Nonparametric competing risks analysis using Bayesian Additive Regression Trees (BART)

Journal Title
XX(X):1–1

Reprints and permission:
sagepub.co.uk/journalsPermissions.nav
DOI: 10.1177/ToBeAssigned
www.sagepub.com/
SAGE

Rodney Sparapani[1], Brent R. Logan[1], Robert E. McCulloch[2], and Purushottam W. Laud[1]

[1]*Division of Biostatistics, Medical College of Wisconsin*
[2]*School of Mathematical and Statistical Sciences, Arizona State University*

Supplement

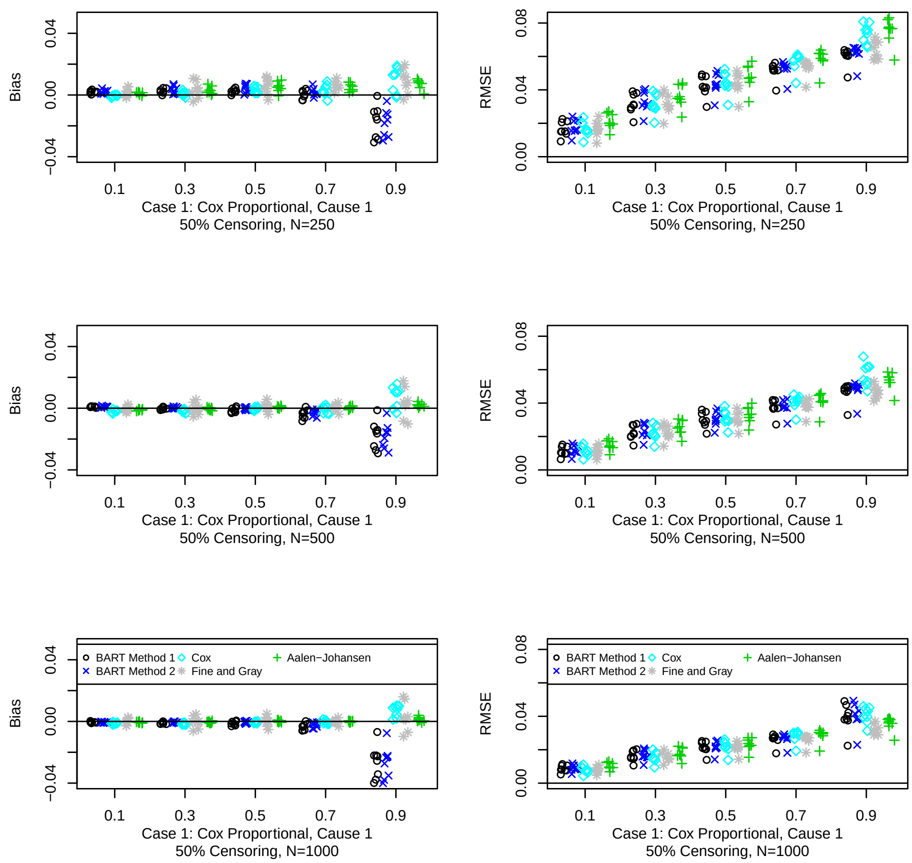

Figure 1. Bias (left) and RMSE (right) for case 1 with 50% censoring: $N = 250$ (first row), $N = 500$ (second row), and $N = 1000$ (third row). Each simulated data set was analyzed with both BART competing risks models, Cox proportional cause specific hazards models, Fine and Gray proportional subdistribution hazards model, and the Aalen-Johansen nonparametric estimator. For brevity, we only consider cause 1 which is generally the cause of interest. For each scenario, we examined the prediction performance in terms of bias and Root Mean Square Error (RMSE), at the following quantiles of the cdf: 10%, 30%, 50%, 70% and 90%. Results are plotted as points against quantile for each case and sample combination; note that there are 16 points (8 shown here and 8 in the article) for each case and sample combination: 2 groups as targets for prediction, $x = 0, 1$; 4 parameter configurations, $a = 1, 2, 3, 4$ (shown in Table 2 of the article); and 2 censoring rates, 20% (shown in the article) and 50% (shown here), $b = 0.2, 0.5$. The bias and RMSE metrics were assessed at the five chosen quantiles, Q, e.g., $\text{bias}_{Nxab} = H^{-1} \sum_h \left[\hat{F}_{1,abh}(t_Q, x) - F_{1,ab}(t_Q, x)\right]$ where t_Q is such that $Q = F_{1,ab}(t_Q, x) + F_{2,ab}(t_Q, x)$; N is the sample size; and $h = 1, \ldots, H$ are the simulated data sets.

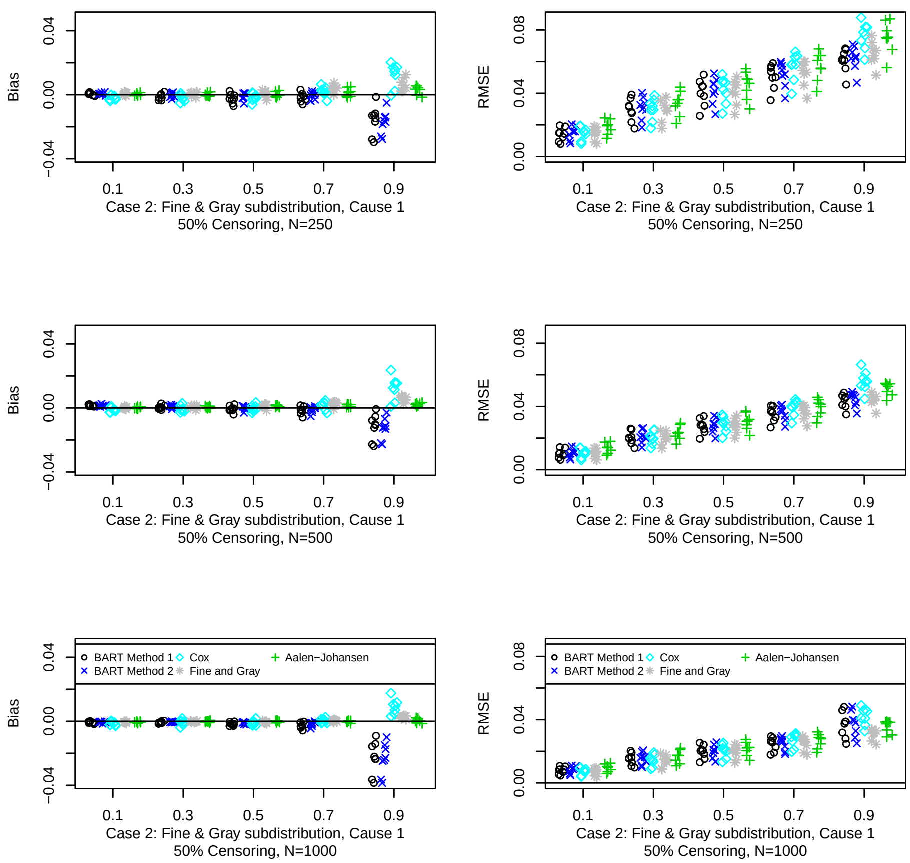

Figure 2. Bias (left) and RMSE (right) for case 2 with 50% censoring: $N = 250$ (first row), $N = 500$ (second row), and $N = 1000$ (third row). Each simulated data set was analyzed with both BART competing risks models, Cox proportional cause specific hazards models, Fine and Gray proportional subdistribution hazards model, and the Aalen-Johansen nonparametric estimator. For brevity, we only consider cause 1 which is generally the cause of interest. For each scenario, we examined the prediction performance in terms of bias and Root Mean Square Error (RMSE), at the following quantiles of the cdf: 10%, 30%, 50%, 70% and 90%. Results are plotted as points against quantile for each case and sample combination; note that there are 16 points (8 shown here and 8 in the article) for each case and sample combination: 2 groups as targets for prediction, $x = 0, 1$; 4 parameter configurations, $a = 1, 2, 3, 4$ (shown in Table 2 of the article); and 2 censoring rates, 20% (shown in the article) and 50% (shown here), $b = 0.2, 0.5$. The bias and RMSE metrics were assessed at the five chosen quantiles, Q, e.g., $\text{bias}_{Nxab} = H^{-1} \sum_h \left[\hat{F}_{1,abh}(t_Q, x) - F_{1,ab}(t_Q, x)\right]$ where t_Q is such that $Q = F_{1,ab}(t_Q, x) + F_{2,ab}(t_Q, x)$; N is the sample size; and $h = 1, \ldots, H$ are the simulated data sets.

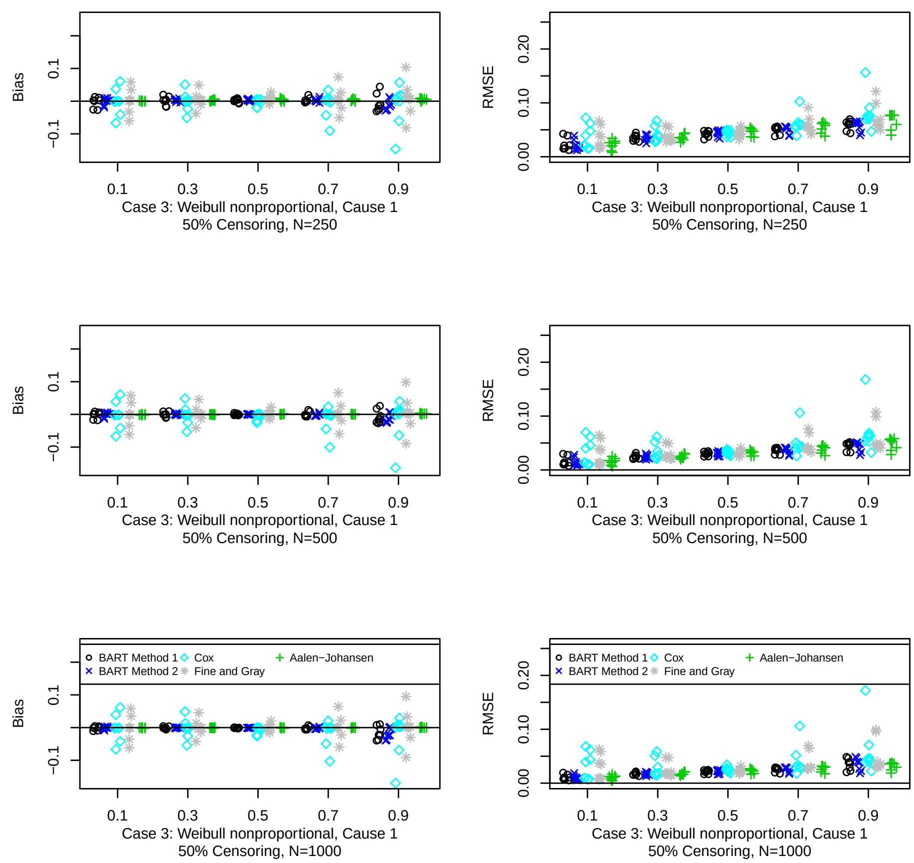

Figure 3. Bias (left) and RMSE (right) for case 3 with 50% censoring: $N = 250$ (first row), $N = 500$ (second row), and $N = 1000$ (third row). Each simulated data set was analyzed with both BART competing risks models, Cox proportional cause specific hazards models, Fine and Gray proportional subdistribution hazards model, and the Aalen-Johansen nonparametric estimator. For brevity, we only consider cause 1 which is generally the cause of interest. For each scenario, we examined the prediction performance in terms of bias and Root Mean Square Error (RMSE), at the following quantiles of the cdf: 10%, 30%, 50%, 70% and 90%. Results are plotted as points against quantile for each case and sample combination; note that there are 16 points (8 shown here and 8 in the article) for each case and sample combination: 2 groups as targets for prediction, $x = 0, 1$; 4 parameter configurations, $a = 1, 2, 3, 4$ (shown in Table 2 of the article); and 2 censoring rates, 20% (shown in the article) and 50% (shown here), $b = 0.2, 0.5$. The bias and RMSE metrics were assessed at the five chosen quantiles, Q, e.g., $\text{bias}_{Nxab} = H^{-1} \sum_h \left[\hat{F}_{1,abh}(t_Q, x) - F_{1,ab}(t_Q, x)\right]$ where t_Q is such that $Q = F_{1,ab}(t_Q, x) + F_{2,ab}(t_Q, x)$; N is the sample size; and $h = 1, \ldots, H$ are the simulated data sets.

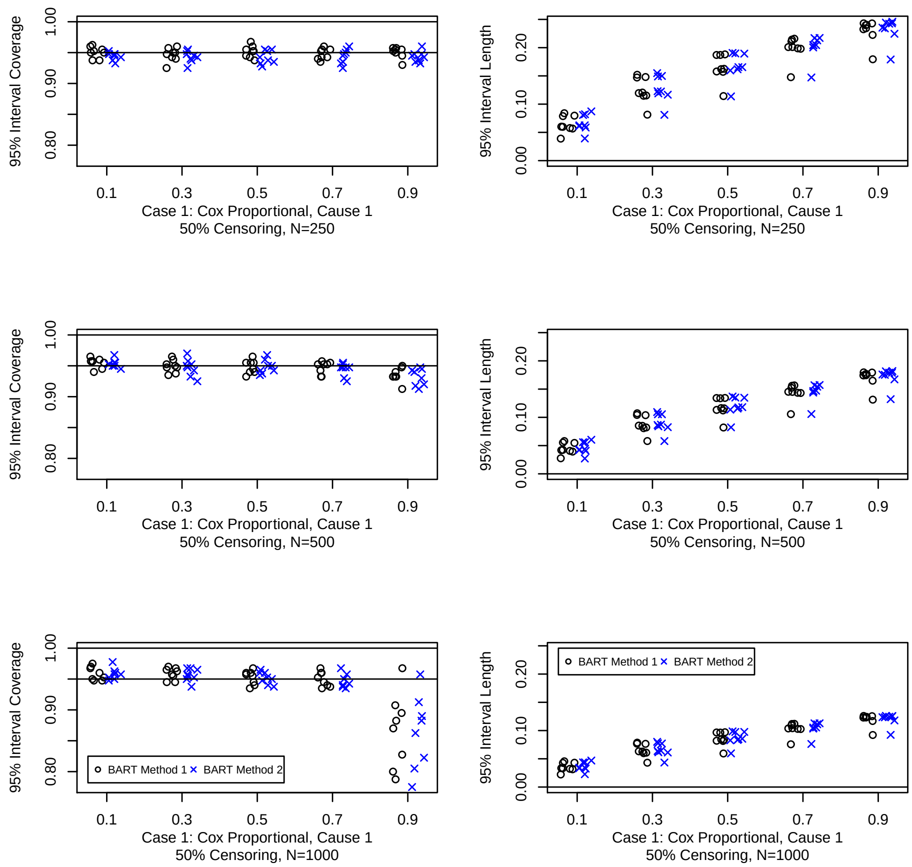

Figure 4. Coverage (left) and width (right) of 95% posterior intervals for case 1 with 50% censoring: $N = 250$ (first row), $N = 500$ (second row), and $N = 1000$ (third row). Each simulated data set was analyzed with both BART competing risks models. For brevity, we only consider cause 1 which is generally the cause of interest. For each scenario, we compare the 95% interval coverage probability and 95% interval length for the two BART methods. Results are plotted as points against quantile for each case and sample combination; note that there are 16 points (8 shown here and 8 in the article) for each case and sample combination: 2 groups as targets for prediction, $x = 0, 1$; 4 parameter configurations, $a = 1, 2, 3, 4$ (shown in Table 2 of the article); and 2 censoring rates, 20% (shown in the article) and 50% (shown here), $b = 0.2, 0.5$. The 95% interval coverage and length was assessed at the five chosen quantiles, e.g., coverage$_{Nxab} = H^{-1} \sum_h \mathrm{I}\left(\hat{F}_{1,abh,0.025}(t_Q, x) \le F_{1,ab}(t_Q, x) \le \hat{F}_{1,abh,0.975}(t_Q, x)\right)$ where t_Q is such that $Q = F_{1,ab}(t_Q, x) + F_{2,ab}(t_Q, x)$; N is the sample size; and $h = 1, \ldots, H$ are the simulated data sets.

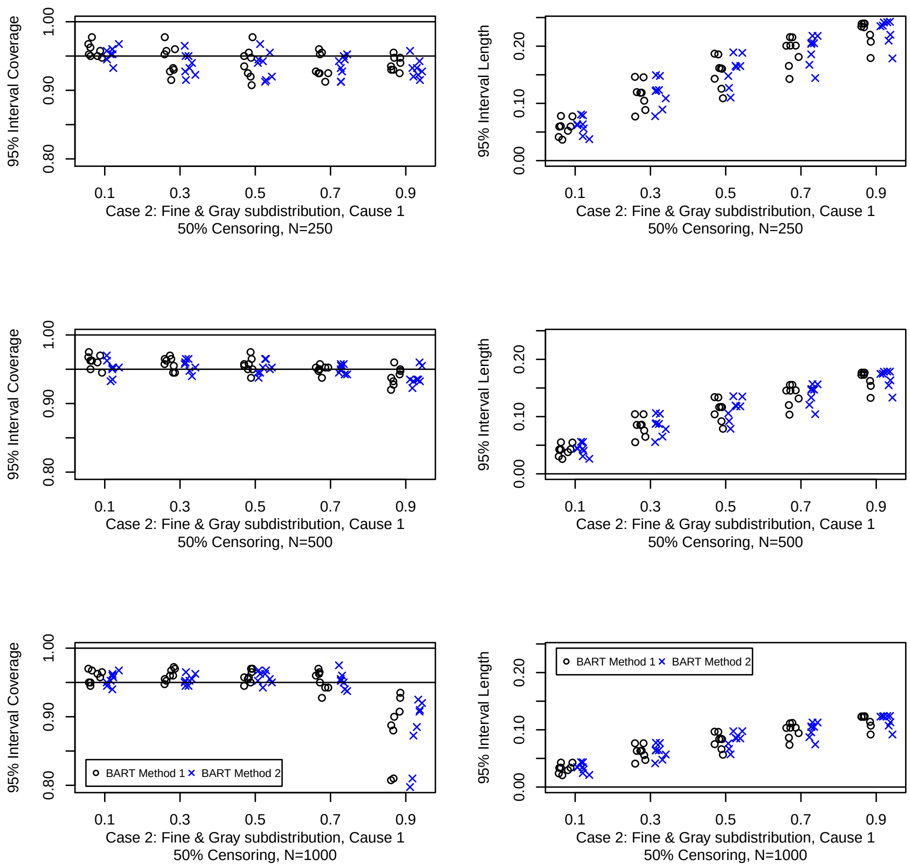

Figure 5. Coverage (left) and width (right) of 95% posterior intervals for case 2 with 50% censoring: $N = 250$ (first row), $N = 500$ (second row), and $N = 1000$ (third row). Each simulated data set was analyzed with both BART competing risks models. For brevity, we only consider cause 1 which is generally the cause of interest. For each scenario, we compare the 95% interval coverage probability and 95% interval length for the two BART methods. Results are plotted as points against quantile for each case and sample combination; note that there are 16 points (8 shown here and 8 in the article) for each case and sample combination: 2 groups as targets for prediction, $x = 0, 1$; 4 parameter configurations, $a = 1, 2, 3, 4$ (shown in Table 2 of the article); and 2 censoring rates, 20% (shown in the article) and 50% (shown here), $b = 0.2, 0.5$. The 95% interval coverage and length was assessed at the five chosen quantiles, e.g.,
$\text{coverage}_{Nxab} = H^{-1} \sum_h \mathrm{I}\Big(\hat{F}_{1,abh,0.025}(t_Q, x) \le F_{1,ab}(t_Q, x) \le \hat{F}_{1,abh,0.975}(t_Q, x)\Big)$
where t_Q is such that $Q = F_{1,ab}(t_Q, x) + F_{2,ab}(t_Q, x)$; N is the sample size; and $h = 1, \ldots, H$ are the simulated data sets.

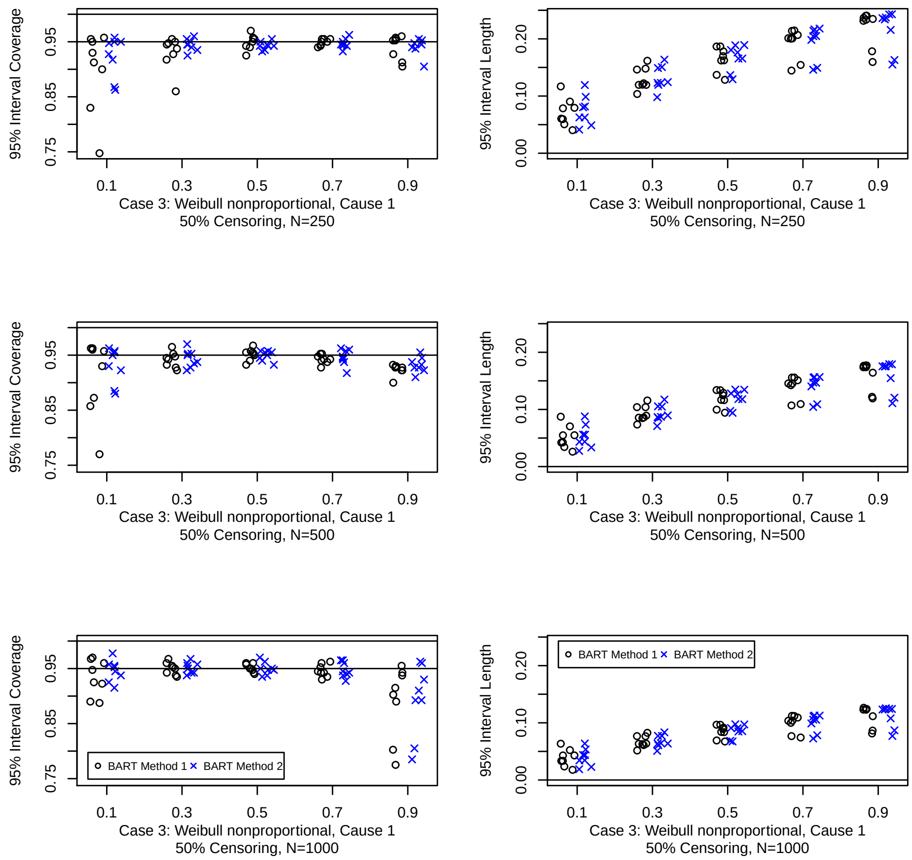

Figure 6. Coverage (left) and width (right) of 95% posterior intervals for case 3 with 50% censoring: $N = 250$ (first row), $N = 500$ (second row), and $N = 1000$ (third row). Each simulated data set was analyzed with both BART competing risks models. For brevity, we only consider cause 1 which is generally the cause of interest. For each scenario, we compare the 95% interval coverage probability and 95% interval length for the two BART methods. Results are plotted as points against quantile for each case and sample combination; note that there are 16 points (8 shown here and 8 in the article) for each case and sample combination: 2 groups as targets for prediction, $x = 0, 1$; 4 parameter configurations, $a = 1, 2, 3, 4$ (shown in Table 2 of the article); and 2 censoring rates, 20% (shown in the article) and 50% (shown here), $b = 0.2, 0.5$. The 95% interval coverage and length was assessed at the five chosen quantiles, e.g.,
coverage$_{Nxab} = H^{-1} \sum_h \mathrm{I}\Big(\hat{F}_{1,abh,0.025}(t_Q, x) \leq F_{1,ab}(t_Q, x) \leq \hat{F}_{1,abh,0.975}(t_Q, x)\Big)$
where t_Q is such that $Q = F_{1,ab}(t_Q, x) + F_{2,ab}(t_Q, x)$; N is the sample size; and $h = 1, \ldots, H$ are the simulated data sets.