

# Long-term memory effects of an incremental blood pressure intervention in a mortal cohort

Maria Josefsson<sup>1,\*</sup>, Nina Karalija<sup>2,3</sup>, Michael J. Daniels<sup>4</sup>

<sup>1</sup>Department of Statistics, USBE, Umeå University, SE-90187, Umeå, Sweden, <sup>2</sup>Center for Functional Brain Imaging, Umeå University, SE-90187, Umeå, Sweden, <sup>3</sup>Department of Integrative Medical Biology, Umeå University, SE-90187, Umeå, Sweden, <sup>4</sup>Department of Statistics, University of Florida, Gainesville, FL 32611, United States

\*Corresponding author: Maria Josefsson, Department of Statistics, USBE, Umeå University, SE-90187, Umeå, Sweden ([maria.josefsson@umu.se](mailto:maria.josefsson@umu.se)).

## ABSTRACT

In the present study, we examine long-term population-level effects on episodic memory of an intervention over 15 years that reduces systolic blood pressure in individuals with hypertension. A limitation with previous research on the potential risk reduction of such interventions is that they do not properly account for the reduction of mortality rates. Hence, one can only speculate whether the effect is due to changes in memory or changes in mortality. Therefore, we extend previous research by providing both an etiological and a prognostic effect estimate. To do this, we propose a Bayesian semi-parametric estimation approach for an incremental threshold intervention, using the extended G-formula. Additionally, we introduce a novel sparsity-inducing Dirichlet prior for longitudinal data, that exploits the longitudinal structure of the data. We demonstrate the usefulness of our approach in simulations, and compare its performance to other Bayesian decision tree ensemble approaches. In our analysis of the data from the Betula cohort, we found no significant prognostic or etiological effects across all ages. This suggests that systolic blood pressure interventions likely do not strongly affect memory, either at the overall population level or among individuals who would remain alive under both the natural course and the intervention (the always survivor stratum).

**KEYWORDS:** BART; dropout; G-computation; LDART; survival; truncation by death.

## 1 INTRODUCTION

Hypertension is a well-established risk factor for cognitive impairment and dementia (Livingston et al., 2020). However, the extent to which antihypertensive treatment can enhance cognitive abilities remains poorly understood, partly due to its long-term nature. Cognitive changes caused by hypertension likely develop over decades rather than years, making randomized clinical trials (RCTs) difficult to conduct (Iadecola et al., 2016). In this study, we extend previous research by examining the long-term population-level effects of an intervention that reduces systolic blood pressure (sBP) in individuals with hypertension. Our data come from the Betula Study, a representative prospective cohort study on aging, memory, and dementia (Nyberg et al., 2020). A key limitation of previous research is the insufficient consideration of altered mortality rates when estimating risk reduction. To address this, our study provides both etiological and prognostic intervention effect estimates. These estimates are critical for public health practice, offering a more comprehensive understanding of the intervention's impact on cognitive health.

The G-formula (Robins, 1986) is commonly used to estimate causal effects in longitudinal settings where time-varying confounders influence both treatment and outcome. It simulates an intervention and, under a set of assumptions, yields causal effect estimates when these assumptions hold. This approach can provide valuable insights when RCTs are not feasible, and have been

used previously for epidemiological research (eg, Zhou et al., 2020; Taubman et al., 2009). For a binary risk factor, this often involves setting the risk factor to a fixed value (Yes/No) (eg, Zhou et al., 2020), or by modifying a subject's odds of receiving treatment through a hypothetical intervention based on incremental propensity scores (Kennedy, 2019). However, for a continuous risk factor, setting the treatment dose to a fixed value for all subjects is often considered to result in highly unrealistic scenarios with limited practical use. For example, a hypothetical intervention where all subjects in the population fix their sBP levels at 140 mm Hg would not take into account the natural variation that is also seen among non-hypertensive subjects. Another complicating factor for a continuous risk factor arises when only subjects meeting a diagnostic criteria are offered treatment. For instance, only subjects with hypertension would be offered antihypertensive agents to control their blood pressure levels.

Two approaches for hypothetical interventions on a continuous risk factor, which only intervene on subjects that do not reach a prespecified threshold, are the *threshold intervention* (Taubman et al., 2009) and the *representative intervention* (eg, Picciotto et al., 2012; Young et al., 2019). The former assigns the threshold value to all individuals who do not meet the prespecified target, while the latter assigns a value drawn from the distribution among those in the population meeting the target level. Hence, both approaches ensure treatment remains within a prespecified interval. Nevertheless, it is common for a significant

Received: March 20, 2024; Revised: December 12, 2025; Accepted: December 22, 2025

© The Author(s) 2026. Published by Oxford University Press on behalf of The International Biometric Society. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.

proportion of the population to be unable to control their sBP at optimal levels, due to factors such as poor response to medications. One alternative approach in these settings is an *incremental intervention*, where a subject's treatment is instead shifted downward (or upwards) by some prespecified function (Haneuse and Rotnitzky, 2013; Muñoz and Van Der Laan, 2012). However, these approaches are not valid for longitudinal data with dropout and deaths.

Attrition due to dropout and deaths is common in longitudinal aging studies, especially when subjects are followed for an extended duration. The G-formula can address both missing response and missing covariate data under an assumption of missing at random (MAR; Robins et al., 1995; Young et al., 2019; Kim et al., 2021). Further strategies for sensitivity analyses have been proposed in situations where the missingness is thought to be missing not at random (MNAR; Josefsson and Daniels, 2021; Josefsson et al., 2023). However, studying cognition after death is typically not defined and is not of interest, leading to consideration of a *mortal cohort* approach (Dufouil et al., 2004). Mortal cohort inferences truncate subjects' outcomes after death. The first approach, *partly conditional inference* (PCI), conditions on the sub-population that is still alive at the end of follow-up (Kurland and Heagerty, 2005; Wen et al., 2018; Wen and Seaman, 2018). The second approach, *principal stratification*, conditions on the subpopulation that would survive through the end of follow-up regardless of treatment assignment, known as *the always survivor stratum* (Frangakis and Rubin, 2002; Frangakis et al., 2007). Although the two approaches have similarities, they differ with respect to the type of information they generate. PCI has been proposed for estimating non-causal associations and is particularly useful for prognostication, that is when the interest is in the health outcome among all individuals who are alive (eg, Taubman et al., 2009). In contrast, principal stratification is useful in situations where survival differs between treated and control groups and has been applied when the interest is in etiologic effect estimates (eg, Hayden et al., 2005; Tchetgen Tchetgen, 2014; Wang et al., 2017).

A drawback of many approaches for estimating the Survivor Average Causal Effect is that they rely on the assumption of monotonicity for valid inference (see, eg, Josefsson and Daniels, 2021; Tan et al., 2021; Shardell and Ferrucci, 2018). This assumption implies that for a beneficial treatment, individuals who survive under treatment would also have survived in the absence of treatment. However, we consider this to be an implausible assumption in most, if not all, real-world settings. This is particularly true for populations with heterogeneous health profiles or competing risks of mortality. To address this limitation, Zaidi et al. (2019) proposed methods for sensitivity analyses that characterize causal effects at the population level while allowing for violations of monotonicity. In the context of clinical trials, Roy et al. (2008) and Lee et al. (2010) developed methods based on the less restrictive *stochastic monotonicity* assumption. This assumption relaxes standard monotonicity by allowing the probability that survival is higher under treatment than under control, conditional on covariates, to be less than one. However, these approaches are not valid for prospective data with time-varying treatments and time-varying confounding.

Our study extends previous research by providing both etiological and prognostic intervention effect estimates, which are both critical for public health practice, offering a more comprehensive understanding of the intervention's impact. For estimation, we propose a Bayesian semi-parametric approach based on the extended G-formula (Robins et al., 2004). In addition, we introduce a novel sparsity-inducing Dirichlet prior designed for regression with longitudinal data, and demonstrate the usefulness of our approach in a simulation relative to other Bayesian tree ensemble approaches.

The paper is organized as follows. Section 2 introduces the intervention in a Bayesian setting; Section 3, the two types of mortal cohort inferences; Section 4, computational details and the sparsity-inducing Dirichlet prior; Section 5, a simulation study; Section 6, the Betula application; and Section 7, conclusions.

## 2 METHODS

Consider a prospective cohort study where measurements on a set of variables are collected on each of  $i = 1, 2, \dots, n$  subjects over a specified follow-up period at  $t = 1, 2, \dots, T$  fixed time points. At time  $t$ , let  $X_{it}$  denote the vector of time-varying confounders. We define  $A_{it}$  as a continuous prognostic variable reflecting actionable treatment recommendations; for simplicity, we refer to  $A_{it}$  as the "treatment" throughout. Denote the continuous outcome measure by  $Y_{it}$ , where interest in this study is the outcome measure at the last time point  $T$ . Furthermore, let  $S_{it} = 1$  indicate if a subject is alive at time point  $t$ ; and  $S_{it} = 0$  otherwise. Let  $R_{it} = 1$  indicate if subject  $i$  has completed the cognitive testing at time  $t$ ; and  $R_{it} = 0$  otherwise. Let the observed history of a time-varying variable be denoted with an  $\leftarrow$  and the future with an  $\rightarrow$ . At each time point  $t$ , we assume data are observed in the temporal ordering  $S_{it}, R_{it}, X_{it}, A_{it}, Y_{it}$ , and for each subject survival status is observed throughout follow-up, while the remaining variables are observed only during the period the subject actively participates in the study. Lower-case letters denote possible realizations of a random variable. We also write  $H_{it}$  to denote the observed past history prior to the outcome variable at time point  $t$ . Specifically,  $H_{it} = (\overleftarrow{Y}_{it-1} = \overleftarrow{y}_{it-1}, \overleftarrow{X}_{it} = \overleftarrow{x}_{it}, \overleftarrow{R}_{it} = \overleftarrow{r}_{it}, \overleftarrow{S}_{it} = 1)$ , with support on the set of possible histories  $\mathcal{H}_{it}$ . Note,  $\overleftarrow{A}_{it} = \overleftarrow{a}_{it}$  is not included in  $H_{it}$ . In what follows, we suppress the subscript  $i$  to simplify notation.

### 2.1 A hypothetical incremental threshold intervention

For the Betula data, we are interested in the long-term population-level effects of a 15-year intervention that reduces sBP in individuals with hypertension. As such, we consider two contrasting regimes:  $g_*$ , represents a setting where all individuals follow the intervention during the studied period. The natural course,  $g_0 = \{\overleftarrow{A}_T = \overleftarrow{a}_T\}$ , represents a setting in which all individuals follow their natural progression in the absence of an intervention. At each follow-up time point, we consider an intervention strategy in which treatment assignment is determined by a fixed rule based on the so-called the natural value of treatment (NVT), denoted by  $A_t^*$ . Following Young et al. (2014),  $A_t^*$  is defined as the treatment measure that would have been observed

at time  $t$  if the intervention had been discontinued right before  $t$ . In particular, for the Betula data, we consider an intervention in which we intervene only on subjects whose  $A_t^* > \tau$ . The threshold  $\tau$  is set by the researcher. We consider three cut-offs using values of 130 mmHg, 140 mmHg, and 150 mmHg for  $\tau$ .

We propose an intervention that assigns a new treatment value to individuals whose  $A_t^* > \tau$ . The new value is obtained by shifting  $A_t^*$  downward by  $\delta_t$ . To quantify uncertainty in the intervention, we treat the shift  $\delta_t$  as an unknown parameter and impose a prior distribution on it. The prior reflects the researcher's beliefs about the plausible values that treatment can take under the intervention. Specifically, if  $a_t^* \leq \tau$ , then the conditional intervention density is  $p_*(a_t | \bar{y}_{t-1}, a_t^*, \bar{a}_{t-1}, \bar{x}_t, \bar{r}_t, \bar{s}_t = 1) = I\{a_t = a_t^*\}$ . However, if  $a_t^* > \tau$ , then  $p_*(a_t | \bar{y}_{t-1}, a_t^*, \bar{a}_{t-1}, \bar{x}_t, \bar{r}_t, \bar{s}_t = 1) = p(\delta_t)I\{a_t = a_t^* - \delta_t\}$ .

### 3 INTERVENTION EFFECTS FOR MORTAL COHORTS

In this section, we describe two types of mortal cohort inferences under the proposed intervention: PCI and inference on the always survivor stratum. Let  $\bar{Y}_T(g)$ ,  $\bar{A}_T(g)$ ,  $\bar{R}_T(g)$ ,  $\bar{S}_T(g)$ , and  $\bar{X}_T(g)$  represent the counterfactual outcome, treatment, missingness, survival, and covariate histories associated with the regime  $g$ .

#### 3.1 Partly conditional inference

Partly conditional models have been proposed to address truncation by death under hypothetical interventions. For these models, inference is conditional on the sub-population alive at a specific time-point, in our application being alive at end of follow-up (time  $T$ ). Hence, inferences focus on what we call a *partly conditional intervention effect*:  $PCIE = E[Y_T(g_*) | S_T(g_*) = 1] - E[Y_T(g_0) | S_T(g_0) = 1]$ .

Although the PCIE gives a crude comparison of differences in potential outcomes under the two regimes, which is useful for prognostication, it often implicitly compares potential outcomes for different underlying populations. For example, it is likely that an sBP intervention may affect survival rates; as such, the population surviving under the natural course will differ from the one surviving under the intervention.

The identification of the PCIE from longitudinal observational data relies on standard assumptions in causal inference: C1, consistency, C2, conditional exchangeability, and C3, sequential positivity. Additionally, we assume: C4, missing at random conditional on survival (MARS) for the outcome, treatment, and time-varying confounders. For a detailed discussion of these assumptions and their implications, see [Supplementary Appendix A](#). Note that, for  $g_0$ , assumption C2 implies that, at each time point, treatment assignment is independent of potential outcomes given the observed past, which we denote as C2a. For  $g_*$ , this assumption states that the NVT at time  $t$  has no direct effect on the outcome (or survival) except through future measurements of the treatment. An implication of this assumption is that the NVT is not itself a time-varying confounder (Richardson and Robins, 2013; Young et al., 2014). We denote this as-

sumption as C2b. For the Betula data we perform sensitivity analyses to investigate robustness of results to violations of assumptions C2, C3, and C4, (see Section 6).

Under the assumptions C1, C2a, C3, and C4, the mean outcome at time  $T$  under regime  $g_0$ , can be computed using the *g-formula* (Robins, 1986) conditioning on survival,

$$\int_{\bar{r}_T} \int_{\bar{x}_T} \int_{\bar{a}_T} \int_{\bar{y}_{T-1}} E[Y_T | \bar{y}_{T-1}, \bar{a}_T, \bar{x}_T, \bar{r}_T, \bar{s}_T = 1] \times \prod_{k=1}^T p(a_k | \bar{y}_{k-1}, \bar{a}_{k-1}, \bar{x}_k, \bar{r}_k, \bar{s}_k = 1) \times p(x_k | \bar{y}_{k-1}, \bar{a}_{k-1}, \bar{x}_{k-1}, \bar{r}_k, \bar{s}_k = 1) \times p(r_k | \bar{y}_{k-1}, \bar{a}_{k-1}, \bar{x}_{k-1}, \bar{r}_{k-1}, \bar{s}_k = 1) \times p(s_k | \bar{y}_{k-1}, \bar{a}_{k-1}, \bar{x}_{k-1}, \bar{r}_{k-1}, \bar{s}_{k-1} = 1) \times p(y_{k-1} | \bar{y}_{k-2}, \bar{a}_{k-1}, \bar{x}_{k-1}, \bar{r}_{k-1}, \bar{s}_{k-1} = 1) \times d\bar{y}_{T-1} d\bar{a}_T d\bar{x}_T d\bar{r}_T. \tag{1}$$

In contrast, for the intervention,  $g_*$ , under assumptions C1, C2b, C3, and C4, the mean outcome at time  $T$  can be computed using the *Extended g-formula* (Robins et al., 2004). The key difference is that  $p(a_k | \bar{y}_{k-1}, \bar{a}_{k-1}, \bar{x}_k, \bar{r}_k, \bar{s}_k = 1)$  in (1) is replaced by  $\int_{a_t^*} p_*(a_k | \bar{y}_{k-1}, a_t^*, \bar{a}_{k-1}, \bar{x}_k, \bar{s}_k = 1) \times p(a_k^* | \bar{y}_{k-1}, \bar{a}_{k-1}, \bar{x}_k, \bar{s}_k = 1)$ ; details are given in [Supplementary Appendix B](#).

The PCIE is then computed as the differences in expected outcomes under the two regimes  $g_0$  and  $g_*$ . More details are given in [Supplementary Appendix C](#).

#### 3.2 Inference on the always survivor stratum

Conditioning on survival as described in Section 3.1 may introduce bias due to the fact that survival is a post-randomization event (Zhang and Rubin, 2003). Hence, the partly conditional approach in Section 3.1 cannot address etiological research questions in mortal cohorts. An alternative estimand is the causal effect on the subpopulation who would survive through the end of follow-up regardless of treatment regime (Frangakis and Rubin, 2002; Frangakis et al., 2007). Our goal here is to estimate the difference in expected potential outcomes between  $g_0$  and  $g_*$  in the always survivor stratum. That is, what we call a *survivors average intervention effect*:  $SAIE = E[Y(g_*) - Y(g_0) | S_T(g_0) = S_T(g_*) = 1]$ . Note, previous studies have shown a link between reduced sBP and reduced all-cause mortality (Bundy et al., 2017). Hence, it is likely that an intervention targeting sBP not only may have an effect on cognition but also on mortality rates. In such situations, the PCIE would compare potential outcomes for different underlying populations. In particular, if subjects live longer under the risk-reducing intervention this population may include less healthy subjects who would die under the natural course, masking a possible intervention effect. The SAIE on the other hand, compare potential outcomes from the same population, namely subjects who would survive both when intervening on sBP and under the natural course. As such, it provides an estimate that adjusts for differences in mortality rates between the two regimes. In our application, we are interested in two types of effect estimates: a prognostic effect, the PCIE, and an etiological effect, the SAIE.

For identification of the SAIE, we need two additional assumptions.

**C5 Stochastic Monotonicity:**  $\Pr[S_T(g_*) = 1 \mid S_T(g_0) = 1] \geq \Pr[S_T(g_*) = 1 \mid S_T(g_0) = 0]$ . In the current application, we impose an intervention that improves individuals' health. As such, we assume the probability of survival under the intervention  $g_*$  is higher among those subjects that would survive under the natural course  $g_0$ , compared to those who would not. We denote  $\Pr[S_T(g) = 1]$  by  $\psi^g$  for  $g$  in  $\{g_0, g_*\}$ , and we assume  $\psi^{g_*} \geq \psi^{g_0}$ . We further assume  $\Pr[S_T(g_*) = 1 \mid S_T(g_0) = 1] = \psi^{g_*} + \lambda \left[ \min \left\{ 1, \frac{\psi^{g_*}}{\psi^{g_0}} \right\} - \psi^{g_*} \right]$ , where  $\lambda$  is a sensitivity parameter (Roy et al., 2008; Lee et al., 2010). Note,  $0 \leq \lambda \leq 1$ . If  $\lambda = 1$  this assumption corresponds to a *deterministic monotonicity* assumption where subjects who were to be alive under the observed regime  $g_0$  would also be alive under the intervention  $g_*$ , i.e.,  $S_T(g_0) \leq S_T(g_*)$ . In contrast, when  $\lambda = 0$ , this corresponds to assuming independence between  $S_T(g_0)$  and  $S_T(g_*)$ ; we consider this assumption unlikely in the present application.

**C6 Difference in expectations of outcomes when comparing different strata.**  $g \in \{g_*, g_0\}$ ,  $g' \in \{g_*, g_0\}$  and  $g \neq g'$ , we assume,  $E[Y_T(g) \mid S_T(g') = S_T(g) = 1] - E[Y_T(g) \mid S_T(g) = 1, S_T(g') = 0] = \Delta^g$ , where  $\Delta^g$  is a sensitivity parameter. We further assume that  $\Delta^g = \Delta^{g'} = \Delta$ , where the difference in expectations of potential outcomes when comparing different strata is the same for the two contrasting regimes. Note, if  $\Delta = 0$  there is no difference in potential outcomes when comparing the always survivor strata to the strata where individuals were to live under regime  $g$  but not under the contrasting regime  $g'$ . In contrast, if  $\Delta > 0$  this implies a higher potential outcome is observed for the always survivor stratum. In the current application, we expect higher memory scores (better memory) for the always survivor stratum.

The identification details of the SAIE are given in [Supplementary Appendix C](#). Briefly, under assumptions C1-C6, the SAIE is given by:  $\text{PCIE} + \Delta \left\{ \psi^{g_*} + \lambda \left( U - \psi^{g_*} \right) \right\} \left( 1 - \frac{1}{U} \right)$ , where  $U = \min \left\{ 1, \frac{\psi^{g_*}}{\psi^{g_0}} \right\}$ . The equation represents a location shift upwards from the PCIE for a beneficial intervention where we expect the participants to improve their health under the intervention compared to the natural course. Note,  $\psi^g = E_{(H_T, g)} \{ \Pr[S_T = 1 \mid H_T, \bar{A}_T = g] \}$ , for  $g \in \{g_*, g_0\}$ , are computed by marginalizing over the distributions of the set of temporally preceding variables using the  $g$ -formula and the extended  $g$ -formula.

#### 4 A SEMI-PARAMETRIC APPROACH FOR ESTIMATION

In this section, we develop a Bayesian semi-parametric (BSP) modeling approach using *soft* Bayesian additive regression trees (BART) models to estimate each of the conditional distributions in (1) and (??) in [Supplementary Appendix B](#). The *soft* BART model improves upon the traditional *hard* BART model by addressing the lack of smoothness inherent in ensemble approaches. To enhance valid inference, especially in scenarios with multiple time points and a large set of regressors, we extend

the current *soft* BART algorithm by introducing a novel sparsity-inducing prior for longitudinal data. This longitudinal Dirichlet prior promotes parsimony by penalizing groups of predictors less likely to be important in longitudinal prediction settings. We refer to our model as a longitudinal Dirichlet additive regression trees (LDART) approach.

As in traditional (*hard*) BART models, the conditional mean function is given by:  $\mu = \mu_0 + \sum_{b=1}^B g(\mathbf{d}; \mathcal{T}_b, \mathcal{M}_b)$ , where  $\mu_0$  is a global intercept, and the BART function is represented as a sum of regression trees. Here,  $\mathbf{d} = \{d_1, \dots, d_p\}$  denotes the set of predictors,  $\mathcal{T}_b$  the structure of the  $b$ th tree, and  $\mathcal{M}_b$  the corresponding terminal node parameters. Continuous outcomes are typically modeled with a Gaussian distribution, whereas binary outcomes are modeled using a probit link.

The regularization prior, controlling the complexity of the sum-of-trees, differ in two ways between the standard BART algorithm and *soft* BART. First, for *soft* BART, the path down the decision tree is probabilistic, where the sharpness of the decision is controlled by the bandwidth parameter. Thereby adapting to the smoothness level of the regression function even for low dimensional tree ensembles, a known shortcoming for (*hard*) BART models. Second, the  $j$ th predictor, from the set of  $P$  predictors, is chosen according to a probability vector  $\mathbf{q} = (q_1, \dots, q_p)$ , typically specified as uniform. In contrast, *soft* BART models incorporate a Dirichlet distribution for  $\mathbf{q}$ , such that  $\mathbf{q} \sim \mathcal{D}(\alpha/P, \dots, \alpha/P)$ . The sparsity-inducing prior was first introduced for high dimensional settings (Lineró, 2018), and can be extended to allow sparsity both within and between groups (Lineró and Yang, 2018; Du and Linero, 2019). Determining the degree of sparsity in the model is done by placing a prior on  $\alpha$ , as such,  $\frac{\alpha}{\alpha + \rho} \sim \text{Beta}(a, b)$ . Default specifications for  $a$ ,  $b$ , and  $\rho$  is 0.5, 1, and  $P$ , respectively. This sparsity inducing prior is implemented for both *hard* and *soft* BART. We refer to BART models with this prior as Dirichlet additive regression trees (DART).

##### 4.1 A grouping prior for longitudinal data

In a longitudinal setting, one may want to introduce a larger degree of sparsity for predictors collected at earlier time points, if they are thought to be less strongly associated with the response. Below we introduce a novel sparsity-inducing prior.

We consider a hierarchical structure where the predictors are divided into two groups: predictors measured at the *current* time point  $t$  and predictors measured at *past* time points  $k = 1, \dots, t - 1$ . Moreover, baseline predictors, eg, age and sex, are grouped with the *current* predictors into one group, since we want to impose the same degree of sparsity on these predictors. Let  $w$  refer to the probability of a predictor being from the set of *current* predictors, versus from the set of *past* predictors; we specify a beta prior,  $w \sim \text{Beta}(a, b)$ . For the *current* predictors, let  $v_j^t$  refer to the probability of selecting the  $j$ th predictor from the set of  $P_t$  *current* predictors  $\mathbf{v}^t = (v_1^t, \dots, v_{P_t}^t)$ ;  $\mathbf{v}^t$  is given a Dirichlet prior  $\mathbf{v}^t \sim \mathcal{D}(\eta/P_t, \dots, \eta/P_t)$ . So a priori, the probability of selecting the  $j$ th *current* predictor to construct a given split is equal to  $w \times v_j^t$ . The *past* predictors are divided into  $t - 1$  groups, where  $P_k$  is the size of group associated with time  $k$ . Let

$u^k$  denote the probability of selecting the  $k$ th set of time-varying predictors, and let  $v_j^k$  denote the probability of selecting the  $j$ th predictor in the  $k$ th set of time-varying predictors. To introduce sparsity related to proximity in time to the response we give  $\mathbf{u}$  a grouping Dirichlet prior  $\mathbf{u} \sim \mathcal{D}(\alpha_1/t - 1, \dots, \alpha_{t-1}/t - 1)$ , where we determine the degree of sparsity in the model by placing a prior on  $\alpha_k$  such that a larger degree of sparsity is obtained for smaller  $k$ 's (see details below). Similarly  $\mathbf{v}^k$  is also given Dirichlet prior  $\mathbf{v}^k \sim \mathcal{D}(\varphi^k/P_k, \dots, \varphi^k/P_k)$ . So a priori, the probability of selecting the  $j$ th predictor from the  $k$ th time point to construct a given split is equal to  $(1 - w) \times u^k \times v_j^k$ .

For the hyperparameters  $\eta$  and  $\varphi^k$  we use the default Beta hyperprior as implemented in DART. For  $\alpha_j$  we specify a Beta hyperprior  $\frac{\alpha_j}{\alpha_j + \rho_j} \sim \text{Beta}(c_j, d_j)$ ; we let  $c_j$  vary as a linear function of time, set  $d_j = 1$ , and  $\rho_j = t$  for all  $j$ . To introduce sparsity related to proximity in time to the response we set  $c_j = 1 - \frac{(t-j)}{t-1} \times 0.5$ . This results in ordered hyperpriors for the  $\alpha_j$ s, where  $c_1 = 0.5$  for predictors measured at the first time point and  $0.5 < c_{j-1} < c_j < 1$  for  $j > 1$ . That is, a higher degree of sparsity is introduced for predictors measured earlier in time and less sparsity for predictors measured closer in time to the response. The simulations in Section 5 shows the strong performance of these prior choices. Details about the posterior computations of the L-Dirichlet (LDART) prior are obtained in the [Supplementary Appendix D](#).

The PCIE and SAIE use the algorithm described in [Supplementary Appendix D](#), implemented in the R package `GCompBART` (Josefsson, 2025).

### 5 SIMULATION STUDY

In this simulation study, we evaluate the performance of LDART and compare it to other BART specifications.

Data were generated according to a modified version of the Friedman's five-dimensional test function (Friedman, 1991), mirroring a longitudinal study with five time-varying predictors and a continuous outcome. Specifically,  $f(\overleftarrow{\mathbf{x}}_i) = \sum_{t=1}^T c_t [10 \sin(\pi x_{1,it} x_{2,it}) + 20(x_{3,it} - 0.5)^2 + 10x_{4,it} + 5x_{5,it}]$ , where

$Y_{iT} \sim \text{Normal} \left\{ f(\overleftarrow{\mathbf{x}}_i), \sigma_T \right\}$ , with  $T = 4$ ,  $\sigma_T = 10$ , and  $c_t$  determines the time-specific association between predictors and outcome.

The longitudinal predictors,  $\overleftarrow{\mathbf{X}}_i = \{\mathbf{X}_{i1}, \dots, \mathbf{X}_{iT}\}$ , with  $\mathbf{X}_{it} = (x_{1,it}, \dots, x_{5,it})'$ , were generated as correlated uniform random variables. Specifically, for each individual  $i$  and predictor  $p = 1, \dots, 5$ , the sequence  $(x_{p,i1}, \dots, x_{p,iT})$  was generated to be correlated over time according to a multivariate normal distribution with mean zero and covariance matrix  $\Sigma$ . Here, the  $(t, k)$  element of  $\Sigma$  is given by  $d_{tk}^2$ . We specified  $d_{t,t-1} = 0.4$ ,  $d_{t,t-2} = 0.2$ , and  $d_{t,t-3} = 0.1$ , reflecting stronger correlations for measurements closer in time. The resulting normal variables were then transformed to uniform margins using the standard normal cumulative distribution function, such that  $x_{p,it} \sim \text{Uniform}(0, 1)$ . Correlations were introduced only across time within each predictor, and all between-

predictor correlations ( $x_{p,it}$  and  $x_{p',it}$  for  $p \neq p'$ ) were set to zero.

We considered missingness in the form of dropout and assumed that approximately 15% of participants dropped out at each follow-up time point  $t = 2, \dots, T$ . Dropout at time  $j$  was generated as  $\pi_j(\overleftarrow{\mathbf{x}}_i) = \text{expit} \left[ h_j \sum_{t=1}^{j-1} (x_{1,it} + x_{2,it} + x_{3,it} + x_{4,it} + x_{5,it}) \right]$ , where  $\text{expit}(x) = 1/(1 + \exp(-x))$ ,  $R_{ij} \sim \text{Binomial}\{\pi_j(\overleftarrow{\mathbf{x}}_i)\}$ , and  $\overleftarrow{h} = (0.75, 0.4, 0.25)$ .

We simulated 2000 datasets of size  $n = (250, 1000, 2000)$ , and three different scenarios by varying  $\overleftarrow{c} = (c_T, \dots, c_1)$ . *Scenario A*:  $\overleftarrow{c} = (1, 0.75, 0.5, 0.25)$ , reflecting a decreasing association over time between the predictors and the outcome; *Scenario B*:  $\overleftarrow{c} = (1, 0, 0, 0)$ , where only predictors at the closest time point are associated with the outcome; *Scenario C*:  $\overleftarrow{c} = (0.25, 0.25, 0.25, 0.25)$ , where a stable association with the outcome for all time points. For all scenarios we estimated  $E_{(H_T, g_0)}\{E[Y_T(g_0) | H_T, g_0]\}$  using equation (1).

We compared the performance of our proposed approach (LDART) with two *hard* BART models: the standard BART algorithm (HBART) and BART with the sparsity-inducing Dirichlet prior (HDART) (as implemented in the R package BART; Sparapani et al., 2021), as well as two *soft* BART models: the standard SoftBart algorithm (SDART; as implemented in the R package SoftBart; Linero, 2022) and SoftBart incorporating only the predictors from the last time point in the model (SDART-11). We specified 50 regression trees for each model. To evaluate their performance, we compared relative absolute bias, root-mean-squared error (RMSE), and 95% coverage probabilities (CPs). Unless stated otherwise, models used default priors, with 1200 warm-up and 2000 sampling iterations (thinning = 4) across 48 chains. Simulations were run in R 4.1 at the *High Performance Computing Center North*.

#### 5.1 Simulation results

The simulation results are presented in detail in Table 1. For all scenarios and models, the absolute relative biases and RMSEs are smaller for the larger sample sizes, as expected. Moreover, the coverage probabilities are close to the nominal values for most scenarios, with CPs (%): 90.2–97.5. The *soft* BART models generally exhibited the smallest biases and RMSEs across the three scenarios. The best-performing model varied by scenario: in Scenario A, which reflects a decreasing association over time, LDART performed best; in Scenario B, where only predictors at the closest time point are associated with the outcome, SBART-11 was the winner; and in Scenario C, characterized by a weak and stable association across all time points, SBART performed best. These results highlight the influence of the underlying data-generating mechanisms. Although LDART was not the winner in Scenarios B and C, it was highly competitive, and that SBART-11 achieved markedly smaller CPs in Scenarios A and C. The two *hard* BART models produces biases and RMSEs that are notably larger than the other models. In addition, they produce overly conservative credible intervals in many scenarios, with CPs close to 97.0%.

**Table 1** Simulation results for estimating  $E_{(H_T, g_0)}\{E[Y_T(g_0) | (H_T, g_0)]\}$  using (1) and the simulated data as described in Section 5 with 50 regression trees. For scenario A:  $c = (1, 0.75, 0.5, 0.25)$ , B:  $c = (1, 0, 0, 0)$ , and C:  $c = (0.25, 0.25, 0.25, 0.25)$ . The results are presented in terms of absolute relative bias  $\times 100$ , root-mean-squared-error (RMSE) and 95% coverage probabilities as percentages [cp (%)]. The best-performing model for each scenario is highlighted.

	A			B			C		
	Bias	RMSE	CP (%)	Bias	RMSE	CP (%)	Bias	RMSE	CP (%)
<i>n</i> = 2000									
LBART	<b>0.86</b>	<b>0.39</b>	96.1	1.85	0.34	96.0	1.73	0.32	<b>95.4</b>
SBART	0.89	0.41	<b>95.4</b>	1.88	0.34	<b>95.0</b>	<b>1.70</b>	<b>0.31</b>	95.8
SBART-l1	0.98	0.44	92.1	<b>1.81</b>	<b>0.33</b>	94.7	1.85	0.34	92.8
HDART	0.99	0.46	95.8	1.99	0.37	96.1	2.02	0.37	96.4
HBART	1.00	0.45	96.4	2.12	0.39	97.0	1.94	0.35	97.5
<i>n</i> = 1000									
LBART	<b>1.20</b>	<b>0.55</b>	96.1	2.61	0.48	95.5	2.31	<b>0.42</b>	96.0
SBART	1.24	0.56	<b>95.2</b>	2.58	0.47	95.2	<b>2.28</b>	0.42	<b>95.6</b>
SBART-l1	1.22	0.56	94.0	<b>2.55</b>	<b>0.46</b>	<b>95.0</b>	2.47	0.45	92.6
HDART	1.32	0.60	96.6	2.79	0.51	96.4	2.66	0.49	96.2
HBART	1.33	0.61	96.0	2.87	0.53	96.5	2.71	0.49	97.3
<i>n</i> = 250									
LBART	<b>2.25</b>	<b>1.03</b>	<b>95.0</b>	4.97	0.91	<b>94.7</b>	4.67	0.85	92.6
SBART	2.33	1.06	95.0	4.84	0.88	94.5	<b>4.47</b>	<b>0.82</b>	<b>93.1</b>
SBART-l1	2.33	1.07	93.8	<b>4.70</b>	<b>0.86</b>	96.0	4.65	0.85	90.2
HDART	2.38	1.09	95.4	5.21	0.94	96.3	4.85	0.89	96.4
HBART	2.37	1.08	95.3	5.27	0.96	96.3	4.87	0.89	96.6

## 6 ANALYSIS OF THE BETULA DATA

In the Betula study, we analyzed the long-term population-level effects of a 15-year intervention reducing sBP in individuals with hypertension, using the incremental threshold approach described in Sections 2–3. The analysis includes 800 participants from eight age cohorts (35–70 years at baseline) followed across four waves at five-year intervals.

Episodic memory was assessed using a composite score ranging from 0 to 76 (Josefsson et al., 2012). Baseline confounders were age, sex, and education, while time-varying covariates included body mass index, smoking, serum cholesterol, self-reported type 2 diabetes, and prior memory performance. Over the 15-year follow-up, 38% of participants dropped out, including 18% who either died or developed dementia. Dementia and death were treated jointly as a single terminating event, and results therefore reflect intervention effects among individuals who remained alive and cognitively healthy. See Table A1 for baseline summary statistics.

We consider an intervention where we assign sBP using the *incremental threshold intervention* described in Sections 2 and 3. The intervention loosely imply that, at each follow-up time point, sBP levels is shifted downwards by a prespecified value in individuals with hypertension, while no intervention is imposed on individuals without hypertension. See Fig. A1 for an illustration of the shift intervention using the Betula data. Specifically, we consider three cut-offs using values of 130 mmHg, 140 mmHg, and 150 mmHg for  $\tau$ . If a subject's NVT at times  $t = 2, \dots, T$  is greater than  $\tau$ , then sBP is assigned by imposing a incremental shift on the NVT, and otherwise we do not intervene.

### 6.1 Priors for the shift and sensitivity parameters

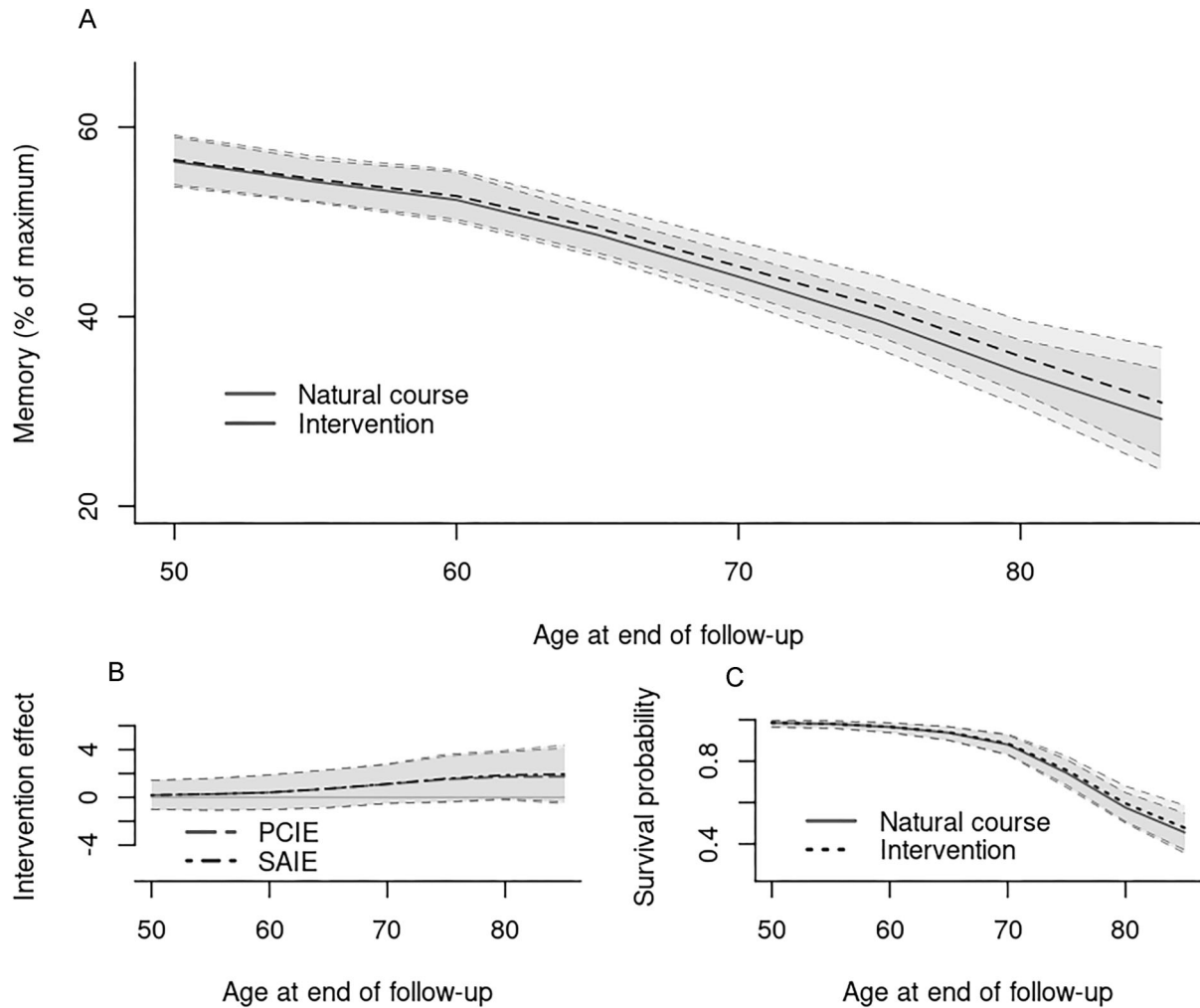
Priors for the shift and sensitivity parameters are specified by the analyst and should reflect plausible values. Here, we spec-

ify a triangular prior for the shift parameter  $\delta_t$ , such that  $\delta_t \sim \text{Triangular}(L_\tau, 0, L_\tau)$ . The lower bound and mode are given by  $L_\tau \approx -44$  mm Hg, corresponding to two standard deviations of baseline sBP, and the upper bound is 0. The intervention implies a negative shift from the NVT with between 0 and 44 mm Hg for individuals with a  $NVT > 140$  at each follow-up time point.

To estimate the SAIE, we also specify priors on the sensitivity parameters in Assumptions C5 and C6. We specify triangular priors for the sensitivity parameters  $\Delta$  and  $\lambda$ . For  $\Delta$ , we assume  $\Delta \sim \text{Triangular}(0, U_\Delta, U_\Delta)$ , with upper bound (and mode) set to 4.6%; 4.6% corresponds to the estimated difference in baseline memory between those who are alive at time point  $T$  and those who are dead, adjusting for age and sex. It is unlikely the difference between the strata would be bigger than this. We further assume  $\lambda \sim \text{Triangular}(0.5, 1, 1)$ , where the upper bound corresponds to the deterministic monotonicity assumption while the lower bound reflects a moderate correlation between survival under the two regimes.

### 6.2 Results

We estimated the PCIE and the SAIE using LDART with priors on the sensitivity parameters as described above. For each of the 16 chains, we used a thinning of 100, discarded the first 300 iterations as burn-in, and kept 2400 posterior samples for inference. Note, with a thinning of 100, we ensured precision in terms of close to independent sampling for all age-cohorts. However, a much smaller number (10–20) would have been sufficient in most cases. We sampled pseudo data of size  $N^* = 10,000$  at each iteration for the MC integration. Convergence of the posterior samples was monitored using trace plots. Note, the memory outcome, a test score, is presented in %.



**Figure 1** Results from the main analysis of Betula data using 130 mmHg as cut-off and default assumptions and sensitivity parameters as described in Section 6. Panel (A) Age-specific posterior mean memory at end of follow-up under the natural course (plain line) and under the hypothetical intervention (dotted line). Panel (B) Results for the PCIE and the SAIE. Panel (C) Posterior mean survival probabilities at end of follow-up under the natural course and the hypothetical intervention. All plots include 95% pointwise credible intervals.

Results for the different age-cohorts using 130 mmHg as cut-off are shown in Figure 1. We observe a consistent attenuation in effects as the threshold increases from 130 mmHg to 150 mmHg (see Supplementary Figure A2). In summary, posterior mean memory score at end of follow-up was attenuated with advancing age. This occurred both under the natural course and under the intervention, although to a lesser extent for the latter. Among subjects 70 and older, the shift in memory from the intervention corresponds to postponing memory decline approximately 2 years on average; however this effect is not significant. The results for the PCIE suggest that there is no long-term cognitive effects of monitoring sBP in this population (credible intervals cover 0 for all age-cohorts). We note however, that the effect is notably stronger for the older cohorts compared to middle-aged cohorts. The PCIE and SAIE estimates were similar. For the SAIE, this suggests no long-term effects within the always-survivor stratum, comprising individuals who would survive and remain free of dementia under both the natural course and the intervention. Hence, the non-significant results for the PCIE are

not driven by differential mortality rates. We also note that only slightly higher survival probabilities were obtained under the intervention across age cohorts compared to the natural course, leading to rather small differences between the PCIE and the SAIE.

We performed sensitivity analyses to assess robustness to violations of assumptions C2–C6 (Supplementary Appendix E; Table A2). For clarity, analyses were stratified into middle-aged ( $age_{IT} < 65$ ) and older ( $age_{IT} \geq 65$ ) groups. Results were largely insensitive to alternative prior specifications for  $\lambda$  and  $\Delta$  (data not shown) and remained consistent when accounting for an unmeasured time-varying confounder. Excluding 61 subjects with limited propensity overlap slightly attenuated effects but did not alter conclusions. Assuming missing-not-at-random (MNARS) outcomes reduced posterior mean memory scores under both regimes, yet PCIE estimates remained stable. Results were similar when fitting alternative soft BART variants (S-DART, S-DART-11) instead of LDART.

## 7 CONCLUSIONS

Using data from a prospective cohort study on aging, cognition, and dementia, we investigated population effects on episodic memory of an hypothetical incremental threshold intervention over 15 years that reduced sBP in individuals with hypertension. In our study, there were no significant prognostic or etiological effects, although effects were markedly stronger in older individuals. Nevertheless, these findings provide important insights into long-term blood pressure control and cognitive aging, suggesting that sBP interventions are unlikely to have a strong effect on memory, whether at the overall population level or among individuals who would remain alive and dementia free under both the natural course and the intervention. Importantly, the methodology applied in this study has broad applicability to interventional research in contexts where mortality is a major concern, including conditions as diverse as cancer, COVID-19, and congestive heart failure.

We propose a Bayesian semi-parametric approach for the proposed incremental threshold intervention, extending prior research in several ways. First, our approach introduces an incremental intervention for prospective cohort data with attrition, offering a viable alternative when clinical trials are not available or feasible to conduct. Second, we offer two types of mortal cohort inferences, providing both prognostic and etiological effect estimates, both of which are critical for public health practice, especially in older populations. Last, our modeling approach employs *soft* BART models with a novel sparsity-inducing prior for longitudinal data, promoting parsimony in longitudinal prediction settings and demonstrating good performance in simulations.

Concerns have been raised about principal stratification methods relying on assumptions that are often hard to justify in practice (eg, Feller et al., 2017). Monotonicity is one such assumption. Our approach relies on the less strong *stochastic monotonicity* assumption and uses sensitivity parameters that allow researcher to investigate sensitivity of results to this unverifiable assumption. However, a limitation with proposed principal stratification approach is that we only included age as a survival-predictive covariate for partial identification of the SAIE. In future work we will explore the possibility of incorporating more baseline covariates (see for example Roy et al., 2008; Tchetgen Tchetgen, 2014; Wang et al., 2017).

## ACKNOWLEDGMENTS

We thank an anonymous referee for helpful comments, which improved the paper.

## SUPPLEMENTARY MATERIALS

Supplementary material is available at *Biometrics* online.

Web Appendices, Supplementary results, and R code referenced in Sections 3, 4, and 6, are available with this paper at the *Biometrics* website on Oxford Academic. **Web Appendices A–E** provide details on assumptions, the extended G-formula, SAIE identification, computation (Dirichlet prior, G-computation), and supplementary Betula results. The R package *Gcomp*–

BART is available on <https://github.com/m4ryjo/GcompBART> (Josefsson, 2025).

## FUNDING

This work was supported by Forte (Dnr 2019-01064 to M.J.), NIH (RO1 HL166324 and HL158963 to M.J.D.), and used resources from High Performance Computing Center North. The Betula Project is funded by the Knut and Alice Wallenberg Foundation and the Swedish Research Council (K2010-61X-21446-01).

## CONFLICTS OF INTEREST

None declared.

## DATA AVAILABILITY

Betula data are available upon request and approval by the project's Steering Group. See <https://www.umu.se/en/research/projects/betula---aging-memory-and-dementia/>.

## REFERENCES

- Bundy, J. D., Li, C., Stuchlik, P., Bu, X., Kelly, T. N., Mills, K. T. et al. (2017). Systolic blood pressure reduction and risk of cardiovascular disease and mortality: a systematic review and network meta-analysis. *JAMA Cardiology*, 2, 775–781.
- Du, J. and Linero, A. (2019). Incorporating grouping information into Bayesian decision tree ensembles. In *Proceedings of the 36th International Conference on Machine Learning*, pp. 1686–1695, Long Beach, California, USA: PMLR.
- Dufouil, C., Brayne, C. and Clayton, D. (2004). Analysis of longitudinal studies with death and drop-out: a case study. *Statistics in Medicine*, 23, 2215–2226.
- Feller, A., Mealli, F. and Miratrix, L. (2017). Principal score methods: Assumptions, extensions, and practical considerations. *Journal of Educational and Behavioral Statistics*, 42, 726–758.
- Frangakis, C. E. and Rubin, D. B. (2002). Principal stratification in causal inference. *Biometrics*, 58, 21–29.
- Frangakis, C. E., Rubin, D. B., An, M.-W. and MacKenzie, E. (2007). Principal stratification designs to estimate input data missing due to death. *Biometrics*, 63, 641–649.
- Friedman, J. H. (1991). Multivariate adaptive regression splines. *The Annals of Statistics*, 19, 1–67.
- Haneuse, S. and Rotnitzky, A. (2013). Estimation of the effect of interventions that modify the received treatment. *Statistics in Medicine*, 32, 5260–5277.
- Hayden, D., Pauler, D. K. and Schoenfeld, D. (2005). An estimator for treatment comparisons among survivors in randomized trials. *Biometrics*, 61, 305–310.
- Iadecola, C., Yaffe, K., Biller, J., Bratzke, L. C., Faraci, F. M., Gorelick, P. B. et al. (2016). Impact of hypertension on cognitive function: a scientific statement from the American Heart Association. *Hypertension*, 68, e67–e94.
- Josefsson, M. (2025). *GcompBART: G-computation using Soft Bayesian Additive Regression Trees*, R package version 0.1.0, available from <https://github.com/m4ryjo/GcompBART>. [Accessed 10 December 2025].
- Josefsson, M. and Daniels, M. J. (2021). Bayesian semi-parametric G-computation for causal inference in a cohort study with MNAR dropout and death. *Journal of the Royal Statistical Society, Series C*, 70, 398.

- Josefsson, M., Daniels, M. J. and Pudas, S. (2023). A Bayesian semiparametric approach for inference on the population partly conditional mean from longitudinal data with dropout. *Biostatistics*, 24, 372–387.
- Josefsson, M., de Luna, X., Pudas, S., Nilsson, L.-G. and Nyberg, L. (2012). Genetic and lifestyle predictors of 15-year longitudinal change in episodic memory. *Journal of the American Geriatrics Society*, 60, 2308–2312.
- Kennedy, E. H. (2019). Nonparametric causal effects based on incremental propensity score interventions. *Journal of the American Statistical Association*, 114, 645–656.
- Kim, K., Kennedy, E. H. and Naimi, A. I. (2021). Incremental intervention effects in studies with dropout and many timepoints. *Journal of Causal Inference*, 9, 302–344.
- Kurland, B. F. and Heagerty, P. J. (2005). Directly parameterized regression conditioning on being alive: analysis of longitudinal data truncated by deaths. *Biostatistics*, 6, 241–258.
- Lee, K., Daniels, M. J. and Sargent, D. J. (2010). Causal effects of treatments for informative missing data due to progression/death. *Journal of the American Statistical Association*, 105, 912–929.
- Linero, A. R. (2018). Bayesian regression trees for high-dimensional prediction and variable selection. *Journal of the American Statistical Association*, 113, 626–636.
- Linero, A. R. (2022). Softbart: Soft Bayesian additive regression trees. R package version 1.0.2.
- Linero, A. R. and Yang, Y. (2018). Bayesian regression tree ensembles that adapt to smoothness and sparsity. *Journal of the Royal Statistical Society, Series B*, 80, 1087–1110.
- Livingston, G., Huntley, J., Sommerlad, A., Ames, D., Ballard, C., Banerjee, S. et al. (2020). Dementia prevention, intervention, and care: 2020 report of the Lancet Commission. *Lancet*, 396, 413–446.
- Muñoz, I. D. and Van Der Laan, M. (2012). Population intervention causal effects based on stochastic interventions. *Biometrics*, 68, 541–549.
- Nyberg, L., Boraxbekk, C.-J., Sörman, D. E., Hansson, P., Herlitz, A., Kauppi, K. et al. (2020). Biological and environmental predictors of heterogeneity in neurocognitive ageing: Evidence from Betula and other longitudinal studies. *Ageing Research Reviews*, 64, 101184.
- Picciotto, S., Hernán, M. A., Page, J. H., Young, J. G. and Robins, J. M. (2012). Structural nested cumulative failure time models to estimate the effects of interventions. *Journal of the American Statistical Association*, 107, 886–900.
- Richardson, T. S. and Robins, J. M. (2013). Single world intervention graphs (swigs): A unification of the counterfactual and graphical approaches to causality. *Center for the Statistics and the Social Sciences, University of Washington Series. Working Paper*, 128, 2013.
- Robins, J. (1986). A new approach to causal inference in mortality studies with a sustained exposure period—application to control of the healthy worker survivor effect. *Math Model*, 7, 1393–1512.
- Robins, J. M., Hernán, M. A. and Siebert, U. (2004). Effects of multiple interventions. *Comparative Quantification of Health Risks*, 1, 2191–2230.
- Robins, J. M., Rotnitzky, A. and Zhao, L. P. (1995). Analysis of semiparametric regression models for repeated outcomes in the presence of missing data. *Journal of the American Statistical Association*, 90, 106–121.
- Roy, J., Hogan, J. W. and Marcus, B. H. (2008). Principal stratification with predictors of compliance for randomized trials with 2 active treatments. *Biostatistics*, 9, 277–289.
- Shardell, M. and Ferrucci, L. (2018). Joint mixed-effects models for causal inference with longitudinal data. *Statistics in Medicine*, 37, 829–846.
- Sparapani, R., Spanbauer, C. and McCulloch, R. (2021). Nonparametric machine learning and efficient computation with Bayesian additive regression trees: The BART R package. *Journal of Statistical Software*, 97, 1–66.
- Tan, Y. V., Flannagan, C. A., Pool, L. R. and Elliott, M. R. (2021). Accounting for selection bias due to death in estimating the effect of wealth shock on cognition for the health and retirement study. *Statistics in Medicine*, 40, 2613–2625.
- Taubman, S. L., Robins, J. M., Mittleman, M. A. and Hernán, M. A. (2009). Intervening on risk factors for coronary heart disease: an application of the parametric G-formula. *International Journal of Epidemiology*, 38, 1599–1611.
- Tchetgen Tchetgen, E. J. (2014). Identification and estimation of survivor average causal effects. *Statistics in Medicine*, 33, 3601–3628.
- Wang, L., Zhou, X.-H. and Richardson, T. S. (2017). Identification and estimation of causal effects with outcomes truncated by death. *Biometrika*, 104, 597–612.
- Wen, L. and Seaman, S. R. (2018). Semi-parametric methods of handling missing data in mortal cohorts under non-ignorable missingness. *Biometrics*, 74, 1427–1437.
- Wen, L., Terrera, G. M. and Seaman, S. R. (2018). Methods for handling longitudinal outcome processes truncated by dropout and death. *Biostatistics*, 19, 407–425.
- Young, J. G., Hernán, M. A. and Robins, J. M. (2014). Identification, estimation and approximation of risk under interventions that depend on the natural value of treatment using observational data. *Epidemiol Methods*, 3, 1–19.
- Young, J. G., Logan, R. W., Robins, J. M. and Hernán, M. A. (2019). Inverse probability weighted estimation of risk under representative interventions in observational studies. *Journal of the American Statistical Association*, 114, 938–947.
- Zaidi, J. M., Tchetgen Tchetgen, E. J. and VanderWeele, T. J. (2019). Quantifying and detecting individual level ‘always survivor’ causal effects under ‘truncation by death’ and censoring through time. *arXiv preprint arXiv:1905.11300*.
- Zhang, J. L. and Rubin, D. B. (2003). Estimation of causal effects via principal stratification when some outcomes are truncated by “death”. *Journal of Educational and Behavioral Statistics*, 28, 353–368.
- Zhou, Z., Cai, L., Fu, J., Han, Y. and Fang, Y. (2020). Hypothetical interventions on risk factors for cognitive impairment among chinese older adults: An application of the parametric G-formula. *International Journal of Environmental Research and Public Health*, 17, 1021.