	exp(coef)	se(coef)	coef lower 95%	coef upper 95%	exp(coef) lower 95%	exp(coef) upper 95%	cmp to	z	p	-log2(p)
x1	0.09	1.09	0.01	0.07	0.11	1.07	1.12	0.00	7.76	<0.005	46.71
x2	0.20	1.22	0.15	-0.09	0.50	0.91	1.64	0.00	1.34	0.18	2.46
x3	0.04	1.04	0.02	0.00	0.07	1.00	1.08	0.00	2.06	0.04	4.67

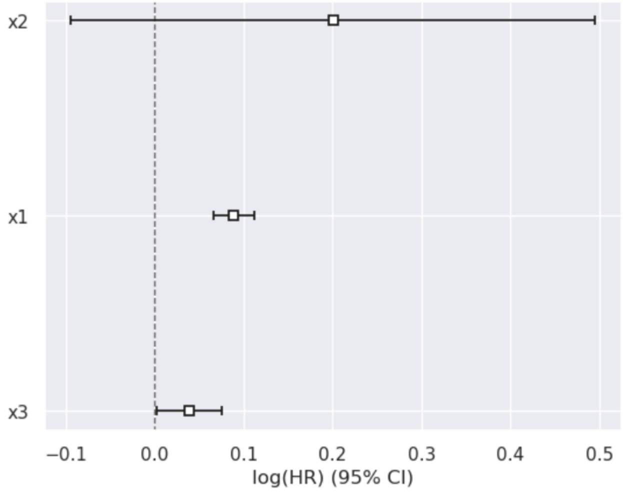
Concordance	0.67
Partial AIC	2502.65
log-likelihood ratio test	72.86 on 3 df
-log2(p) of ll-ratio test	49.77

lifelines summary

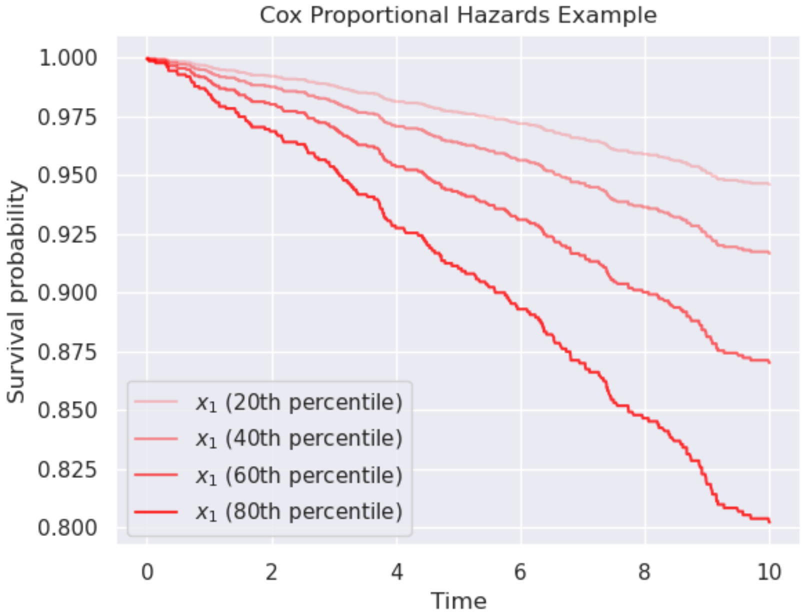


Cox regression (the most common survival model)

$$h(t|\boldsymbol{x}) = h_0(t)\exp(\beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_M x_M)$$



lifelines forest plot



A bit of public health history

Lester Breslow (1915-2012), the father of Norman Breslow after whom the method was named

Dr. **Lester Breslow**, a former dean of the UCLA Jonathan and Karin Fielding School of Public Health, professor emeritus of health services, and one of the leading figures in public health for seven decades, died Monday. He was 97.

Breslow was a visionary public health figure with a well-established track record for being ahead of his time. As early as the 1940s, he linked tobacco use to disease in three studies that were later cited in the U.S. Surgeon General's landmark 1964 report.

He is widely known for his early advocacy and research into health promotion and disease prevention. Breslow's pioneering Alameda County studies beginning in the early 1960s were among the first to show that simple health practices — such as getting regular exercise and sleep, not drinking excessively, not smoking, and maintaining a healthy weight — add both years and quality to life.

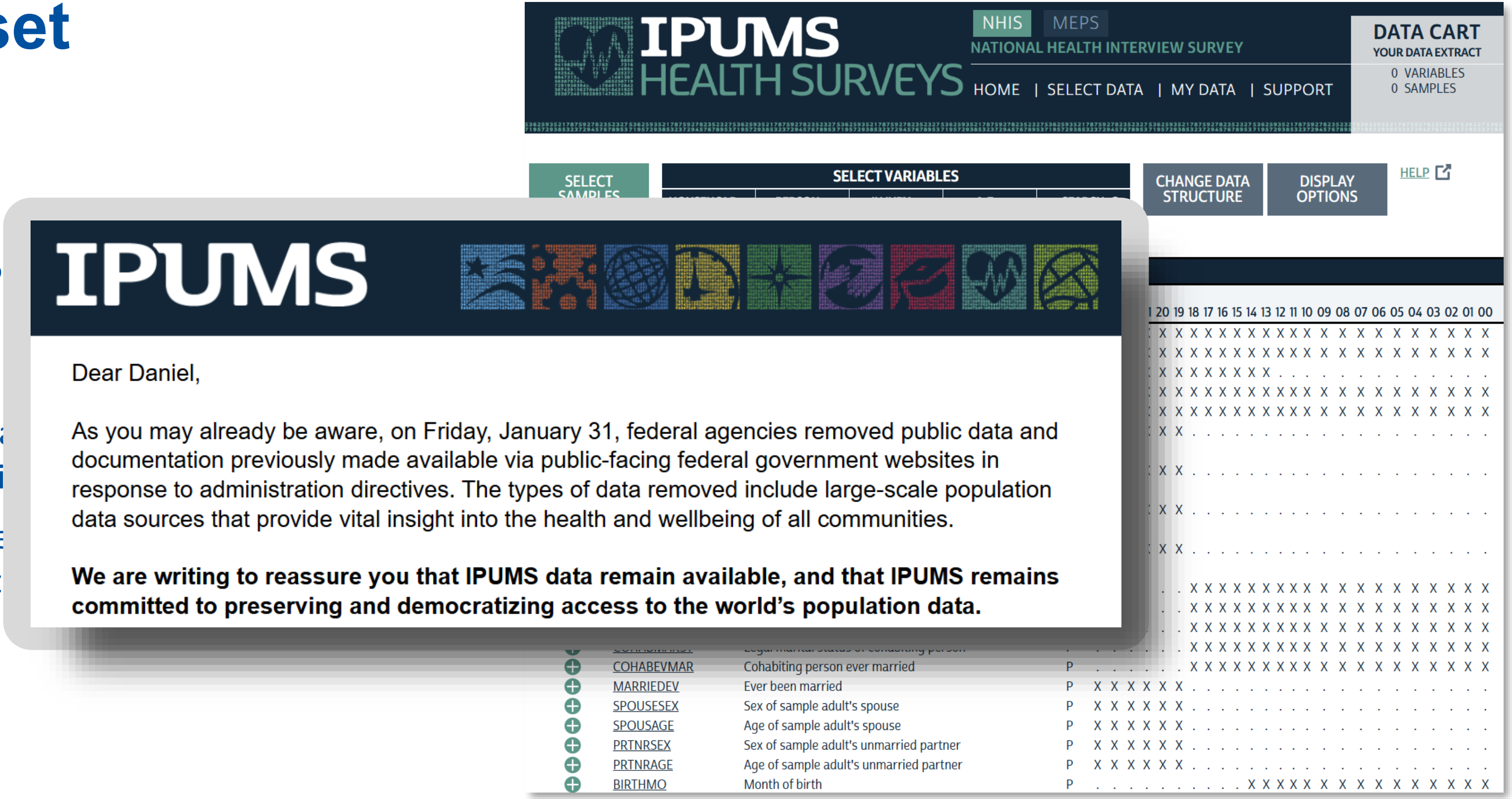
While these conclusions are taken for granted today, the idea of such a strong connection between lifestyle and health was seen as "bizarre" at the time, Breslow noted decades later. He would smile when recalling the response of the National Institutes of Health panel of scientists that reviewed the initial study proposal: "Unanimous rejection." When the study was completed, even Breslow was shocked at the magnitude of the results, which helped usher in current thinking about health and fitness.

Source: https://ph.ucla.edu/news-events/news/memoriam-dr-lester-breslow-public-health-visionary

DAV Deutsche Aktuarvereinigung e.V. DGVFM Deutsche Gesellschaft für Versicherungs- und Finanzmathematik e.V.

The dataset

- Data since 196
- 100k individua
- 500+ features: economics, hea health, mortali
- Part of our dire datasets at act



IPUMS

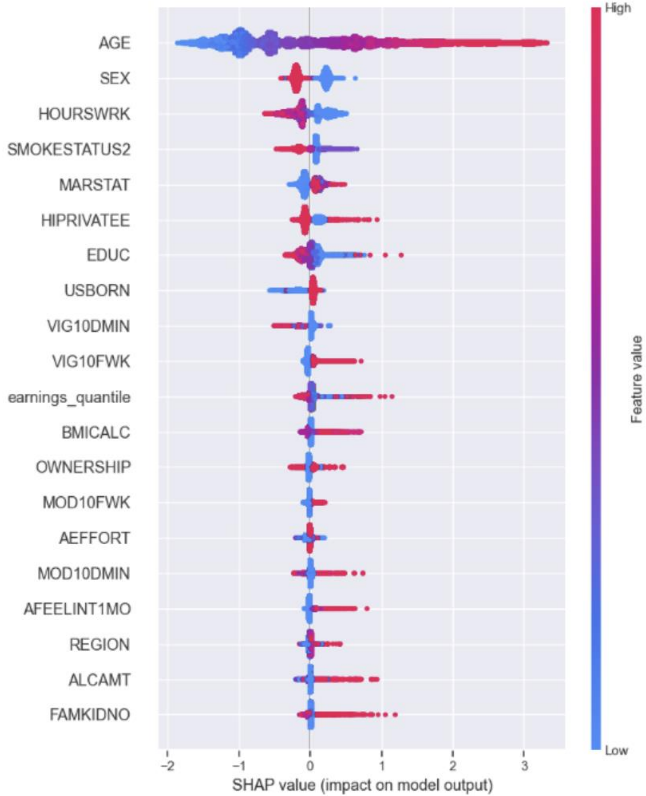
Dear Daniel,

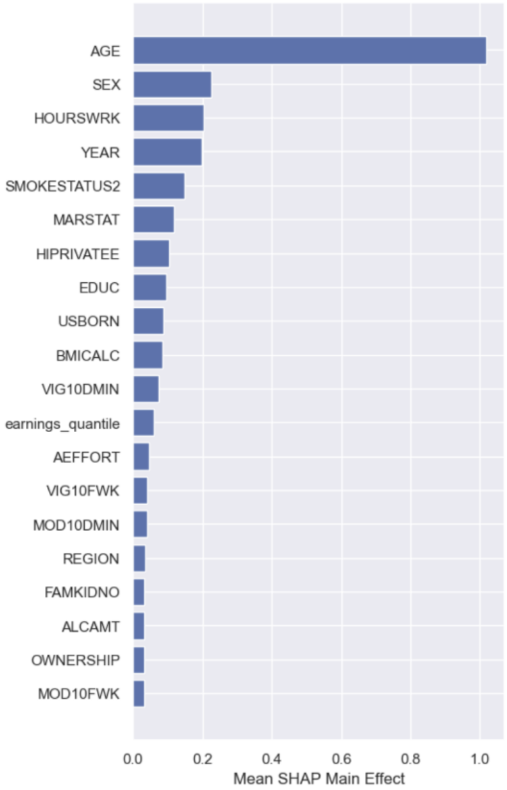
As you may already be aware, on Friday, January 31, federal agencies removed public data and documentation previously made available via public-facing federal government websites in response to administration directives. The types of data removed include large-scale population data sources that provide vital insight into the health and wellbeing of all communities.

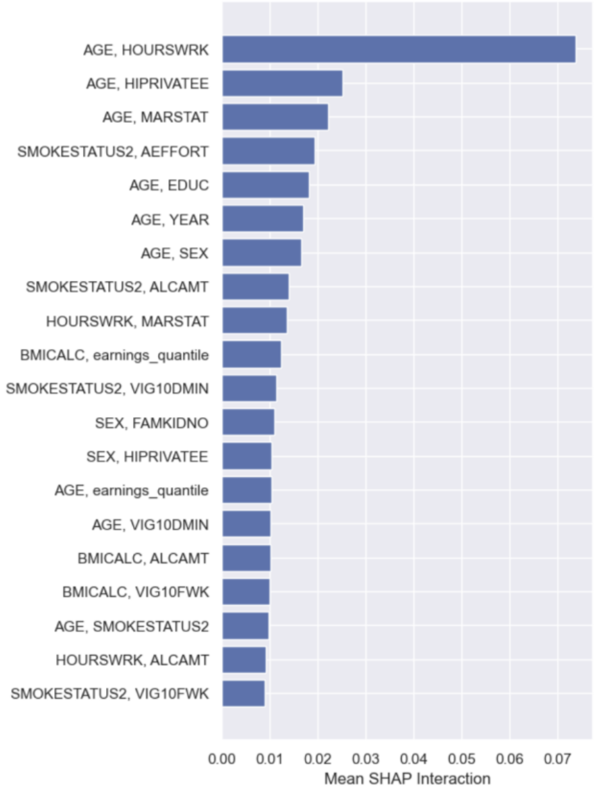
We are writing to reassure you that IPUMS data remain available, and that IPUMS remains committed to preserving and democratizing access to the world's population data.

Source: nhis.ipums.org

SHAP values


AGE
SEX
HOURSWRK
SMOKESTATUS2
MARSTAT
HIPRIVATEE
EDUC
USBORN
VIG10DMIN
VIG10FWK
earnings_quantile
BMICALC
OWNERSHIP
MOD10FWK
AEFFORT
MOD10DMIN
AFEELINT1MO
REGION
ALCAMT
FAMKIDNO
−2
−1
0
1
2
3
SHAP value (impact on model output)
High
Feature value
Low


AGE
SEX
HOURSWRK
YEAR
SMOKESTATUS2
MARSTAT
HIPRIVATEE
EDUC
USBORN
BMICALC
VIG10DMIN
earnings_quantile
AEFFORT
VIG10FWK
MOD10DMIN
REGION
FAMKIDNO
ALCAMT
OWNERSHIP
MOD10FWK
0.0
0.2
0.4
0.6
0.8
1.0
Mean SHAP Main Effect


AGE, HOURSWRK
AGE, HIPRIVATEE
AGE, MARSTAT
SMOKESTATUS2, AEFFORT
AGE, EDUC
AGE, YEAR
AGE, SEX
SMOKESTATUS2, ALCAMT
HOURSWRK, MARSTAT
BMICALC, earnings_quantile
SMOKESTATUS2, VIG10DMIN
SEX, FAMKIDNO
SEX, HIPRIVATEE
AGE, earnings_quantile
AGE, VIG10DMIN
BMICALC, ALCAMT
BMICALC, VIG10FWK
AGE, SMOKESTATUS2
HOURSWRK, ALCAMT
SMOKESTATUS2, VIG10FWK
0.00
0.01
0.02
0.03
0.04
0.05
0.06
0.07
Mean SHAP Interaction

Models

Model	Parameterization	Comments	Perform.	Speed
Base q_x	$h(t \mid \boldsymbol{x}) = q_x$	Only considering age x and sex	★☆☆☆☆	★★★★★

Models

Light GBM

Model	Parameterization	Comments	Perform.	Speed
Base q_x	$h(t \mid \boldsymbol{x}) = q_x$	Only considering age x and sex	★☆☆☆☆	★★★★★
LGBM	$S(t_1 \mid \boldsymbol{x}) - S(t_2 \mid \boldsymbol{x}) = f^{(m)}(\boldsymbol{x})$	Only for interval prediction in $[t_1, t_2)$	★★★★☆	★★★★★

Models

Cox regression

Model	Parameterization	Comments	Perform.	Speed
Base q_x	$h(t \mid \boldsymbol{x}) = q_x$	Only considering age x and sex	★☆☆☆☆	★★★★★
LGBM	$S(t_1 \mid \boldsymbol{x}) - S(t_2 \mid \boldsymbol{x}) = f^{(m)}(\boldsymbol{x})$	Only for interval prediction in $[t_1, t_2)$	★★★★☆	★★★★★
Cox	$h(t \mid \boldsymbol{x}) = h_0(t) \exp(\boldsymbol{\beta}^\intercal \boldsymbol{x})$	Proportional hazards, de facto standard	★★☆☆☆	★★★★★

Models

Accelerated failure time

Model	Parameterization	Comments	Perform.	Speed
Base q_x	$h(t \mid \boldsymbol{x}) = q_x$	Only considering age x and sex	★☆☆☆☆	★★★★★
LGBM	$S(t_1 \mid \boldsymbol{x}) - S(t_2 \mid \boldsymbol{x}) = f^{(m)}(\boldsymbol{x})$	Only for interval prediction in $[t_1, t_2)$	★★★★☆	★★★★★
Cox	$h(t \mid \boldsymbol{x}) = h_0(t) \exp(\boldsymbol{\beta}^\intercal \boldsymbol{x})$	Proportional hazards, de facto standard	★★☆☆☆	★★★★★
AFT	$T = \exp(\varepsilon) \exp(\boldsymbol{\beta}^\intercal \boldsymbol{x})$	Incl. Weibull (AFT and PH), scales T	★★☆☆☆	★★★★★

Models

Model	Parameterization	Comments	Perform.	Speed
Base q_x	$h(t \mid \boldsymbol{x}) = q_x$	Only considering age x and sex	★☆☆☆☆	★★★★★
LGBM	$S(t_1 \mid \boldsymbol{x}) - S(t_2 \mid \boldsymbol{x}) = f^{(m)}(\boldsymbol{x})$	Only for interval prediction in $[t_1, t_2)$	★★★★☆	★★★★★
Cox	$h(t \mid \boldsymbol{x}) = h_0(t) \exp(\boldsymbol{\beta}^\intercal \boldsymbol{x})$	Proportional hazards, de facto standard	★★☆☆☆	★★★★★
AFT	$T = \exp(\varepsilon) \exp(\boldsymbol{\beta}^\intercal \boldsymbol{x})$	Incl. Weibull (AFT and PH), scales T	★★☆☆☆	★★★★★
ST	$S(t \mid \boldsymbol{x}) = S_{\ell(\boldsymbol{x})}(t)$	Kaplan–Meier curves at tree leaves ℓ	★☆☆☆☆	★★☆☆☆

Survival trees

Models

Random survival forests

Model	Parameterization	Comments	Perform.	Speed
Base q_x	$h(t \mid \boldsymbol{x}) = q_x$	Only considering age x and sex	★☆☆☆☆	★★★★★
LGBM	$S(t_1 \mid \boldsymbol{x}) - S(t_2 \mid \boldsymbol{x}) = f^{(m)}(\boldsymbol{x})$	Only for interval prediction in $[t_1, t_2)$	★★★★☆	★★★★★
Cox	$h(t \mid \boldsymbol{x}) = h_0(t) \exp(\boldsymbol{\beta}^\intercal \boldsymbol{x})$	Proportional hazards, de facto standard	★★☆☆☆	★★★★★
AFT	$T = \exp(\varepsilon) \exp(\boldsymbol{\beta}^\intercal \boldsymbol{x})$	Incl. Weibull (AFT and PH), scales T	★★☆☆☆	★★★★★
ST	$S(t \mid \boldsymbol{x}) = S_{\ell(\boldsymbol{x})}(t)$	Kaplan–Meier curves at tree leaves ℓ	★☆☆☆☆	★★☆☆☆
RSF	$S(t \mid \boldsymbol{x}) := \frac{1}{B} \sum_{b=1}^{B} S_{\ell^{(b)}(\boldsymbol{x})}(t)$	Ensemble of trees, smooth predictions	★★★☆☆	★★☆☆☆

Models

Model	Parameterization	Comments	Perform.	Speed
Base q_x	$h(t \mid \boldsymbol{x}) = q_x$	Only considering age x and sex	★☆☆☆☆	★★★★★
LGBM	$S(t_1 \mid \boldsymbol{x}) - S(t_2 \mid \boldsymbol{x}) = f^{(m)}(\boldsymbol{x})$	Only for interval prediction in $[t_1, t_2)$	★★★★☆	★★★★★
Cox	$h(t \mid \boldsymbol{x}) = h_0(t) \exp(\boldsymbol{\beta}^\intercal \boldsymbol{x})$	Proportional hazards, de facto standard	★★☆☆☆	★★★★★
AFT	$T = \exp(\varepsilon) \exp(\boldsymbol{\beta}^\intercal \boldsymbol{x})$	Incl. Weibull (AFT and PH), scales T	★★☆☆☆	★★★★★
ST	$S(t \mid \boldsymbol{x}) = S_{\ell(\boldsymbol{x})}(t)$	Kaplan–Meier curves at tree leaves ℓ	★☆☆☆☆	★★☆☆☆
RSF	$S(t \mid \boldsymbol{x}) := \frac{1}{B} \sum_{b=1}^{B} S_{\ell^{(b)}(\boldsymbol{x})}(t)$	Ensemble of trees, smooth predictions	★★★☆☆	★★☆☆☆
GBST	$h(t \mid \boldsymbol{x}) = h_0(t) \exp(f^{(m)}(\boldsymbol{x}))$	Refining trees series $f^{(0)}, f^{(1)}, \ldots, f^{(m)}$	★★☆☆☆	★☆☆☆☆

Gradient boosted survival trees (scikit)

Models

Model	Parameterization	Comments	Perform.	Speed
Base q_x	$h(t \mid \boldsymbol{x}) = q_x$	Only considering age x and sex	★☆☆☆☆	★★★★★
LGBM	$S(t_1 \mid \boldsymbol{x}) - S(t_2 \mid \boldsymbol{x}) = f^{(m)}(\boldsymbol{x})$	Only for interval prediction in $[t_1, t_2)$	★★★★☆	★★★★★
Cox	$h(t \mid \boldsymbol{x}) = h_0(t) \exp(\boldsymbol{\beta}^\intercal \boldsymbol{x})$	Proportional hazards, de facto standard	★★☆☆☆	★★★★★
AFT	$T = \exp(\varepsilon) \exp(\boldsymbol{\beta}^\intercal \boldsymbol{x})$	Incl. Weibull (AFT and PH), scales T	★★☆☆☆	★★★★★
ST	$S(t \mid \boldsymbol{x}) = S_{\ell(\boldsymbol{x})}(t)$	Kaplan–Meier curves at tree leaves ℓ	★☆☆☆☆	★★☆☆☆
RSF	$S(t \mid \boldsymbol{x}) := \frac{1}{B} \sum_{b=1}^{B} S_{\ell^{(b)}(\boldsymbol{x})}(t)$	Ensemble of trees, smooth predictions	★★★☆☆	★★☆☆☆
GBST	$h(t \mid \boldsymbol{x}) = h_0(t) \exp(f^{(m)}(\boldsymbol{x}))$	Refining trees series $f^{(0)}, f^{(1)}, \ldots, f^{(m)}$	★★☆☆☆	★☆☆☆☆
XGB	$h(t \mid \boldsymbol{x}) = h_0(t) \exp(f^{(m)}(\boldsymbol{x}))$	Refining trees series $f^{(0)}, f^{(1)}, \ldots, f^{(m)}$	★★★★☆	★★★★★

Gradient boosted survival trees (XGB)

Models

Model	Parameterization	Comments	Perform.	Speed
Base q_x	$h(t \mid \boldsymbol{x}) = q_x$	Only considering age x and sex	★☆☆☆☆	★★★★★
LGBM	$S(t_1 \mid \boldsymbol{x}) - S(t_2 \mid \boldsymbol{x}) = f^{(m)}(\boldsymbol{x})$	Only for interval prediction in $[t_1, t_2)$	★★★★☆	★★★★★
Cox	$h(t \mid \boldsymbol{x}) = h_0(t) \exp(\boldsymbol{\beta}^\intercal \boldsymbol{x})$	Proportional hazards, de facto standard	★★☆☆☆	★★★★★
AFT	$T = \exp(\varepsilon) \exp(\boldsymbol{\beta}^\intercal \boldsymbol{x})$	Incl. Weibull (AFT and PH), scales T	★★☆☆☆	★★★★★
ST	$S(t \mid \boldsymbol{x}) = S_{\ell(\boldsymbol{x})}(t)$	Kaplan–Meier curves at tree leaves ℓ	★☆☆☆☆	★★☆☆☆
RSF	$S(t \mid \boldsymbol{x}) := \frac{1}{B} \sum_{b=1}^{B} S_{\ell^{(b)}(\boldsymbol{x})}(t)$	Ensemble of trees, smooth predictions	★★★☆☆	★★☆☆☆
GBST	$h(t \mid \boldsymbol{x}) = h_0(t) \exp(f^{(m)}(\boldsymbol{x}))$	Refining trees series $f^{(0)}, f^{(1)}, \ldots, f^{(m)}$	★★☆☆☆	★☆☆☆☆
XGB	$h(t \mid \boldsymbol{x}) = h_0(t) \exp(f^{(m)}(\boldsymbol{x}))$	Refining trees series $f^{(0)}, f^{(1)}, \ldots, f^{(m)}$	★★★★☆	★★★★★
DeepSurv	$h(t \mid \boldsymbol{x}) = h_0(t) \exp(z_\theta(\boldsymbol{x}))$	Replacing $f^{(m)}$ by a neural network z_θ	★★★★☆	★★★★☆

Deep survival

Models

Model	Parameterization	Comments	Perform.	Speed
Base q_x	$h(t \mid \boldsymbol{x}) = q_x$	Only considering age x and sex	★☆☆☆☆	★★★★★
LGBM	$S(t_1 \mid \boldsymbol{x}) - S(t_2 \mid \boldsymbol{x}) = f^{(m)}(\boldsymbol{x})$	Only for interval prediction in $[t_1, t_2)$	★★★★☆	★★★★★
Cox	$h(t \mid \boldsymbol{x}) = h_0(t) \exp(\boldsymbol{\beta}^\intercal \boldsymbol{x})$	Proportional hazards, de facto standard	★★☆☆☆	★★★★★
AFT	$T = \exp(\varepsilon) \exp(\boldsymbol{\beta}^\intercal \boldsymbol{x})$	Incl. Weibull (AFT and PH), scales T	★★☆☆☆	★★★★★
ST	$S(t \mid \boldsymbol{x}) = S_{\ell(\boldsymbol{x})}(t)$	Kaplan–Meier curves at tree leaves ℓ	★☆☆☆☆	★★☆☆☆
RSF	$S(t \mid \boldsymbol{x}) := \frac{1}{B} \sum_{b=1}^{B} S_{\ell^{(b)}(\boldsymbol{x})}(t)$	Ensemble of trees, smooth predictions	★★★☆☆	★★☆☆☆
GBST	$h(t \mid \boldsymbol{x}) = h_0(t) \exp(f^{(m)}(\boldsymbol{x}))$	Refining trees series $f^{(0)}, f^{(1)}, \dots, f^{(m)}$	★★☆☆☆	★☆☆☆☆
XGB	$h(t \mid \boldsymbol{x}) = h_0(t) \exp(f^{(m)}(\boldsymbol{x}))$	Refining trees series $f^{(0)}, f^{(1)}, \dots, f^{(m)}$	★★★★☆	★★★★★
DeepSurv	$h(t \mid \boldsymbol{x}) = h_0(t) \exp(z_\theta(\boldsymbol{x}))$	Replacing $f^{(m)}$ by a neural network z_θ	★★★★☆	★★★★☆
DeepHit	Time-discretized $f(t \mid \boldsymbol{x})$	Neural network, allows competing risks	★★★★★	★★★☆☆

DeepHit

Models

Model	Parameterization	Comments	Perform.	Speed
Base q_x	$h(t \mid \boldsymbol{x}) = q_x$	Only considering age x and sex	★☆☆☆☆	★★★★★
LGBM	$S(t_1 \mid \boldsymbol{x}) - S(t_2 \mid \boldsymbol{x}) = f^{(m)}(\boldsymbol{x})$	Only for interval prediction in $[t_1, t_2)$	★★★★☆	★★★★★
Cox	$h(t \mid \boldsymbol{x}) = h_0(t) \exp(\boldsymbol{\beta}^\intercal \boldsymbol{x})$	Proportional hazards, de facto standard	★★☆☆☆	★★★★★
AFT	$T = \exp(\varepsilon) \exp(\boldsymbol{\beta}^\intercal \boldsymbol{x})$	Incl. Weibull (AFT and PH), scales T	★★☆☆☆	★★★★★
ST	$S(t \mid \boldsymbol{x}) = S_{\ell(\boldsymbol{x})}(t)$	Kaplan–Meier curves at tree leaves ℓ	★☆☆☆☆	★★☆☆☆
RSF	$S(t \mid \boldsymbol{x}) := \frac{1}{B} \sum_{b=1}^{B} S_{\ell^{(b)}(\boldsymbol{x})}(t)$	Ensemble of trees, smooth predictions	★★★☆☆	★★☆☆☆
GBST	$h(t \mid \boldsymbol{x}) = h_0(t) \exp(f^{(m)}(\boldsymbol{x}))$	Refining trees series $f^{(0)}, f^{(1)}, \ldots, f^{(m)}$	★★☆☆☆	★☆☆☆☆
XGB	$h(t \mid \boldsymbol{x}) = h_0(t) \exp(f^{(m)}(\boldsymbol{x}))$	Refining trees series $f^{(0)}, f^{(1)}, \ldots, f^{(m)}$	★★★★☆	★★★★★
DeepSurv	$h(t \mid \boldsymbol{x}) = h_0(t) \exp(z_\theta(\boldsymbol{x}))$	Replacing $f^{(m)}$ by a neural network z_θ	★★★★☆	★★★★☆
DeepHit	Time-discretized $f(t \mid \boldsymbol{x})$	Neural network, allows competing risks	★★★★★	★★★☆☆
DSM	Mixture of, e.g., Weibull	Neural network, allows competing risks	★★★★☆	★★★☆☆

Deep survival machines

Models

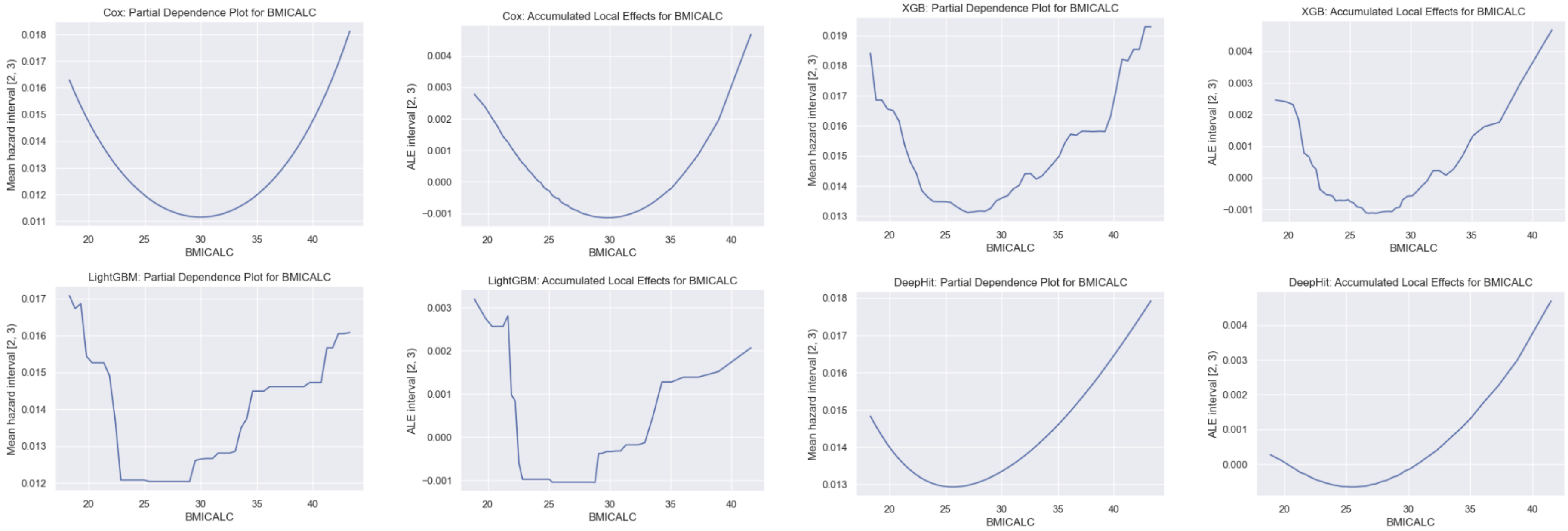
Model	Parameterization	Comments	Perform.	Speed
Base q_x	$h(t \mid \boldsymbol{x}) = q_x$	Only considering age x and sex	★☆☆☆☆	★★★★★
LGBM	$S(t_1 \mid \boldsymbol{x}) - S(t_2 \mid \boldsymbol{x}) = f^{(m)}(\boldsymbol{x})$	Only for interval prediction in $[t_1, t_2)$	★★★★☆	★★★★★
Cox	$h(t \mid \boldsymbol{x}) = h_0(t) \exp(\boldsymbol{\beta}^\intercal \boldsymbol{x})$	Proportional hazards, de facto standard	★★☆☆☆	★★★★★
AFT	$T = \exp(\varepsilon) \exp(\boldsymbol{\beta}^\intercal \boldsymbol{x})$	Incl. Weibull (AFT and PH), scales T	★★☆☆☆	★★★★★
ST	$S(t \mid \boldsymbol{x}) = S_{\ell(\boldsymbol{x})}(t)$	Kaplan–Meier curves at tree leaves ℓ	★☆☆☆☆	★★☆☆☆
RSF	$S(t \mid \boldsymbol{x}) := \frac{1}{B} \sum_{b=1}^{B} S_{\ell^{(b)}(\boldsymbol{x})}(t)$	Ensemble of trees, smooth predictions	★★★☆☆	★★☆☆☆
GBST	$h(t \mid \boldsymbol{x}) = h_0(t) \exp(f^{(m)}(\boldsymbol{x}))$	Refining trees series $f^{(0)}, f^{(1)}, \ldots, f^{(m)}$	★★☆☆☆	★☆☆☆☆
XGB	$h(t \mid \boldsymbol{x}) = h_0(t) \exp(f^{(m)}(\boldsymbol{x}))$	Refining trees series $f^{(0)}, f^{(1)}, \ldots, f^{(m)}$	★★★★☆	★★★★★
DeepSurv	$h(t \mid \boldsymbol{x}) = h_0(t) \exp(z_\theta(\boldsymbol{x}))$	Replacing $f^{(m)}$ by a neural network z_θ	★★★★☆	★★★★☆
DeepHit	Time-discretized $f(t \mid \boldsymbol{x})$	Neural network, allows competing risks	★★★★★	★★★☆☆
DSM	Mixture of, e.g., Weibull	Neural network, allows competing risks	★★★★☆	★★★☆☆
TSM	Time-discretized $f(t \mid \boldsymbol{x})$	Neural network, use of covariates' history	★★★★☆	★★★☆☆

Transformer-based survival models

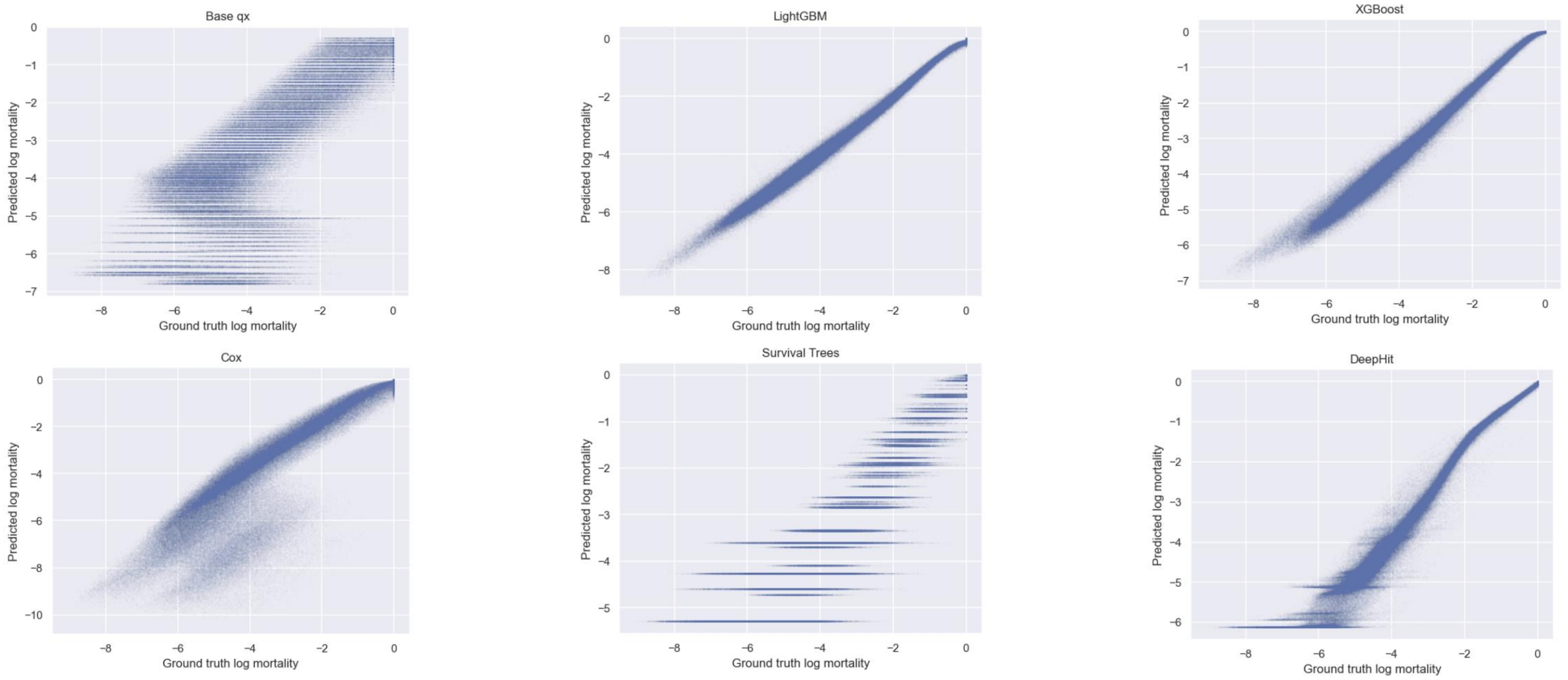
DAV Deutsche Aktuarvereinigung e.V.

DGVFM Deutsche Gesellschaft für Versicherungs- und Finanzmathematik e.V.

Partial dependence plots + accumulated local effects



Ground truth vs. predictions on a synthetic dataset



Strong clinical trials, long-term effectiveness pending

	Semaglutide (Ozempic/ Wegovy)	**Tirzepatide** (Mounjaro/ Zepbound)
In the head-to-head study over 72 weeks:		
Weight lost (kg)	↓ 15.0	22.8
Weight loss (%)	↓ 13.7	20.2
Waist circumference (cm)	↓ 13.0	18.4
BMI points	↓ 5.3	8.0

Tirzepatide beat semaglutide in every category

? Between 39% (semaglutide) and 18% (tirzepatide) participants lost <10% weight

Additional improvements in blood pressure, HbA1c and cholesterol markers were seen in both drugs but superior for tirzepatide.

Clinical trial results do not directly translate into real-world results. Longer-term considerations include drug adherence and sustained weight loss.

Started as niche diabetes treatment, became big business

Heavy guy, light appetite: this gila monster eats a few times a year, slowly digesting his food thanks to his naturally long-lasting GLP-1 hormone

Future trends in drug development

More targets: semaglutide: 1 target. Tirzepatide: 2 targets. New drugs: 2-3+ targets and/or combinations

Different doses: today the standard is weekly injections. Trials are for fortnightly or monthly injections

Fast weight loss: companies are targeting headline large losses of weight over as little time as possible

CURRENTLY USED

Previous generation medications

Ozempic/Wegovy (novo nordisk)
Semaglutide (2021)
Format: once-weekly injectable

Zepbound/Mounjaro (Lilly)
Tirzepatide (2023)
Format: once-weekly injectable

IN DEVELOPMENT

Rybelsus (novo nordisk)
Year: 2025 – approved for diabetes, pending approval for weight loss
Format: daily oral tablet

Orforglipron (Lilly)
Year: 2025 – phase 3 clinical trial
Format: daily oral tablet

Others
- multitude of developers trialing different receptors, dosages, regimes and delivery mechanisms

MariTide (AMGEN)
Year: 2024 – phase 2 clinical trial
Format: monthly injectable

AZD5004 (AstraZeneca)
Year: 2024 – phase 1 clinical trial
Format: daily oral tablet

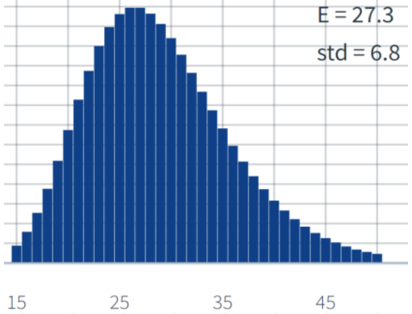
CT-996 (Roche)
Year: 2025 – phase 1 clinical trial
Format: daily oral tablet

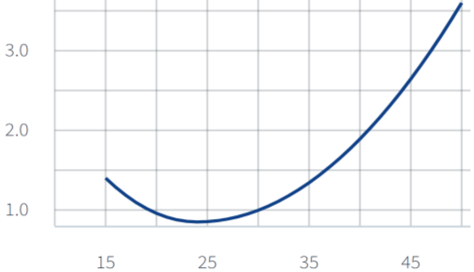
How they work: slow gastric emptying, feeling of satiety, reduce calories

Metabolic health model

Data inputs

BMI **distribution** and **relative mortality** risk by age, SES, country





Analysis process

Expert elicitation on BMI & GLP-1

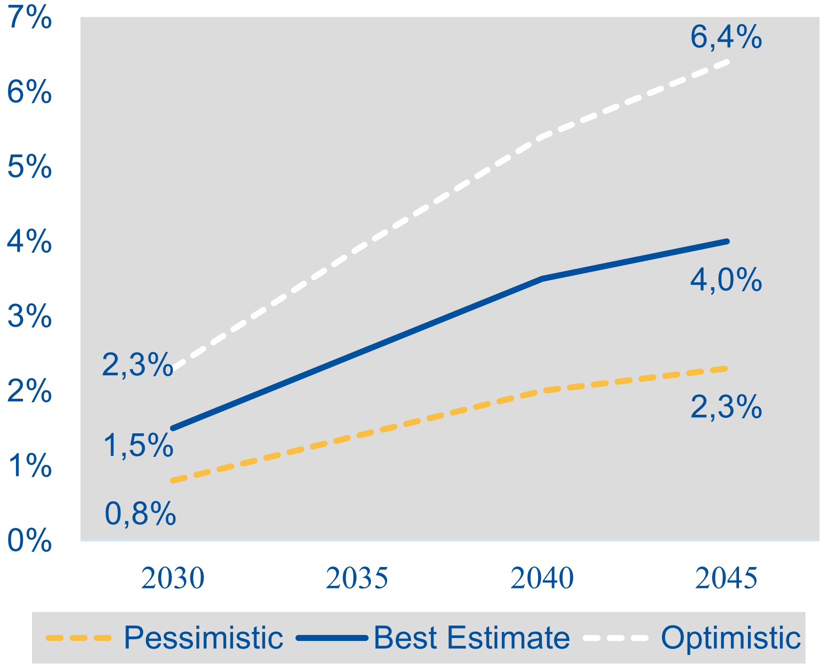
Literature review on all factors

Modelling output

- **Population**: general and insured
- **Countries**: US, China, UK
- **Timeframe**: 2025 – 2045
- **Scenarios**:
  1. Best estimate
  2. Optimistic
  3. Pessimistic
- **Methodology**: simulates 100k individuals with a given age, SES, BMI, SBP, etc.
- **Output**: aggregates total relative mortality risk distributions

Mortality reductions in US general population

US general population projection



Rationale

Broad starting health: US general population has a wide range of health statuses

Lifestyle changes: behavioural adaptions support medium- and long-term mortality benefits

Age: middle-aged population is expected to see the largest mortality benefits

Greatest opportunity: widespread metabolic health changes would see substantial public health improvements

Vielen Dank für Ihre Aufmerksamkeit.

Daniel Meier, Swiss Re
daniel_meier@swissre.com

Besuchen Sie unsere Webseite

www.aktuar.de