Assessing non-linear and time-dependent effects of a sparsely measured continuous time-varying covariate

Yishu Wang * **& Michal Abrahamowicz**

Department of Epidemiology & Biostatistics, McGill University, Montreal, CANADA

Support: Canadian Institutes for Health Research (CIHR) grant PJT-148946

Background: links to STRATOS goals

Long-term objectives of the STRATOS initiative include (among other)
[Sauerbrei et al, Stats Med 2014]:

- **Identifying "unmet (analytical) needs"** i.e. those challenges, relevant for the analyses of complex observational studies, that **need further methodological developments**
- **Stimulating collaboration between different Topic Groups (TG) and/or Panels** whose **Joint Expertise** will be necessary to address such new analytical challenges
- **Providing evidence-based guidance regarding (new or existing) Methods** that may help addressing specific challenges **and Conditions for their use** (e.g. assumptions re: data structure etc. under which specific methods may be appropriate or not)

===

Our talk illustrates these 3 points in the context of:

- **Flexible modeling of the effects of continuous time-varying covariates in survival analysis**

Background

- Prospective (Prognostic and Etiologic) cohort studies increasingly include repeated measures of important risk factors/exposures/treatments

- Typically, these repeated measures are adjusted as time-varying covariates (TVCs) in the Cox Proportional Hazards (PH) model, and the TVC is represented by its most recently observed value *[Andersen & Liesol, 2003]*

- Conventional Cox PH model imposes the restrictive assumptions that:
  i. the Hazard Ratios are constant during the follow-up (PH assumption)
  ii. the effect of a Continuous TVC on the log hazard is Linear (Linearity assumption)

- In the past 25 years, **for time-fixed covariates**, many studies *[e.g. Gray JASA 1992; Grambsch & Therneau 1994; Abrahamowicz & MacKenzie 2007]* :
  i. proposed **flexible extension of the Cox model that avoid such restrictive assumptions**
  ii. yielded **evidence of frequent violations of the PH and/or Linearity assumptions**

- **In contrast, to date, flexible modeling of the effects of continuous time-varying covariates received little attention**

Objectives of our work & Links with specific STRATOS TG's/Panels

We propose a new flexible model to estimate the effect of a continuous time varying covariate (TVC) in time-to-event analyses (TG8: Survival Analysis) that simultaneously accounts for:

i. possibly **Non-Linear (NL) functional form** of its association with log hazard **(TG2: Functional Forms & Variable Selection)**

ii. potential **Time-Dependent (TD) effect** i.e. changes over time in the strength of this association **(TG8: Survival Analysis)**

iii. **specific Measurement Errors** induced when the previously observed TVC value is used as a 'proxy' for its un-observed current value **(TG4: Measurement Errors)**

iv. In addition, the **New Methods must be Evaluated in Simulations**, under a variety of clinically plausible assumptions **(Simulation Panel)**

Flexible modeling of NL & TD effects of a continuous TVC

- We adapt the flexible "product model" *[Abrahamowicz & MacKenzie 2007]* to the modeling of a Time-Varying Covariate (TVC):
- Hazard at time *u*, conditional on the current TVC value *X(u)*, is defined as:

$$\lambda[u|X(u)] = \lambda_0(u)\,\exp\{\beta(u)g(X(u))\} \quad (1)$$

  - ***β*(·)** is a smooth function of follow-up time representing the **TD effect,** which describes how the strength of the TVC effect changes over time
  - ***g*(·)** is a smooth function representing **the functional form** for (possibly **NL) effect** of the 'current' value of a continuous TVC on the log hazard at a time *u*

Extending flexible model (1) to account for Time Elapsed since Last observation

- For a continuous TVC *X(t)*, at any time *u* during the follow-up, we define:

$$u^* = \max_{t \le u}(t)\ \{t : \text{time when measurements of } X(t) \text{ are available}\},$$

i.e. *u* = time of the most recent TVC measurement Before*

Time Elapsed since Last observation (TEL) *[de Bruijne 2001]* **: *TEL*(*u*) = *u* – *u***

- TEL is considered an Effect Modifier for the Strength of the Joint NL/TD effect of the TVC on the hazard
- **Accordingly, we extend the flexible product model in (1) to propose *simultaneous* estimation of *TEL*, *TD, and NL* effects:**

$$\lambda[u|X(u^*)] = \lambda_0(u)\ \exp\{\gamma(u\text{-}u^*)\ \ [\beta(u)g(X(u^*))]\} \quad (2)$$

  - $\gamma(u - u^*)$ represents the effect of TEL,
  - $\beta(\cdot)$ represents the TD effect
  - $g(\cdot)$ represents the NL effect

Regression spline modeling

- All three functions are modeled with (un-penalized) quadratic regression B-splines, so the model (2) can be written as:

$$\lambda[u|X(u^*)] = \lambda_0(u)\exp\left\{\sum_{\mathrm{j}=1}^{\mathrm{J}} a_j A_j(u-u^*) * \left[\sum_{k=1}^{K} b_k B_k(u) * \sum_{l=1}^{L} c_l C_l\left(X(u^*)\right)\right]\right\} \qquad (3)$$

  - A_j, B_k, and C_l are the quadratic B-spline basis
  - a_j, b_k, and c_l (j=1,...,J; l=1,...,L; k=1,...,K) are the coefficients of the corresponding splines to be estimated

Alternative Conditional Estimation

- We estimate the 3 functions through the 3-step iterative Alternating Conditional Estimation (ACE) algorithm, that extends the previous 2-step ACE *[Wynant & Abrahamowicz 2014]*:
  - **Step 1:** Estimate the NL effect $\hat{c}_l$, l=2,...,L, of the continuous TVC, conditional on TD and TEL effects $\hat{\boldsymbol{a}}$ and $\boldsymbol{b}$ estimated in steps 2 and 3, respectively, of the previous iteration
  - **Step 2:** Estimate the TD effect b_k, k=1,...,K, conditional on the NL effect $\hat{\boldsymbol{c}}$ estimated in step 1 of the same iteration and the TEL effect $\hat{\boldsymbol{a}}$ estimated in step 3 of the previous iteration
  - **Step 3:** Estimate the TEL effect $\hat{a}_j$, j=1,...,J, conditional on the NL and TD effects estimated in, respectively, steps 1 and 2 of the same iteration

Selection of NL and/or TD effects into the final model

- In most applications, it is unknown whether the TVC of interest has truly NL and/or TD effects
- TD and NL estimates depend on what other effects are adjusted for *[Abrahamowicz & MacKenzie 2007]*
- Thus, **we adapt Backward Elimination to select specific effects into the final model** *[Royston & Sauerbrei 2008; Wynant & Abrahamowicz 2014]*
- Selection starts with full model and at each step the least significant TD or NL effect is eliminated until all remaining effects are significant ($p<0.05$)
- Likelihood Ratio Tests are used to test each effect, while adjusting for all other effects in a given model

Simulation studies: general design

- Hypothetical long-term cohort study of an association between a **single continuous TVC** and the time to a clinical event
- **N=1,000** subjects followed for up to 300 months
- 40% administrative censoring at the end of follow-up (**~ 600 un-censored events** per sample)
- **4 different data generating scenarios** (A1-A4), each corresponding to a different true model **for relationship between the current TVC value *X(u)* and the log hazard at time *u***:
  - (A1) Constant-over-time (as per the PH assumption) and Linear
  - (A2) Constant-over-time and NL
  - (A3) TD and Linear
  - (A4) TD and NL

Simulations: data generation & objectives

- The **hazard at time *u* depends only on the current TVC value *X*(*u*)**, updated every month (*u*=1,...300*)*
- **Event times were generated using the Permutation Algorithm**, specifically designed for time-to-event simulations involving TVC's *[Sylvestre & Abrahamowicz Stat Med 2008; CRAN R package: Sylvestre et al 2015]*
- **100 samples** independently generated & analyzed for each scenario
- In analyses, **we assumed TVC was measured every: 12 OR 36 months**
- **2 different Simulation Studies:**

  i. Preliminary Simulations: to assess the **ability of the Backward Elimination to identify the "Correct" final model**

  ii. **Main Simulations:** to assess the **accuracy of the TD and NL estimates under a correctly specified model**

Preliminary simulations (results of backward selection): % samples with a given effect selected

Scenario (true model)	Time elapsed between consecutive TVC measurements					
	12 months			36 months		
	TD effect selected	NL effect selected	True model Selected	TD effect selected	NL effect selected	True model selected
A1 (PH + linear)	4*	1*	95	5*	5*	92
A2 (PH + NL)	5*	99	94	4*	82	78
A3 (TD + linear)	100	4*	96	99	3*	96
A4 (TD + NL)	95	100	95	79	88	70

- Type I error rate

- Conclusions:
  - **Correct Type I Error** (**≤ 5%** in all cases)
  - good **~ ≥80% Power (with 600 events), even with Very Sparse data** (36 months) & **70%-96% probabilities of Selecting the "True model"**

Recommendations from the STRATOS Simulation Panel

Main simulations were designed in an attempt to "comply" with the recommendations of the STRATOS Simulation Panel, based on the recent Letter by **Boulesteix, Binder, Abrahamowicz & Sauerbrei [*Biometrical J* 2018]:**

The over-arching Goal of the Simulation Panel is to advocate more wide-spread use of *neutral* (un-biased) and '*realistic*' comparison studies evaluating the performance of existing and new statistical methods using (mostly) Simulated or Real-life data.

In analogy to the guidelines for RCT's [e.g. CONSORT] evaluating new treatments, **statistical simulations should meet several criteria, and provide answers re:**

(i) **How to simulate data in a realistic way**, inspired from real datasets?

(ii) How to ensure the **reproducibility and transparency of the methods used for data generation and analyses ?**

(iii) What are typical **sources of potential biases** and how can they be avoided?

(iv) How can the results be interpreted, without **the risk of over-interpretation**?

(v) **What parameters & assumptions should be varied** across simulated scenarios?

(vi) What **range of sample sizes** should be considered?

(vii) **Which "competing methods"** should be considered?

(viii) **etc. ...**

Design main simulations to "comply" with Simulation Panel criteria

- **General Objective**: to assess the accuracy of the effect estimates assuming ** the model was correctly specified

  (** consistent with encouraging results of Preliminary Simulations)

- **Main Simulations were designed so as to comply with most of the aforementioned criteria:**

  i. <u>re: Realistic Assumptions:</u> results of analyses involving sparse TVC measurements may depend critically on the stochastic nature of the repeated measurements and their auto-correlation structure *[Andersen & Liestol 2003]* . Thus, to ensure Clinical Plausibility of simulated data, and results, **the repeated TVC measurements were generated using Empirical Data on biennial Systolic Blood Pressure (SBP) measurements observed over 50 yrs of follow-up in a Random Sample of 1,000 participants of the Framingham Heart Study (FHS)** *(details below)*

Design main simulations (cont-d) to "comply" with Simulation Panel criteria

i. re: Reproducibility & Transparency: all details of assumptions and data generation/analyses are provided in Yishu Wang's PhD thesis (McGill University) and R programs are available on request;

ii. re: Exploring/Correcting for sources of Bias and (iv) Interpretation: as sparsity of TVC measurements is the main source of bias, we analyzed each simulated sample twice, using (a) complete TVC history (values updated each month) *vs.* (b) actual observed values in the FHS data; we then explored if TEL estimates may help interpreting the latter results & correcting for bias due to sparsity;

iii. re: Varying the essential assumptions: because results of flexible modeling depend largely on the Complexity of the "True Data-Generating Model' *[Binder et al, Stat in Med 2013]*, in 4 scenarios (A1-A4) we generated data assuming various combinations of (a) either Linear or NL, and (b) either constant (PH) or TD effects;

iv. re: Sample Size: we assumed a Realistic Effective Sample Size of ~ 600 events, typical of moderately sized clinical/epi Cohort Studies.

Figure 1. Estimated NL & TD effects in scenarios A2-A4, assuming the complete history of TVC measurements is available and the models were correctly specified. Grey curves represent individual estimates from 100 simulated samples, the white curve represents the mean estimate, and the black curve is the true effect.

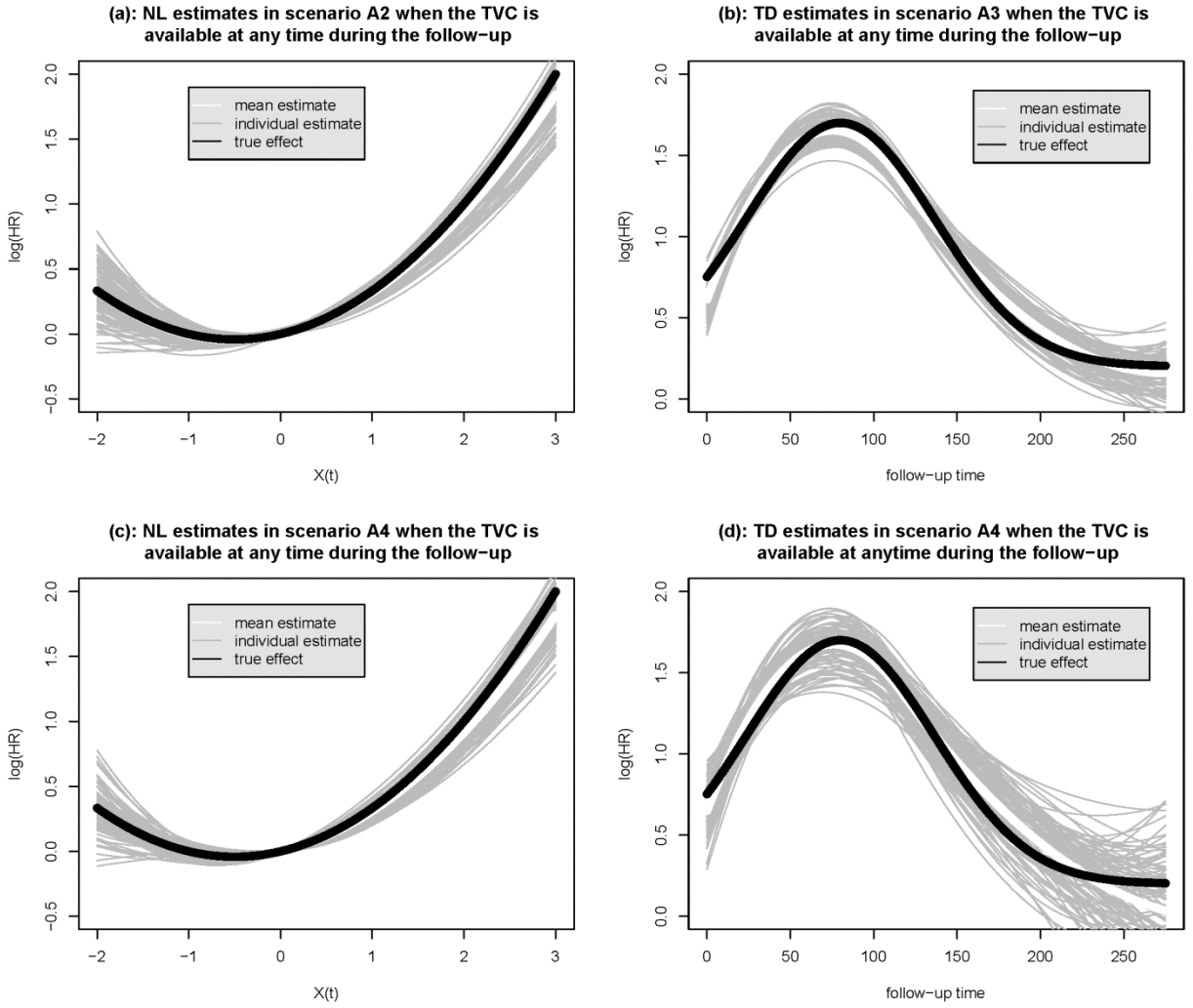
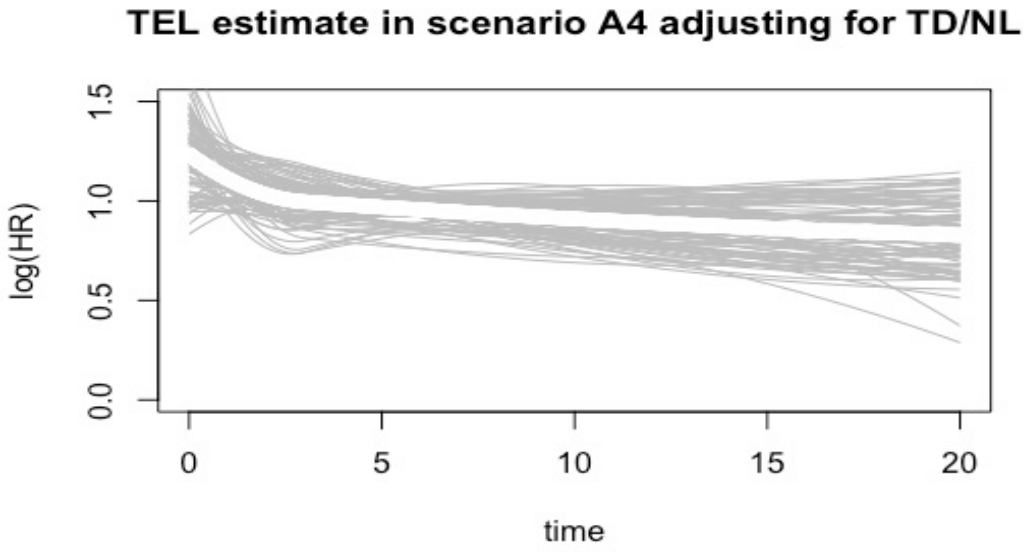


Figure 2. Estimated TEL effects in scenario A4



To correct for the attenuation due to “measurement error” specific to sparse TVC setting *[Andersen & Liestol 2003]*:

- we propose to re-construct what would be the strength of the effect if TVC values were always available (implying TEL=0) by calculating:

$$\hat{\gamma}(0)*\hat{\beta}(u)*\hat{g}(x(u))$$

Figure 3. Estimated NL & TD effects in scenario A4 when the model was correctly specified but the measurements of TVC were sparse. (c) and (d) NL effects at 12 months and 60 months of follow-up: the mean of uncorrected estimates (solid gray curve) and the mean of the TEL-corrected estimates (dotted curve) vs. the true effect (solid black curve).

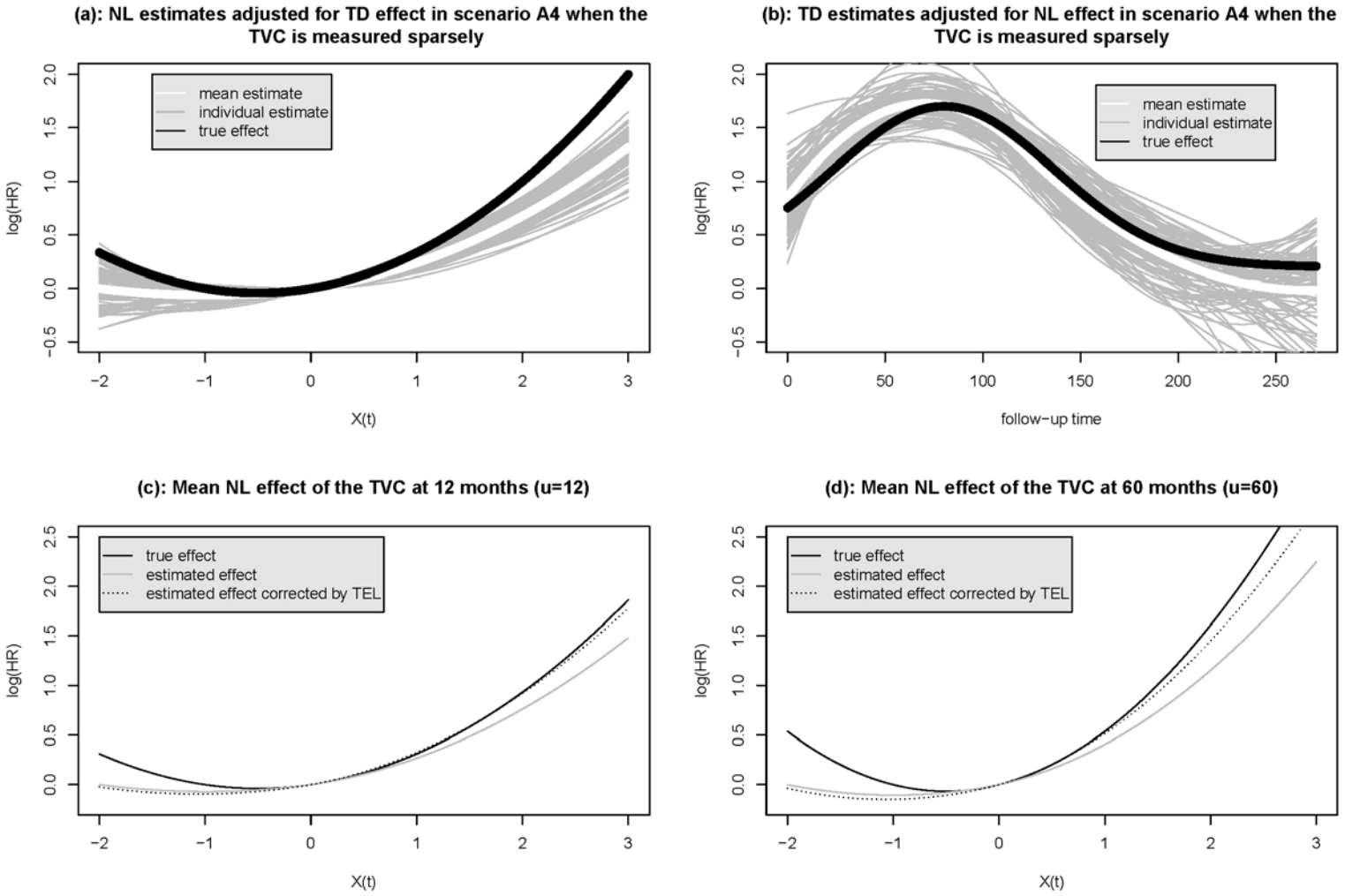
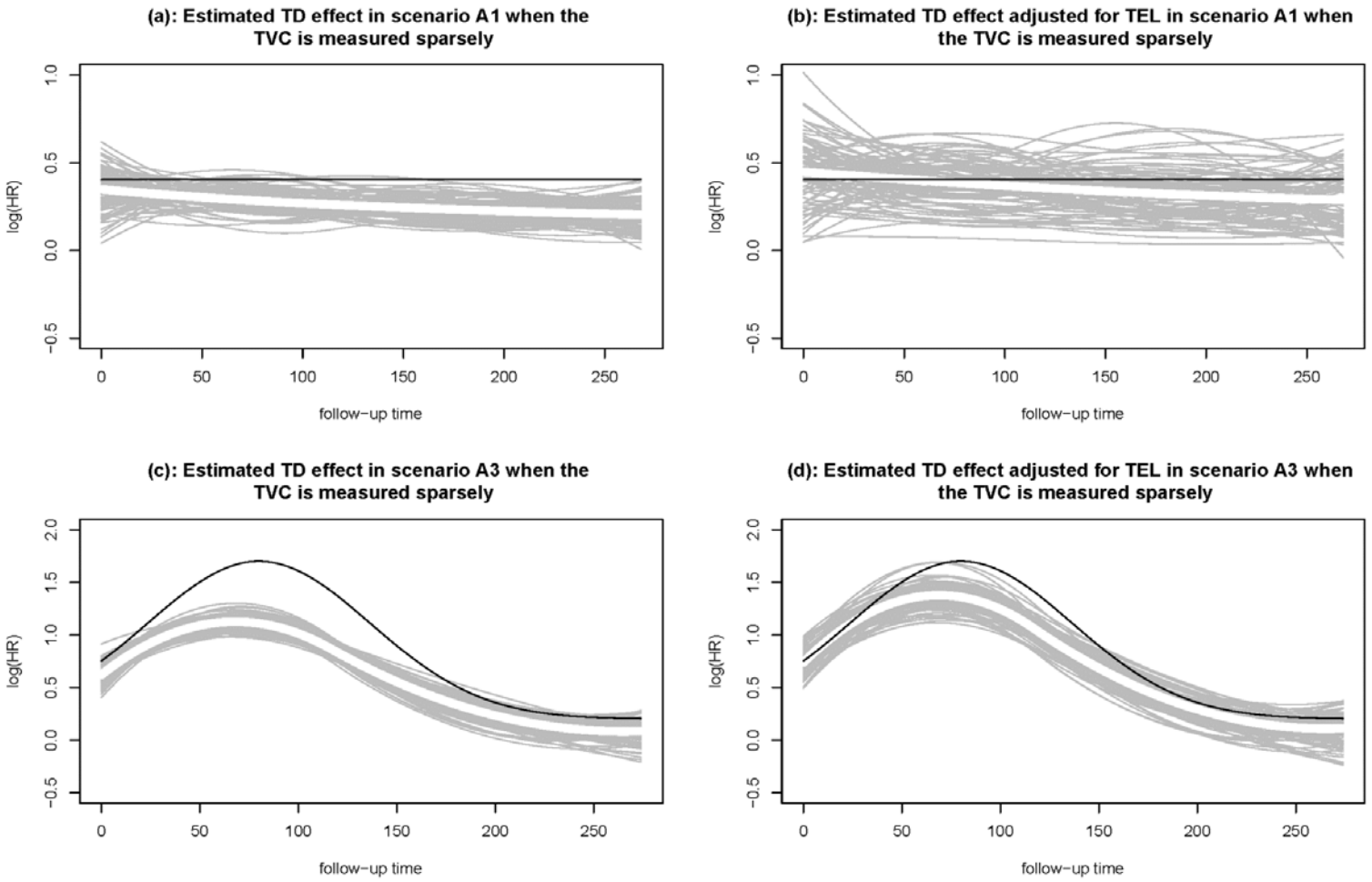


Figure 4. Estimated TD effects with a sparsely measured TVC in scenario A1 when there was no true TD effect (panels a and c), and in scenario A2 when there was a true (non-monotone) TD effect (panels b and d). (a) and (c) only TD effect, (b) and (d) TD effect adjusted for the TEL effect.



Application: re-assessing the effects of Systolic Blood Pressure and total cholesterol on the Cardiovascular Disease risks in women in the Framingham Heart Study (FHS)

- Objective: **to re-assess the effects of the updated values of time-varying, approximately biennial, measures of systolic blood pressure (SBP) and total cholesterol on the hazard of the first major cardiovascular disease (CVD) event in women from the original FHS cohort** *[Dawber et al 1951]*
- CVD was defined as a composite of *[Cupples & D'Agostino 1987]* :
  - coronary heart disease (coronary death, myocardial infarction, coronary insufficient, and angina)
  - cerebrovascular events (ischemic and hemorrhagic stroke, and transient ischemic attack)
  - peripheral artery disease and heart failure
- **Four time-varying risk factors** were included in all multivariable models**:**
  - systolic blood pressure (SBP)
  - total Cholesterol
  - current Age
  - current Smoking status (binary)

Application: Basic characteristics of the FHS data

- **N = 2,675** women, aged between 29-68 years at the cohort entry
- **Up to 28 biennial exam** visits, **median follow up time: 28.6 years**
- **1,547 (58%) developed CVD events** during the follow up (**incidence rate of 1.90/100 person-years**)
- Across all subjects and all visits:

  10% of SBP values, 32% of current smoking status, and 44% of total cholesterol values were missing.

  All the missing data were imputed using the last observation carried forward (LOCF) approach.
- median TEL: 16 months for SBP, 24 months for Cholesterol

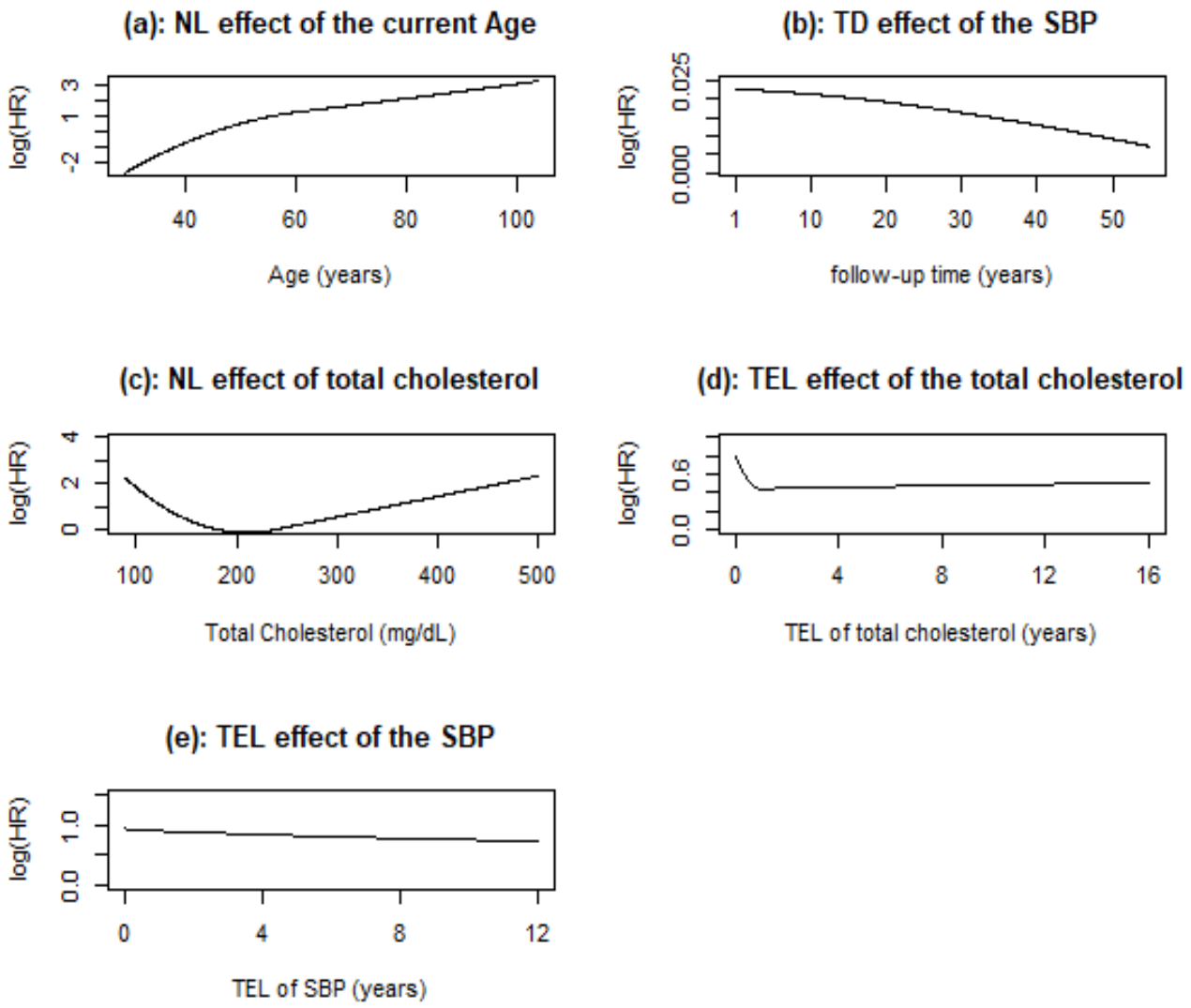
Summary of FHS analyses

- Final model selected by backward elimination (when TEL effects were not considered):
  - NL effect of Age (*p*=0.006)
  - TD effect of SBP (*p*=0.032)
  - TD (p=0.04) and NL (p<0.001) effects of total Cholesterol
  - current smoking status (HR=1.32, 95% CI: 1.20-1.45)

- **The flexible model fitted the data very significantly better than the conventional Cox PH model** that imposed the PH and linearity assumptions for all four risk factors (***p*<0.00001, for a 11-df LRT**), and had a much better **AIC (21,558.8 vs. 21,601.2**)

- When this model was expanded to include the TEL effects of total cholesterol and SBP, both TEL effects were significant (p<0.05), but adjustment for the TEL effect of total cholesterol made its TD effect non-significant (p=0.37)

Figure5. Estimated TD and/or NL effects of current age, SBP and total cholesterol, and the TEL effects of total cholesterol and SBP, on the hazard of the first CVD event in women. The NL effect of current age is plotted relative to the mean baseline age of 44.8 years, and the NL effect of current cholesterol is plotted relative to its mean value (241 mg/dL).



Conclusions

- Simulation results indicated that both TD & NL effects of a continuous TVC could be correctly estimated in the ideal case when the complete TVC history was available.

- When TVC measurements were sparse, the NL effect of the TVC was systematically underestimated compared to the true effect, but the TEL estimate could be used to help reduce this bias.

- The real-life **FHS application illustrates the insights offered by flexible modeling of potential TD and NL effects of the TVCs**, e.g.:
  i. the TD estimate suggests **impact of SBP decreases with aging**
  ii. the NL estimate indicates **CVD risks increase for women with either very Low or High for Cholesterol values**
  iii. TEL estimates suggested that, whereas the current value of total **cholesterol has an acute impact** on the hazard of CVD events, the **effect of SBP may be lagged or may cumulate** over values observed in the past several years.

Limitations

- **More simulation scenarios need to be investigated** in future research with *different realistic assumptions* about **(STRATOS Simulation Panel):**
  - number of TVCs included in the model
  - true covariate effects
  - frequencies of observed TVC measurements
  - sample sizes
  - stochastic structure of the longitudinal data (auto-correlation between consecutive TVC values)
  - comparison with alternative methods
- We assumed:
  - the *hazard is associated only with the* ***current TVC value***, but it may also depend on *past TVC values,* through their Cumulative/Lagged effects *[Wang & Abrahamowicz, "Flexible modeling of cumulative effects of a sparsely measured continuous TVC" ISCB 2018]*
  - the TVC is *not a mediator* of the effects of another time-varying exposure or risk factor
  - the *TVC values were measured without errors*

Present & Future links to other STRATOS-related developments

- Our study has close links to work by other members of STRATOS:
  - **TG8: Survival Analysis** (P.K. Andersen, T.Therneau, H. van Houwelingen)
  - **TG2: Functional forms & Variable Selection** (H. Binder, P. Royston & W. Sauerbrei)
  - **Simulation Panel** (A.L. Boulesteix)

- **Need for Future Collaboration with:**
  - **TG4 (Measurement Errors)** in 2 aspects:
    i. Specific measurement errors due to the *sparse measurements of TVC* during the follow-up;
    ii. *Actual measurement errors* of the observed TVC measurements
  - **TG1 (Missing Data)** to find more refined approaches to impute missing over-time TVC values than the LOCF

References

- Sauerbrei, W., Abrahamowicz, M., Altman, D. G., le Cessie, S., Carpenter, J., & STRATOS initiative. (2014). STRengthening analytical thinking for observational studies: the STRATOS initiative. Statistics in medicine, 33(30), 5413-5432.
- Abrahamowicz M, MacKenzie T. Joint estimation of time-dependent and non-linear effects of continuous covariates on survival. *Stat Med.* 2007;26(2):392-408.
- Maas AH, Appelman YE. Gender differences in coronary heart disease. *Neth Heart J.* 2010;18(12):598-602.
- Sylvestre MP, Abrahamowicz M. Comparison of algorithms to generate event times conditional on time-dependent covariates. *Stat Med.* 2008;27(14):2618-2634.
- Sylvestre MP, Evans T, MacKenzie T, Abrahamowicz M. PermAlgo: Permutational algorithm to generate event times conditional on a covariate matrix including time-dependent covariates. **R package version 1.1, 2015. (CRAN)**
- Towfighi A, Zheng L, Ovbiagele B. Sex-specific trends in midlife coronary heart disease risk and prevalence. *Arch Intern Med.* 2009;169(19):1762-1766.
- Andersen, P. K., & Liestøl, K. (2003). Attenuation caused by infrequently updated covariates in survival analysis. Biostatistics, 4(4), 633-649.
- Grambsch, P. M., & Therneau, T. M. (1994). Proportional hazards tests and diagnostics based on weighted residuals. Biometrika, 81(3), 515-526.

References

- de Bruijne, Mattheus HJ, et al. "On the use of Cox regression in the presence of an irregularly observed time-dependent covariate." Statistics in medicine 20.24 (2001): 3817-3829.
- Boulesteix, A. L., Binder, H., Abrahamowicz, M., Sauerbrei, W., & Simulation Panel of the STRATOS Initiative. (2018). On the necessity and design of studies comparing statistical methods. Biometrical Journal, 60(1), 216-218.
- Cupples LA, D'Agostino RB. Some risk factors related to the annual incidence of cardiovascular disease and death using pooled repeated biennial measurements: Framingham study, 30-year follow-up. *The Framingham Heart Study: An Epidemiological Investigation of Cardiovascular Disease.* Washington, DC: Government Printing Office; 1987:87-203.
- Dawber TR, Meadors GF, Moore FE. Epidemiological approaches to heart disease: the Framingham Study. *Am J Public Health Nations Health* 1951;41(3):279-286.
- Gray, R. J. (1992). Flexible methods for analyzing survival data using splines, with applications to breast cancer prognosis. Journal of the American Statistical Association, 87(420), 942-951.

THANK YOU

yishu.wang@mail.mcgill.ca

&

michal.abrahamowicz@mcgill.ca

Main Simulations: details of TVC generation based on real-life SBP values in Framingham Study

Generate the TVC based on the SBP measurements from the Framingham Heart Study (FHS):

i. randomly selected a subset of 1,000 / 5,209 participants of the original FHS cohort, and retrieved all their available SBP measurements

ii. imputed un-observed SBP values for each two-month interval (time unit) through linear interpolation between the two adjacent actual measurements for the same subject

iii. added normally distributed random errors to each imputed value

- Each sample data were analyzed twice, assuming, respectively:

i. Complete TVC history were available (for each month)

ii. Sparse TVC measurements, available only every 12 time units (~8% of all monthly values used to generate data)

Figure 6. Estimated TD effects based only on baseline values of the TVC: (a) estimated spurious TD effect when the true effect meets the PH assumption, i.e. is constant-over-time (black line), (b) systematically biased TD estimates (monotonically decreasing gray curves) when the true TD effect is non-monotone (black curve).

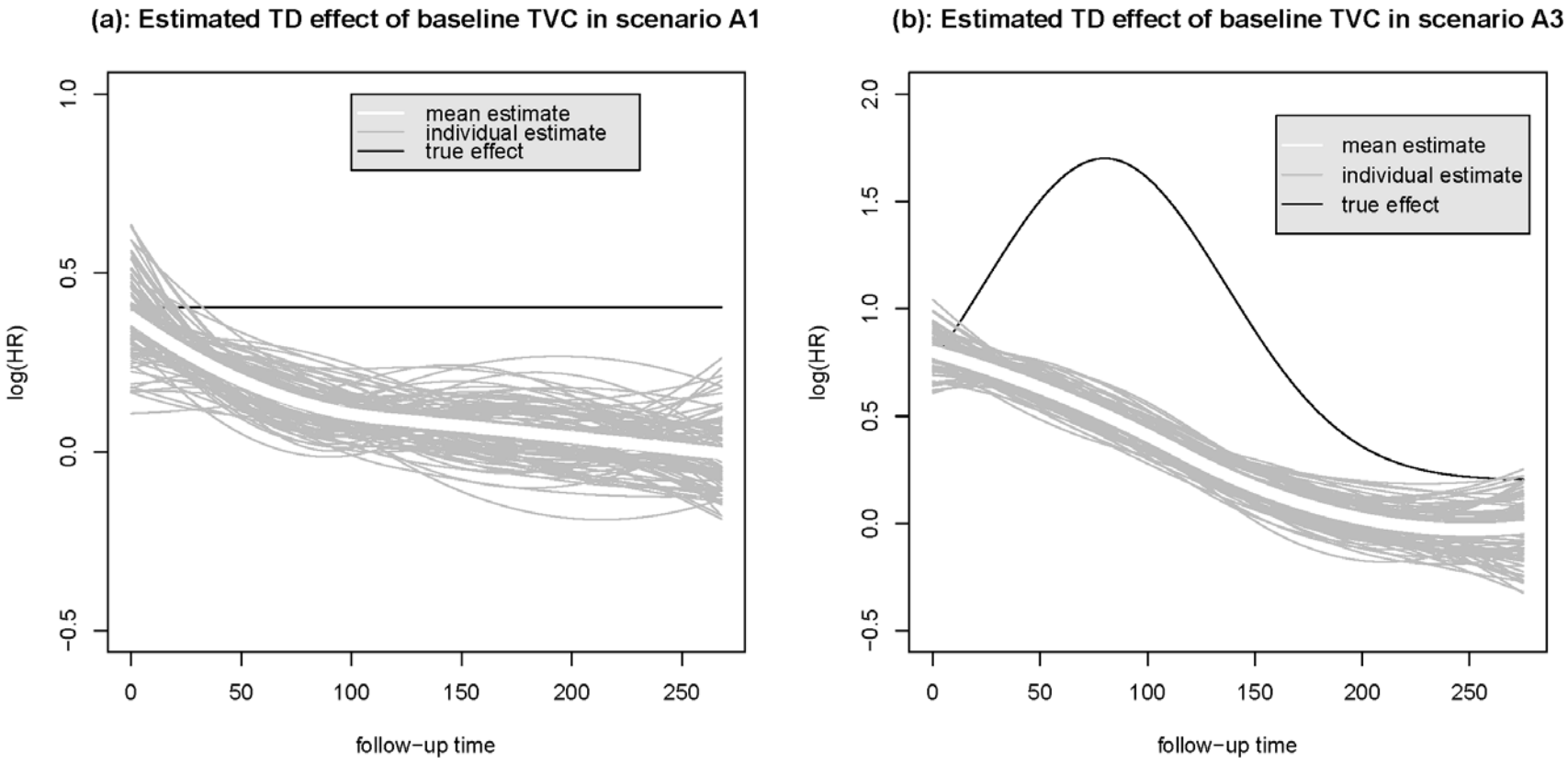


Figure 7. Estimated TD and/or NL effects of current age, SBP and total cholesterol on the risk of first CVD event in women, when individual woman's follow-up is censored at the time when the most recent measurement of total cholesterol has been observed more than 6 years ago

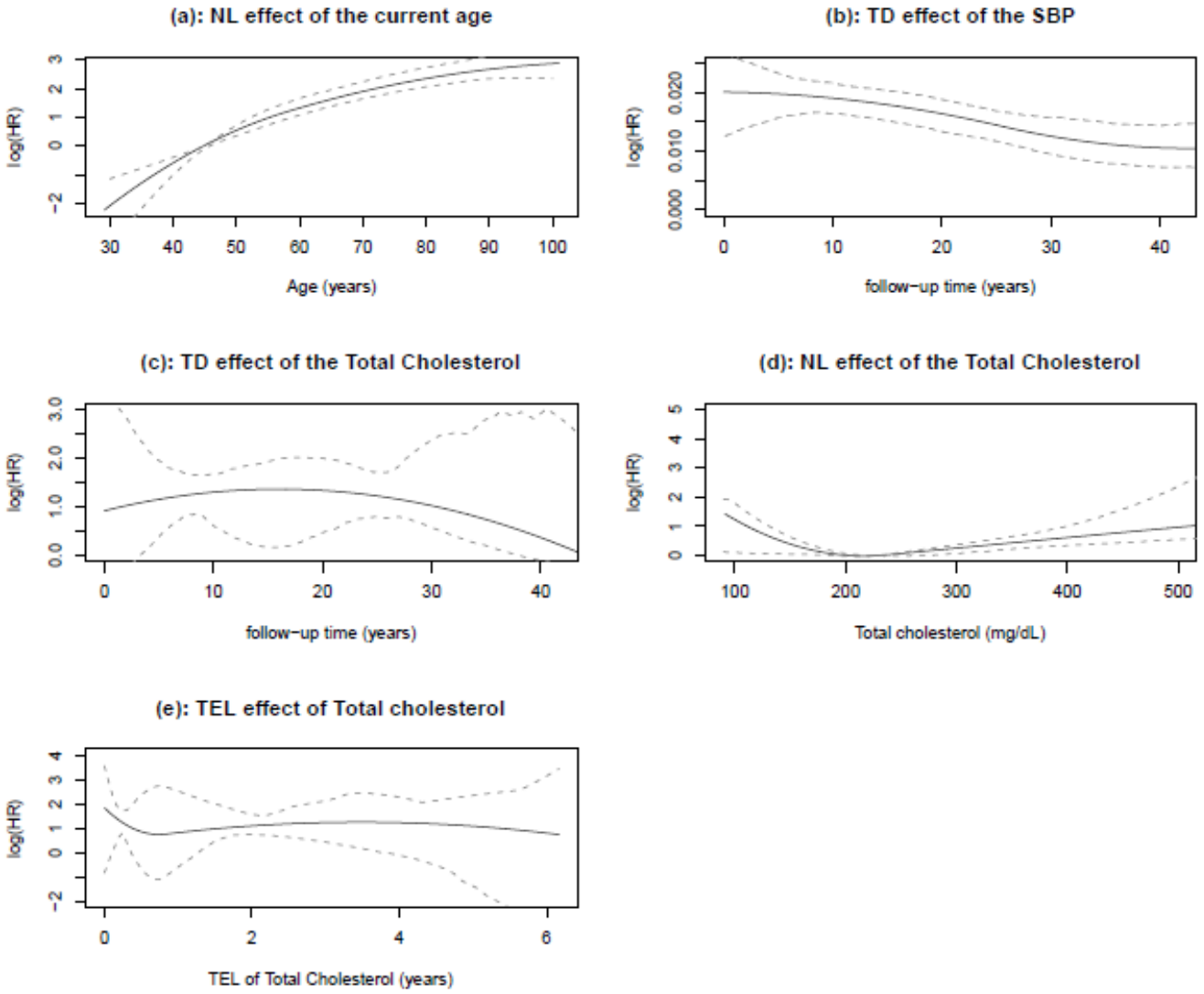


Figure 8. Estimated TEL effect when the hazard is associated with different temporal relationship of the TVC