Modelling non-proportional hazards: time-dependent coefficients, parametric “double Cox” regression and Landmark analysis

Elena Kulinskaya
School of Computing Sciences,
University of East Anglia

The **‘Use of Big Health and Actuarial Data for understanding Longevity and Morbidity Risks’** research programme is being funded by the Actuarial Research Centre.

Actuarial Research Centre
Institute and Faculty of Actuaries

Disclaimer

What if the proportional hazards assumption is not met?

- For a Cox model $\mu(t|\beta, Z) = \mu_0(t)\exp(Z^{\mathrm{T}}\beta)$ we discussed two ways to cope with non-proportionality:
- Stratify the analysis on violating variable: $\mu_s(t|\beta, Z) = \mu_{0s}(t)e^{Z^{\mathrm{T}}\beta}$

  - baseline hazards vary by strata s;

  - Here we add an option of modelling shape of baseline hazards
- Include time-varying effects: $\mu(t, |\beta, Z) = \mu_0(t)e^{Z^{\mathrm{T}}\beta(t)}$

  - Coefficients $\beta(t)$ are continuous functions of time

  - Use landmark analysis

Actuarial Research Centre

Institute and Faculty of Actuaries

Parametric "Double-Cox" regression

Components:

- A baseline hazard function (which changes over time).
- The risk factors Z have a log-linear contribution to the force of mortality which does not depend on time t.

The Cox parametric regression model

$$\mu(t|Z) = \mu_0(t|Z)\ \exp(Z^{\mathrm{T}}\beta)$$

Baseline hazard function

β is a vector of unknown parameters for scale and Z is a vector of covariates

Weibull or Gompertz baseline hazard function with scale λ and shape k. Shape k is modelled as k=k(Z).

$$\mu_0(t|Z) = \frac{k(Z)}{\lambda}\left(\frac{t}{\lambda}\right)^{k(Z)-1}$$

$$\mu_0(t|Z) = \lambda \exp(k(Z)t)$$

$$\mathrm{k(Z)} = k_0 e^{Z^{\mathrm{T}}\beta_k}$$

Additional regression model to allow varying shape depending on covariates

Actuarial Research Centre
Institute and Faculty of Actuaries

Cox model with shared frailty

Proportional hazards model with frailty:

$$\boldsymbol{\mu}(t|U,Z) = \boldsymbol{\mu}_0(t)Ue^{Z^{\mathrm{T}}\beta},$$

For mathematical convenience, it is frequently assumed that frailty U is gamma-distributed with mean 1 and unknown variance σ^2:

$$U \sim Gamma(\sigma^{-2}, \sigma^{-2}).$$

The frailty variance σ^2 characterizes heterogeneity in the population.

Shared frailty assumption:

All patients from the same **unit** /clients from the same company are in the same cluster j, j=1,…,J and share the same frailty U_j.

Actuarial Research Centre
Institute and Faculty of Actuaries

“Double-Cox” model with shared frailty

- Standard shared frailty Cox model : $\boldsymbol{\mu}(t|U,Z) = \boldsymbol{\mu}_0(t)Ue^{Z^{\mathrm{T}}\beta}$;
- Baseline hazard $\boldsymbol{\mu}_0(t)$=$\boldsymbol{\mu}_0(t;\lambda, k)$;
- Cox-like parameterization for the shape of the baseline hazard function: k(Z)=$k_0 e^{Z^{\mathrm{T}}\beta_k}$;
- Frailty U ~ Gamma with mean 1 and variance σ^2.
- If needed, competing risks can be introduced through correlated shared frailty components.

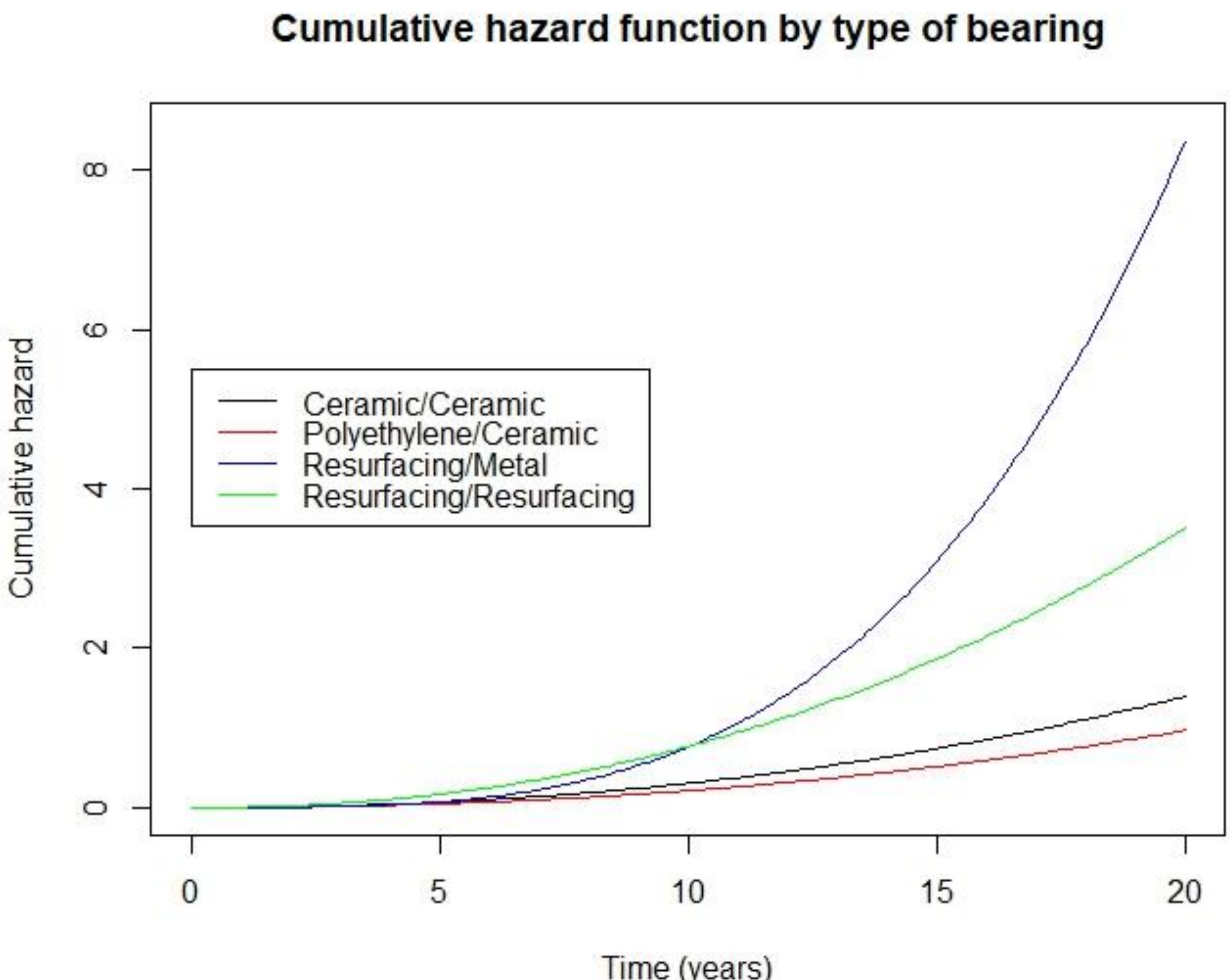
Find MLE of the vector of unknown parameters θ=(λ, k_0, σ^2, β, β_k).

This model was introduced in [1] for analysis of time to revision/

time to death after hip replacement.

Actuarial Research Centre
Institute and Faculty of Actuaries

Different shapes of cumulative hazards for revision surgery after hip replacement



Actuarial Research Centre
Institute and Faculty of Actuaries

Extended Cox regression with time-varying covariates and regression effects

A model may include both constant and time-varying effects:

$$\mu(t, |\beta, Z) = \mu_0(t) e^{Z(t)^{\mathrm{T}}\beta(t) + X(t)^{\mathrm{T}}\gamma}$$

- Here Z(t) and X(t) are time-varying covariates (updated over time).
- Z(t) are covariates with time-varying hazards β(t), and X(t) covariates have constant hazards γ.
- See Ch. 6 in the book by Martinussen&Scheike [2] and their R package *timereg* for analysis of extended multiplicative hazards models.
- Their program *timecox* can test for and fit models with both constant and time-varying effects.

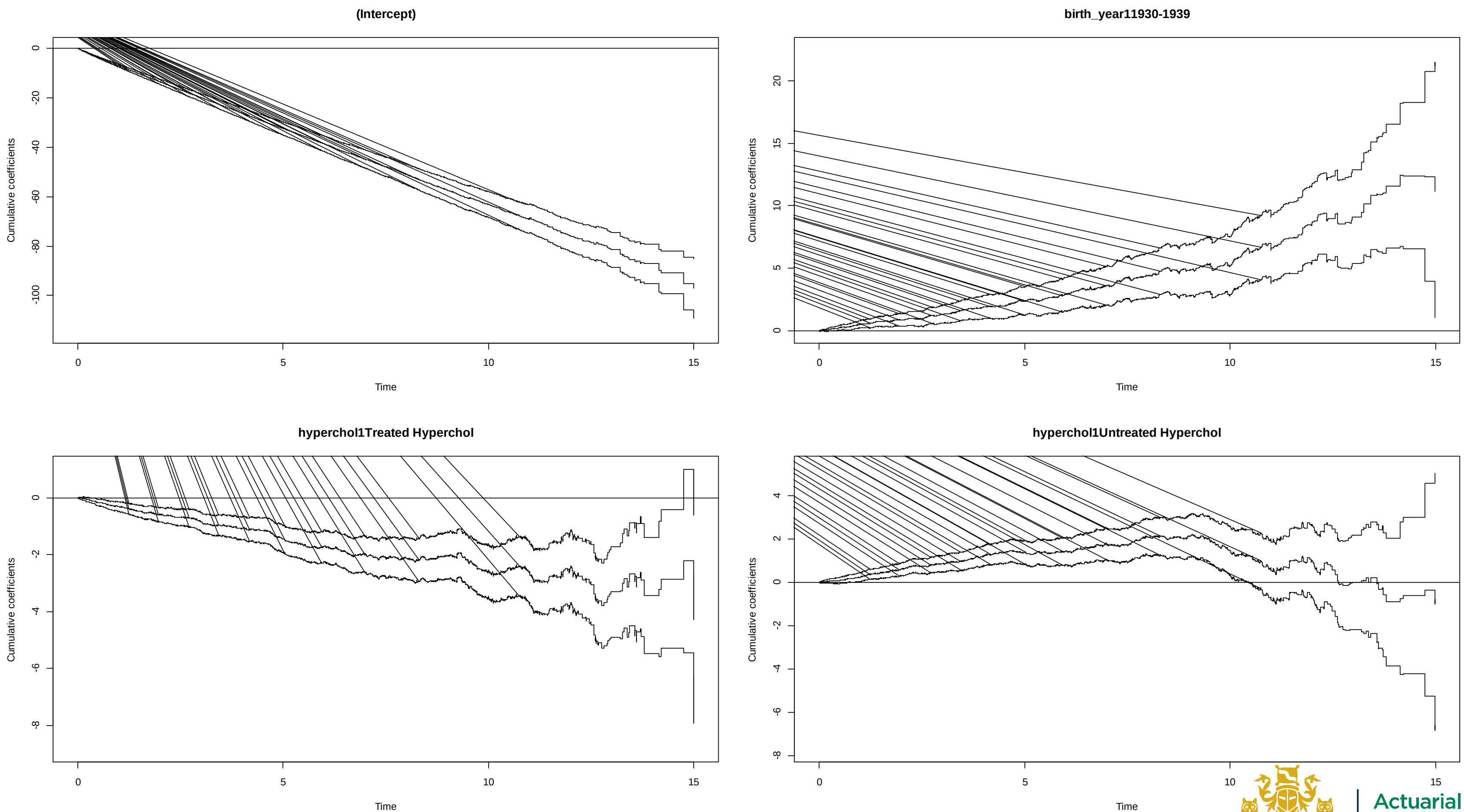
Actuarial Research Centre
Institute and Faculty of Actuaries

Inference in extended Cox model

- It is easier to estimate cumulative regression coefficients $B(t)=\int_0^t \beta(s)ds$, their estimates are $n^{1/2}$-consistent and asymptotically Normal.
- This allows to draw confidence bands for B(t) and to test hypotheses about them.
- A simple test of $\beta_p(t)=\beta_p$ is based on maximum deviation of the cumulative coefficient $B_p(t)$ from a straight line over an interval [0,T].
- Similarly, cumulative residuals are used for various diagnostic purposes.

Actuarial Research Centre
Institute and Faculty of Actuaries

Plots of cumulative coefficients for DM2 study



Actuarial Research Centre
Institute and Faculty of Actuaries

Robustness of the Cox model

Consider once more the extended Cox model

$$\mu(t|\beta, Z) = \mu_0(t)e^{Z^{\mathrm{T}}\beta(t)}.$$

The cumulative hazard *M(t|Z)=-ln(S(t|Z)*. The ratio

$$\frac{M(t|Z)}{M_0(t)} = \frac{\int \mu_0(s)\exp(\mathrm{Z}^{\mathrm{T}}\beta(s))\mathrm{ds}}{\int \mu_0(s)\mathrm{ds}} \approx \frac{\exp\int \mu_0(s)\,(\mathrm{Z}^{\mathrm{T}}\beta(s))\mathrm{ds}}{\int \mu_0(s)\mathrm{ds}} = \exp(\mathrm{Z}^{\mathrm{T}}\bar{\beta}(\mathrm{t})),$$

where $\bar{\beta}(\mathrm{t}) = \frac{\int \mu_0(s)\beta(s)\mathrm{ds}}{\int \mu_0(s)\mathrm{ds}}$, if the variance $\frac{\int \mu_0(s)\,(\mathrm{Z}^{\mathrm{T}}\,(\beta(s) - \bar{\beta}(\mathrm{t})))2\mathrm{ds}}{\int \mu_0(s)\mathrm{ds}}$ is small. This means that the Cox model gives approximately correct predictions of surviving up to time t.

Actuarial Research Centre
Institute and Faculty of Actuaries

What is landmark analysis

In the landmarking approach, dynamic predictions for the conditional survival after t=t_{LM} is used on current information for all patients still alive just prior to t_{LM}. [Van Houwelingen, H. and Putter, H., 2011]

The sliding landmark model is the simple Cox model

$$h(t|x, t_{LM}, w) = h_0(t|t_{LM}, w)\exp(x^T \beta_{LM}), \qquad s \le t \le s + w$$

for the data set obtained by truncation at $s = t_{LM}$ and administrative censoring at t_{LM}+w.

$h_0(t|t_{LM}, w)$ is the baseline hazard or force of mortality.

This is a convenient way to obtain a dynamic prediction without fitting a complicated model with time-varying effects.

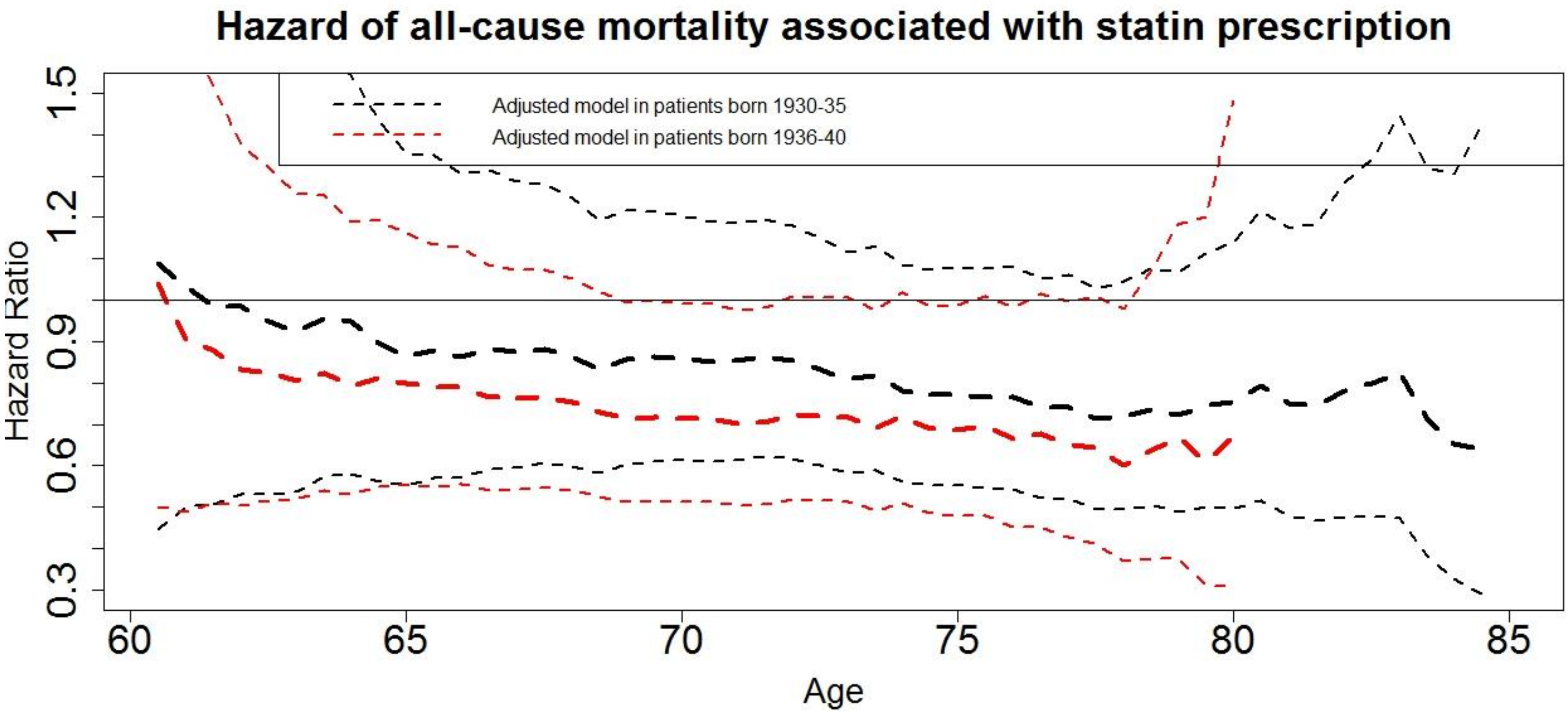
Actuarial Research Centre
Institute and Faculty of Actuaries

Super-prediction data set

- Fix the prediction window w; [say, w=5 years]
- Select a set of prediction time points $\{s_1, \ldots, s_L\}$, 20 ≤ L ≤ 100; [say, every 6 months.]
- Create a prediction data set for each $t_{LM}=s_l$ by truncation and administrative censoring;
- Stack all these data into a single "Super-prediction data set". The subsets corresponding to a given prediction time $t_{LM}=s_l$ are "strata".
- The risk set R(t_i) for an event time t_i is present in all strata with $s \leq t_i \leq s + w$. Passing from stratum s to s+1 corresponds to sliding the window over the time range.
- Individual life *j* contributes up to w/$|s_{l+1} - s_l|$ times in each prediction window. [10 times when w=5 and the time shift $s_{l+1} - s_l$ is 6 months.]

Actuarial Research Centre
Institute and Faculty of Actuaries

Sliding Cox model results (crude model)



Integrated partial log-likehood landmark model - ipl

The landmark super prediction model with window *w* and letting the regression coefficients β_{LM} depend on time t_{LM} is given by

$$h(t|x, t_{LM} = s, w) = h_0(t|s, w) \exp\left(x^T \beta_{LM}(s)\right), \qquad s \leq t \leq s + w$$

where $$\beta_{LM}(s) = \sum_{j=1}^{m} \gamma_j \; f_j(s).$$

- $f_j(s)$ are the basis functions, $f_1(s)$=1, $f_j(0)$=0 for j>1, and γ_j are the parameters, with $\beta_{LM}(0) = \gamma_1$.
- The parameters of this model are estimated by maximizing the integrated (over s) partial log-likelihood introduced by van Houwelingen (2007).
- This approach is based on a stratified (on s) analysis with smooth landmark dependent effect $\beta_{LM}(s)$ and separate estimated baseline hazards for each stratum.

Actuarial Research Centre
Institute and Faculty of Actuaries

Pseudo-partial log-likelihood landmark model - ipl^*

In the $\boldsymbol{ipl^*}$ model, the baseline hazard is modelled directly as

$$h_0(t|s,w) = h_0(t)\exp\big(\theta(s)\big),$$

where $\theta(s)=\sum_{j=1}^{m} \eta_j \; g_j(s)$

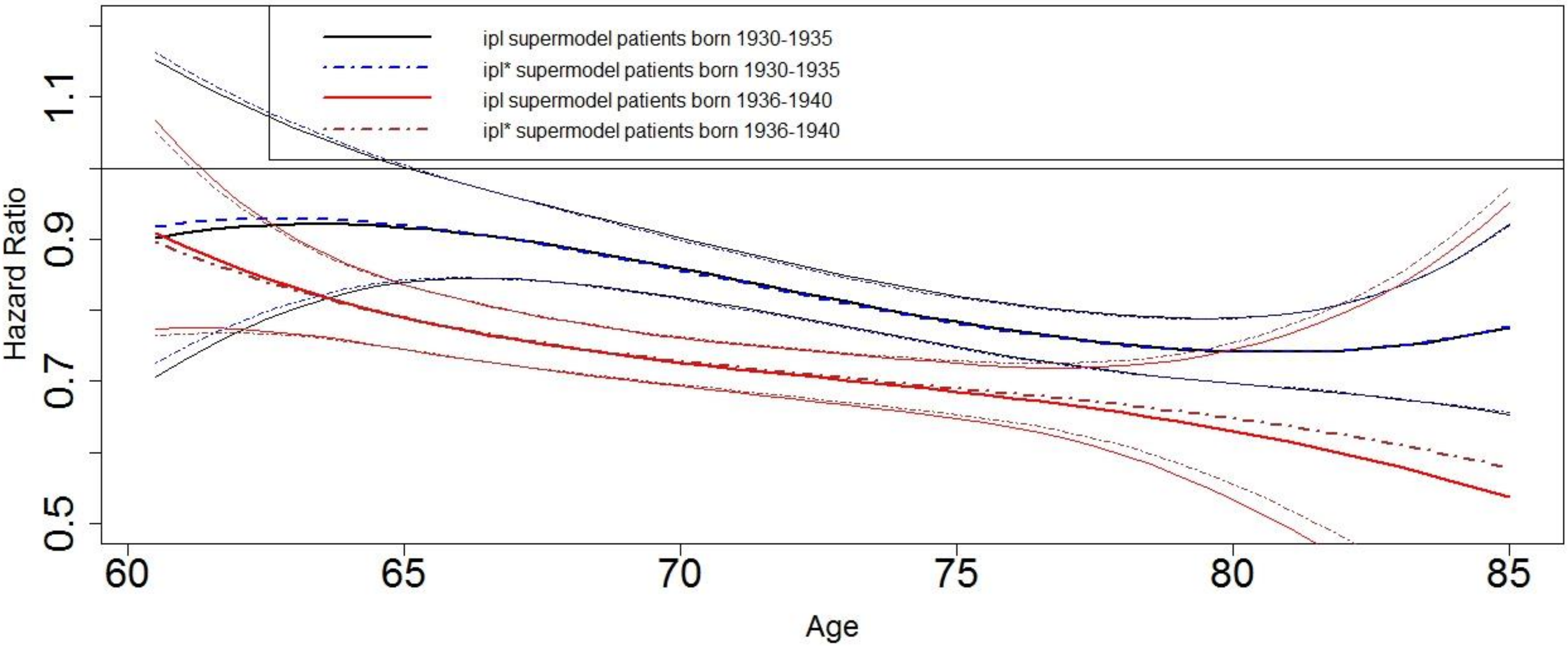
for proper basis functions $g_j(s)$ with $g_j(s_1)=0$, resulting in

$$h(t|x,t_{LM}=s,w) = h_0(t)\exp\big(x^T\beta_{LM}(s)+\theta(s)\big), \qquad s \le t \le s+w,$$

where $\beta_{LM}(s)$ and $\theta(s)$ are the mth degree polynomials in s.

Actuarial Research Centre
Institute and Faculty of Actuaries

Adjusted hazard of all-cause mortality associated with current statin prescription



Predicted probabilities of survival in a window

Predictions for all s ϵ $[s_1, s_L]$ in the $\boldsymbol{ipl}^*$ model are obtained from estimated cumulative hazards

$$H(s+w|x, t_{LM}=s) = \exp(x^T \beta_{LM}(s) + \theta(s))\ (H_0^*(s+w) - H_0^*(s-))$$

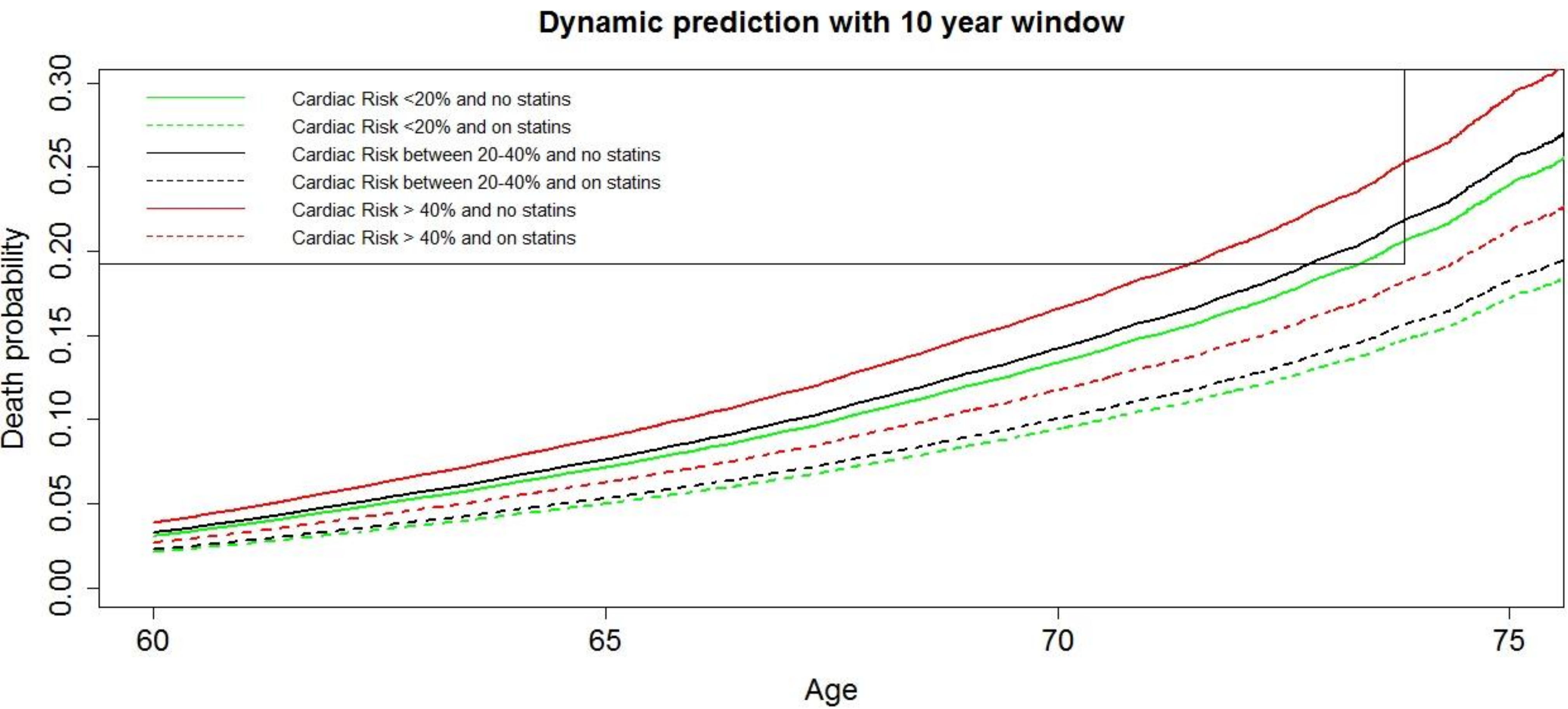
This is because in the $\boldsymbol{ipl}^*$ model

$$h(t|x, t_{LM}=s, w) = h_0(t) \exp\left(x^T \beta_{LM}(s) + \theta(s)\right), \qquad s \leq t \leq s+w,$$

only the baseline hazard $h_0(t)$ depends on t.

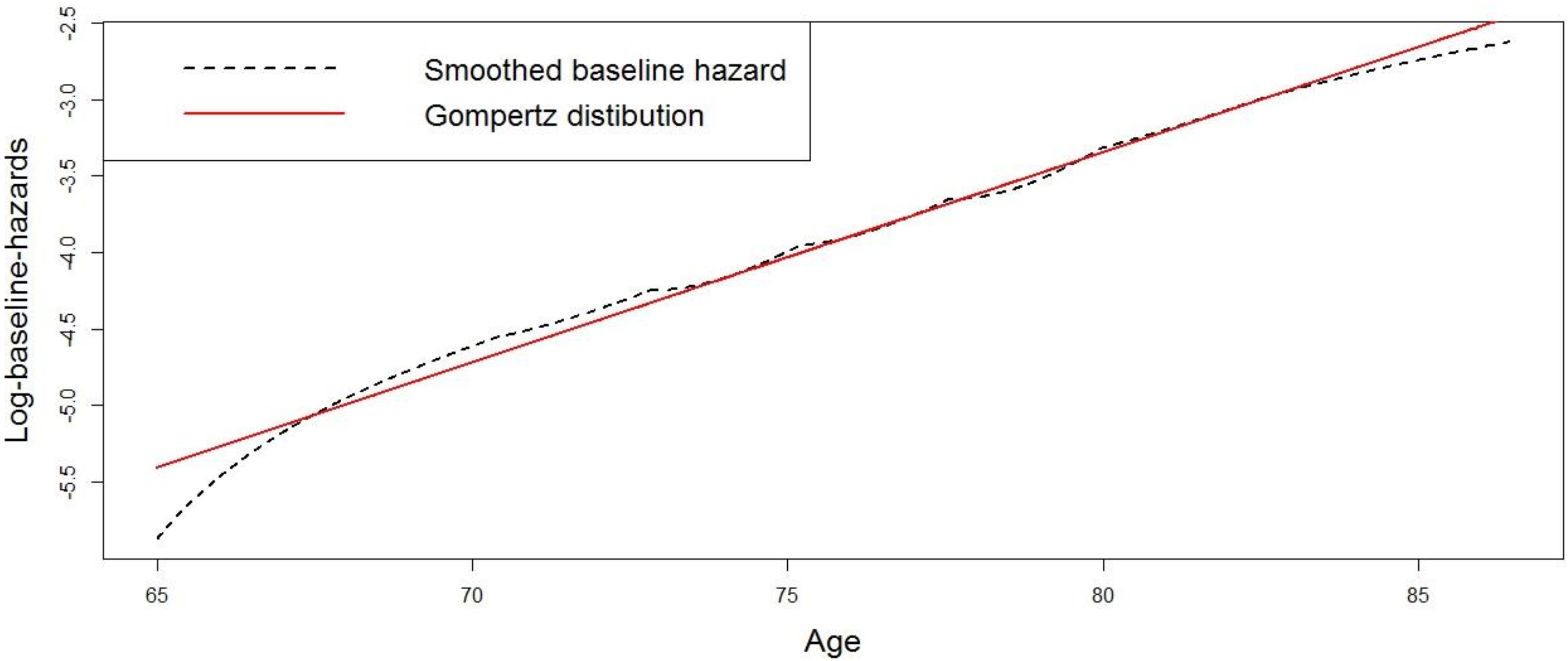
Actuarial Research Centre
Institute and Faculty of Actuaries

Probabilities of death for 1936-1940 cohort



Actuarial Research Centre
Institute and Faculty of Actuaries

Baseline hazard in the statins landmark model



The baseline hazard is well approximated by the Gompertz hazard

Actuarial Research Centre
Institute and Faculty of Actuaries

The $\boldsymbol{ipl^*}$ landmark model in actuarial research

In the $\boldsymbol{ipl^*}$ model, the hazards are modelled as

$$h(t|x, t_{LM} = s, w) = h_0(t)\exp\left(x^T(s)\beta_{LM}(s) + \theta(s)\right), \qquad s \le t \le s + w,$$

where $\beta_{LM}(s)$ and $\theta(s)$ are the *k*th and the (*k-1*)th degree polynomials of $s = t - t_0$.

The log-hazards are $\lambda(t|x, t_0)$= λ_0(t)+$x^T(s)\beta_{LM}(s)$ +$\theta(s)$.

For Gompertz baseline hazard, λ_0(t)=*a*+*bt*.

Values of *a* and *b* can be estimated from the estimated baseline hazard or substituted for a particular population. Next, we can obtain cumulative hazards, survival and period life expectancy for various scenarios of changing risks x(s).

Discussion and conclusions

- The most general form of extended Cox regression with time-dependent effects is difficult to use. To make it relevant to actuarial research we also need to consider the shape of the baseline hazards.
- Parametric "double-Cox" model is a useful replacement for the stratified Cox model which also models shape of baseline hazards and can be easily used for actuarial purposes.
- Landmark analysis is a convenient way to model dynamically changing survival data. The ipl* model conveniently lends itself to actuarial modelling.
- Extra development is required to use the results for population LE projections using methodology similar to that in Kulinskaya et al. (2019) .

References:

1. Begun A., Kulinskaya E. and MacGregor A. , 2019. Risk-adjusted CUSUM control charts for shared frailty survival models with application to hip replacement outcomes: a study using the NJR dataset. *BMC Medical Research Methodology, in print.*
2. Martinussen T, Scheike TH. *Dynamic Regression Models for Survival Data*. Springer: New York, 2006.
3. Van Houwelingen, H.C., 2007. Dynamic prediction by landmarking in event history analysis. *Scandinavian Journal of Statistics*, *34*(1), pp.70-85.
4. Van Houwelingen, H. and Putter, H., 2011. *Dynamic prediction in clinical survival analysis*. CRC Press.
5. Gitsels, L.A., Bakbergenuly, I., Steel, N. and Kulinskaya, E., 2019. Dynamic prediction of long-term survival effects of statin therapy: a landmarking analysis in the general population. *British Medical Journal* (under review)
6. Kulinskaya E., Gitsels, LA. and Bakbergenuly, I., 2019. Calculation of changes in individual and period life expectancy based on proportional hazards model of an intervention. *Insurance Mathematics and Economics,* under review.

Actuarial Research Centre
Institute and Faculty of Actuaries

Questions

Comments

Actuarial Research Centre
Institute and Faculty of Actuaries