Biometrics, 2025, 81(3), ujaf098
https://doi.org/10.1093/biomtc/ujaf098
Biometric Methodology

OXFORD

Causal machine learning for heterogeneous treatment effects in the presence of missing outcome data

Matthew Pryce [1,*], Karla Diaz-Ordaz [2], Ruth H. Keogh [1], Stijn Vansteelandt [3]

[1]Department of Medical Statistics, London School of Hygiene & Tropical Medicine, London WC1E 7HT, United Kingdom, [2]Department of Statistical Science, University College London, London WC1E7HB, United Kingdom, [3]Department of Applied Mathematics, Computer Science, and Statistics, Ghent University, Ghent, 9000, Belgium

[*]Corresponding author: Matthew Pryce, Department of Medical Statistics, London School of Hygiene & Tropical Medicine, London, WC1E 7HT, United Kingdom (matthew.pryce@lshtm.ac.uk).

ABSTRACT

When estimating heterogeneous treatment effects, missing outcome data can complicate treatment effect estimation, causing certain subgroups of the population to be poorly represented. In this work, we discuss this commonly overlooked problem and consider the impact that missing at random outcome data has on causal machine learning estimators for the conditional average treatment effect (CATE). We propose 2 de-biased machine learning estimators for the CATE, the mDR-learner, and mEP-learner, which address the issue of under-representation by integrating inverse probability of censoring weights into the DR-learner and EP-learner, respectively. We show that under reasonable conditions, these estimators are oracle efficient and illustrate their favorable performance through simulated data settings, comparing them to existing CATE estimators, including comparison to estimators that use common missing data techniques. We present an example of their application using the GBSG2 trial, exploring treatment effect heterogeneity when comparing hormonal therapies to non-hormonal therapies among breast cancer patients post surgery, and offer guidance on the decisions a practitioner must make when implementing these estimators.

KEYWORDS: causal machine learning; heterogeneous treatment effects; influence functions; missing outcome data.

1 INTRODUCTION

When evaluating the effect of an intervention, investigators are often interested in how the effect may vary within a target population. One approach used to explore treatment effect heterogeneity for a binary intervention is to estimate the conditional average treatment effect (CATE), defined as $\theta(x) = \mathbb{E}[Y(1)|X = x] - \mathbb{E}[Y(0)|X = x]$, where $Y(0)$ and $Y(1)$ are potential outcomes under the 2 levels of the treatment (Rubin, 2005) and X represents the individual (pre-treatment) characteristics in which heterogeneity is of interest.

CATEs can be used to explore treatment effect heterogeneity or to derive individualized treatment rules, aiding in the development of precision medicine (VanderWeele et al., 2019). Many estimators of the CATE have been proposed, with the focus turning toward non-parametric estimators, using machine learning (ML) to estimate complex functions of high dimensional data (Künzel et al., 2019; Nie and Wager, 2021; Kennedy, 2023; van der Laan et al., 2024). Of these estimators, each requires that the training data be fully observed and no data be missing. In this paper, we relax this requirement and propose 2 novel CATE estimators, the mDR-learner and mEP-learner, which demonstrate how causal ML estimators can be constructed when outcome data is missing at random (MAR).

MAR outcome data occur frequently in practice, typically arising when individuals are lost to follow-up. When it occurs, the observed data may no longer represent the target population, and subgroups that have high levels of drop-out can be under-represented. This presents a challenge for existing non-parametric CATE estimators, which do not address this under-representation and are prone to producing biased estimates of the CATE within these under-represented subgroups. To overcome this, some authors propose using CATE estimators in combination with established missing data techniques, such as imputing missing outcomes (Groenwold et al., 2014; Berrevoets et al., 2023), or re-weighting the population using inverse probability of censoring weights (IPCW) (Robins et al., 1994; Gonzalez Ginestet et al., 2021). However, when implementing these approaches using non-parametric, ML techniques, the inherent slow convergence of ML algorithms can introduce errors into the IPCW/imputation predictions, which then propagate through to the CATE estimates.

Our work aims to overcome this issue, with the mDR-learner and mEP-learner robustly incorporating IPCWs into the DR-learner (Kennedy, 2023) and EP-learner (van der Laan et al., 2024), respectively. We show that these estimators are oracle efficient under reasonable conditions and demonstrate their empirical performance through a simulation study (Section 4). We then illustrate their application, exploring treatment effect heterogeneity within the GBSG2 trial (Section 5), and finish by discussing potential extensions (Section 6).

Received: December 24, 2024; Revised: June 20, 2025; Accepted: July 10, 2025

2 BACKGROUND

2.1 Setting

We define a sample of n independent and identically distributed (i.i.d.) observations drawn from $O = (Z, A, CY)$, where A are binary treatments, $A \in \{0, 1\}$ and Y are continuous outcomes. In this setting, we allow outcome data to be missing and define C as an indicator of outcome missingness, $C \in \{0, 1\}$, with $C = 1$ indicating that an individual's outcome is non-missing. We then define Z to be a set of baseline covariates that contains all confounders between the treatment A and the outcome Y, and sufficient information for missing outcomes to be assumed MAR (ie, after controlling for Z and A, C, and Y are independent). Finally, we focus on learning the CATE, $\theta(x) = \mathbb{E}[Y(1)|X = x] - \mathbb{E}[Y(0)|X = x]$, which is conditional on $X \subseteq Z$, where X are the covariates for which heterogeneity is of interest.

2.2 Estimating the CATE using ML—no missing outcome data

We first consider how causal ML estimators are constructed when no outcome data is missing. Under the standard causal assumptions of (A1) consistency, (A2) no unmeasured confounding, $Y(a) \perp\!\!\!\perp A|Z$ for $a \in \{0, 1\}$, and (A3) positivity of treatment exposure, $0 < P(A = 1|Z) < 1$ (with probability 1) (Pearl et al., 2016), the CATE can be identified as:

$$\theta(x) = \mathbb{E}[\mathbb{E}[Y|A = 1, Z] - \mathbb{E}[Y|A = 0, Z]|X = x], \quad (1)$$

or

$$\theta(x) = \mathbb{E}\left[\frac{AY}{P[A = 1|Z]} - \frac{(1 - A)Y}{1 - P[A = 1|Z]}\middle| X = x\right]. \quad (2)$$

ML can be used to estimate the conditional expectations/probabilities (nuisance functions) found in Equations (1) or (2), with these estimators commonly referred to as plug-in estimators. One common example of a plug-in estimator is the T-learner (Künzel et al., 2019), which requires $X = Z$ and estimates the conditional expectations from Equation (1) in the subsets of individuals who are treated/untreated, $\mu^1(Z) = \mathbb{E}[Y|A = 1, Z]$ and $\mu^0(Z) = \mathbb{E}[Y|A = 0, Z]$, taking their difference to obtain CATE estimates. Alternatively, plug-in estimators based on Equation (2) are often referred to as inverse probability of treatment weight (IPTW) estimators, as they require the estimation of the propensity score, $\pi(Z) = P[A = 1|Z]$ (Kennedy, 2023).

While such ML-based plug-in estimators are simple to understand and implement, they are also prone to bias. This arises as ML algorithms use regularization to avoid over-fitting to the training data, reducing the rate at which these estimators converge toward the true parameter. This introduces non-negligible bias when fit using finite samples, with the errors in these nuisance function estimates (eg, $\hat{\mu}^1(Z)$, $\hat{\mu}^0(Z)$ or $\hat{\pi}(Z)$) directly propagating through to the estimates of the CATE, known as "plug-in bias" (Morzywolek et al., 2023). Furthermore, the T-learner does not ensure an optimal bias-variance trade-off is made for the CATE, as it optimizes predictions for their outcome functions rather than for the CATE itself (Kennedy, 2023). See Web Appendix 1 for an illustrative example.

Because the CATE is often much smoother than the nuisance functions in practice, it should be easier to estimate, provided the estimator is shielded from the slow convergence rates that affect the given nuisance-function estimators. Estimators that have this property are typically constructed using the efficient influence function (EIF) of the estimand of interest, where the EIF represents how sensitive a measure of prediction error is to changes in the data generating distribution. The EIF offers a useful tool for constructing estimators, allowing estimators to be approximately insensitive to small changes in their nuisance functions, meaning estimation errors in the outcome functions or propensity score affect errors in the estimates of the target function only through their product. Unfortunately, the EIF of the CATE is generally not well defined (ie, has infinite variance) whenever it depends on continuous variables (Hines et al., 2022). Instead, estimators that achieve these properties can be constructed based on the EIF of a well-chosen loss function, in this case, a measure of counterfactual prediction error (Foster and Syrgkanis, 2023; Morzywolek et al., 2023). Two examples include the DR-learner (Kennedy, 2023) and EP-learner (van der Laan et al., 2024).

The DR-learner is a model-agnostic CATE estimator, meaning the user can choose any estimation strategy, including data adaptive methods, when estimating any functions within it. It is constructed using a 2-step procedure (Kennedy, 2023), with the first step calculating pseudo-outcomes, Y_{DR}, using the EIF of the mean square error (MSE) for the CATE:

$$Y_{\text{DR}} = \frac{(A - \pi(Z))}{\pi(Z)(1 - \pi(Z))}\{Y - \mu^A(Z)\} + \mu^1(Z) - \mu^0(Z), \quad (3)$$

where $\hat{\mu}^A(Z) = A \cdot \hat{\mu}^1(Z) + (1 - A) \cdot \hat{\mu}^0(Z)$. The second step then learns the CATE by regressing the pseudo-outcomes on the covariate set in which heterogeneity is of interest, X. See Web Appendix 2 for the full algorithm. By using a pseudo-outcome regression, the DR-learner targets the CATE directly and benefits from faster convergence rates when the CATE is smoother than the baseline function. Additionally, these pseudo-outcomes are derived by considering the components of the MSE of the CATE that depend on $\theta(X)$, $\psi_{\text{CATE}} = \mathbb{E}\left[\theta(X)^2 - 2\theta(X)(\mu^1(Z) - \mu^0(Z))\right]$, using the EIF of this risk function,

$$\phi = \left(\theta^2(X) - \theta(X)\left(\mu^1(Z) - \mu^0(Z)\right)\right) + \frac{2\theta(X)(A - \pi(Z))}{\pi(Z)(1 - \pi(Z))}\{Y - \mu^A(Z)\} - \psi_{\text{CATE}}, \quad (4)$$

to define its pseudo-outcomes such that the sample average of the drift term (second term in Equation (4)) goes to 0. These pseudo-outcomes resemble the EIF from the average treatment effect and consist of the plug-in estimates and a weighted error term. By using these pseudo-outcomes, the DR-learner ensures that the gradient of the MSE risk function, with respect to the CATE, is less sensitive to errors in its nuisance functions. It also allows the DR-learner to achieve oracle efficiency, meaning that when the product of the convergence rates for the outcome predictions and propensity score estimates is faster than the rate of the oracle learner, it performs asymptotically as if the nuisance functions were known (Kennedy, 2023). However, the convergence rate of the DR-learner still depends on the convergence

rate of its pseudo-outcome regression, which depends on the complexity of the CATE.

Despite its desirable properties, the DR-learner can be sensitive to extreme propensity scores, with the IPTWs used in its pseudo-outcomes causing the pseudo-outcomes to grow infinitely large when propensity score estimates are near 0/1. To prevent this, some trim the propensity score estimates, therefore stopping the pseudo-outcomes from growing too large. However, this introduces bias to the propensity score estimates, which can propagate through to the CATE estimates themselves. For this reason, we discuss an alternative approach known as infinite-dimensional targeting (iTMLE) (Luedtke et al., 2017; Vansteelandt and Morzywołek, 2023; van der Laan et al., 2024). This technique has been developed for counterfactual outcome prediction (Vansteelandt and Morzywołek, 2023) and tends to have less sensitivity to extreme propensity scores, as its targeted learning framework can moderate the impact of outlying propensity score estimates. In this paper, we discuss a very similar approach known as the EP-learner (efficient plug-in learner) (van der Laan et al., 2024), which uses iTMLE in the context of causal contrasts, including CATE estimation.

The EP-learner is a model-agnostic estimator that also uses a 2 step procedure, first deriving pseudo-outcomes that are later regressed on X. However, the EP-learner uses an iTMLE procedure to generate its pseudo-outcomes. This iTMLE procedure plays a similar role to the one-step correction used in the DR-learner and is motivated by the EIF of the MSE risk function for the CATE, Equation (4). Using this EIF, the iTMLE procedure aims to update the initial plug-in outcome estimates, $\hat{\mu}^0(Z)$ and $\hat{\mu}^1(Z)$ using targeted learning such that the sample average of the drift term (second term) in Equation (4) goes to 0. Yet, since the drift term in Equation (4) contains an infinite dimensional $\theta(X)$ (when any variables in X are continuous), the classical targeted learning procedure, which regresses outcomes against a scalar (known as a clever covariate), would fail to set the sample average of the drift term to 0. To resolve this, the iTMLE procedure defines a vector of univariate basis functions, referred to as a sieve, $\varphi(X)$, which it uses within the targeting step, approximating $\theta(X)$ and allowing the updates to the plug-in estimates to vary by X.

The targeting step then works by regressing the outcomes Y on the sieve basis, $\varphi(X)$, in a weighted linear regression, with an offset $\mu^A(Z)$, and weight, $\dot{H}(A,Z) = \frac{A}{\hat{\pi}(Z)} + \frac{1-A}{(1-\hat{\pi}(Z))}$, suggested by the EIF of the risk function, Equation (4). After fitting this model, efficient plug-in estimates, $\hat{\mu}^{0*}(Z)$ and $\hat{\mu}^{1*}(Z)$ are obtained by adding/subtracting the estimated linear predictor from this model to the plug-in outcome estimates (see Web Appendix 3 for the full algorithm). As the updated outcome predictions are defined such that the sample average of the drift term in Equation (4) converges to 0, the EP-learner achieves the same oracle efficiency properties as the DR-learner. Additionally, by using iTMLE, extreme pseudo-outcomes are less common, and the CATE estimates should be more stable (van der Laan et al., 2024).

In addition, both learners require the data that is used in the second stage optimization is i.i.d. To achieve this, sample splitting is used when estimating the nuisance functions, preventing correlations from being introduced into the pseudo-outcomes. In this paper, we implement a K-fold cross-fitting procedure found in Web Appendix 2 and 3.

2.3 Estimating the CATE using ML—with missing outcome data

Now we consider how MAR outcome data impacts the existing estimators. To identify the CATE when outcome data is MAR, we require assumptions A1-A3 from Section 2.2 and 2 additional assumptions on the missingness mechanism: (A4) Outcomes are MAR, $Y \perp\!\!\!\perp C|A, Z$; (A5) Positivity of outcomes being non-missing, $0 < P(C = 1|A, Z)$, with probability 1. Under assumptions A1-A5, the CATE can be identified as Equations (5) and (6).

$$\theta(x) = \mathbb{E}[\mathbb{E}[Y|A = 1, C = 1, Z] - \mathbb{E}[Y|A = 0, C = 1, Z]|X = x], \quad (5)$$

or

$$\theta(x) = \mathbb{E}\left[\frac{\text{CAY}}{P[C = 1|A, Z]P[A = 1|Z]} - \frac{C(1-A)Y}{P[C = 1|A, Z](1 - P[A = 1|Z])}\middle| X = x\right]. \quad (6)$$

The no unmeasured confounding assumption (A2) and MAR assumption (A4) need not be conditional on the same covariates; however, for simplicity, we define both assumptions to be conditional on the same set of covariates, Z. Using Equations (5) and (6), we see that when all of the covariates in Z are present, the CATE can be estimated using the observed data. Consequently, the existing CATE estimators, such as the T-learner, DR-learner, and EP-learner can produce asymptotically unbiased estimates of the CATE by restricting their analyses to complete cases and adjusting for Z through their outcome models $\mu^a(Z) = \mathbb{E}[Y|A = a, C = 1, Z]$, $A \in \{0, 1\}$. We refer to this approach as an "available case analysis," and note that while the outcome regressions are limited to complete cases, the propensity score models should still be estimated using the full dataset.

Available case analyses offer the simplest way of estimating the CATE in the presence of MAR outcome data; however, their asymptotic properties do not assure that they perform well when fit using finite samples. Instead, available case analyses can often be inefficient, as their outcome regressions restrict the population to those with complete cases, with information being ignored for individuals who have a missing outcome. Equally, when the outcome models are fit using ML, they will be prone to over-smoothing in the subsets of the population that experience high levels of outcome missingness. This can cause complex non-linear CATEs to be missed, with these approaches over-smoothing their outcome predictions due to the missing data, or can lead to estimators identifying heterogeneity where non-exists, for instance, when the missingness only occurs within one treatment arm. Specific examples of data generating processes (DGPs) where this can occur are presented in Section 4.

Because of these limitations, some authors choose to address missing outcome data by utilizing missing data techniques. A common missing data technique is to impute outcomes, replacing the missing outcomes with outcome predictions gained from an imputation model, $\hat{\mathbb{E}}[Y|C = 1, A, Z]$ (Groenwold et al.,

2014; Berrevoets et al., 2023). These are easy to implement when using the existing non-parametric CATE estimators, as the estimators can be run using the imputed, complete dataset. However, as outcome imputations require the estimation of an additional nuisance model, plug-in bias can be introduced, with errors in the outcome imputations propagating through to the estimates of the CATE. Alternatively, other authors suggest addressing missing outcome data by re-weighting observed individuals based on their probability of having a non-missing outcome, $G(A, Z) = P[C = 1|A, Z]$ (Robins et al., 1994; Gonzalez Ginestet et al., 2021). These weights can be used to estimate the CATE using a similar approach to the one Kennedy (2023) uses to incorporate IPTWs, weighting observed individuals to create pseudo-outcomes, then regressing the pseudo-outcomes against X to estimate the CATE. However, this estimator, which we will refer to as the IPTW-IPCW estimator, also suffers from plug-in bias, as errors in the IPTWs and IPCWs will propagate through to the CATE estimates. We therefore construct IF based estimators that incorporate these weights naturally and which offer greater robustness to errors in the missingness predictions $G(A, Z)$.

3 DR-LEARNER/EP-LEARNER EXTENSIONS

In this section, we extend the DR-learner (Kennedy, 2023) and EP-learner (van der Laan et al., 2024) to handle missing outcome data, leading to 2 new estimators, the mDR-learner (missing outcome DR learner) and mEP-learner (missing outcome EP learner), respectively.

3.1 mDR-learner

We begin by considering an extension of the DR-learner to the MAR outcome data setting. Recall that the DR-learner does not account for the under-representation that occurs as a result of MAR outcomes. This is because the DR-learner's pseudo-outcomes are derived using a risk function that assumes complete data. However, when outcomes are MAR, the risk function takes a new form, $\psi_{\text{CATE}} = \mathbb{E}[\theta(X)^2 - 2\theta(X)(\mathbb{E}[Y|C = 1, A = 1, Z] - \mathbb{E}[Y|C = 1, A = 0, Z])]$, which now involves the indicator for outcomes being non-missing. It can be shown (Web Appendix 4) that the EIF is

$$\phi = \left(\theta^2(X) - 2\theta(X)\left(\mu^1(Z) - \mu^0(Z)\right)\right) - \frac{2\theta(X)(A - \pi(Z))C}{\pi(Z)(1 - \pi(Z))G(A, Z)}\{Y - \mu^A(Z)\} - \psi_{\text{CATE}}. \tag{7}$$

Using this EIF, we can construct pseudo-outcomes for the mDR-learner as:

$$Y_{\text{mDR}} = \frac{(A - \pi(Z))\,C}{\pi(Z)\,(1 - \pi(Z))\,G(A, Z)}\left\{Y - \mu^A(Z)\right\} + \mu^1(Z) - \mu^0(Z). \tag{8}$$

We note that IPCWs now appear alongside IPTWs, allowing these pseudo-outcomes to account for the shift in covariate distribution caused by both missing outcome data and confounding. The mDR-learner then proceeds by regressing the estimates of these pseudo-outcomes Y_{mDR} against covariates X to obtain estimates of the CATE (see Figure 1).

As the mDR-learner defines its pseudo-outcomes using the EIF of the MSE for the CATE (under outcome missingness), it not only minimizes the MSE risk function but also experiences less sensitivity to errors in its nuisance functions (including the IPCWs). We demonstrate this by exploring the excess risk of the mDR-learner, defined as the difference in MSE risk function, $L(.)$, when evaluated at $\hat{\theta}(X)$ and at $\theta^*(X)$, where $L(\theta^*) = \inf_\theta L(\theta)$. Using the approach laid out by Foster and Syrgkanis (2023), we provide an upper bound for the excess risk (see Web Appendix 5). This bound demonstrates how the gradient of the MSE risk function for the mDR-learner has reduced sensitivity to errors in the nuisance functions, which now include the missingness model, $G(A, Z)$. It also shows how the mDR-learner can obtain oracle efficiency under similar assumptions to those of the DR-learner.

For the mDR-learner to obtain oracle efficiency, it requires that 2 conditions hold; firstly, that the product of the convergence rates for the outcome predictions and propensity score estimates is faster than the rate for the oracle estimator, and secondly, that the product of the convergence rates for the outcome predictions and missingness estimates is also faster than the rate for the oracle estimator. For instance, consider an example where the oracle CATE estimator converges at a $\sqrt{n}$ rate. In this example, if we wish to estimate the nuisance functions within the mDR-learner without impacting the overall convergence rate of our CATE estimator, we require that the estimates of each of the nuisance functions converge at rates faster than $\sqrt[4]{n}$, hence allowing the 2 products of these rates to converge faster than $\sqrt{n}$. Equally, when the outcome predictions converge at rates slower than this, oracle efficiency can still be obtained as long as the propensity score estimates and missingness function estimates converge at rates that are sufficiently fast for the above conditions to hold. Under these weakened convergence requirements, ML can then be used when estimating nuisance functions within the mDR-learner without errors propagating through to the CATE.

3.2 mEP-learner

We now demonstrate how the EP-learner, a targeted-learning-based framework, can be extended for the setting with MAR outcome data. When outcomes are MAR, the EP-learner fails to account for under-representation introduced by missing outcomes. If we wish to account for this, we must recalculate the EIF which it uses to derive its pseudo-outcomes, with this EIF taking a new form, Equation (7). This EIF contains IPCWs within its drift term (second term), and hence, if we wish to set the sample average of the drift term to 0, removing plug-in bias, we must update the iTMLE process used within the EP-learner (see Figure 2).

To do so, we redefine the weight used in the iTMLE algorithm by considering Equation (7), defining the weight to be $\hat{H}(A, C, Z) = \frac{CA}{\hat{G}(A,Z)\hat{\pi}(Z)} + \frac{C(1-A)}{\hat{G}(A,Z)(1-\hat{\pi}(Z))}$. This weight ensures that the sample average of the second term in Equation (7) converges to 0, and hence the mEP-learner will also be oracle efficient when its nuisance function estimates converge sufficiently fast.

Algorithm 1 mDR-learner algorithm

1: Split the data randomly into K (e.g., 10) equal sized folds of n observations from $O = (Z, A, YC)$, denoted $D_1, \ldots, D_K$.

2: For $j \in 1, ..., K$ and using all folds $\{D_i, i = 1, \ldots, K, i \neq j\}$ except D_j, train models for

$$\pi(Z) = P[A = 1|Z], \text{(propensity score)} \quad (11)$$

$$G(A, Z) = P[C = 1|A, Z], \text{(missingness model)} \quad (12)$$

$$\mu^0(Z) = \mathbb{E}[Y|A = 0, C = 1, Z], \text{(conditional untreated outcome model)} \quad (13)$$

$$\mu^1(Z) = \mathbb{E}[Y|A = 1, C = 1, Z], \text{(conditional treated outcome model).} \quad (14)$$

3: For all individuals in D_j $(j \in 1, ..., K)$, obtain predictions of $\hat{\pi}$, $\hat{G}$, $\hat{\mu}^0$ and $\hat{\mu}^1$, based on the models fitted in the remaining folds.

4: Construct the pseudo outcomes for each individual in the data using

$$Y_{mDR} = \frac{(A - \hat{\pi}(Z))\, C}{\hat{\pi}(Z)\,(1 - \hat{\pi}(Z))\,\hat{G}(A, Z)} \left\{Y - \hat{\mu}^A(Z)\right\} + \hat{\mu}^1(Z) - \hat{\mu}^0(Z)$$

5: Regress the pseudo outcomes Y_{mDR} on covariates X, and obtain predictions of $\theta(X)$:

$$\hat{\theta}_{mDR}(X) = \hat{\mathbb{E}}[Y_{mDR}|X]$$

FIGURE 1 mDR-learner algorithm.

3.3 Implementation

As both the mDR-learner and mEP-learner are general frameworks for estimating the CATE with MAR outcome data, their implementation requires the user to make several key decisions. In this section, we break these decisions down into 2 groups; (a) decisions required for obtaining CATE estimates and (b) decisions required for assessing CATE performance.

3.3.1 Decisions required for obtaining CATE estimates

As both learners are model-agnostic, the user must first decide how to estimate the nuisance functions/pseudo-outcome model. Data adaptive techniques can be chosen; however, if oracle performance is to be achieved, the estimates from the nuisance models must converge sufficiently fast to the truth. For this reason, we illustrate their implementation using the Super Learner (Van der Laan et al., 2007), an ensemble learner that allows a range of data-adaptive algorithms to be implemented and which performs asymptotically as well as its best candidate learner.

After algorithm choice, users are required to decide which type of sample splitting they will implement within the mDR or mEP learners, ensuring the data used in the pseudo-outcome regression is i.i.d. Various options exist for achieving this, including K-fold cross-fitting (see Section 2.2) or independent sample splitting (Kennedy, 2023); however, estimators that allocate fully independent data for each nuisance/target model are typically less efficient, introducing finite sample bias by reducing sample sizes. For this reason, we demonstrate the mDR-learner and mEP-learner using a 10-fold cross-fitting process (see Figure 1), but note that the appropriate number of folds will depend on the complexity and smoothness of the underlying nuisance/target functions. If users wish to explore variations of cross-fitting approaches, sensitivity analyses could be run using an alternative number of splits.

We also highlight that when cross-fitting is used, reduced training sample sizes for each model can introduce positivity violations, leaving certain subgroups poorly represented within some folds. To overcome this, the mDR-learner and mEP-learner can be run multiple times (J times), using a different seed for the cross-fitting split. This results in a vector of CATE estimates for each individual, $\hat{\theta}_s(X)$, $s = 1, .., J$. Final CATE estimates are then obtained by taking the median across these estimates, $\widehat{\theta}(X) = \text{median}\{\widehat{\theta}_s(X)\}$ (Jacob, 2020).

Finally, when implementing the mEP-learner, the sieve basis used within the iTMLE process must be specified. The existing iTMLE implementations use a univariate trigonometric cosine polynomial basis, as it offers strong approximation guarantees under smoothness assumptions (Zhang and Simon, 2023). We also implement this sieve basis, following the guidance of Zhang and Simon (2023) to define the dimension of the sieve and its interaction order. However, alternative options exist, such

Algorithm 2 mEP-learner algorithm

1: Split the data randomly into K (e.g., 10) equal sized folds of n observations from $O = (Z, A, YC)$, denoted $D_1, \ldots, D_K$.

2: For $j \in 1, ..., K$ and using all folds $\{D_i; i = 1, \ldots, K, i \neq j\}$ except D_j, train models for

$$\pi(Z) = P[A = 1|Z], \text{(propensity score)} \quad (15)$$

$$G(A, Z) = P[C = 1|A, Z], \text{(missingness model)} \quad (16)$$

$$\mu^0(Z) = \mathbb{E}[Y|A = 0, C = 1, Z], \text{(conditional untreated outcome model)} \quad (17)$$

$$\mu^1(Z) = \mathbb{E}[Y|A = 1, C = 1, Z], \text{(conditional treated outcome model).} \quad (18)$$

3: For all individuals in D_j $(j \in 1, ..., K)$, obtain predictions of $\hat{\pi}$, $\hat{G}$, $\hat{\mu}^0$ and $\hat{\mu}^1$, based on the models fitted in the remaining folds.

4: For all individuals in the data, update the outcome predictions

a) Calculate the clever covariate, $\hat{H}(A, C, Z) = \frac{CA}{\hat{G}(A,Z)\hat{\pi}(Z)} + \frac{C(1-A)}{\hat{G}(A,Z)(1-\hat{\pi}(Z))}$.

b) Choose a sieve basis, $\hat{\varphi}(X)$

c) Run a linear regression of outcomes Y on feature vector $\hat{\varphi}(X)$ with offset $\hat{\mu}^A(Z)$ and weight $\hat{H}(A, C, Z)$ in the complete cases.

d) Estimate the coefficients $\hat{\epsilon}$ from c) and use these to update $\hat{\mu}^0(Z)$ and $\hat{\mu}^1(Z)$

$$\hat{\mu}^{1*}(Z) = \hat{\mu}^1(Z) + \hat{\epsilon} \cdot \hat{\varphi}(X), \quad \hat{\mu}^{0*}(Z) = \hat{\mu}^0(Z) - \hat{\epsilon} \cdot \hat{\varphi}(X)$$

5: Construct the pseudo outcome for all individuals in the data

$$Y_{mEP} = \hat{\mu}^{1*}(Z) - \hat{\mu}^{0*}(Z) \quad (19)$$

6: Regress the pseudo outcomes Y_{mEP} on covariates X, and obtain predictions of $\theta(X)$:

$$\hat{\theta}_{mEP}(X) = \hat{\mathbb{E}}[Y_{mEP}|X]$$

FIGURE 2 mEP-learner algorithm.

as a cross-validation option for choosing the sieves tuning parameters (van der Laan et al., 2024) and a penalized iTMLE implementation, which can improve performance for large sieves (Vansteelandt and Morzywołek, 2023).

3.3.2 *Decisions required for assessing CATE performance*

In addition to obtaining CATE estimates, users may also want to assess the accuracy of these estimates by obtaining measures of uncertainty or calculating evaluation metrics. Calculating measurements of uncertainty for CATE estimates generated using non-parametric estimators is challenging, as the theoretical convergence guarantees required for confidence intervals (CIs) to be derived are often not met. However, recent work by Bonvini et al. (2023); Takatsu and Westling (2025); and Ritzwoller and Syrgkanis (2024) offers potential solutions. In this work we explore how one of these techniques can be used with our estimators, focusing on the half-sample bootstrap approach (Ritzwoller and Syrgkanis, 2024), which can provide CIs when kernel based approaches are used to estimate the pseudo-outcome regression. Further details on this technique are outlined in Web Appendix 7, and we evaluate the performance of this approach within our simulation study in Web Appendix 12.

Additionally, we note that evaluating the performance of CATE estimators when using real world data is challenging, with

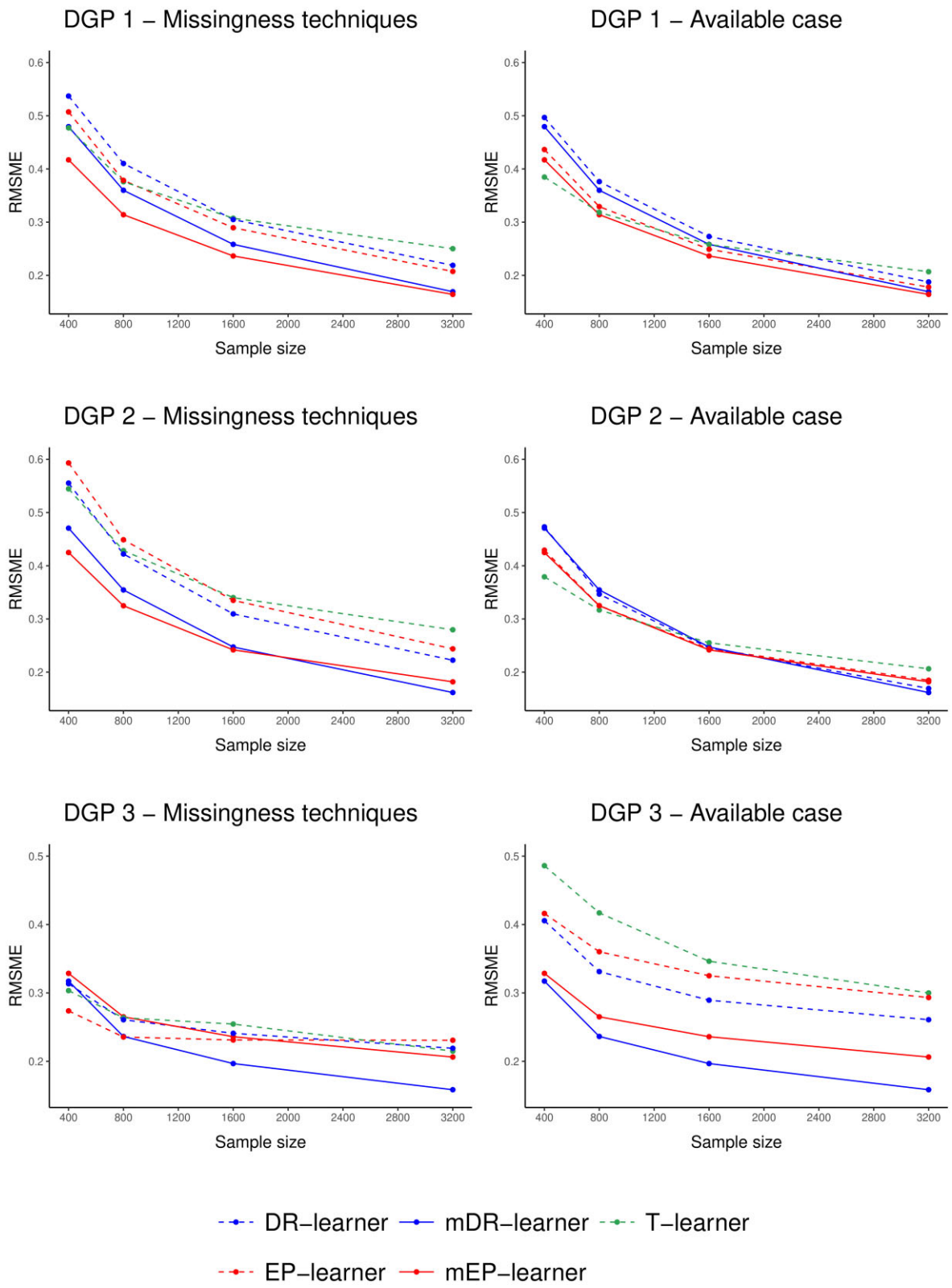

FIGURE 3 Root mean square median error (RMSME) for mDR-learner, mEP-learner, DR-learner, EP-learner, and T-learner in 3 data generating processes (DGPs) plotted by training sample size. Plots in the left column compare the mDR-learner and mEP-learner to the DR-learner, EP-learner, and T-learner when used in combination with an outcome imputation model in DGP 1, 2, and 3, respectively. Plots in the right column compare the mDR-learner and mEP-learner to the available case versions of the DR-learner, EP-learner, and T-learner in DGP 1, 2, and 3, respectively.

only one of the 2 potential outcomes, $Y(0)$ and $Y(1)$, ever observed for each individual. This means standard metrics such as the MSE cannot be calculated and used for estimator comparison and equally makes deriving EIF based estimators for these metrics tricky, as by definition, the MSE metric and its EIF will be 0 at the true CATE. Alternatively, model performance can be assessed by reviewing the stability of model estimates across different sample splitting seeds, enabling users to understand the variability in individual CATE estimates. Additionally, if choosing between the mDR/mEP learners, users should focus on the strengths/weaknesses of each learner, with the mDR-learner performing best when the mEP-learner's sieve poorly approximates the CATE, that is, when the CATE is non-smooth or sparse, while the mEP-learner may perform best when the mDR-learner's weights and pseudo-outcomes are highly variable/unstable.

4 SIMULATION STUDY

4.1 Set up

We study the finite sample performance of the mDR-learner and mEP-learner across 3 DGPs, where each DGP corresponds to a setting in which missing outcome data can complicate the estimation of the CATE. In each setting, we generate 6 uniformly distributed covariates Z, a binary treatment A, and a continuous outcome Y. In the first 2 DGPs, we define a simple unexposed outcome function $\mu^0(Z)$, a complex CATE $\theta(X)$ and define outcome missingness such that it occurs with high probability in only the treatment arm (DGP 1), or in both arms (DGP 2). This makes the complex CATE challenging to learn. In the third DGP, we define a complex unexposed outcome function, a simple CATE and define outcome missingness to occur with high probability in only the treatment arm, making the simple CATE difficult to learn (Web Appendix 8).

We vary the training data sample size from 400 to 3200 and use 500 replicates for each scenario. We compare the mDR-learner and mEP-learner to 4 alternative CATE estimators; the IPTW-IPCW learner, the DR-learner, EP-learner, and the T-learner, with the latter 3 implemented using (1) available cases (Section 2.3) and (2) imputed outcomes. All estimators were implemented using 10 fold cross-fitting, with the nuisance models fit using the *Super Learner*. Additionally, the pseudo-outcome models were fit using random forests (with 500 bootstrap half samples), enabling the generation of half-sample bootstrap CIs. To assess the performance of each estimator, we generated one test dataset with sample size $n = 10\,000$ per DGP and obtained the CATE estimates for each individual using each estimator. Performance was measured by calculating the root mean square *median* error (RMSME) of each learner (Web Appendix 10), as mean root mean square error (RMSE) estimates were found to be skewed when using 500 replications. Conditional CI coverage was calculated/reported in Web Appendix 12. For comparisons made using the mean RMSE, see Web Appendix 11.

4.2 Findings

When comparing the mDR-learner and mEP-learner with the DR-learner and EP-learner using available cases (Figure 3—right column), we see the mDR-learner and mEP-learner outperform the DR/EP learners, respectively, across all 3 DGPs. Equally, when comparing the mDR-learner and mEP-learner to the DR-learner and EP-learner fit using imputed outcomes (Figure 3—left column) both learners outperformed their corresponding imputed outcome version when the CATE was complex (DGP 1 and 2), while when the CATE was simple, the imputed outcome DR-learner and EP-learner performed well. We also note how the IPTW-IPCW learner and available case/imputed outcome T-learner were sensitive to nuisance function complexity, with their performance depending heavily on the complexity of the outcome functions and propensity score/censoring functions, respectively. IPTW-IPCW learner results are excluded from Figure 3 to aid interpretability (see Web Appendix 11).

Finally, we note how the CATE estimates obtained across simulations were more stable for the mEP-learner, EP-learner, and T-learner compared to those obtained by the mDR-learner, DR-learner, or IPTW-IPCW learner. This demonstrates how these estimators are prone to producing extreme CATE estimates when their weights are unstable. This can be seen more clearly when performance is measured using mean RMSE (Web Appendix 11).

5 GBSG2 TRIAL ANALYSIS

5.1 Background and methods

We illustrate the use of the mDR-learner and mEP-learner by applying them to the German Breast Cancer Study (GBSG2) randomized trial (Schumacher et al., 1994). This trial randomly assigned patients to a hormonal therapy ($n = 440$) or no hormonal therapy ($n = 246$) after surgery and recorded baseline covariates on demographics, medical history, and disease progression. Treatment efficacy was explored by reviewing a binary indicator of having breast cancer recurrence or death within 3 years of surgery. As some patients leave the study before making it to 3 years, missing outcome data is present, with 158 (46.5%) and 66 (26.8%) of the randomized patients lost to follow-up in each treatment arm. We conduct an intention-to-treat analysis and estimate 2 CATEs: one conditioned on all baseline covariates and one conditioned solely on progesterone receptor levels (fmol/l), where higher levels are associated with greater benefits from hormonal therapies.

In this trial, patients with non-missing outcomes had higher average progesterone receptor levels at baseline than the full randomized population, with a greater increase seen in amongst patients in the hormonal therapy arm. If left unaccounted for, this may result in CATE estimates that suggest hormonal therapies have a greater benefit than is true. Instead, we estimate CATEs using the mDR-learner and mEP-learner and compare these to estimates from the the DR-learner, EP-learner, T-learner, and IPTW-IPCW learner, with the first 3 fit using available case analyses as well as in combination with imputed outcomes. All nuisance models were fit using all baseline covariates, and all models, including the pseudo-outcome models, were estimated using a *Super Learner*, with the focus on obtaining accurate point estimates rather than CIs. The DR, EP, and IPTW-IPCW

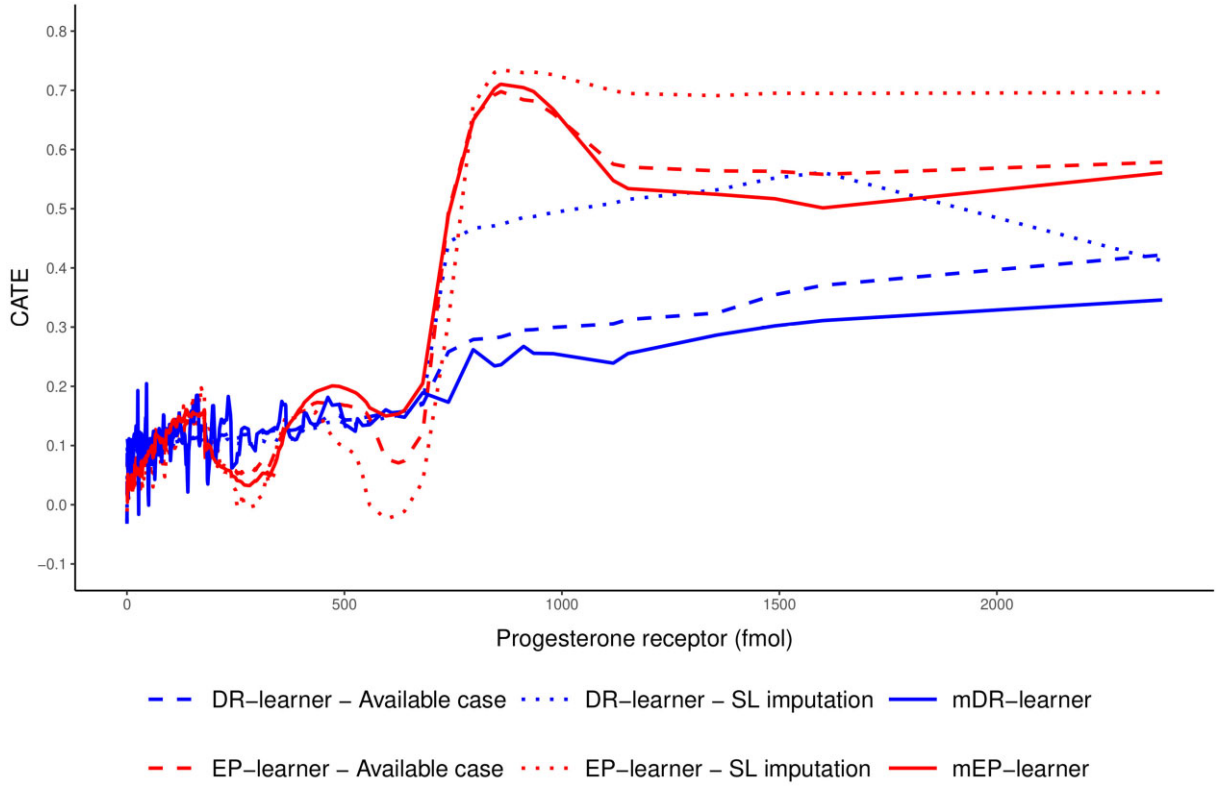

FIGURE 4 Median conditional average treatment effect (CATE) estimates plotted by progesterone receptor (fmol/l).

TABLE 1 Mean conditional average treatment effect (CATE) estimates by progesterone receptor groups when allowing the CATE to be conditional on all baseline covariates.

Progesterone receptor (Categorical)	mDR	mEP	DR		EP		T		IPTW-IPCW
			AC	Imputation	AC	Imputation	AC	Imputation	
< 500	0.095	0.080	0.114	0.109	0.081	0.067	0.075	0.061	0.069
$500 - 999$	0.051	0.397	−0.074	−0.084	0.402	0.157	0.102	0.149	−0.785
$1000 - 1499$	0.388	0.405	0.401	0.393	0.408	0.345	0.292	0.282	0.144
$1500 - 1999$	0.452	0.336	0.502	0.517	0.358	0.285	0.333	0.304	0.857
≥ 2000	0.290	0.239	0.283	0.282	0.244	0.195	0.179	0.166	0.299

AC = available case.

learners are implemented using 10 fold cross-fitting, and we report the median CATE estimate over 10 different sample splitting random seeds. Estimates with CIs that were obtained using untuned random forests (and 500 bootstrap samples) are reported in Web Appendix 13. Additionally, as the GBSG2 dataset contains event times, we provide an additional comparison to estimates obtained using causal survival forests in Web Appendix 14.

5.2 Findings

Figure 4 shows the median CATE estimates conditional on progesterone receptor levels only. The available case DR and EP learners produce higher CATE estimates than the mDR and mEP learners, respectively. This is expected, as the observed hormonal therapy arm has higher progesterone receptor levels, and once adjusted for, we would expect to see smaller CATE estimates. Additionally, we note that the estimates from the DR and EP learners fit using imputed outcomes increase rather than decrease. We also review the CATE estimates, which are conditional on all baseline covariates, with Table 1 reporting the mean CATE estimates for individuals in 5 progesterone receptor groups. Similar trends are seen in "1500-1999" receptor level category; however, trends are less obvious in areas of the population with good representation. Table 1 also reports estimates from the IPTW-IPCW learner, which are highly unstable, and estimates from the T-learner, which suggest smaller treatment effects. Finally, greater stability is observed for estimates obtained from the EP-learner variations than the DR-learner variations (Figure 5), highlighting the DR/mDR-learner's instability.

6 DISCUSSION

In this paper, we discussed the commonly overlooked problem of estimating the CATE when outcome data is MAR. Our work

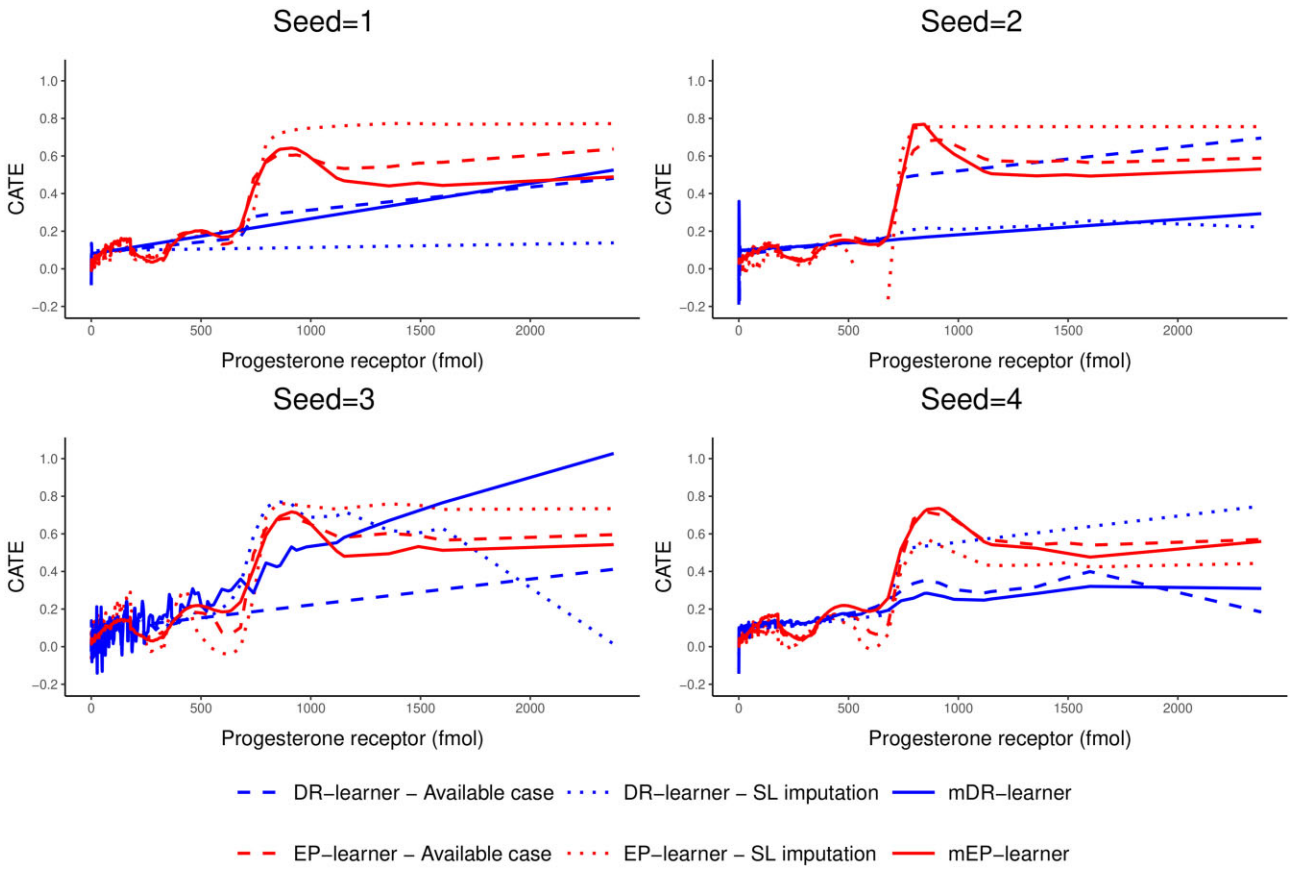

FIGURE 5 Conditional average treatment effect (CATE) estimates from single cross-fitting seeds plotted by progesterone receptor (fmol/l).

proposes 2 robust oracle efficient estimators, the mDR-learner and mEP-learner, which address the population imbalances introduced by missing outcome data by robustly incorporating IPCWs into the DR-learner and EP-learner, respectively. Our proposed approaches and implementation guidance have the potential to help improve CATE estimation in real-world data settings where outcome data is MAR. However, there remains considerable scope for further developments in this area.

Firstly, the existing tools for obtaining CIs for non-parametric CATEs either restrict the form of the CATE or limit the estimation tools that can be used. We think further development of these techniques to allow for a wider variation of estimation techniques would greatly improve the utility of these approaches. Additionally, when generating half sample bootstrap CIs, we observed very poor coverage for certain individuals, along with very wide CIs for others. For these CIs to have utility in practical examples, improved conditional coverage will be required. We also think there is great scope for further extensions of these techniques to handle more complex data, for example, post-baseline covariate information or missing covariate data. We outline an example of one of these extensions in Web Appendix 15, where we discuss how the mDR-learner could be extended to handle post-baseline covariates. Finally, we highlight that although the CATE can be used to construct individualized treatment rules, estimators that directly target such rules often prove more efficient than CATE-then-threshold pipelines (Qian and Murphy, 2011; Luedtke and Chambaz, 2020). For this reason, missing outcome data extensions of existing estimators, which directly estimate individualized treatment rules, would also be of great interest.

SUPPLEMENTARY MATERIALS

Supplementary material is available at *Biometrics* online.

Web Appendices referenced in Sections 2-6 and the R code used to implement the learners in this paper are available with this paper at the Biometrics website on Oxford Academic. Additional R code can be found at https://github.com/Matt-Pryce/mDR-learner_mEP-learner.

FUNDING

This work was supported by the Medical Research Council [grant number MR/N013638/1]. K.D.O. was funded by a Royal Society-Welcome Trust Sir Henry Dale fellowship, grant number 218554/Z/19/Z. R.H.K. was funded by UK Research and Innovation (Future Leaders Fellowship MR/X015017/1). S.V. was supported by Advanced ERC grant ACME (101141305).

CONFLICT OF INTEREST

None declared.

DATA AVAILABILITY

The GBSG2 dataset used in this paper can be accessed through the TH.data package in R: https://cran.r-project.org/web/packages/TH.data/TH.data.pdf.

REFERENCES

Berrevoets, J., Imrie, F., Kyono, T., Jordon, J. and van der Schaar, M. (2023). To impute or not to impute? Missing data in treatment effect estimation. In: *International Conference on Artificial Intelligence and Statistics*. PMLR.

Bonvini, M., Zeng, Z., Yu, M., Kennedy, E. H. and Keele, L. (2023 .). Flexibly estimating and interpreting heterogeneous treatment effects of laparoscopic surgery for cholecystitis patients, arXiv, arXiv:2311.04359, preprint: not peer reviewed.

Foster, D. J. and Syrgkanis, V. (2023). Orthogonal statistical learning. *The Annals of Statistics*, 51, 879–908.

Gonzalez Ginestet, P., Kotalik, A., Vock, D. M., Wolfson, J. and Gabriel, E. E. (2021). Stacked inverse probability of censoring weighted bagging: a case study in the InfCareHIV Register. *Journal of the Royal Statistical Society Series C: Applied Statistics*, 70, 51–65.

Groenwold, R. H., Moons, K. G. and Vandenbroucke, J. P. (2014). Randomized trials with missing outcome data: how to analyze and what to report. *CMAJ*, 186, 1153–1157.

Hines, O., Dukes, O., Diaz-Ordaz, K. and Vansteelandt, S. (2022). Demystifying statistical learning based on efficient influence functions. *The American Statistician*, 76, 1–48.

Jacob, D. (2020). Cross-fitting and averaging for machine learning estimation of heterogeneous treatment effects. Report, IRTG 1792 Discussion Paper.

Kennedy, E. H. (2023). Towards optimal doubly robust estimation of heterogeneous causal effects. *Electronic Journal of Statistics*, 17, 3008–3049.

Künzel, S. R., Sekhon, J. S., Bickel, P. J. and Yu, B. (2019). Metalearners for estimating heterogeneous treatment effects using machine learning. *Proceedings of the National Academy of Sciences*, 116, 4156–4165.

Luedtke, A. and Chambaz, A. (2020). Performance guarantees for policy learning. *Annales de l'IHP Probabilites et statistiques*, 56, 2162.

Luedtke, A. R., Sofrygin, O., van der Laan, M. J. and Carone, M. (2017). Sequential double robustness in right-censored longitudinal models, arXiv, arXiv:1705.02459, preprint: not peer reviewed.

Morzywolek, P., Decruyenaere, J. and Vansteelandt, S. (2023). On a general class of orthogonal learners for the estimation of heterogeneous treatment effects. arXiv, arXiv:2303.12687, preprint: not peer reviewed.

Nie, X. and Wager, S. (2021). Quasi-oracle estimation of heterogeneous treatment effects. *Biometrika*, 108, 299–319.

Pearl, J., Glymour, M. and Jewell, N. P. (2016). *Causal Inference in Statistics: A Primer*. Hoboken, NJ: John Wiley and Sons.

Qian, M. and Murphy, S. A. (2011). Performance guarantees for individualized treatment rules. *Annals of Statistics*, 39, 1180.

Ritzwoller, D. M. and Syrgkanis, V. (2024). Uniform inference for subsampled moment regression. arXiv, arXiv:2405.07860, preprint: not peer reviewed.

Robins, J. M., Rotnitzky, A. and Zhao, L. P. (1994). Estimation of regression coefficients when some regressors are not always observed. *Journal of the American Statistical Association*, 89, 846–866.

Rubin, D. B. (2005). Causal inference using potential outcomes: Design, modeling, decisions. *Journal of the American Statistical Association*, 100, 322–331.

Schumacher, M., Bastert, G., Bojar, H., Hübner, K., Olschewski, M. and Sauerbrei, e. a. (1994). Randomized 2 x 2 trial evaluating hormonal treatment and the duration of chemotherapy in node-positive breast cancer patients. *Journal of Clinical Oncology*, 12, 2086–2093.

Takatsu, K. and Westling, T. (2025). Debiased inference for a covariate-adjusted regression function. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 87, 33–55.

van der Laan, L., Carone, M. and Luedtke, A. (2024). Combining T-learning and DR-learning: a framework for oracle-efficient estimation of causal contrasts, arXiv, arXiv:2402.01972, preprint: not peer reviewed.

Van der Laan, M. J., Polley, E. C. and Hubbard, A. E. (2007). Super learner. *Statistical Applications in Genetics and Molecular Biology*, 6, Article25.

VanderWeele, T. J., Luedtke, A. R., van der Laan, M. J. and Kessler, R. C. (2019). Selecting optimal subgroups for treatment using many covariates. *Epidemiology*, 30, 334–341.

Vansteelandt, S. and Morzywołek, P. (2023). Orthogonal prediction of counterfactual outcomes. arXiv, arXiv:2311.09423, preprint: not peer reviewed.

Zhang, T. and Simon, N. (2023). Regression in tensor product spaces by the method of sieves. *Electronic Journal of Statistics*, 17, 3660–3727.

Received: December 24, 2024; Revised: June 20, 2025; Accepted: July 10, 2025