Published in final edited form as:

Epidemiology. 2017 November ; 28(6): 771–779. doi:10.1097/EDE.0000000000000742.

Doubly Robust Additive Hazards Models to Estimate Effects of a Continuous Exposure on Survival

Yan Wang^{a,b}, Mihye Lee^a, Pengfei Liu^c, Liuhua Shi^a, Zhi Yu^d, Yara Abu Awad^a, Antonella Zanobetti^a, and Joel D. Schwartz^a

^aDepartment of Environmental Health, Harvard T.H. Chan School of Public Health, Boston, MA

^bDepartment of Biostatistics, Harvard T.H. Chan School of Public Health, Boston, MA

^cHarvard John A. Paulson School of Engineering and Applied Sciences, Cambridge, MA

^dDivision of Rheumatology, Allergy and Immunology, Section of Clinical Sciences, Brigham and Women's Hospital, Boston, MA

Abstract

Background—The effect of an exposure on survival can be biased when the regression model is misspecified. Hazard difference is easier to use in risk assessment than hazard ratio and has a clearer interpretation in the assessment of effect modifications.

Methods—We proposed two doubly robust additive hazards models to estimate the causal hazard difference of a continuous exposure on survival. The first model is an inverse probability-weighted additive hazards regression. The second model is an extension of the doubly robust estimator for binary exposures by categorizing the continuous exposure. We compared these with the marginal structural model and outcome regression with correct and incorrect model specifications using simulations. We applied doubly robust additive hazard models to the estimation of hazard difference of long-term exposure to PM_{2.5} (particulate matter with an aerodynamic diameter less than or equal to 2.5 microns) on survival using a large cohort of 13 million older adults residing in seven states of the Southeastern United States.

Results—We showed that the proposed approaches are doubly robust. We found that each 1 $\mu\text{g m}^{-3}$ increase in annual PM_{2.5} exposure was associated with a causal hazard difference in mortality of 8.0×10^{-4} (95% confidence interval 7.4×10^{-4} , 8.7×10^{-4}), which was modified by age, medical history, socioeconomic status, and urbanicity. The overall hazard difference translates to approximately 5.5 (5.1, 6.0) thousand deaths per year in the study population.

Correspondence: Joel D. Schwartz, Department of Environmental Health, Harvard T.H. Chan School of Public Health, 401 Park Drive, PO Box 15677, Boston, MA 02215. jschwartz@hsph.harvard.edu.

The authors report no conflicts of interest.

Data and code availability: R and SAS code for doubly robust additive hazards model A is given in eAppendix. The simulation code is available upon request. The Medicare dataset was used under a data use agreement and cannot be shared.

Supplemental digital content is available through direct URL citations in the HTML and PDF versions of this article (www.epidem.com).

Conclusions—The proposed approaches improve the robustness of the additive hazards model and produce a novel additive causal estimate of PM_{2.5} on survival and several additive effect modifications, including social inequality.

Multiplicative hazards models are the most commonly used approach for survival data analysis.¹ They assume that the same increment in exposure will have larger impacts in populations with higher baseline hazards. In some circumstances, it may be more reasonable to assume that the effect of the exposure is to increase mortality rates by the same increment regardless of baseline and to test additive effect modifications to identify characteristics of people at higher risk. Additive hazards models, which aims to estimate hazard differences, can address these issues.^{2,3} Compared with hazard ratios, hazard differences can be used in risk assessment without requiring additional data to estimate the baseline rates and are more comprehensible to general audience. Effect modification on additive scale incorporates information about the potential differences in baseline hazards between subgroups, whereas effect modification on multiplicative scale does not. A lower ratio for a common cause of death may imply a greater increase in risk of dying than does a larger ratio for a rare cause of death. Therefore, ratios need to be interpreted more cautiously, particularly when they are not presented along with baseline rates or hazards. In addition, additive models may also be more useful to assess underlying biologic interactions.^{4,5}

In observational studies where the exposure is not randomly assigned, several approaches can be used to estimate causal effects. The first widely used approach seeks to correctly specify an outcome regression (OR) model to adjust for confounding. The second approach tries to mimic randomized trials using propensity score (PS) for either matching or reweighting the observations using inverse probability weights (IPW), which makes the exposure independent of covariates. An example is marginal structural models (MSMs).^{6–8} The third approach, doubly robust estimators,^{9,10} combines the idea of outcome regression and IPW and has the advantage that the estimator is valid when either the propensity score model or the outcome regression model is misspecified, i.e., robust to the misspecification of one of the two models. In practice, it offers the users two chances to correctly or approximately correctly specify one of the models and get a causal estimate with little loss in efficiency compared with a correctly specified the outcome regression model.⁹ This is an appealing property, because it may be easier to nearly correctly specify the propensity score model in some cases, while sometimes it may be easier to nearly correctly specify the outcome regression.

The present study proposed two doubly robust additive hazards models, which is motivated by estimating the effect of long-term PM_{2.5} on survival. The World Health Organization estimates that PM_{2.5} is associated with over three million deaths per year, based on multiple cohort studies.¹¹ To date, most studies have used Cox proportional hazards model or Poisson regressions and reported conditional hazard ratio or relative risk estimates.^{12–20} However, it is unclear whether a given increment of PM_{2.5} will increase mortality rates more in populations where baseline rates are greater, and comparisons of additive to multiplicative models can inform this question. Additive hazards may allow easier identification of characteristics that modify risks²¹ and provide a clearer interpretation than hazard ratios.

METHODS

Additive Hazards Model

We consider a counting process formulation of individual survival data.²² Each subject i is followed over time from the start of year 1 to the end of year t_j , which is the year of censoring or death whichever occurs first. Each record in the dataset represents one person-year of follow-up and contains covariates for person i in year j . We assume the covariates for each subject are constant within each follow-up year but can vary from year to year. Although some variables are likely to change within a year, it may be impractical to update them too frequently in many cohort studies (e.g. once per year in Medicare data; once per two years for some variables in the Nurses' Health Study). We assume noninformative right censoring, which is a common assumption made in many survival analyses and the epidemiology of long-term air pollution on survival. For example, administrative censoring is often considered noninformative because the censoring time is unrelated to the susceptibility of an individual to the exposure.

Additive hazards models consider that the exposure has an additive effect on hazard for each unit increase. Given the exposure x_{ij} (the exposure of individual i in year j) and covariate vector $\mathbf{v}_{ij}$, we assume a parametric form of an additive hazards model according to Farrington (1996)²

$$\lambda_{ij}(t|x_{ij}, \mathbf{v}_{ij}) = \beta_0 + \beta x_{ij} + s(\mathbf{v}_{ij}; \boldsymbol{\gamma}), \quad j-1 < t \leq j \quad (1)$$

where λ_{ij} is the hazard for person i in year j , and $s(\mathbf{v}_{ij}; \boldsymbol{\gamma})$ is a function of $\mathbf{v}_{ij}$ parametrized by $\boldsymbol{\gamma}$. The likelihood of the model is given in eAppendix 1; <http://links.lww.com/EDE/B260>. The advantage of this formulation is that it preserves the interpretability of the effect of the exposure and considers a general functional form for the covariates. The disadvantage of the formulation is that it does not naturally bound the hazard to be nonnegative without a constraint on the parameter space which is a common issue of additive hazards models, and it considers an additive, rather than a more flexible, structure between the exposure and the function for the covariates.

Using the relationship between hazard function (λ_{ij}) and survival function for individual i up

until time j ($S_i(j)$) that $\frac{S_i(j)}{S_i(j-1)} = \exp\left(-\int_{j-1}^j \lambda_{ij} dt\right) = \exp(-1 \times \lambda_{ij})$, we could rewrite the model as

$$\begin{aligned} -\ln[1 - E(D_{ij}|x_{ij}, \mathbf{v}_{ij})] &= -\ln \frac{S_i(j)}{S_i(j-1)} \\ &= 1 \times (\beta_0 + \beta x_{ij} + s(\mathbf{v}_{ij}; \boldsymbol{\gamma})) \end{aligned} \quad (2)$$

where D_{ij} is an indicator of outcome for person i in year j , and “1×” indicates that each follow-up year is the unit of time and that all variables are assumed constant within a follow-up year.

Doubly Robust Additive Hazards Model

To estimate the causal effect of the exposure, we can adjust for variables that are the common causes of the exposure and outcome in the regression model. We can also mimic randomized trials and make the exposure independent of covariates through weighting. A simultaneous usage of regression adjustment and IPW will give the property of double robustness. A correct form of outcome regression removes the confounding between the exposure and survival through adjusting for covariates. A correct IPW creates a pseudo-population where the exposure becomes independent of all confounders.

Two doubly robust additive hazards models are proposed. We start the descriptions of the methods with the calculation of the IPW. We calculated the stabilized IPW under a normality assumption. Briefly, we fit a linear regression (propensity score model) on X_{ij} against a function of covariates V_{ij} which explains the variation in the exposure. Assuming a normal distribution for the residuals $N(0, \sigma^2)$ (this assumption can be relaxed using kernel density estimation²³), we calculated the conditional density of being exposed to the observed

exposure $\hat{f}(x_{ij}|v_{ij};\hat{\alpha}) = \frac{1}{\sqrt{2\pi\hat{\sigma}^2}} \exp(-\frac{\hat{\epsilon}_{ij}^2}{2\hat{\sigma}^2})$, where $\hat{\alpha} = (\hat{\sigma}^2, \hat{\epsilon}_{ij})$ is the maximum likelihood estimators for the residual variance σ^2 and the residual of exposure for individual i in year j . Similarly, assuming a marginally normal distribution of the exposure (this assumption can be relaxed by marginalizing the kernel density estimation for the conditional distribution²³), we calculated the marginal density of the exposure using

$\hat{f}(x_{ij}) = \frac{1}{\sqrt{2\pi\hat{\sigma}'^2}} \exp(-\frac{(x_{ij}-\hat{\mu})^2}{2\hat{\sigma}'^2})$, where $\hat{\mu}$ and $\hat{\sigma}'^2$ are the maximum likelihood estimators for the marginal mean and variance of the exposure. The IPW is $\hat{f}(x_{ij})/\hat{f}(x_{ij}|v_{ij}, \hat{\alpha})$. Influential values of IPW may be truncated.²⁴

Doubly robust additive hazards model A (DRAHM-A) is fitted using an IPW-weighted Poisson regression on the indicator of survival against a function of covariates and a linear term of exposure (Equation 2 weighting each observation by IPW). Doubly robust additive hazards model B (DRAHM-B) categorizes the continuous exposure and extends the traditional doubly robust model for binary exposure.

Technicalities of DRAHM-A

DRAHM-A Models

$$\begin{aligned} -\ln [1 - E(D_{ij}|x_{ij}, v_{ij})] &= \beta_0 + \beta x_{ij} + s(v_{ij}; \gamma), \\ \text{weight}_{ij} &= \text{IPW}_{ij} \end{aligned} \quad (3)$$

This can be fitted using a weighted Poisson regression with a logarithm link function on an outcome of $1 - d_{ij}$ (1 if individual i survived in year j) and a design matrix that has the opposite sign to the original variables. Technically, $\hat{\beta}_0, \hat{\beta}, \hat{\gamma}$ were estimated by solving the following score equations $\mathbf{U}(\theta) = (\mathbf{U}(\beta_0), \mathbf{U}(\beta), \mathbf{U}(\gamma))'$, where $\theta = (\beta_0, \beta, \gamma)'$.

$$\begin{aligned} \mathbf{U}_{\beta_0} &= \sum_{i=1}^m \sum_{j=1}^{t_i} \frac{1(1-d_{ij} - \exp(-\beta_0 - \beta x_{ij} - s(\mathbf{v}_{ij}; \gamma)))}{\hat{f}(x_{ij} | \mathbf{v}_{ij}; \hat{\alpha})} \hat{f}(x_{ij}) = 0 \\ \mathbf{U}_{\beta} &= \sum_{i=1}^m \sum_{j=1}^{t_i} x_{ij} \frac{(1-d_{ij} - \exp(-\beta_0 - \beta x_{ij} - s(\mathbf{v}_{ij}; \gamma)))}{\hat{f}(x_{ij} | \mathbf{v}_{ij}; \hat{\alpha})} \hat{f}(x_{ij}) = 0 \\ \mathbf{U}_{\gamma} &= \sum_{i=1}^m \sum_{j=1}^{t_i} \frac{\partial s(\mathbf{v}_{ij}; \gamma)}{\partial \gamma} \frac{(1-d_{ij} - \exp(-\beta_0 - \beta x_{ij} - s(\mathbf{v}_{ij}; \gamma)))}{\hat{f}(x_{ij} | \mathbf{v}_{ij}; \hat{\alpha})} \hat{f}(x_{ij}) = 0. \end{aligned} \quad (4)$$

$\hat{\beta}$ is doubly robust to the misspecification of the propensity score and outcome regression (eAppendix 2; <http://links.lww.com/EDE/B260>). It is consistent and asymptotically normal when at least one of the PS and OR is correctly specified. The asymptotic variance-covariance matrix of the regression coefficients could be estimated using the heteroscedasticity consistent sandwich estimator

$$\left[E \left(\frac{\partial \mathbf{U}(\theta)}{\partial \theta'} \right) \right]^{-1} \text{Var}(\mathbf{U}(\theta)) \left[E \left(\frac{\partial \mathbf{U}(\theta)}{\partial \theta'} \right) \right]^{-1'}. \text{ A consistent variance estimator is obtained by plugging in the empirically estimated } E \left(\frac{\partial \mathbf{U}(\theta)}{\partial \theta'} \right) \text{ and } \text{Var}(\mathbf{U}(\theta)).$$

Marginal structural model is a special case of DRAHM-A, where IPW is used to make the exposure independent of the covariates and the OR model does not contain any other covariates but the exposure ($s(\cdot)$ dropped). The Farrington's additive hazards model (i.e., OR) is also a special case of DRAHM-A where the weight for each observation is set to be equal.

A general description of the algorithm, an R function, and a SAS macro for DRAHM-A are given in the eAppendix 3; <http://links.lww.com/EDE/B260>.

Technicalities of DRAHM-B

DRAHM-B extends the scheme of the doubly robust estimators for binary treatment by categorizing a continuous exposure.⁹ We first split the data into K partitions ($\mathcal{P}_k$) where $k = 1, 2, \dots, K$ with $K - 1$ cutoffs at the $100/K^{\text{th}}$ quantile of the exposure. In each of the partitions, the median exposure is denoted by x_k . In the first stage, we fit the following model to the subset $\mathcal{P}_k$:

$$\begin{aligned} \ln[1 - E(D_{ij} | x_{ij} \in \mathcal{P}_k, \mathbf{v}_{ij})] &= 1 \times (\beta_{0k}^{**} + s(\mathbf{v}_{ij}; \gamma_k)), \\ \text{weight}_{ij} &= \text{IPW}_{ij} \end{aligned} \quad (5)$$

The model is also fitted by an IPW-weighted Poisson regression with a logarithm link function. Using the fitted model for subset $\mathcal{P}_k$, we predicted the probabilities of survival for each follow-up year in the entire population, had they all been in exposure category k . The average of those predicted probabilities is denoted using p_k , which has an expectation of $\int [1 - E(D_{ij} | x_{ij} = x_k, \mathbf{v}_{ij})] dF(\mathbf{v}_{ij})$. In the second stage, we model

$$-\ln E(p_k) = -\ln \int [1 - E(D_{ij} | x_{ij} = x_k, \mathbf{v}_{ij})] dF(\mathbf{v}_{ij}) - \beta_0^{**} + \beta x_k \quad (6)$$

using an unweighted Poisson regression with a logarithm link function. The variance is similarly estimated using sandwich estimator

$\left[E \left(\frac{\partial U(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}'} \right) \right]^{-1} \text{Var}(U(\boldsymbol{\theta})) \left[E \left(\frac{\partial U(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}'} \right) \right]^{-1'}$, where $\boldsymbol{\theta} = (\beta_0^{**}, \boldsymbol{\beta})'$. This approach is also doubly robust to misspecification of either the propensity score or the outcome regression and approximates the nonbinning approach as the number of bins increases (eAppendix 4; <http://links.lww.com/EDE/B260>).

Interpretations

The coefficient of the exposure represents the hazard difference for each one unit increase in the exposure. Because the logarithm link function is collapsible, the marginal effect and the conditional effect are numerically the same (eAppendix 5; <http://links.lww.com/EDE/B260>). Given a survival probability close to 1 per year, $-\ln[1 - E(D_{ij} | x_{ij}, \mathbf{v}_{ij})]$ is approximately $E(D_{ij} | x_{ij}, \mathbf{v}_{ij})$, which is the probability of dying in each year given covariates. So the hazard difference is also approximately the risk difference (difference in the probability) of dying for each one unit increase in the exposure. This approximation may have a large bias for susceptible populations whose probability of survival is low. For example, for a positive effect, the estimated hazard difference underestimates the true risk difference. Multiplying the hazard difference by the number of people at risk per year will give the number of death attributable to each unit increase in exposure.

If we further assume no unmeasured confounders, positivity, and consistency, the coefficient for the exposure will estimate the causal hazard marginal or conditional difference for each one unit increase in exposure. The assumption of no unmeasured confounders is untestable. The estimates would still be an approximate to the true causal effect, if the specified propensity score and outcome regression models are close to the true model and the omitted variables are not influential. Sometimes it may be easier to collect information that explains the variation in exposure and correctly specify a propensity score, whereas in other cases, it may be easier to correctly specify an outcome regression.

The framework of doubly robust additive hazards models is flexible. Several extensions that relax the aforementioned assumptions are discussed in eAppendix 6; <http://links.lww.com/EDE/B260>. Because there are multiple choices of model, sensitivity analyses would be helpful to assess the robustness of the result.

Simulation Study

We assumed a 14-year follow-up with a constant baseline hazard, where $j = 1, 2, \dots, 14$. We simulated a binary covariate V_1 following Bernoulli (0.5), four time-invariant covariates V_2, V_3, V_4 , and V_5 for each person i following uniform(-0.5, 0.5), and five time-varying covariates V_6, V_7, V_8, V_9 , and V_{10} within each person-year ij following uniform(-0.5, 0.5). In the following three scenarios, we varied the number of subjects (n) as follows: 200, 500, 1,000, 2,000, 5,000, 10,000, 20,000, 50,000, and 100,000. All methods were tested using 1,000 randomly generated datasets.

Simulation 1: Double Robustness of DRAHM—The aim of simulation 1 is to show the double robustness of DRAHM. The exposure X_{ij} was constructed by $7 + 0.8 \times V_{7,ij} V_{8,ij} + V_{9,ij} V_{10,ij} + N(0,1)$. The probability of dying per person-year was constructed by

$$\max(0, 1 - \exp(\beta X_{ij} + 0.18 \times (V_{1,ij} + V_{2,ij}^2 + \exp(V_{3,ij}/3) V_{4,ij} + \cos V_{5,ij} + V_{5,ij} V_{2,ij} V_{1,ij} + 2V_{6,ij}^2 V_{7,ij}^2 + 3V_{7,ij}^3 + V_{8,ij} V_{9,ij} V_{10,ij})).$$

. With nonlinearity and interactive relationships, this formulation allows us to explore the consequences of model misspecification. β was either 0 or 0.005. The outcome variable (survival, $1 - D_{ij}$) was constructed using a Bernoulli random variable with a probability of dying constructed in the former equation. Censoring time for each subject was generated from a uniform distribution of (0, 15). If an individual has a simulated censoring time greater than or equal to j and smaller than $j + 1$, person-years from $j+1$ onward will be censored. We compared DRAHM-A, DRAHM-B, marginal structural models, and outcome regression. The details are given in Table 1.

Simulation 2: Dummy Variables in DRAHM—The aim is to assess the performance of DRAHM in the existence of dummy variables. We considered the misspecification of the functional form of follow-up year. The exposure X_{ij} was constructed by $7 + 0.8 \times V_{7,ij} V_{8,ij} + V_{9,ij} V_{10,ij} + 0.1 \times \text{Year}_{ij} + N(0,1)$. Year_{ij} represents the year of follow-up. The probability of dying in person-year p_{ij} was constructed by

$$1 - \exp(\beta X_{ij} - 0.1 \times \sin\left(2\pi \frac{\text{Year}_{ij} - 1}{13}\right) + 0.18 \times (V_{1,ij} + V_{2,ij}^2 + \exp(V_{3,ij}/3) V_{4,ij} + \cos V_{5,ij} + V_{5,ij} V_{2,ij} V_{1,ij} + 2V_{6,ij}^2 V_{7,ij}^2 + 3V_{7,ij}^3 + V_{8,ij} V_{9,ij} V_{10,ij} + V_{6,ij} V_{7,ij} V_{10,ij})).$$

. β was either 0 or 0.005. So, β is confounded by follow-up year. Other settings are the same as simulation 1. We specifically compared the performance of DRAHM-A and DRAHM-B to model year as a dummy variable. Details for simulation 2 are given in Table 2.

Simulation 3: Truncation of IPW—The aim is to assess the utility of truncating IPW (0, 0.005, 0.01, 0.02, 0.05, 0.10, and 0.20) from each tail of the distribution of IPW, when the error distribution has a heavier tail than normal distribution. We modified simulation 1 by replacing the residuals, $N(0,1)$, in the true propensity score model with a standard t-distribution with three degrees of freedom. We used DRAHM-A, DRAHM-B, and marginal structural models where the functional form of propensity score was correct, but the error distribution was assumed normal, and the functional form of outcome regression was a linear combination (LC) of covariates (misspecified).

Motivating Example Data Analysis

We estimated the hazard difference of long-term exposure to $PM_{2.5}$ on the survival of Medicare beneficiaries (age ≥ 65) using doubly robust additive hazards models. We obtained the enrollment records of beneficiaries between 1 January 2000, and 31 December 2013, and who resided in one of the seven states in the Southeastern United States (Alabama, Florida, Georgia, Mississippi, North Carolina, South Carolina, and Tennessee).

We constructed a counting process dataset starting from the maximum of 2000 or year of entry up to the year of death or censoring whichever occurred first. Most censoring is administrative at the end of the study period, which may make it plausible to assume that the censoring was noninformative. Note that the same Andersen-Gill data format can be used to fit both doubly robust additive hazards models and Cox proportional hazards regression.

The exposure was the annual $PM_{2.5}$ in the ZIP code tabulation area (ZCTA) that each beneficiary resided in, which was estimated using a hybrid satellite-based prediction model.²⁵ The confounders include individual-, ZCTA-, and county-level covariates (age, race, sex, socioeconomic status, medical history, and area-level health variables), which were assumed constant within a follow-up year but could vary between follow-up years. A detailed description of the covariates is published elsewhere¹⁷ and is also given in eAppendix 7; <http://links.lww.com/EDE/B260>.

We applied DRAHM-A, DRAHM-B, marginal structural models, and outcome regression to the dataset. The propensity score was constructed using a linear regression of $PM_{2.5}$ against calendar year, state, linear terms of the covariates described above, and (1) a constant time trend or (2) dummy variables for each follow-up year. $s(v_{ij}; \gamma)$ in the outcome regression was the same as the propensity score model. In this motivating example, the outcome regression model may be less likely to represent the underlying biologic mechanism for probability of death, because the dataset does not measure a number of important individual-level covariates such as smoking, alcohol consumption, obesity, and family history of diseases. For example, individual-level smoking may be predictive of all-cause mortality and is not included in our outcome model. However, any association between smoking and $PM_{2.5}$ cannot be causal, i.e., smoking does not cause ambient air pollution to go up. Rather it is likely a result of some other variables such as poverty, which is predictive both of smoking and of living in a more highly polluted area. Hence, if poverty and other socioeconomic status factors are controlled for in the propensity score model, unbiased estimates are likely to be derived for $PM_{2.5}$ even in the absence of individual-level smoking in the outcome regression model. We conducted several stratified analyses to explore additive effect modifications by sex, race, age group, medical history, dual eligibility, and urbanicity. The hazard differences are compared with the estimated hazard ratios for the same set of subgroups reported in a previous publication.¹⁷

The study protocol was approved by the Harvard T.H. Chan School of Public Health Human Subjects Committee.

RESULTS

Simulation Study

In simulation 1, the misspecified outcome regression models produce biased estimates (Figure 1A) and higher mean squared errors (MSE) (Figure 1B) than the correct model. When either or both of propensity score and outcome regression is correctly specified, DRAHM-A and DRAHM-B have little bias (C and E) and a MSE lower than the model with a misspecified propensity score and outcome regression (D and F). The marginal structural model also has little bias when the propensity score is correctly specified but is biased when it is misspecified (Figure 1G). The results are similar when the true β is 0.005 (eFigure 1; <http://links.lww.com/EDE/B260>). DRAHM-B with 10 bins has a larger bias than DRAHM-A (eFigure 1; <http://links.lww.com/EDE/B260> and eTable 1; <http://links.lww.com/EDE/B260>): for example, when both propensity score and outcome regression are correct and $n = 1 \times 10^5$, the bias for DRAHM-A is 1×10^{-5} and DRAHM-B with 10 bins is 1×10^{-4} . eFigure 2; <http://links.lww.com/EDE/B260> and eTable 1; <http://links.lww.com/EDE/B260> show that when the number of subjects (similarly sample size) is large ($n > 5000$), the bias decreases with the number of bins (e.g., when both propensity score and outcome regression are correct and $n = 1 \times 10^5$, the bias is 2.7×10^{-4} with 10 bins, 1.1×10^{-4} with 20 bins, 4.1×10^{-5} with 50 bins, and 1.8×10^{-5} with 100 bins). When the number of subject (similarly sample size) is small but the number of bins is large (e.g., $n < 500$ for 100 bins), the binning makes the approach unstable and the coefficient not estimable.

In simulation 2, DRAHM-A is doubly robust (eFigure 3a, b; <http://links.lww.com/EDE/B260>). DRAHM-B, however, is unable to predict the average probability of survival for the entire data (no estimates in [C] and [D]), when the subset does not contain data for all follow-up years. In this case, DRAHM-B requires a correct specification of the propensity score to estimate the causal effect. The marginal structural model produces an estimate with little bias when the propensity score is correctly specified. In addition, the results are similar when the true β is 0.005 (eFigure 4; <http://links.lww.com/EDE/B260>).

When either or both of propensity score and outcome regression is correctly specified, the coverage of the sandwich estimator is close to the nominal value (0.95 in this case) in both simulations 1 and 2 (eFigure 5; <http://links.lww.com/EDE/B260>). In simulation 1, the coverage of the confidence intervals for DRAHM-B approximates the nominal value as the number of bins increases (eFigure 6; <http://links.lww.com/EDE/B260>).

When the residual of the true propensity score is t-distributed (simulation 3), a small percentage of truncation (0.5% or 1%) from each tail reduces the variance of the estimator (eFigures 7 and 8; <http://links.lww.com/EDE/B260>).

PM_{2.5} and survival in the Southeast United States

The dataset for the motivating example contains 13 million Medicare beneficiaries that were followed-up a maximum of 14 years of follow-up from 2000 to 2013 in the seven Southeastern United States.¹⁷

Using the propensity score model described above, we found that 0.002% of the IPW were influential and comprised of 99.99% of the IPW weights. To stabilize the estimates, we truncated the lowest and highest 1% of the IPW weights in the following analysis.

Using a common intercept for each follow-up year in the propensity score model and using DRAHM-A (Figure 2A), we found an increase of 8.3×10^{-4} (95% confidence interval = 7.8×10^{-4} , 8.9×10^{-4}) in mortality hazard for each $1 \mu\text{g m}^{-3}$ increase in $\text{PM}_{2.5}$. Using DRAHM-B, the hazard difference varied with the number of bins. We found an increase of 3.7×10^{-4} in hazard using five bins and 1×10^{-3} for 1000 bins. By comparison, the hazard difference was 1.4×10^{-3} using the marginal structural model and the hazard difference was 7×10^{-4} using the outcome regression.

Using a dummy variable for each follow-up year in the propensity score and outcome regression and using DRAHM-A (Figure 2B), we found that the hazard, which approximates the probability of dying, increases 8.0×10^{-4} (7.4×10^{-4} , 8.7×10^{-4}) for each $1 \mu\text{g m}^{-3}$ increase in $\text{PM}_{2.5}$. The estimate is causal if one of the propensity score model or the outcome regression model is correctly specified. The estimate of DRAHM-A is closer to that of the outcome regression without IPW weights and is lower than the marginal structural model (Figure 2B). Similar to simulation 2, DRAHM-B fails with the use of dummy variables for follow-up years. By comparison, we found a hazard ratio of 1.021 (1.019, 1.022) using Cox proportional hazard model adjusting for the same set of covariates.¹⁷ We further found that an increase in the percentage of truncation from 1% to 5% resulted in a decrease in the effect estimate for $\text{PM}_{2.5}$ (eFigure 9; <http://links.lww.com/EDE/B260>).

Given that the average Medicare population per year, the results suggest that a $1 \mu\text{g m}^{-3}$ reduction of annual $\text{PM}_{2.5}$ exposure in these states could prevent 5.5 (5.1, 6.0) thousand premature deaths each year based on the approximate risk difference of 8.0×10^{-4} . This estimate may underestimate of the true mortality attributable for $1 \mu\text{g m}^{-3}$ reduction in annual $\text{PM}_{2.5}$ because of the approximation of risk difference using hazard difference.

The results for subgroup analyses are presented in Figure 3 (DRAHM-A, dummy variable for baseline hazard, 1% IPW truncation). We found that the hazard difference was substantially higher as age increases, among individuals with previous hospital admissions for congestive heart failure (CHF), chronic obstructive pulmonary disease (COPD), diabetes, or myocardial infarction (MI); dual eligibility for Medicaid; and who reside in urban areas. By comparison, some effect modification using hazard difference (e.g., hospital admissions, dual eligibility) showed the same direction as hazard ratios, whereas some (e.g., age) does not.¹⁷ In addition, we found that black people with dual eligibility had a hazard difference of 1.60×10^{-3} (1.21 , 1.99) $\times 10^{-3}$. Black people age 75–84 years old had a hazard difference of 2.28×10^{-3} (1.80 , 2.75) $\times 10^{-3}$, suggesting an increase in susceptibility among these populations.

DISCUSSIONS

We proposed two DRAHMs in survival analysis to estimate hazard difference (approximately risk difference). Both approaches are robust to the misspecification of either

the propensity score or outcome regression model, whereas DRAHM-A is superior in terms of bias, stability, and computational cost.

Doubly robust additive hazards models supplement the Cox proportional hazards model to estimate the associations between $PM_{2.5}$ and survival.^{12,13,16,26,27} The additive nature of the proposed approach makes the interpretation of the association clearer and easier. The hazard difference associated with the long-term effect of $PM_{2.5}$ is increased substantially with age: 0.02% for age 65–74, 0.25% increase per $1 \mu g m^{-3}$ for age 75–84, and 0.47% for age > 84. By comparison, the Cox model adjusting for the same set of variables suggested that the hazard ratio decreased with age.¹⁷ This suggests that the baseline hazard is much higher among the oldest, and a larger increase on the additive scale may still correspond to a smaller increase on the multiplicative scale. The additive and multiplicative effect modifications by dual eligibility suggest that beneficiaries with a lower socioeconomic status (e.g., lower income) had a higher hazard (or approximately the probability) of death compared with beneficiaries with a higher socioeconomic status. Here, we found that people with any previous hospitalizations due to CHF, COPD, diabetes, or MI had a substantially higher hazard (or approximately the probability) of death associated with $PM_{2.5}$. The direction of effect modification is the same for hazard difference and ratio. The effect estimate for people living in rural areas in the southeastern United States was also smaller.

CONCLUSIONS

Doubly robust additive hazards models can potentially be a useful approach to estimating hazard difference and is doubly robust to model misspecification. The applications to $PM_{2.5}$ added to the literature a hazard difference estimate with a novel and clearer interpretation.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This publication was made possible by USEPA grant RD-83587201. Its contents are solely the responsibility of the grantee and do not necessarily represent the official views of the USEPA. Further, USEPA does not endorse the purchase of any commercial products or services mentioned in the publication. This publication was also made possible by National Institute of Environmental Health Sciences (NIEHS) grants R01 ES024332, and ES000002, and Health Effects Institute grant 4953.

References

1. David CR. Regression models and life tables (with discussion). *J R Stat Soc.* 1972; 34:187–220.
2. Farrington CP. Interval censored survival data: A generalized linear modelling approach. *Stat Med.* 1996; 15:283–292. [PubMed: 8643886]
3. Aalen, O. A model for nonparametric regression analysis of counting processes. In: Klonecki, W.Kozek, A., Rosi ski, J., editors. *Mathematical Statistics and Probability Theory: Proceedings, Sixth International Conference; Wisla (Poland).* 1978; New York, NY: Springer New York; 1980. p. 1-25.
4. Knol MJ, VanderWeele TJ, Groenwold RH, Klungel OH, Rovers MM, Grobbee DE. Estimating measures of interaction on an additive scale for preventive exposures. *Eur J Epidemiol.* 2011; 26:433–438. [PubMed: 21344323]

5. Knol MJ, VanderWeele TJ. Recommendations for presenting analyses of effect modification and interaction. *Int J Epidemiol*. 2012; 41:514–520. [PubMed: 22253321]
6. Hernán MA, Brumback B, Robins JM. Marginal structural models to estimate the causal effect of zidovudine on the survival of HIV-positive men. *Epidemiology*. 2000; 11:561–570. [PubMed: 10955409]
7. Hernán MA, Cole SR, Margolick J, Cohen M, Robins JM. Structural accelerated failure time models for survival analysis in studies with time-varying treatments. *Pharmacoepidemiol Drug Saf*. 2005; 14:477–491. [PubMed: 15660442]
8. Robins JM, Hernán MA, Brumback B. Marginal structural models and causal inference in epidemiology. *Epidemiology*. 2000; 11:550–560. [PubMed: 10955408]
9. Bang H, Robins JM. Doubly robust estimation in missing data and causal inference models. *Biometrics*. 2005; 61:962–973. [PubMed: 16401269]
10. Rotnitzky, A., Robins, JM. *Encyclopedia of Biostatistics*. Hoboken, NJ: John Wiley & Sons, Ltd; 2005. Inverse probability weighting in survival analysis.
11. Forouzanfar MH, Alexander L, Anderson HR, et al. Global, regional, and national comparative risk assessment of 79 behavioural, environmental and occupational, and metabolic risks or clusters of risks in 188 countries, 1990–2013: A systematic analysis for the Global Burden of Disease Study. *Lancet*. 2013; 386:2287–2323.
12. Dockery DW, Pope CA 3rd, Xu X, et al. An association between air pollution and mortality in six U.S. cities. *N Engl J Med*. 1993; 329:1753–1759. [PubMed: 8179653]
13. Lepeule J, Laden F, Dockery D, Schwartz J. Chronic exposure to fine particles and mortality: An extended follow-up of the Harvard Six Cities study from 1974 to 2009. *Environ Health Perspect*. 2012; 120:965–970. [PubMed: 22456598]
14. Pope CA 3rd, Ezzati M, Dockery DW. Fine-particulate air pollution and life expectancy in the United States. *N Engl J Med*. 2009; 360:376–386. [PubMed: 19164188]
15. Schwartz J, Coull B, Laden F, Ryan L. The effect of dose and timing of dose on the association between airborne particles and survival. *Environ Health Perspect*. 2008; 116:64–69. [PubMed: 18197301]
16. Hoek G, Krishnan RM, Beelen R, et al. Long-term air pollution exposure and cardio- respiratory mortality: a review. *Environ Health*. 2013; 12:43. [PubMed: 23714370]
17. Wang Y, Shi L, Lee M, et al. Long-term Exposure to PM_{2.5} and Mortality Among Older Adults in the Southeastern US. *Epidemiology*. 2017; 28:207–214. [PubMed: 28005571]
18. Di Q, Wang Y, Zanobetti A, et al. Air pollution and mortality in the Medicare population. *N Engl J Med*. 2017; 376:2513–2522. [PubMed: 28657878]
19. Wang Y, Kloog I, Coull BA, Kosheleva A, Zanobetti A, Schwartz JD. Estimating causal effects of long-term PM_{2.5} exposure on mortality in New Jersey. *Environ Health Perspect*. 2016; 124:1182–1188. [PubMed: 27082965]
20. Kioumourtzoglou MA, Austin E, Koutrakis P, Dominici F, Schwartz J, Zanobetti A. PM_{2.5} and survival among older adults: Effect modification by particulate composition. *Epidemiology*. 2015; 26:321–327. [PubMed: 25738903]
21. VanderWeele TJ, Vansteelandt S. Invited commentary: Some advantages of the relative excess risk due to interaction (RERI)—towards better estimators of additive interaction. *Am J Epidemiol*. 2014; 179:670–671. [PubMed: 24488514]
22. Andersen PK, Gill RD. Cox's regression model for counting processes: A large sample study. *Ann Stat*. 1982; 10:1100–1120.
23. Kennedy EH, Ma Z, McHugh MD, Small DS. Non-parametric methods for doubly robust estimation of continuous treatment effects. *J R Stat Soc Series B Stat Methodol*. 2016; 79:1229–1245. [PubMed: 28989320]
24. Cole SR, Hernán MA. Constructing inverse probability weights for marginal structural models. *Am J Epidemiol*. 2008; 168:656–664. [PubMed: 18682488]
25. Lee M, Kloog I, Chudnovsky A, et al. Spatiotemporal prediction of fine particulate matter using high-resolution satellite images in the Southeastern US 2003–2011. *J Expo Sci Environ Epidemiol*. 2016; 26:377–384. [PubMed: 26082149]

26. Puett RC, Hart JE, Yanosky JD, et al. Chronic fine and coarse particulate exposure, mortality, and coronary heart disease in the Nurses' Health Study. *Environ Health Perspect.* 2009; 117:1697–1701. [PubMed: 20049120]
27. Pope CA 3rd, Thun MJ, Namboodiri MM, et al. Particulate air pollution as a predictor of mortality in a prospective study of U.S. adults. *Am J Respir Crit Care Med.* 1995; 151(3 Pt 1):669–674. [PubMed: 7881654]

FIGURE 1.

Bias (A, C, E, and G) and the mean squared error (MSE) (B, D, F, and H) of the estimated coefficient for the exposure in simulation 1 where the true β is zero. Each method corresponds to a method listed under the first scenario in Table 1.

FIGURE 2.

Hazard difference of death for each $1 \mu\text{g m}^{-3}$ increase in $\text{PM}_{2.5}$ using double robust additive hazard models (DRAHM-A, DRAHM-B with 5, 10, 20, 50, 100, 200, 500, and 1000 bins, marginal structural modeling, and outcome regression using (A) an intercept for follow-up year in the propensity score and outcome regression and (B) dummy variables for follow-up year in the propensity score and outcome regression. The upper 1% and lower 1% of the inverse probability weights were truncated.

FIGURE 3.

Hazard difference of death for each $1 \mu\text{g m}^{-3}$ increase in $\text{PM}_{2.5}$ for each of the subgroups using doubly robust additive hazard model DRAHM-A. The upper 1% and lower 1% of the inverse probability weights were truncated.

TABLE 1

Model Specification for Simulation 1

Method	Abbreviation	PS ^a	IPW Truncation (%)	OR ^b
Unadjusted model	Unadjusted			x
Misspecified outcome regression	LC			x and a linear combination of $V_1 - V_{10}$
Correctly specified outcome regression	OR+			The functional form as specified in the data generating procedure
DRAHM-A with misspecified PS and OR	PS- OR-, DRAHM-A	A linear combination of $V_1 - V_{10}$	0	x and a linear combination of $V_1 - V_{10}$
DRAHM-A with misspecified PS and correctly specified OR	PS- OR+, DRAHM-A	A linear combination of $V_1 - V_{10}$	0	The functional form as specified in the data generating procedure
DRAHM-A with correctly specified PS and misspecified OR	PS+ OR-, DRAHM-A	The functional form as specified in the data generating procedure	0	x and a linear combination of $V_1 - V_{10}$
DRAHM-A with misspecified PS and correctly specified OR	PS+ OR+, DRAHM-A	The functional form as specified in the data generating procedure	0	The functional form as specified in the data generating procedure
DRAHM-B with misspecified PS and OR with 10 bins	PS- OR-, 10 bins	A linear combination of $V_1 - V_{10}$	0	A linear combination of $V_1 - V_{10}$
DRAHM-B with misspecified PS and correctly specified OR with 10 bins	PS- OR+, 10 bins	A linear combination of $V_1 - V_{10}$	0	The functional form as specified in the data generating procedure leaving out x
DRAHM-B with correctly specified PS and misspecified OR with 10 bins	PS+ OR-, 10 bins	The functional form as specified in the data generating procedure	0	A linear combination of $V_1 - V_{10}$
DRAHM-B with misspecified PS and correctly specified OR with 10 bins	PS+ OR+, 10 bins	The functional form as specified in the data generating procedure	0	The functional form as specified in the data generating procedure leaving out x
MSM with a misspecified PS	PS-, MSM	A linear combination of $V_1 - V_{10}$	0	x only
MSM with a correctly specified PS	PS+, MSM	The functional form as specified in the data generating procedure	0	x only

OR+ indicates correctly specified outcome regression; OR-, misspecified outcome regression; PS-, misspecified propensity score; PS+, correctly specified propensity score.

^aThe functional form of the denominator model for the propensity score.

^bThe functional form of the outcome regression (the left-hand side of the outcome regression is $-\ln(1 - E(D_{ij}|X_{ij}, V_{ij}))$). x represents a linear term for the exposure variable in the simulation.

TABLE 2

Model Specification for Simulation 2

Method	Abbreviation	Part of PS to Model Year ^a	IPW Truncation (%)	Part of OR to Model Year and Exposure ^b
DRAHM-A with misspecified PS and OR for year	PS- OR-, DRAHM-A	a constant	0	x and a constant for year
DRAHM-A with misspecified PS for year and correctly specified OR	PS- OR+, DRAHM-A	a constant	0	x and a constant for year
DRAHM-A with correctly specified PS and misspecified OR for year	PS+ OR-, DRAHM-A	Dummy variables for each year	0	x and dummy variables for each year
DRAHM-A with misspecified PS for year and correctly specified OR	PS+ OR+, DRAHM-A	Dummy variables for each year	0	x and dummy variables for each year
DRAHM-B with misspecified PS and OR for year with 10 bins	PS- OR-, 10 bins	a constant	0	A constant for year
DRAHM-B with misspecified PS for year and correctly specified OR with 10 bins	PS- OR+, 10 bins	a constant	0	A constant for year
DRAHM-B with correctly specified PS and misspecified OR for year with 10 bins	PS+ OR-, 10 bins	Dummy variables for each year	0	Dummy variables for each year
DRAHM-B with misspecified PS for year and correctly specified OR with 10 bins	PS+ OR+, 10 bins	Dummy variables for each year	0	Dummy variables for each year
MSM with a misspecified PS for year	PS-, MSM	a constant	0	x only
MSM with a correctly specified PS	PS+, MSM	Dummy variables for each year	0	x only

OR+ indicates correctly specified outcome regression; OR-, misspecified outcome regression; PS-, misspecified propensity score; PS+, correctly specified propensity score.

^aThe functional form for year in the denominator model of the propensity score. In addition, we added the correct functional form for $V_1 - V_{10}$ as specified in the data generating procedure.

^bThe functional form for year in the outcome regression (the left-hand side of the outcome regression is $-\ln(1 - E(D_{ij} | X_{ij}, V_{ij}))$). x represents a linear term for the exposure variable in the simulation. In addition to the modeling of year, we included the correct functional form for $V_1 - V_{10}$ as specified in the data generating procedure.