

Flexible modeling of the cumulative effects of time-dependent exposures on the hazard

Marie-Pierre Sylvestre and Michal Abrahamowicz*,†

Department of Epidemiology, Biostatistics, and Occupational Health, McGill University, 1020 Pine Avenue West, Montreal, Que., Canada H3A 1A2

SUMMARY

Many epidemiological studies assess the effects of time-dependent exposures, where both the exposure status and its intensity vary over time. One example that attracts public attention concerns pharmacoepidemiological studies of the adverse effects of medications. The analysis of such studies poses challenges for modeling the impact of complex time-dependent drug exposure, especially given the uncertainty about the way effects cumulate over time and about the etiological relevance of doses taken in different time periods. We present a flexible method for modeling cumulative effects of time-varying exposures, weighted by recency, represented by time-dependent covariates in the Cox proportional hazards model. The function that assigns weights to doses taken in the past is estimated using cubic regression splines. We validated the method in simulations and applied it to re-assess the association between exposure to a psychotropic drug and fall-related injuries in the elderly. Copyright © 2009 John Wiley & Sons, Ltd.

KEY WORDS: survival analysis; proportional hazards model; regression splines; simulations; time-dependent covariates; pharmacoepidemiology

1. INTRODUCTION

The aim of many epidemiological studies is to assess the effects of time-dependent exposures, where both the exposure status and its intensity vary over time. Adequate modeling of such complex time-dependent exposures poses challenges because of the uncertainty about (i) whether and how

*Correspondence to: Michal Abrahamowicz, McGill University Health Centre, 687 Pine Avenue West, V building, Montreal, Que., Canada H3A 1A1.

†E-mail: michal.abrahamowicz@mcgill.ca, michal@epimgh.mcgill.ca

Contract/grant sponsor: National Sciences and Engineering Research Council of Canada; contract/grant number: 228203

Contract/grant sponsor: Canadian Institutes for Health Research; contract/grant number: MOP-8127

Contract/grant sponsor: Doctoral Fellowship (MPS)

Contract/grant sponsor: Fonds de recherche en santé du Québec (FRSQ)

Received 3 June 2008

Accepted 6 July 2009

exposure effects cumulate over time and (ii) the etiological relevance of exposures that occurred in different time periods [1–3].

Ignoring the complex time-varying nature of exposure or restricting its modeling to a subset of its components, such as current dose or duration of exposure, may miss more complex forms of association between exposure and outcome or lead to etiologically incorrect conclusions [2]. A more comprehensive representation of exposure history involves the cumulative exposure, calculated as un-weighted sum of past exposures [4, 5]. Yet, conventional cumulative exposure assumes that the impact of past exposures on the current risk does not depend on their timing, which may be debatable in many instances.

Both Breslow *et al.* [6] and Thomas [7] have discussed the concept of weighted cumulative exposure (WCE) that combines information about duration, intensity and timing of exposure into a summary measure. The WCE at time u is calculated by assigning weights to past exposures up to time u , and then summing up over time. Consider $x(t)$, the instantaneous intensity of exposure at time t . Let $w(u-t)$ be a function that assigns weights to exposures in the past, based on their impact on the risk of event occurring after the time interval $u-t$. If exposure is measured at discrete time units, then the WCE can be expressed as:

$$\sum_{t=1}^u w(u-t)x(t) \quad (1)$$

Notice that equation (1) represents a convolution and the weight function $w()$ can be considered as a specific example of the general concept of a familiar kernel function.

Vacek proposed parametric modeling of the weight function in case-control studies, where u was fixed at the index date and a single value of the WCE for each subject was incorporated in a conditional logistic regression model [2]. Four alternative parametric forms for the weight function were selected *a priori* and the fit of the resulting models were compared to select the specification most consistent with the data [2]. In the analysis of a case-control study of the impact of smoking on lung cancer, Hauptmann *et al.* used a more flexible WCE framework to summarize the individual smoking histories up to the date of cancer diagnosis for cases or index date for controls [3]. They estimated the weight function using constrained regression splines within a generalized linear model using a constrained maximization algorithm.

Abrahamowicz *et al.* used the parametric WCE framework within the Cox's proportional hazards (PH) model to refine the assessment of the associations between exposure to three benzodiazepines and fall-related injuries in the elderly [8]. The parametric form of the weight function was *a priori* selected to reflect assumptions about the pharmacokinetics properties of benzodiazepines. Two alternative parameter values were considered to represent different windows of clinically relevant exposure. The WCE models provided better fit to the data than time-dependent indicators of current dose or exposure duration [8]. The best-fitting WCE model was different for each of the three benzodiazepines, possibly reflecting the differences in their elimination half-life and/or withdrawal effects [8].

However, modeling strategies that impose a specific parametric form of the weight function, in the absence of solid prior knowledge about its shape, may lead to invalid results if the function is incorrectly specified [2]. An alternative approach consists in estimating the functional form of the weight function from the data, using flexible non- or quasi-parametric methods [3]. In the current article, we propose the regression spline-based method for modeling the WCE as a time-dependent covariate in the Cox's PH regression analysis of cohort studies. This involves calculating the WCE at each time u during the follow-up, in contrast to computing its value only once at the end of

follow-up, as it is done in the case-control framework [2, 3]. We propose a parametrization of the WCE in the Cox's PH model that does not require constrained optimization, so that it can be readily estimated using most statistical software packages.

Next section discusses the estimation of our model, the pointwise confidence bands for the weight function, and hypothesis-testing procedures. Section 3 presents the design and results of the simulations, which evaluated the performance of our method. Section 4 illustrates the application of the method to study the association between exposure to Flurazepam (a psychotropic drug) and fall-related injuries in the elderly. The paper ends with concluding remarks in Section 5.

2. METHODS

2.1. Cox's PH model with time-dependent WCE

Consider a cohort, in which individual exposure intensity or dose $X(t)$ varies over time. At any time u during the follow-up, we represent the joint effect of the past exposures by the WCE metric

$$\text{WCE}(u) = \sum_t^u w(u-t)X(t) \quad (2)$$

where $t \leq u$ indexes times of exposure preceding u , and $w(u-t)$ is the function assigning weights to past exposures, based on the time elapsed since the exposure occurred. The WCE is then modeled as a time-dependent covariate in Cox's PH model [9]

$$h(u|\mathbf{X}(u), Z(u)) = h_0(u) \exp \left[\beta \sum_t^u w(u-t)X(t) + \sum_{s=1}^q \eta_s Z_s(u) \right] \quad (3)$$

where $h_0(u)$ is the (un-specified) baseline hazard, $\mathbf{X}(u) = \{X(t), 0 \leq t \leq u\}$ represents the time-vector of the past exposures and $Z_s(u)$, $s = 1, \dots, q$ are the values of fixed-in-time or time-dependent covariates relevant for time u .

We assume that, in most clinical and epidemiological applications, the analytical form of the weight function $w(u-t)$ is unknown and its shape has to be estimated without imposing *a priori* constraints, except for smoothness. Therefore, we propose to estimate the weight function $w(u-t)$ using cubic regression B-splines [10]:

$$w(u-t) = \sum_{j=1}^m \theta_j B_j(u-t) \quad (4)$$

where B_j , $j = 1, \dots, m$, represent the m functions in the cubic spline basis and θ_j , $j = 1, \dots, m$, represent the estimable coefficients of the linear combination of the basis splines.

Conceptually, $w(u-t)$ in (3) estimates the shape of the weight function and β estimates the magnitude of the effect of cumulative exposure on the hazard. However, both β in (3) and θ_j in (4) reflect the strength of the association between the exposure and the risk. Therefore, to avoid identifiability problem when jointly estimating the parameters β and θ_j , we define:

$$\gamma_j = \beta \theta_j, \quad j = 1, \dots, m \quad (5)$$

Furthermore, estimation of model in (3) is largely facilitated by introducing artificial time-dependent covariates [11]:

$$D_j(u) = \sum_t^u B_j(u-t)X(t), \quad j=1, \dots, m \quad (6)$$

Given (5) and (6), the model in (3) is redefined into a familiar form of the PH regression model with time-dependent covariates $D_j(u)$:

$$h(u|\mathbf{X}(u), Z(u)) = h_0(u) \exp \left[\sum_{j=1}^m \gamma_j D_j(u) + \sum_{s=1}^q \eta_s Z_s(u) \right] \quad (7)$$

Once the m artificial time-dependent covariates $D_j(u)$ are calculated, any standard statistical software for Cox's PH regression with time-dependent covariates can be used to estimate the parameters of the model (7).

It should be noted that as β is absorbed in the γ vector, the magnitude of the effect of cumulative exposure on the risk of an event cannot be summarized by a single parameter. However, measures of relative risks, such as hazard ratios (HR), may easily be calculated from model (7) by comparing estimates for different exposure patterns, while keeping other variables constant. Specifically, at any time u during follow-up, the HR for individuals with the past exposure vectors of, respectively, $X_1(t)$ and $X_0(t)$, $t \leq u$, can be estimated as:

$$\exp \left[\sum_{j=1}^m \hat{\gamma}_j \sum_t^u B_j(u-t)[X_1(t) - X_0(t)] \right] \quad (8)$$

where $\hat{\gamma}_j$ are parameters estimated from model (7). Depending on the selected exposure histories, one can estimate the HR resulting from, for example, (i) changing the dose, (ii) changing the duration and timing of exposure or (iii) any combination of (i) and (ii). The corresponding 95 per cent confidence interval may be obtained via bootstrap, as described below.

2.2. Choice of the cubic regression spline basis

We estimate the weight function in (4) using cubic regression splines because they provide smooth estimates with continuous first two derivatives that are flexible enough to represent a variety of clinically plausible shapes [12], and can easily be computed by any statistical package. The number of knots determines the flexibility of the estimated spline function and the model's degrees of freedom. A cubic B-spline basis with $m-4$ interior knots consists of m curves, whose linear combination provides the estimated spline function. In general, not more than five knots are enough to model a smooth uni- or bi-modal curve while reducing the risk of major over-fitting bias [13].

Regression splines have a finite support interval, which has to be defined by the user [14]. In our framework, the support interval corresponds to the time window $[u-a, u]$ of exposures that are considered potentially etiologically relevant at time u . Outside this interval, that is, for $t < u-a$, the exposures are *a priori* believed to be too remote in time to influence the risk of an outcome at time u so that $w(u-t)$ is *a priori* set to 0 for $t < u-a$. Therefore, the weight function in (4) is estimated over the limited interval $[0, a]$ only. Accordingly, we place four exterior knots at both 0 and a . In the absence of prior knowledge about the plausible shape of the weight function, we place the interior knots at equal distances within the time window $[0, a]$. If, based on prior knowledge, the user expects that the weight function might have a local extremum or an inflection

point near a particular time t^* , then it will be advisable to place a knot at or near t^* . In addition, in many applications, it may be desirable to consider a few alternative time windows of etiologically relevant exposures and to estimate a separate WCE models for each of these windows. The WCE model with the best fit can be then selected. In Sections 2.5 and 2.6, we describe how the inference about the WCE estimates should account for *a posteriori* model selection including the selection of the support interval and/or the number of interior knots.

2.3. Constraining the weight function

In many, but not all applications, it may be *a priori* evident that the weights should smoothly decrease to zero in the right end of the support interval. Therefore, in addition to the unconstrained weight function models, in which the coefficients γ_j for all m artificial time-dependent variables in (6) are estimated, we also considered constrained models where $\gamma_{m-1} = \gamma_m = 0$, which results in constraining the weight function, and its first derivative, to reach 0 at $t = u - a$ [15].

Furthermore, in some applications, it may be deemed necessary to ensure that the estimated weight function is non-negative across the entire etiologically relevant time window $[u - a, u]$. Implementation of an *a priori* non-negativity constraint $w(u - t) \geq 0$ for all $0 < u - t < a$ directly on the (partial) maximum likelihood estimate of the polynomial spline function would require a customized, complex and computationally intense procedure [15]. On the other hand, a simpler non-negativity constraint on the estimated spline coefficients, corresponding to $\gamma_j \geq 0$ for $j = 1, \dots, m$ in our model (7), represents only a sufficient but not a necessary condition for the non-negativity of the resulting function $w(u - t)$ [15]. Yet, imposing the above constraint *a priori* would still require using specialized software for constrained optimization and, thus, would preclude implementation of our method with widely available statistical software for Cox's regression with time-dependent covariates.

To avoid aforementioned difficulties, we propose a simplified *a posteriori* approach to be considered in those applications where it is essential to ensure the non-negativity of the estimated weight function. The approach is described here and then briefly illustrated in Section 4 in the context of the empirical analyses of the Flurazepam example. First, the user fits the model (7), without any constraints on the estimable coefficients. Then, the values of the resulting weight function in (4) are computed, throughout the etiologically relevant window $[0, a]$. If any of the computed values is negative, this implies that at least one of the estimated coefficients γ_j is also negative.

Assume the only negative coefficient(s) correspond to the boundary artificial covariates $D_j(u)$, that is, represent the first ($j = 1, 2, \dots$) and/or the last ($j = \dots, m-1, m$) in the sequence of m estimable coefficients. Then, the new model should be re-fitted while simply constraining all those originally negative coefficients to 0, which will be sufficient and necessary to ensure the non-negativity of the new, constrained weight function [15]. However, it is also possible that one or more negative coefficients are assigned to the artificial covariates $D_j(u)$ in the middle of the sequence, that is, the negative coefficient(s) fall in the middle of a sequence of non-negative estimated coefficients. In this case, constraining the initially negative coefficient(s) to 0 and re-fitting the constrained model may not be sufficient to obtain uniformly non-negative weight function because—as indicated by limited simulation experiments—some of the re-estimated adjacent, initially positive, spline coefficient(s) may now be assigned negative values (data not shown). Thus, in the most complex situations, several gradually more constrained models may have to be fitted until the entire weight function becomes uniformly non-negative. Still, the fact that all

splines in the B-spline basis are non-negative for all argument values [10] ensures that such a non-negativity is always achievable through non-negativity constraints on an adequate number of artificial covariates $D_j(u)$ in (7).

To assess the impact of the additional constraints on the model's fit to data, the user may use either the Bayesian information criterion (BIC) criterion or the likelihood ratio test (LRT), because each additionally constrained model is nested within a less constrained model. However, because the decisions regarding which coefficients should be constrained to 0 are based on data-dependent *a posteriori* criteria, the resulting comparisons will be conservatively biased and will tend to favor the more constrained models.

In most applications, there may not be enough substantive knowledge about the shape of the weight function to constrain it to be strictly positive (or negative). Still, if the user decides *a priori* to impose non-negativity constraints on the estimated weight function, through either of the approaches we outline above, or another procedure, then we suggest that the inference about the estimates from the final model should account for these additional steps in the model selection. Accordingly, the procedures for hypothesis testing and confidence bands estimation, described in Sections 2.5 and 2.6, should include the *a posteriori* constrained models among the list of candidate models considered.

2.4. Model selection

We *a priori* limit the number of the interior knots to be between one and three, which implies that five to seven spline coefficients are estimated in the unconstrained model, and three to five in the constrained model. We then rely on the BIC [16] to select the best-fitting unconstrained and/or constrained model(s). Following [17], we define the BIC for a given model as:

$$\text{BIC} = -2 \ln(\text{PL}) + p \ln(d) \quad (9)$$

where PL is the model's partial likelihood, p is the number of estimable parameters in (7) and d the number of un-censored events.

2.5. Hypothesis-testing procedures

Two null hypotheses of primary interest in the WCE framework are: (i) no association between the exposure and the outcome and (ii) constant weights, corresponding to the un-weighted $\text{WCE} = \sum X(t)$. Because the unconstrained spline model in (4) includes both (i) the null model $w(u-t)=0$ and (ii) the constant weight function $w(u-t)=1$, both hypotheses can be tested using LRT [11]. However, when significance tests are performed on the results of the model selected by some data-dependent criteria based on goodness of fit, standard inference procedures are invalid because of inflated type I error [18–20]. The distribution of the test statistics conditional on an *a posteriori* selected model and its unconditional distribution may be significantly different [18]. In the context of our method, the inference about the effects of exposure based on the BIC-optimal model needs to take into account the uncertainty at the model selection stage, and the interdependence of tests conditional on alternative WCE models with different numbers of knots and/or different time windows of clinically relevant exposure $[u-a, u]$ [11, 20, 21].

To obtain valid LRT for the association between the WCE represented by $\sum \gamma_j D_j(u)$ in (7), and the hazard, we propose a simulation-based approach. Our approach is similar to that employed and validated by Mahmud *et al.* to correct the significance level α^* of the LRT of the effect of a continuous covariate on the hazard for *a posteriori* selection of the functional form of the

dose-response curve [20]. We first generate 1000 data sets under the assumption of no association between exposure and outcome (data generation procedures are discussed in Section 3 and in the online material[†]) and then select the BIC-optimal WCE model for each simulated sample. Next, for the i th simulated sample, $i = 1, \dots, 1000$, we compute the LRT statistic with m_i degrees of freedom, where m_i denotes the number of splines coefficients γ estimated in the BIC-optimal model for this specific sample, and establish the corresponding, un-corrected, p -value. Next, we construct the empirical distribution of the 1000 resulting conditional p -values, across 1000 samples generated under the null hypothesis. This distribution will be shifted to the left, that is, toward lower p -values, relative to the uniform distribution that would be obtained if a single model was selected *a priori* [20]. Therefore, using an uncorrected significance level, for example, $\alpha=0.05$, would lead to inflated type I error, with the magnitude of inflation increasing as the sampling variation in the results of BIC-based model selection increases [11]. However, the empirical distribution of conditional p -values directly accounts for such an inflation [20]. Therefore, we use the corrected 5th percentile of the empirical distribution of the conditional p -values as the significance level α_0^* in LRT, conditional on the BIC-optimal model, to ensure an appropriate corrected type I error rate of 0.05. Indeed, under the null hypothesis, about 5 per cent of the conditional p -values will be expected to fall below the corrected significance level α_0^* [20].

As the unweighted cumulative dose model, with $w(u-t)=1$, for $t \leq u$ is nested within the unconstrained weighted cumulative dose model (7), a similar procedure can be used to test whether the unconstrained WCE model provides a better fit to the data than the unweighted model. In this case, the corrected critical value α_1^* corresponds to the 5th percentile of the empirical distribution of the conditional p -values obtained from samples simulated from the unweighted model. The LRT-based procedure, however, cannot be used with the constrained weighted cumulative dose model, as the constraint $\gamma_{m-1}=\gamma_m=0$, forces $w(u-t) \rightarrow 0$ as $t \rightarrow (u-a)$ and, thus, precludes the unweighted cumulative dose model to be nested within the set of constrained WCE models. Nevertheless, the fit of non-nested models can be compared using BIC [16]. Therefore, we propose that the simulation-based procedure for testing the constrained WCE model against the unweighted cumulative model $w(u-t)=\sum X(t)$ can rely on the difference in BIC between these models. Accordingly, in analyses where the constrained WCE model is employed, we define the cutoff as the 5th percentile of the empirical distribution of the BIC differences across simulated samples in which the true weight function is constant.

The simulation-based testing procedures outlined above can be easily extended to multivariable analyses where exposure effect is adjusted for a covariate vector Z . In such cases, we propose to first fit the multivariable model corresponding to the null hypothesis of interest. This will imply either (i) excluding the exposure from the model, if the null hypothesis of no association is to be tested or (ii) modeling exposure with the un-weighted cumulative metric $\sum X(t)$. Next, we carry out simulations, in which the relationship between simulated event times and covariate vector corresponds to the estimated model. Finally, the empirical distribution of the relevant test statistics, across simulated samples, is obtained, and its 5th percentile is used as the corrected cutoff.

2.6. Pointwise confidence bands for the estimated weight function

To assess the precision of the estimated weight function $w(u-t)$, we relied on non-parametric bootstrap re-sampling [22]. The bootstrap routine should account for both (i) the sampling variation

[†]Supporting information associated with this article is available online.

of the regression coefficients, and (ii) the uncertainty at the model selection stage, that is, the additional variance due to *a posteriori* selection of the number of knots and/or of the true support interval $[u-a, u]$ [11, 23]. Accordingly, for each of the B bootstrap samples, alternative versions of model (7) with 1–3 interior knots, and possibly different support intervals, are estimated and the BIC is used to select the best estimate of $w(u-t)$ for a given bootstrap sample. Next, for each $u-a \leq t \leq u$, the empirical distribution of the B point estimates of $w(u-t)$, each corresponding to the BIC-optimal model for a given bootstrap sample, is constructed. The percentile method can then be used to compute 95 per cent pointwise confidence bands, for $w(u-t)$ [22].

2.7. Software specification

The algorithms for the data generation and for computing the time-dependent covariates in (6) were coded in R version 2.6.1. Then, model (7) was fit using the R procedure `coxph` with Efron method for tie handling [24]. The R scripts are available from the authors. Details of implementation are described in the supplemental material online.

3. SIMULATION STUDY

To validate our model and investigate its properties, we simulated a hypothetical prospective cohort study of the association between exposure to a single drug and time to an adverse event. A concise description of the simulation study is presented here. A comprehensive description of the design and analysis of the simulations are available online.

We simulated a cohort of 500 new users of the drug, with time zero defined as the first day of the drug use [25]. Individuals could interrupt and resume their treatment repeatedly thereafter. The follow-up was limited to one year. We simulated six different scenarios, each corresponding to a different true weight function, each defined over a one-year interval $[0, 1]$. As in (4), each weight function described how the relative weight of past exposure on the risk at current time u changed with increasing time ($u-t$) since the exposure. We generated event times conditional on the WCE, using the permutational algorithm, specifically developed and validated for simulating event times conditional on time-dependent covariates and/or effects [26, 27]. For each of the 1000 samples simulated for a given scenario, we estimated unconstrained and constrained models with 1–3 interior knots, and used the BIC to select the best-fitting model. We also estimated an alternative simpler model in which the exposure was expressed with a time-dependent cumulative (unweighted) dose $\text{CUMDOSE} = \sum_{t=u-a}^u X(t)$.

Figures 1 and 2 show a random sample of 100 estimated normalized weight functions obtained from, respectively, the unconstrained and constrained spline models. Each estimate corresponds to the BIC-optimal model for a given simulated sample. In each of the figures' panels, the true weight function used to generate the data is plotted in white. In most scenarios, except for panels (e) and (f), the majority of the estimated weight functions were able to capture the shape of the true weight functions, albeit with some variation in the amplitude of the curves (Figures 1 and 2).

In the unconstrained models, the estimates show considerable over-fitting bias in the right tail of the plot, especially in scenarios when the exposures that occurred relatively long ago had little impact on the risk (Figure 1). Instability at the tails is a known feature of B-spline estimates [28]. As Figure 2 shows, constraining the weight functions to smoothly go to zero at the end of the exposure time window considerably reduced the variation of the estimates. In addition, in

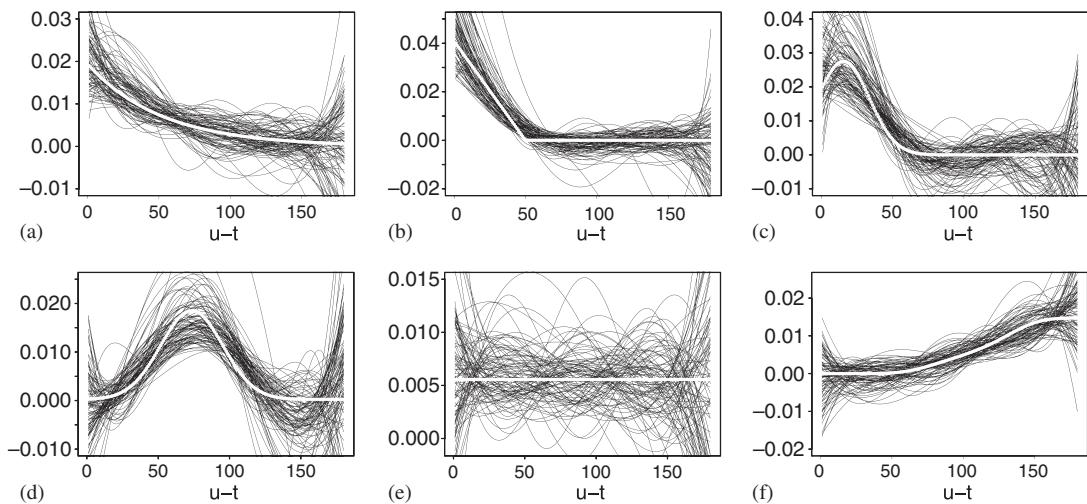


Figure 1. A random sample of 100 normalized estimated weight functions for the unconstrained models with the true weight function in thick white: (a) exponential; (b) bi-linear; (c) early peak; (d) inverted U; (e) constant; and (f) hat. Note that, to make the label of the X-axis readable, we show time in days, while in the text, we use 1 year as the unit of time, so that the values on the axes should be divided by 365.

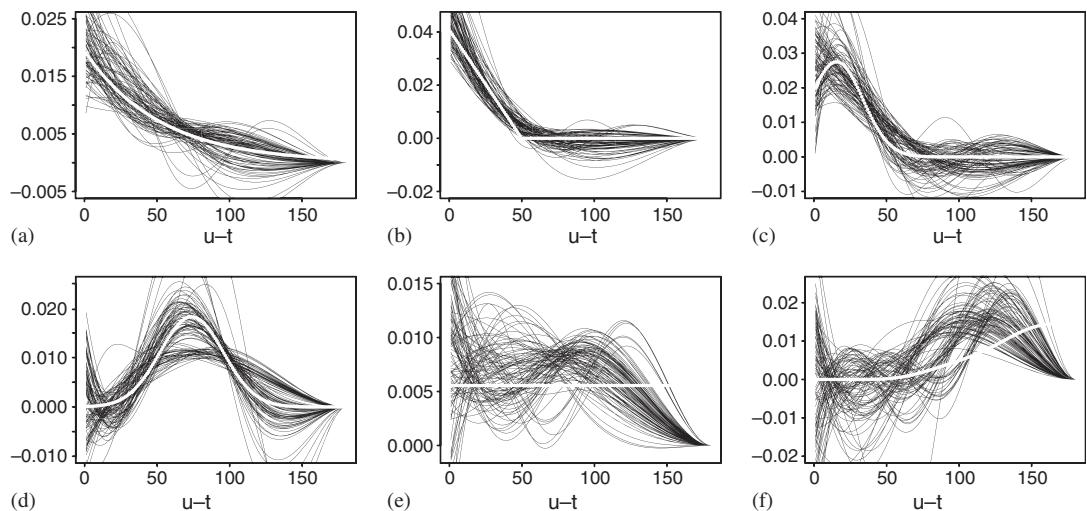


Figure 2. A random sample of 100 normalized estimated weight functions for the constrained models with the true weight function in thick white: (a) exponential; (b) bi-linear; (c) early peak; (d) inverted U; (e) constant; and (f) hat. Note that, to make the label of the X-axis readable, we show time in days, while in the text, we use 1 year as the unit of time, so that the values on the axes should be divided by 365.

panels (a)–(d), where the true weight functions decreased to zero on the right of $u-t = \frac{50}{365}$, the constrained WCE models fitted better than unconstrained models in more than 90 per cent of the simulated samples.

When the true model corresponded to a constant weight function (Figures 1(e) and 2(e)), the correct model with the cumulative unweighted exposure (CUMDOSE) had the best fit in more than 90 per cent of the simulated samples and, thus, should be selected instead of the weighted spline-based WCE model. This largely reduces any potential concerns about the overfit bias shown in Figures 1(e) and 2(e).

In Figures 1(f) and 2(f), the true weight function increased until $u - t = \frac{180}{365}$ and then decreased. This scenario was specifically selected to assess the performance of the method in a difficult situation when the *a priori* selected support interval for the estimated weight function $[u - \frac{180}{365}, u]$, did not cover the entire time window of etiologically relevant exposure. Accordingly, in Figure 2(f), the constrained weight function estimates were forced to go to zero at $u - t = \frac{180}{365}$, where the true function was in fact at its maximum. Therefore, the constrained WCE models provided very biased estimates, and the unconstrained models were performing better in approximating the true shape of the weight function, at least within the $[u - \frac{180}{365}, u]$ window (Figure 1(f)).

The unconstrained WCE estimates shown Figure 2(f) suggest that even those exposures near the right end of the support window have a marked impact on the current risk. Moreover, in this scenario, the shapes of almost all unconstrained weight functions suggest that exposures that occurred about half a year ago may be much more important than the more recent exposures. This should provide a strong suggestion that the support interval $[u - \frac{180}{365}, u]$, *a priori* set by the hypothetical user, is actually much too short. At the end of Section 2.2 of the online material, we discuss how extending the exposure window improved the estimated WCE estimates.

Finally, we investigated whether the estimated effect of the WCE on the risk of an event accurately reflected the true value of β in equation (3), that is, the regression coefficient used to generate the data. The supplemental material online describes in details how the $\hat{\beta}$'s were reconstructed and presents the results. In general, for estimates corresponding to panels (a)–(d) in which the WCE had the best fit, the relative bias for $\hat{\beta}$ was relatively small, mostly below 5 per cent, for both unconstrained and constrained models.

4. APPLICATION

We applied our WCE method to re-analyze the association between exposure to Flurazepam, a psychotropic drug prescribed to treat insomnia [29], and fall-related injuries. Our initial cohort consisted of 4666 elderly subjects from the province of Québec, Canada, first prescribed Flurazepam between January 1, 1990, and December 31, 1994 [25]. The cohort was assembled from provincial health administrative databases that provided information on the drug, dose, duration, and date of each prescription dispensed [25]. A standardized average daily dose was computed [25] and converted into percentage of the World Health Organization (WHO) recommended maximum adult daily dose for Flurazepam [30].

In our analyses, time zero corresponded to January 1, 1990, so that the time axis corresponded to the calendar time. The use of calendar time eliminated potential concerns about unobserved confounding due to seasonal variation and possible secular trends in the frequency of, and reasons for, prescribing Flurazepam, as well as in the frequency of hospitalizations due to falls and injuries [31]. To avoid the immortal time or survival bias [32–34], individuals entered the risk sets only once they initiated their first Flurazepam use. Subjects were followed up until the first fall-related injury, death, or the end of follow-up (December 30, 1994), whichever came first.

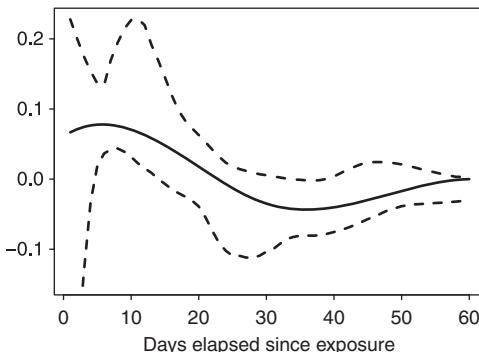


Figure 3. Estimated weight function (solid curve) and pointwise 95 per cent bootstrap CI (dotted curves) for the WCE model of the association between exposure to Flurazepam and fall-related injuries.

Individuals who died, moved out of the province, or were institutionalized before their first fall-related injury were censored.

A total of 252 subjects (5.4 per cent) experienced an event during the follow-up. We were interested in estimating the cumulative impact of past exposures to Flurazepam on the current risk of fall-related injuries. We estimated three different constrained WCE models: (1) a model with the weighted cumulative dose; (2) a model with the weighted cumulative duration of treatment; and (3) a model with the current dose and the weighted cumulative duration of treatment [8]. In each model, the weight function was constrained to smoothly reach zero at the end of the time window. We considered three different windows of etiologically relevant exposure: 60, 90, and 180 days. For each time window, we estimated alternative models, with 1–3 interior knots uniformly placed across the time window.

The fit of the WCE models was compared with four simpler models, in which exposure to Flurazepam was modeled using the following time-dependent variables: (i) a current use indicator, taking the value of 1 on the day when the individual was exposed and 0 otherwise; (ii) a current dose variable; (iii) a cumulative, unweighted, dose; and (iv) a cumulative, unweighted, duration of past treatment. All models were adjusted for three fixed-in-time covariates: age at first benzodiazepine prescription, sex, and a binary indicator of the history of injury in the baseline year (1989) [8]. Finally, we estimated 95 per cent pointwise confidence bands around the weight function for the model with the best fit, using the bootstrap procedure proposed in Section 2.6.

The best fit was obtained with the model that represented exposure with the weighted cumulative duration of treatment, with the weight function estimated with a 1-knot spline function. This result was consistent regardless of the time window considered (60, 90, or 180 days), but the 60-day time window yielded the best fit. The difference in BIC between the best and second best spline model for each of the three time windows considered was of at least 3 units. The corresponding estimated weight function, with bootstrap-based 95 per cent confidence bands, is shown in Figure 3. The estimated weight function was relatively flat in the first 10 days but decreased thereafter, reaching approximately zero after 20 days. This suggested that the effects of recent exposures to Flurazepam cumulate during about 10 days, and do not last beyond three weeks after exposure.

The estimated weight function in Figure 3 assigned negative weights to Flurazepam exposures that occurred about 30–40 days ago. Yet, the 95 per cent confidence bands around the negative portion of the weight curve included 0 across the exposure window (Figure 3), which may suggest

Table I. Adjusted hazard ratios*, with 95 per cent confidence intervals, for the association between various patterns of Flurazepam and fall-related injuries.

Pattern of use	Reference	Hazard ratio	95 per cent confidence interval	BIC [‡]
<i>BIC-optimal WCE model[†]</i>				
Current user, dose=1, for 1 day	Non-user	1.07	(0.70–1.26)	3763.5
Current user, dose=1, for 7 days	Non-user	1.68	(0.68–3.05)	
Past user, dose=1, 8–14 days ago	Non-user	1.59	(1.19–4.35)	
Past user, dose=1, 15–21 days ago	Non-user	1.23	(0.85–1.77)	
Past user, dose=1, 22–28 days ago	Non-user	0.92	(0.49–1.14)	
Current user, dose=1, for 30 days	Non-user	2.83	(1.45–4.34)	
Current user, dose=1, for 7 days	Past user, dose=1, for 7 days, between 14 and 7 days ago	1.37	(0.43–3.45)	
<i>Conventional models</i>				
Current use	Non-user	1.34	(0.97–1.85)	3769.9
Current dose	Non-user	1.67	(1.14–2.46)	3766.7
Unweighted duration of use (30 days)	Non-user	0.99	(0.97–1.02)	3772.4
Unweighted cumulative dose (30 days)	Non-user	0.99	(0.97–1.03)	3772.7

*All models were adjusted for age at first benzodiazepine prescription, sex, and a binary indicator of the history of injury in the baseline year (1989).

†1-knot constrained spline-based WCE model.

‡BIC of the corresponding model, with lower BIC indicating better fit to data.

that such local negativity reflected the overfit bias rather than a truly protective effect of exposures that occurred more than a month ago. To further explore this issue, we applied the approach suggested in Section 2.3 to constrain the estimated weight function to be uniformly non-negative. The negative portion of the original estimate reflected the negative value of the coefficient of the 2nd artificial $D_j(u)$ variable in equation (7). Re-analyzing the data, constraining this coefficient to 0, did not improve the fit ($BIC = 3763.5$ for the original model vs $BIC = 3764.4$ for the new, constrained model). Moreover, the LRT test comparing the two models showed a significant improvement in the fit of the un-constrained (original) model ($PL = -1865.167$ vs $PL = -1868.376$, $p = 0.011$ for a 1-df chi-square test). Therefore, the original estimate, presented in Figure 3, seems to be more consistent with the data, even if some values of the weight function are negative.

Table I shows the adjusted HR corresponding to different patterns of Flurazepam use, estimated with either the best-fitting WCE model (upper part of the table) or one of the conventional models (lower part). The HR obtained from the best-fitting weighted cumulative dose model for comparing the risk associated with the use of Flurazepam for the last 30 days, relative to non-use during the same period was 2.83 (95 per cent CI 1.45–4.34). Shorter use, during last 7 days, yielded an HR of 1.68 (95 per cent CI 0.68–3.05). The uncertainty in the estimation of the early part of the Flurazepam curve, as indicated by the wide 95 per cent pointwise bootstrap confidence bands in Figure 3, explained the lack of precision of this estimate. Comparisons of past use (i) days 8–14 ago, (ii) 15–21 days ago, and (iii) 22–28 days ago with no use at any time indicate that the HR associated with a 1 week exposure to Flurazepam gradually decreases after 1, 2, and 4 weeks since last exposure (rows 3–5 of Table I). In the conventional models, the higher current dose of Flurazepam was associated with an increased risk of fall-related injuries (Table I $HR = 1.67$,

95 per cent CI 1.14–2.46), but the unweighted cumulative dose over 30 days was not (Table I HR=0.99 95 per cent CI 0.97–1.03).

Our findings regarding the cumulative effects of Flurazepam use should be replicated in an independent study with adjustment for additional potential confounders, including the presence of such conditions as osteoporosis, diabetes, heart diseases, bone diseases, blood circulation problems, and use of concurrent use of other medications possibly affecting cognitive functioning in the elderly.

5. DISCUSSION

We proposed a flexible method for modeling cumulative effects of time-dependent exposures, weighted by recency, in the Cox's proportional hazards model. Our method avoids making *a priori* assumptions about the shape of the function that assigns weights to exposures in the past, which is estimated with flexible cubic regression B-splines. In this aspect, the method is similar to that proposed by Hauptmann *et al.* for logistic regression analyses of case–control studies [3]. However, the proposed parametrization of our flexible WCE model avoids the need for constrained optimization required by their approach [3], and allows for time-to-event analyses of prospective and retrospective cohort studies. Indeed, our model (7) can be implemented by any standard statistical package for survival analysis with time-dependent covariates.

Simulations indicated that the WCE model recovers a variety of clinically plausible shapes of the true weight function and provides satisfactory estimates of the strength of the association between exposure and hazard. The proposed strategy used to test if the WCE model detects a statistically significant cumulative exposure effect provides a better fit to the data than the conventional unweighted cumulative exposure model. It accounts for sampling variation at the model selection stage and, thus, can be used for hypothesis testing in other flexible analyses where the test is conditional on the model selected *a posteriori*, based on data-dependent criteria such as BIC [11]. Pointwise confidence bands may be obtained using a bootstrap method that also accounts for data-dependent BIC-based selection of the number of interior knots and/or the length of the etiologically relevant exposure window. The strength of the cumulative effect may be assessed by estimating HR corresponding to contrasts between pre-specified patterns of exposure histories, as illustrated by the Flurazepam example.

While our estimates are generally able to capture the overall shape of the weight function, they may fail to accurately reflect sudden and rapid change in the function, resulting in high local curvature. If such effects are anticipated, the user may improve the model by adding knot(s) in the regions where high curvature is expected. In addition, future research should consider expanding our model to incorporate the non-negativity constraint on the estimated weight function, as such a constraint may be *a priori* biologically plausible in some applications. In this context, a major challenge will be to develop accurate and efficient constrained optimization algorithms and, ultimately, to allow their implementation with widely available statistical software.

Similar to previous studies, we provided evidence of a statistically significant association between the use of Flurazepam and the risk of fall-related injuries [8, 25, 35]. However, the weighted cumulative dose model offered additional insights regarding the mechanism linking Flurazepam use to the risk of falls. It suggested that the effects of recent Flurazepam exposure cumulate over time, but that on a given day, the window of etiologically relevant exposure for Flurazepam is limited to approximately the 10 previous days.

When the exposure does *not* change often over time, the time-invariant metrics of baseline exposure, such as a binary indicator of the baseline exposure status or a quantitative measure of the baseline exposure intensity (dose), may predict the future outcomes quite well. On the other hand, it is also possible that the impact of the past exposures does cumulate over time, so that the cumulative effect does increase with increasing follow-up duration, even if the individual exposure status/dose does not change. To discriminate between alternative ways in which a time-invariant exposure may affect the risk, one can use the BIC criterion to compare the goodness of fit between (i) the model with time-invariant exposure metrics and (ii) the WCE model. However, in our simulations, the individual subjects exposure status and dose varied greatly over time and the true effect of exposure was cumulative over time. Similarly, in the Flurazepam application, subjects switched frequently between the periods of (current) exposure and non-exposure. Therefore, the use of models with time-invariant exposure metrics could not be justified in neither type of our analyses.

Our study focuses on the cumulative effects of time-varying exposures. However, in some applications it may be of interest to investigate potentially cumulative effects of exposures that remain constant in time, either by definition or by study design. In that case, it may be more practical to use one of the several time-varying coefficients models, developed in the last two decades to incorporate time-dependent effects of time-invariant exposures in Cox's model, that is, to account for their non-proportional effects [11, 21, 36–40]. In all these models, the constant-over-time exposure effect β is replaced by a flexible function of time $\beta(t)$, which describes how the corresponding log HR changes with increasing follow-up duration t . To avoid restrictive *a priori* assumptions about the shape of $\beta(t)$, it is estimated using various flexible modeling techniques, including regression splines, smoothing splines, or fractional polynomials. As suggested by Hauptmann *et al.* [3], for a time-invariant exposure, the estimated time-dependent effect will reflect possible cumulative effects and may help discriminating between alternative hypotheses regarding relative weights of exposures that occurred at different points in the past.

Our method follows the tradition of survival analysis in that it focuses on the effect of time-varying exposure on time to a single event. Future studies should attempt to generalize our WCE model to the analysis of competing risks, possibly by adapting the Lunn and McNeil extension of Cox's model [41]. Such an extended model would allow considering, for example, a possible impact of Flurazepam use on other clinical outcomes, not related to falls.

Pharmacoepidemiological studies based on prescription databases have the advantage of providing individual time-dependent measures of exposure over a fine time grid [42]. However, potential applications of the proposed WCE method include any exposure that varies over time and may have cumulative effects, with possibly differential impact of exposures that occurred at different times in the past. Our method can be used regardless of whether the study involves a *delayed entry* [43] or a *new users* [44] design, as illustrated, respectively, by our Flurazepam example and our simulations. Additional simulations indicate that both designs yield similar regression coefficients if there are no temporal trends in exposure–risk association (data not shown). However, if such trends are plausible, the delayed entry design reduces the risk of resulting residual confounding. The model can also be used to assess the weighted cumulative duration of past exposure, where the exposure history $X(t)$ is represented with a vector of time-dependent binary indicators of the status of exposure at various times in the past, instead of quantitative measures of exposure intensity. Models that include both a time-dependent variable representing the current intensity of exposure and the weighted cumulative duration of treatment can also be fitted to disentangle the effect of the current dose from that of the duration of past treatment [8].

The concept of the recency-weighted cumulative dose metric has been present in the literature over more than 20 years [6, 7]. However, its use has been restricted to few studies [2, 3, 8, 45, 46], practically all of which represented individual exposure history with a single value of a time-independent covariate. We hope that our implementation of the flexible modeling of the weighted cumulative dose in the familiar Cox's PHs model will motivate a more widespread use of this metric. We also believe that the WCE model can provide useful insights regarding the mechanisms linking the history of time-dependent exposure with the risk of events investigated in many clinical and epidemiological studies.

ACKNOWLEDGEMENTS

We are thankful to Dr Robyn Tamblyn for providing us with the data set to illustrate our method. The initial data collection was funded by the Fonds de recherche en santé du Québec (FRSQ) and the study was approved by the McGill University Faculty of Medicine Institutional Review Board, the Quebec Access to Information Commission, and the Régie de l'assurance maladie du Québec. We thank Dr Emmanuel Lesaffre for insightful discussions of the methodology proposed in this paper, and Dr Ahmadou Alioum for his helpful comments on the manuscript.

REFERENCES

- Collet JP, Sharpe C, Belzile E, Boivin JF, Hanley J, Abenhaim L. Colorectal cancer prevention by non-steroidal anti-inflammatory drugs: effects of dosage and timing. *British Journal of Cancer* 1999; **81**(1):62–68.
- Vacek PM. Assessing the effect of intensity when exposure varies over time. *Statistics in Medicine* 1997; **16**:505–513.
- Hauptmann M, Wellmann J, Lubin JH, Rosenberg PS, Kreienbrock L. Analysis of exposure-time-response relationships using a spline weight function. *Biometrics* 2000; **56**(4):1105–1108.
- Stranges S, Bonner MR, Fucci F, Cummings KM, Freudenheim JL, Dorn JM, Muti P, Giovino GA, Hyland A, Trevisan M. Lifetime cumulative exposure to secondhand smoke and risk of myocardial infarction in never smokers: results from the Western New York health study, 1995–2001. *Archives of Internal Medicine* 2006; **166**(18):1961–1967.
- van Leeuwen R, Vingerling JR, Hofman A, de Jong PTVM, Stricker BHCh. Cholesterol lowering drugs and risk of age related maculopathy: prospective cohort study with cumulative exposure measurement. *British Medical Journal* 2003; **326**(7383):255–256.
- Breslow NE, Lubin JH, Marek P, Langholz B. Multiplicative models and cohort analysis. *Journal of the American Statistical Society* 1983; **78**(381):1–12.
- Thomas DC. Models for exposure–time–response relationships with applications to cancer epidemiology. *Annual Reviews of Public Health* 1988; **9**:451–482.
- Abrahamowicz M, Bartlett G, Tamblyn R, du Berger R. Modeling cumulative dose and exposure duration provided insights regarding the associations between benzodiazepines and injuries. *Journal of Clinical Epidemiology* 2006; **59**(4):393–403.
- Cox DR. Regression models and life tables (with Discussion). *Journal of the Royal Statistical Society, Series B* 1972; **34**:187–220.
- de Boor C. *A Practical Guide to Splines*. Springer: Berlin, Heidelberg, New York, 1978.
- Abrahamowicz M, MacKenzie T, Esdaile JM. Time-dependent hazard ratio: modeling and hypothesis testing with application in lupus nephritis. *Journal of the Acoustical Society of America* 1996; **91**:1432–1439.
- Smith PL. Splines as a useful and convenient statistical tool. *American Statistician* 1979; **33**(2):57–62.
- Stone CJ. Generalized additive models: comment. *Statistical Science* 1986; **1**(3):312–314.
- Ramsay JO. Monotone regression splines in action. *Statistical Science* 1988; **3**(4):425–441.
- Abrahamowicz M, Ciampi A, Ramsay JO. Nonparametric density estimation for censored survival data: regression-spline approach. *The Canadian Journal of Statistics/La Revue Canadienne de Statistique* 1992; **20**(2):171–185.
- Schwarz G. Estimating the dimension of a model. *Annals of Statistics* 1978; **6**(2):461–464.

17. Volinsky C, Raftery A. Bayesian information criterion for censored survival models. *Biometrics* 2000; **56**:256–262.
18. Hurvich CM, Tsai CH. The impact of model selection on inference in linear regression. *The American Statistician* 1990; **44**(3):214–217.
19. Leeb H, Pötscher BM. Can one estimate the unconditional distribution of post-model-selection estimators? *MPRA Paper*, University Library of Munich, Germany, 2005. Available from: <http://ideas.repec.org/p/pra/mprapa/72.html> [20 March 2008].
20. Mahmud M, Abrahamowicz M, Leffondré K, Chaubey Y. Selecting the optimal transformation of a continuous covariate in Cox's regression: implications for hypothesis testing. *Communications in Statistics: Simulation and Computation* 2006; **35**(1):27–45.
21. Kooperberg C, Stone CJ, Truong YK. Hazard regression. *Journal of the American Statistical Association* 1995; **90**(429):78–94.
22. Efron B, Tibshirani RJ. *An Introduction to the Bootstrap*. Chapman & Hall: London, 1993.
23. Rosenberg PS, Katki H, Swanson CA, Brown LM, Wacholder S, Hoover RN. Quantifying epidemiologic risk factors using non-parametric regression: model selection remains the greatest challenge. *Statistics in Medicine* 2003; **22**(21):3369–3381.
24. R Development Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing: Vienna, 2006.
25. Tamblyn R, Abrahamowicz M, du Berger R, McLeod P, Bartlett G. A 5-year prospective assessment of the risk associated with individual benzodiazepines and doses in new elderly users. *Journal of the American Geriatrics Society* 2005; **53**(2):233–241.
26. MacKenzie T, Abrahamowicz M. Marginal and hazard ratio specific random data generation: applications to semi-parametric bootstrapping. *Statistics and Computing* 2002; **12**(3):245–252.
27. Sylvestre MP, Abrahamowicz M. Comparison of algorithms to generate event times conditional on time-dependent covariates. *Statistics in Medicine* 2008; **27**(14):2618–2634.
28. Hastie TJ, Tibshirani RJ. *Generalized Additive Models*. Chapman & Hall: London, 1990.
29. Flurazepam. *Compendium of Pharmaceuticals and Specialties*. e-cps, [online]. Ottawa (ON). Canadian Pharmacists Association; c2007. Available from: <http://www.e-cps.ca> (Page consulted on October 24, 2007).
30. Anatomical therapeutic chemical classification index with defined daily doses (ATC/DDD). *WHO Collaborating Centre for Drug Statistics Methodology*, World Health Organization, Oslo, 2000.
31. Bartlett G, Abrahamowicz M, Tamblyn R, Grad R, Capek R, du Berger R. Longitudinal patterns of new benzodiazepine use in the elderly. *Pharmacoepidemiology and Drug Safety* 2004; **13**(10):669–682.
32. Suissa S. Effectiveness of inhaled corticosteroids in chronic obstructive pulmonary disease: immortal time bias in observational studies. *American Journal Respiratory and Critical Care Medicine* 2003; **168**(1):49–53.
33. Zhou Z, Rahme E, Abrahamowicz M, Pilote L. Survival bias associated with time-to-treatment initiation in drug effectiveness evaluation: a comparison of methods. *American Journal of Epidemiology* 2005; **162**(10):1–8.
34. Sylvestre MP, Huszti E, Hanley JA. Do oscar winners live longer than less successful peers? A reanalysis of the evidence. *Annals of Internal Medicine* 2006; **145**(5):361–363.
35. Pierfitte C, Macouillard G, Thicoipe M, Chaslerie A, Pehourcq F, Aissou M, Martinez B, Lagnaoui R, Fourrier A, Begaud B, Dangoumau J, Moore N. Benzodiazepines and hip fractures in elderly people: case-control study. *British Medical Journal* 2001; **322**(7288):704–708.
36. Abrahamowicz M, MacKenzie T. Joint estimation of time-dependent and non-linear effects of continuous covariates on survival. *Statistics in Medicine* 2007; **26**(2):392–408.
37. Gray RJ. Flexible methods for analysing survival data using splines with applications to breast cancer prognosis. *Journal of the American Statistical Association* 1992; **87**:942–951.
38. Amorim LD, Cai J, Zeng D, Barreto ML. Regression splines in the time-dependent coefficient rates model for recurrent event data. *Statistics in Medicine* 2008; **27**(28):5890–5906.
39. Hastie T, Tibshirani R. Varying-coefficient models. *Journal of the Royal Statistical Society, Series B* 1993; **55**:757–796.
40. Nan B, Lin X, Lisabeth LD, Harlow SD. A varying-coefficient Cox model for the effect of age at a marker event on age at menopause. *Biometrics* 2005; **61**(2):576–583.
41. Lunn M, McNeil D. Applying Cox regression to competing risks. *Biometrics* 1995; **51**:524–532.
42. Schneeweiss S, Avorn J. A review of uses of health care utilization databases for epidemiologic research on therapeutics. *Journal of Clinical Epidemiology* 2005; **58**(4):323–337.
43. de Luise C, Lanes SF, Jacobsen J, Pedersen L, Sorensen RT. Cardiovascular and respiratory hospitalizations and mortality among users of tiotropium in Denmark. *European Journal of Epidemiology* 2007; **22**(4):267–272.

44. Ray WA. Evaluating medication effects outside of clinical trials: new-user designs. *American Journal of Epidemiology* 2003; **158**:915–920.
45. Berry G, Gilson JC, Holmes S, Lewinsohn HC, Roach SA. Asbestosis: a study of dose–response relationships in an asbestos textile factory. *British Journal of Industrial Medicine* 1979; **36**(2):98–112.
46. Lundin Jr FD, Archer VE, Wagoner JK. In *An Exposure–Time–Response Model for Lung Cancer Mortality in Uranium Miners Effects of Radiation Exposure, Age and Cigarette Smoking*, Breslow NE, Whittemore AS (eds). SIAM: Philadelphia, PA, 1979; 243–246.