Iván Díaz* and Mark J. van der Laan

Targeted Data Adaptive Estimation of the Causal Dose–Response Curve

Abstract: Estimation of the causal dose–response curve is an old problem in statistics. In a non-parametric model, if the treatment is continuous, the dose–response curve is not a pathwise differentiable parameter, and no $\sqrt{n}$-consistent estimator is available. However, the risk of a candidate algorithm for estimation of the dose–response curve is a pathwise differentiable parameter, whose consistent and efficient estimation is possible. In this work, we review the cross-validated augmented inverse probability of treatment weighted estimator (CV A-IPTW) of the risk and present a cross-validated targeted minimum loss–based estimator (CV-TMLE) counterpart. These estimators are proven consistent and efficient under certain consistency and regularity conditions on the initial estimators of the outcome and treatment mechanism. We also present a methodology that uses these estimated risks to select among a library of candidate algorithms. These selectors are proven optimal in the sense that they are asymptotically equivalent to the oracle selector under certain consistency conditions on the estimators of the treatment and outcome mechanisms. Because the CV-TMLE is a substitution estimator, it is more robust than the CV-AIPTW against empirical violations of the positivity assumption. This and other small sample size differences between the CV-TMLE and the CV-A-IPTW are explored in a simulation study.

Keywords: causal inference, dose–response curve, TMLE, asymptotic linearity, oracle inequality

*Corresponding author: Iván Díaz, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD, E-mail: idiaz@jhu.edu
Mark J. van der Laan, University of California – Berkeley, Berkeley, CA 94709, USA, E-mail: laan@berkeley.edu

1 Introduction

Estimating the causal effect of an exposure A on an outcome Y when the relation between them is confounded by a set of covariates is a very common problem in causal inference, of high relevance for applications in epidemiology, medical, and social research, among other fields.

In general, causal effects are defined as parameters of the distribution of the counterfactual outcome process [1, 2] $\{Y_a : a\}$ that contains the variables that would have been observed if, possibly contrary to the fact, the subject would have received level a of the exposure. Computation of causal parameters involves expectations with respect to the distribution of the stochastic process that one would have observed if, for each subject, all the counterfactual outcomes were observed. Since the observed data for subject i contain only one of the counterfactuals, namely $Y_i = Y_{A_i}$ (this is often called the consistency assumption), additional untestable assumptions are needed in order to identify parameters of the counterfactual process distribution as parameters of the observed data distribution. These assumptions are usually described in terms of the so-called no unmeasured confounders assumption, a particular case of the coarsening at random assumption, which roughly states that the censoring or exposure processes cannot depend on unobserved covariates that are also related to the outcome.

In spite of the large number of causal inference problems that are inherently defined in terms of exposures of continuous nature, most of the attention in the field of causal inference has focused in the definition and estimation of parameters for binary treatments, in which it is natural to compare the counterfactual outcome under two possible exposure levels. Estimation of causal parameters for binary exposures has been widely studied [3–9]. The main reason why consistent and efficient estimators of the

causal dose–response curve (CDRC) for continuous treatments in the nonparametric model have not yet been developed is that it is not a pathwise differentiable parameter [10, chapter 3 and 5] and therefore cannot be estimated at a consistency rate of $n^{-1/2}$. Examples of pathwise differentiable parameters that measure the causal effect of a continuous exposure on an outcome of interest are given by the parameters defined in Díaz and van der Laan [11, 12]. These approaches make use of stochastic interventions [13–15] as a means to define a counterfactual outcome in a post-intervened world, which compared to the expectation of the actual outcome defines the causal effect of an intervention.

The most widely known method for estimation of the CDRC for continuous exposures is the so-called marginal structural model (MSM) framework, which was first proposed by Robins et al. [16]. In addition to the structural causal assumptions required for identifiability, the veracity of the statistical claims of MSM analyses relies on the correct specification of a parametric model for the CDRC, that is, the model needs to contain the true dose–response curve. Neugebauer and van der Laan [17] generalize the MSM methodology to avoid dependence on the correct specification of a parametric model by defining the parameter of interest as the projection of the true CDRC on the space of functions defined by the parameterization implied by the MSM, providing robustness against misspecification of the parametric MSM. Their work also includes identification results for this projection parameter, as well as inverse probability weighted estimators (IPTW), G-computation (G-comp) and augmented IPTW (A-IPTW) double robust estimators. Marginal structural models represent only a provisional solution to the problem, because in many instances the interest relies on estimating the actual CDRC and not its projection on some parametric space of functions.

An alternative and widely used method for estimating non–pathwise differentiable parameters is the selection of the best performing candidate among a list of algorithm estimators, where performance is defined in terms of the cross-validated risk. Formal analytical asymptotic arguments backing the use of cross-validation as an estimator selection tool were first given by van der Laan and Dudoit [18], van der Vaart [19], van der Laan et al. [20], among others. The main result of these works is a finite sample size inequality that bounds the risk of the cross-validation selector by the risk of the oracle selector (the selector based on the true distribution), which in turn used to establish, under certain conditions, the asymptotic equivalence between the cross-validation and the oracle selectors. These results are later explored in specific contexts by Dudoit and van der Laan [21], van der Laan et al. [22], among others. Of special interest is the work of van der Vaart et al. [23], in which the cross-validation oracle inequalities are extended to candidate libraries with a continuous index set and unbounded loss functions. van der Laan et al. [22] demonstrates that this oracle property for cross-validation combined with the right library of estimators results in a minimal adaptive optimal estimator. van der Laan et al. [24] use these optimality results in order to define the super learner prediction algorithm, implemented in the SuperLearner R library. van der Laan and Petersen [25] propose a general methodology for estimating the risk of a loss function indexed by nuisance parameters using cross-validated targeted minimum loss–based estimators (CV-TMLE, [26]).

For the particular case of the CDRC, van der Laan and Dudoit [18, p. 52] prove that under convergence of the initial estimators, the candidate selector based on the cross-validated A-IPTW risk is asymptotically equivalent to the oracle selector. An important assumption for identification and estimation of causal effects is the positivity assumption, which loosely ensures that there is "enough experimentation" in the data so that all the subjects have a positive probability of receiving every level of the exposure. Since the general A-IPTW methodology does not provide bounded estimators, the estimates can fall outside the parameter space and be overly sensitive to violations of the positivity assumption, which are very likely to occur when working with continuous exposures.

The main contribution of this article is to present a CV-TMLE of the risk of a CDRC candidate estimator that is endowed with an oracle inequality analogue to that of the A-IPTW. The CV-TMLE we propose is more robust to empirical violations of the positivity assumption than the CV-A-IPTW of Dudoit and van der Laan [21]. The CV-TMLE is also a substitution estimator, which guarantees estimates that are within the bounds of the parameter space. These two estimators have also been proven to be asymptotically linear with influence function equal to the efficient influence function, under certain conditions, which implies that they are consistent and efficient estimators of the risk.

The article is organized as follows. In Section 2, we formally describe the inference problem, define the loss and risk functions, and present the efficient influence function of the parameter of interest. In order to first introduce relevant concepts, in Section 2.1 we present four estimators (G-comp, IPTW, A-IPTW, and TMLE) of the risk when the candidate estimators of the CDRC are assumed fixed functions. In Section 2.2, we generalize these estimators to the case when the candidates are estimated from the sample and present the corresponding cross-validated versions of the A-IPTW and TMLE. In Section 3, we present a theorem describing the conditions under which the CV-TML estimator of the risk is an asymptotically linear estimator, the conditions under which it is consistent and efficient, as well as a discussion on the estimation of its variance. Section 4 presents the main contribution of this article; an oracle inequality for the selector based on the CV-TML estimator of the risk, and the conditions under which it is asymptotically equivalent to the oracle selector. In Section 5, we use Monte Carlo simulation to compare the performance of CV-TMLE and CV-A-IPTW selectors and estimators of the risk in finite sample sizes. We conclude in Section 6 with a summary of the work and a discussion of the limitations of our proposal and directions of future research.

2 Definition and estimation of the risk of an estimator of the CDRC

Consider an experiment in which an exposure variable A, a continuous or binary outcome Y and a set of covariates W are measured for n randomly sampled subjects. Let $O = (W, A, Y)$ represent a random variable with distribution P_0, and $O_1, \ldots, O_n$ represent n i.i.d. observations of O. The range of W, A and Y will be denoted by $\mathcal{W}$, $\mathcal{A}$ and $\mathcal{Y}$, respectively. Assume that the following non-parametric structural equation model (NPSEM) holds:

$$W = f_W(U_W); \ A = f_A(W, U_A); \ Y = f_Y(A, W, U_Y), \tag{1}$$

where U_W, U_A and U_Y are exogenous random variables such that $U_A \perp\!\!\!\perp U_Y | W$ (randomization assumption). The true distribution P_0 of O can be factorized as

$$P_0(O) = P_0(Y|A, W)P_0(A|W)P_0(W),$$

where we denote $g_0(A|W) \equiv P_0(A|W)$, $\bar{Q}_{1,0}(A, W) \equiv E_0(Y|A, W)$, $\bar{Q}_{2,0}(A, W) \equiv E_0(Y^2|A, W)$, $Q_{W,0}(W) \equiv P_0(W)$, and $Pf = \int f dP$ for a given function f. For a given value $a \in \mathcal{A}$, the counterfactual of Y is defined as the value $Y_a = f_Y(a, W, U_Y)$, the counterfactual process of Y is given by $(Y_a : a \in \mathcal{A})$ and the full data is denoted by $X = \{W, (Y_a : a \in \mathcal{A})\} \sim F_0$.

In this article, we will discuss the estimation of the causal dose–response curve within strata of the covariates $Z \subset W$, given by the expression

$$\Psi^f(F_0)(a, Z) = E_{F_0}(Y_a|Z) = \arg\min_{\psi} R^f(\psi, F_0), \tag{2}$$

where $R^f(\psi, F_0) = F_0 L^f(\psi)$, $L^f(\psi)(X) = \int_{\mathcal{A}} \{Y_a - \psi(a, Z)\}^2 h(a, Z) d\mu(a)$, the superscript f stands for full data, and h is a non-negative function such that $\int h d\mu = 1$. The second equality in eq. (2) is true because $E_{F_0}(Y_a|Z)$ is the projection of Y_a into the space of functions of Z, and $F_0 L^f(\psi)$ is the integral over $\mathcal{A}$ of the squared norm of $Y_a - \psi(a, Z)$. The randomization assumption implies that $Y_a \perp\!\!\!\perp A|W$, which allows identification of the full data parameter (2) in terms of a function of the observed data distribution as the mapping

$$\Psi(P)(a, Z) = E_P\{\bar{Q}(a, W)|Z\}, \tag{3}$$

where we denote $\psi_0 = \Psi(P_0)$. If A is continuous, $\Psi(P)$ is not a pathwise differentiable parameter in the non-parametric model, and $\sqrt{n}-$ consistent estimation is not possible [10, chapter 3, 5]. However, the risk of a

given candidate value ψ_k, is a pathwise differentiable parameter for which it is possible to find regular asymptotically linear estimators.

Following the ideas of Wang et al. [27], consider a list of candidates values $\psi_k : k = 1, \ldots, K_n$ for ψ_0^f. Throughout the article, we will make a distinction between candidate values (denoted ψ_k) and candidate estimators (denoted $\hat{\Psi}_k$), where the difference is that the former are given functions, whereas the latter are functions of (a, Z) estimated from the sample.

If the full data X were observed, a general selection procedure would involve computing $R^f(\psi_k, F_0) : k = 1, \ldots, K_n$, and estimating ψ_0^f based on ψ_{k_0}, where $k_0 = \arg\min_k R^f(\psi_k, F_0)$. Of course, this optimization procedure cannot be carried out as described because: (1) only a coarsened version of X denoted by O is observed, (2) the distribution P_0 of O is unknown and (3) in most cases we have a list of candidate estimators $\hat{\Psi}_k$, as opposed to a list of candidate values ψ_k, which raises the issue of over-fitting.

In order to overcome these obstacles one needs to:

1. Find a mapping $R(\psi, \cdot) : \mathcal{M} \to \mathbb{R}$ that identifies R^f, i.e., a mapping such that $R(\psi, P_0)$ equals $R^f(\psi, F_0)$, under certain assumptions. It is common that $R(\psi, P) = PL_{\Gamma(P)}(\cdot, \psi)$ for a loss function $L_{\Gamma(P)}$ that is now indexed by a nuisance parameter $\Gamma : \mathcal{M} \to \mathscr{F}_\gamma$.

2. If P_0 is known, the value $R(\psi, P_0)$ suffices to find a selector among the K_n candidate values. However, since P_0 is unknown, we now need to estimate $R(\psi, P_0)$. At this point, it is worth to note that even though $\Psi(P)$ is not a pathwise differentiable parameter, the mapping $R(\psi, \cdot)$ is pathwise differentiable and can therefore be $\sqrt{n}$-consistently estimated under regularity conditions.

3. If candidate values are not available, it is necessary to estimate the risk of candidate estimators $\hat{\Psi}_k$ that are trained in the sample, which makes necessary the use of cross-validated versions of these estimators.

In the remaining of this section, we will discuss the identification of R^f. The risk of a candidate ψ is given by $R^f(\psi, F) = FL^f(\psi)$ and is identified as a function of the observed data distribution by

$$R(\psi, P) = E_P L_{\bar{Q}(P)}(O, \psi), \tag{4}$$

where

$$L_{\bar{Q}}(O, \psi) = \int_{\mathscr{A}} E_P\left\{(Y - \psi(a, Z))^2 | A = a, W\right\} h(a, Z) d\mu(a).$$
$$= \int_{\mathscr{A}} \left\{\bar{Q}_2(a, W) - 2\bar{Q}_1(a, W)\psi(a, Z) + [\psi(a, Z)]^2\right\} h(a, Z) d\mu(a), \tag{5}$$

given the randomization assumption and the positivity assumption

$$\sup_{a \in \mathscr{A}} \frac{h(a, Z)}{g_0(a, W)} < \infty, \ Q_{W,0} - a.e. \tag{6}$$

Note that the loss function that defines the risk is not unique, since the loss functions

$$L_g(O, \psi) = \frac{(Y - \psi(A, Z))^2}{g(A, W)} h(A, Z), \tag{7}$$

$$L_{\bar{Q},g}(O, \psi) = \frac{h(A, Z)}{g(A, W)}\left[\{Y^2 - \bar{Q}_2(A, W)\} - 2\psi(A, Z)\{Y - \bar{Q}_1(A, W)\}\right]$$
$$+ \int_{\mathscr{A}} \left\{\bar{Q}_2(a, W) - 2\psi(a, Z)\bar{Q}_1(a, W) + \psi^2(a, Z)\right\} h(a, Z) d\mu(a) \tag{8}$$

lead to the same definition of the risk. Loss functions (5) and (7) come from more intuitive definitions of the risk, whereas the loss function (8) comes from efficient estimation theory and is closely related to the efficient influence function of $R(\psi, P)$. This fact is exploited by Wang et al. [27] in order to define estimators

of the risk as a cross-validated average of estimators of these loss functions. We will work toward the definition of a CV-TMLE analogue of those estimators and present similar results to those obtained by van der Laan and Dudoit [18] in terms of an oracle inequality, as well as the conditions under which the estimator of the risk is asymptotically linear.

The loss function (8), referred to as the double robust loss function, defines the efficient influence function of parameter $R(\psi)$ and plays a very important role in double robust and efficient estimation of $R(\psi)$, as explained in the next section.

Parameter (4) is a pathwise differentiable parameter, for which consistent asymptotically linear estimators can be found. Note that $R(\psi, P)$, as defined in eq. (4), depends on P only through $Q = (\bar{Q}, Q_W)$, where $\bar{Q} = (\bar{Q}_1, \bar{Q}_2)$. In an abuse of notation, we will use $R(\psi, P)$ and $R(\psi, Q)$ interchangeably, and the true value $R(\psi, Q_0)$ will be denoted by $R_0(\psi)$. We will also use the notations $R(\psi, Q)$ and $R(\psi)(Q)$ interchangeably. In Section 2.1, we will focus on the estimation of the risk when the candidates are given values. Given candidate values constitute a situation that is not very common in research problems, but provides an easy way to introduce the estimators that are going to be developed in Section 2.2, in which we will generalize these estimators to the case of a candidate estimated from the sample. Cross-validation will be used as a tool to avoid over-fitting and will lead to an oracle inequality presented in Section 4.

The efficient influence function of the risk $R(\psi, Q)$ is given by the expression

$$D(Q, g, \psi)(O) = L_{\bar{Q}, g}(O, \psi) - R(\psi, Q), \tag{9}$$

with $L_{\bar{Q}, g}$ defined in eq. (8).

2.1 Estimators of the risk of a candidate parameter value

In this section, we exploit the definitions of the risk in terms of loss functions given in the previous section in order to define various estimators of the risk. As we will see, the definitions of the risk through the different loss functions previously described lead to the definition of G-comp, IPTW and A-IPTW estimators. We will also use the efficient influence function of $R(\psi, P)$ in order to define a targeted maximum likelihood estimator of $R_0(\psi)$. The A-IPTW loss function is closely related to the efficient influence curve of $R(\psi, P)$, which results in the consistency and efficiency of the A-IPTW and TMLE. Analytical properties of these estimators have been discussed elsewhere [8, 9, 28].

We will assume that ψ is a given function of a and Z in the sense that it is not estimated from the sample. Such scenario is attainable, for example, in situations in which a pilot study is conducted in order to postulate candidate estimators with the objective of assessing their performance with data from a posterior study.

Let $\hat{\bar{Q}} = (\hat{\bar{Q}}_1, \hat{\bar{Q}}_2)$ and $\hat{g}$ be initial estimators of $\bar{Q}_0 = (\bar{Q}_{1,0}, \bar{Q}_{2,0})$ and g_0, respectively. These estimators will be denoted $\hat{\bar{Q}}$ or $\hat{\bar{Q}}(\mathbb{P})$, depending on whether it is necessary to emphasize their dependence on the empirical distribution

$$\mathbb{P} = \frac{1}{n} \sum_{i=1}^{n} \delta_{O_i}$$

with δ_x denoting a Dirac delta with a point mass at x. Estimation of $\bar{Q}_2$ is only necessary when the risk is a parameter of interest in itself, and as we will see in Section 4 it is unnecessary for candidate selection.

2.1.1 G-comp, IPTW and A-IPTW estimators

The equivalent definitions of the risk through G-comp, IPTW and A-IPTW loss functions allow the straightforward definition of three estimators of the risk of a candidate value, given by:

$$\hat{R}^G(\psi) = \frac{1}{n}\sum_{i=1}^n L_{\hat{\bar{Q}}}(O_i, \psi), \quad \hat{R}^I(\psi) = \frac{1}{n}\sum_{i=1}^n L_{\hat{g}}(O_i, \psi), \text{ and } \hat{R}^{DR}(\psi) = \frac{1}{n}\sum_{i=1}^n L_{\hat{\bar{Q}},\hat{g}}(O_i, \psi),$$

which can be seen as solutions in R of the corresponding estimating equations $\mathbb{P}D^I(\cdot|\hat{g}, \psi, R) = 0$, $\mathbb{P}D^G(\cdot|\hat{\bar{Q}}, \psi, R) = 0$, and $\mathbb{P}D^{DR}(\cdot|\hat{\bar{Q}}, \hat{g}, \psi, R) = 0$, where

$$D^I(O|g, \psi, R) = L_g(O, \psi) - R$$

$$D^G(O|\bar{Q}, \psi, R) = L_{\bar{Q}}(O, \psi) - R$$

$$D^{DR}(O|\bar{Q}, g, \psi, R) = L_{\bar{Q},g}(O, \psi) - R.$$

According to theorem 5.11 of van der Vaart [29], if $L_{\hat{g}}$ falls in a Glivenko–Cantelli class $\{L_g : g \in \mathscr{G}\}$ with probability tending to one, and $P_0(L_{\hat{g}} - L_{g_0})^2 \to 0$, then the IPTW estimator is consistent for $R_0(\psi)$. Under an appropriate Donsker condition and consistency of $\hat{g}$, the IPTW estimator is also asymptotically linear with influence function $D^I(O|g_0, \psi, R_0)$, as explained in theorem 6.18 of van der Vaart [29] and the theorems of Chapter 2 of van der Laan and Robins [8]. As a consequence, it is an inefficient estimator of the risk $R_0(\psi)$, and its variance can be estimated with the empirical variance of $D^I(O|\hat{g}, \psi, R_0)$. The G-comp estimator is asymptotically linear and efficient in a parametric model that is correctly specified, but it is inefficient otherwise.

Following similar arguments, the A-IPTW estimator is double robust in the sense that it is consistent if either of $\hat{\bar{Q}}$ or $\hat{g}$ is consistent. It is also efficient if both $\hat{\bar{Q}}$ and $\hat{g}$ are consistent. Even though the A-IPTW represents an important improvement with respect to the G-comp or the IPTW, it suffers from some of the drawbacks inherited from the estimating equation methodology. One of the most important problems of such methodology is the possibility of solutions out of the parameter space, or very unstable estimators if the positivity assumption is practically violated. For this reason, we prefer estimators that are substitution estimators, i.e. estimators that are the result of applying the map $R(\psi)$ to a certain estimated distribution $P^* \in \mathscr{M}$. As we will see, the TMLE is such a substitution estimator.

2.1.2 Targeted minimum loss–based estimator

For a review on TMLE and its properties, we refer the interested reader to van der Laan and Rose [9]. TML estimation requires the specification of three components: a valid loss function for the relevant part of the likelihood, a parametric submodel whose generalized score equals the efficient influence function and initial estimators of the relevant parts of the likelihood.

We will assume that Y is binary or that $P(Y \in [a, b]) = 1$ for known values a and b, in which case we can work with $Y^* = (Y - a)/(b - a)$ and interpret the results accordingly. Consider the loss functions $-L_j\{(\bar{Q}_j)(O)\} = Y^j \log \bar{Q}_j(A, W) + \{1 - Y^j\} \log\{1 - \bar{Q}_j(A, W)\}$; $j = 1, 2$, for $\bar{Q}_j$, and the parametric fluctuations given by $\text{logit}\,\bar{Q}_j(\varepsilon_j) = \text{logit}\,\bar{Q}_j + \varepsilon_j H_j(\psi, g)$, where

$$H_1(\psi, g)(A, W) = -2\psi(A, Z)\frac{h(a, Z)}{\hat{g}(A, W)},$$

$$H_2(\psi, g)(A, W) = \frac{h(a, Z)}{\hat{g}(A, W)}.$$

Note that these loss functions are not related to those in eqs (5), (7) or (8). The generalized scores are equal to

$$\frac{d}{d\varepsilon_1} L_1\{\bar{Q}_1(\varepsilon_1), O\}\Big|_{\varepsilon_1=0} = -2\frac{\psi(A, Z)h(A, Z)}{\hat{g}(A, W)}\{Y - \bar{Q}_1(A, W)\}$$

$$\frac{d}{d\varepsilon_2} L_2\{\bar{Q}_2(\varepsilon_2), O\}|_{\varepsilon_2=0} = \frac{h(A, Z)}{\hat{g}(A, W)}\{Y^2 - \bar{Q}_2(A, W)\},$$

corresponding with the first two parts of the efficient influence curve presented in eq. (9). The marginal distribution of W is estimated with the empirical distribution $Q_W(\mathbb{P})$ of $W_1, \ldots, W_n$. It can be shown that $Q_W(\mathbb{P})$ solves $L_{\bar{Q}}(\psi) - E_{Q_W}L_{\bar{Q}}(\psi)$ (the third part of the efficient influence curve equation) at any $\bar{Q}$.

For initial estimators $\hat{\bar{Q}}$ and $\hat{g}$, the first step TMLE of $\bar{Q}_0$ is given by $\hat{\bar{Q}}_j^* = \hat{\bar{Q}}_j(\hat{\varepsilon}_j)$, where

$$\hat{\varepsilon}_j = \arg\min_{\varepsilon} \mathbb{P} \, L_j\{\hat{\bar{Q}}_j(\varepsilon)\}. \tag{10}$$

The TMLE of $R_0(\psi)$ is now defined as the plug-in estimator $\hat{R}(\hat{\Psi}_k) \equiv R(\psi)(\hat{Q}^*)$, where $\hat{Q}^* = (\hat{\bar{Q}}_1^*, \hat{\bar{Q}}_2^*, Q_W(\mathbb{P}))$.

Under certain conditions explained in detail in van der Laan and Rose [9, appendix A.18], if $\bar{Q}_0$ and g_0 are consistently estimated, this TMLE of $R_0(\psi)$ is asymptotically linear with influence curve $D(O|\bar{Q}_0, g_0, R_0(\psi))$, which means that it is consistent and efficient. If $\hat{g}$ is consistent but $\hat{\bar{Q}}^*$ is not, the TMLE is consistent but inefficient, and its variance can be conservatively estimated by

$$\hat{\sigma}^2 = \frac{1}{n^2}\sum_{i=1}^{n}\left\{D^{DR}\left(O_i|\bar{Q}^*, \hat{g}, \hat{R}(\hat{\Psi}_k)\right)\right\}^2.$$

If one uses data-adaptive estimators in $\hat{\bar{Q}}$ and $\hat{g}$, it is often appropriate to replace the estimate of the variance by a cross-validated estimator.

The conditions needed for asymptotic linearity of the TMLE [9, appendix 18] include a Donsker condition on the class of functions that contains the estimated efficient influence function D. Such Donsker conditions impose certain restrictions on the type of algorithms that can be used for estimation of $\bar{Q}_0$ and g_0, forcing the user to find a trade off between obtaining the best possible prediction algorithms and not using algorithms that are too data-adaptive, because data-adaptive algorithms might lead to estimators that do not belong to a Donsker class (e.g. random forest).

The cross-validated TMLE, whose theoretical properties are discussed in Zheng and van der Laan [26], provides a template for the joint use of cross-validation and TMLE methodology that avoids Donsker conditions and therefore allows the use of very data-adaptive techniques in order to find consistent estimators of $\bar{Q}_0$ and g_0. An additional advantage of CV-TMLE in this setting is that it allows us to have a valid estimator of the risk of an estimated CDRC, solving the issue of over-fitting through the use of cross-validation.

2.2 Estimators of the risk of a candidate estimator

The previous section provided an algorithm to estimate the risk of a candidate value for the causal dose–response curve, when the value is given and not estimated from the sample. That scenario is very rare in real data applications, and it is very common that the CDRC candidates have to be estimated from the sample as well. In such situations, if the algorithms $\hat{\Psi}_k$ are trained in the whole sample, the use of the estimators of the risk presented in the previous sections would lead to the selection of the candidates that overfit the data.

van der Laan and Dudoit [18], van der Vaart [19], van der Laan et al. [20], among others, show that cross-validation is a powerful tool for estimating the risk of a candidate estimator of a non−pathwise differentiable parameter and shows that such cross-validation−based selection endows the selector with an oracle inequality that translates into asymptotic optimality.

Assume now that $\hat{\Psi}_k$ is a mapping that maps elements in a non-parametric statistical model into a space of functions of a and Z ($\hat{\Psi}_k : \mathcal{M} \to \mathcal{H}$). An estimate of $\psi_0 = \Psi(P_0)$ is now seen as such map evaluated in the empirical distribution of $O_1, \ldots, O_n$, i.e., $\hat{\Psi}_k(\mathbb{P})$.

Consider the following cross-validation scheme. Let a random variable S taking values in $\{0, 1\}^n$ index a random sample split into a validation sample $V_S = \{i \in \{1, \ldots, n\} : S_i = 1\}$ and a training sample

$T_S = \{V_S\}^c$, where S has a uniform distribution over a given set $\{s_1, \ldots, s_m\}$ such that $n_s \equiv \sum_j s_{i,j} \to \infty$ as $n \to \infty$ for all $i = 1, \ldots, m$. Note that this restriction excludes certain widely used types of cross-validation (e.g. leave-one-out cross-validation). This restriction is necessary because the asymptotic results of Section 3 and 4 rely on empirical process theory applied to the validation sample and therefore assume that its size converges to infinity. We also note that the union of the validation samples equals the total sample: $\cup_S V_S = \{1, \ldots, n\}$, and the validation samples are disjoint: $V_{s_1} \cap V_{s_2} = \emptyset$ for $s_1 \neq s_2$. Denote $\mathbb{P}_{T_S}$ and $\mathbb{P}_{V_S}$ the empirical distributions of a training and validation sample, respectively. For a function $g\{T_S, V_S\}$, we denote $E_S g(T, V) = \frac{1}{m} \sum_{j=1}^{m} g\{T_{s_j}, V_{s_j}\}$.

Since $\hat{\Psi}_k(\mathbb{P})$ is now a value that depends on the sample, it does not make sense to talk about a parameter $R\{\hat{\Psi}_k(\mathbb{P}), Q_0\}$, because it does not agree with the formal definition of a parameter. Nonetheless, in an abuse of language, we will talk about "estimation" of the "parameter" $E_S R\{\hat{\Psi}_k(\mathbb{P}_T), Q_0\}$, which we call the conditional (on the sample) risk of $\hat{\Psi}_k$.

2.3 Cross-validated augmented IPTW

This estimator is also discussed by Wang et al. [27] and is given by the solution of the cross-validated version of the A-IPTW estimating equation, given by

$$E_S \mathbb{P}_V L_{\hat{Q}(\mathbb{P}_T), \hat{g}(\mathbb{P}_T)} \{\hat{\Psi}_k(\mathbb{P}_T)\}.$$

This estimator is asymptotically linear under the conditions presented in van der Laan and Dudoit [18]. An oracle inequality for the selector based on the A-IPTW risk estimator is also proved in the original paper.

2.4 Cross-validated TMLE

The cross-validated targeted maximum likelihood estimator was introduced by Zheng and van der Laan [30] as an alternative to the TMLE that avoids the Donsker conditions on the efficient influence curve (discussed in Section 2.1.2). Donsker conditions on the class of functions generated by the estimated efficient function D represent an important limitation to the kind of algorithms that can be used in the initial estimators of $\bar{Q}_0$ and g_0: very data-adaptive techniques will give as a result functions that do not belong to a Donsker class. As discussed in Section 2.1.2, the consistency and efficiency of the risk estimator depend on the consistency of the initial estimator of $\bar{Q}_0$ and g_0. It is common practice in statistics to assume parametric models in order to estimate these quantities. Such parametric models are often chosen ad-hoc, based on arbitrary preferences of the researcher, and do not encode legitimate knowledge about the data generating process. Thus, we avoid such parametric assumptions and prefer to use data-adaptive techniques to find the algorithm that best approximates $\bar{Q}_0$ and g_0.

As we will see in the next section, the use of cross-validation also equips the CV-TML selector with an oracle inequality, meaning that such selector performs asymptotically as well as a selector in which the risk is computed based on the true (unknown) probability distribution.

Zheng and van der Laan [30] present two types of CV-TML estimators: one for a general parameter, and a specific CV-TMLE for the case in which Q can be partitioned as (Q_1, Q_2) and the mapping that defines the parameter is linear in Q_1. As discussed in Section 2, the risk $R(\psi, P)$ of a given candidate depends on P only through $Q(P) = \{\bar{Q}(P), Q_W(P)\}$, and it can be easily verified that $R(\psi, Q)$ is linear in Q_W.

The construction of a CV-TML estimator requires the specification of the same three components discussed in Section 2.1.2: a logistic loss function, a logistic parametric fluctuation and an initial estimator of Q. For each S, let

$$\text{logit } \bar{Q}_{j,k}(\mathbb{P}_{T_S})(\varepsilon_{j,k}) = \text{logit } \bar{Q}_j(\mathbb{P}_{T_S}) + \varepsilon_{j,k} H_j\{\hat{\Psi}_k(\mathbb{P}_{T_S}), \hat{g}(\mathbb{P}_{T_S})\},$$

where H_j; $j = 1, 2$ were defined in Section 2.1.2. This is the same fluctuation considered before, but defined only based on the training sample. With this modification, the CV-TMLE is defined analogous to the regular TMLE. Let

$$\hat{\varepsilon}_{j,k} = \arg\min_{\varepsilon} E_S \mathbb{P}_V L_j\{\hat{\bar{Q}}_{j,k}(\mathbb{P}_T)(\varepsilon)\}; j = 1, 2, \tag{11}$$

and for each S define the updates

$$\hat{\bar{Q}}^*_{j,k}(\mathbb{P}_{T_S}) = \hat{\bar{Q}}_j(\mathbb{P}_{T_S})(\hat{\varepsilon}_{j,k}); j = 1, 2, \tag{12}$$

which results in the plug-in estimator of the oracle risk

$$\hat{R}(\hat{\Psi}_k) = E_S R\left\{\hat{\Psi}_k(\mathbb{P}_T), \hat{\bar{Q}}^*_k(\mathbb{P}_T), Q_W(\mathbb{P}_V)\right\}$$

$$= \frac{1}{m} \sum_{s \in \{s_1,\ldots,s_m\}} \frac{1}{n_s} \sum_{i \in V_s} \int_{\mathscr{A}} \left\{\hat{\bar{Q}}^*_{2,k}(\mathbb{P}_{T_s})(a, W_i) - 2\hat{\bar{Q}}^*_{1,k}(\mathbb{P}_{T_s})(a, W_i)\hat{\Psi}_k(\mathbb{P}_{T_s})(a, Z_i) + \left[\hat{\Psi}_k(\mathbb{P}_{T_s})(a, Z_i)\right]^2\right\} h(a, Z_i) d\mu(a), \tag{13}$$

where $\hat{\bar{Q}}^*_k(\mathbb{P}_{T_S}) = \{\hat{\bar{Q}}^*_{1,k}(\mathbb{P}_{T_S}), \hat{\bar{Q}}^*_{2,k}(\mathbb{P}_{T_S})\}$, $Q_W(\mathbb{P}_{V_S})$ denotes the empirical distribution of W in the validation sample S, and n_s denotes the size of V_S.

For a definition of the CV-TMLE for general parameters, the interested reader is referred to the original article. In the next sections, we will present the asymptotic linearity of the previous estimator, as well as an oracle inequality for the selector based on it.

3 Asymptotic linearity of CV-TML estimator of the risk

In this section, we present a theorem establishing asymptotic linearity of the CV-TML estimator of the risk. This theorem is analogue to the theorems presented in Zheng and van der Laan [30], and its proof uses the same ideas presented in that article.

An analogue version of this theorem for the CV-A-IPTW is presented in van der Laan and Dudoit [18]. The CV-TMLE is expected to perform better than the CV-A-IPTW in finite sample sizes, in which practical positivity violations are often present and lead to CV-A-IPTW estimators that are either very unstable or provide solutions out of the range of the parameter of interest.

Theorem 1 (Asymptotic linearity). *Define*

$$\hat{R}_0(\hat{\Psi}_k) = E_S R\left\{\hat{\Psi}_k(\mathbb{P}_T), Q_0\right\} \text{ and } \hat{R}(\hat{\Psi}_k) = E_S R\left\{\hat{\Psi}_k(\mathbb{P}_T), \hat{\bar{Q}}^*_k(\mathbb{P}_T), Q_W(\mathbb{P}_V)\right\}$$

with $R\{\psi, Q\} = Q_W L_{\bar{Q}}(\psi)$. For a function $f(\mathbb{P}_{T_s})$ of O, define the norm $\|f(\mathbb{P}_T)\|_{0,S} = \sqrt{E_S P_0 f(\mathbb{P}_T)^2 h}$. Assume:

1. *There exist constants $\delta_1 > 0$ and $\delta_2 > 0$ such that $P(\hat{g}(\mathbb{P})(A|W) > \delta_1) = 1$ and $g_0(a|w) > \delta_2 \, \forall \, a, w$.*
2. *$\|\hat{g}(\mathbb{P}_T) - g_0\|^2_{0,S} = o_P(1/\sqrt{n})$*
3. *$\hat{\bar{Q}}^*_1(\mathbb{P}_{T_S}), \hat{\bar{Q}}^*_2(\mathbb{P}_{T_S})$ and $\hat{\Psi}_k(\mathbb{P}_{T_S})$ converge to some fixed $\hat{\bar{Q}}^*_1(P_0), \hat{\bar{Q}}^*_2(P_0)$ and $\hat{\Psi}_k(P_0)$ in the sense that*

$$\|\hat{g}(\mathbb{P}_T) - g_0\|_{0,S}\|\hat{\bar{Q}}^*_2(\mathbb{P}_T) - \hat{\bar{Q}}^*_2(P_0)\|_{0,S} = o_P(1/\sqrt{n})$$

$$\|\hat{g}(\mathbb{P}_T) - g_0\|_{0,S}\|\hat{\bar{Q}}^*_1(\mathbb{P}_T) - \hat{\bar{Q}}^*_1(P_0)\|_{0,S} = o_P(1/\sqrt{n})$$

$$\|\hat{g}(\mathbb{P}_T) - g_0\|_{0,S}\|\hat{\Psi}_k(\mathbb{P}_T) - \hat{\Psi}_k(P_0)\|_{0,S} = o_P(1/\sqrt{n})$$

4. *For some mean zero function $IC_g(P_0) \in \mathscr{L}^2_0(P_0)$, we have*

$$P_0 \frac{g_0 - \hat{g}(\mathbb{P})}{g_0^2} h\left[\left\{\bar{Q}_{2,0} - \hat{\bar{Q}}^*_2(P_0)\right\} - 2h\psi_0\left\{\bar{Q}_{1,0} - \hat{\bar{Q}}^*_1(P_0)\right\}\right] = (\mathbb{P} - P_0)IC_g(P_0) + o_P(1/\sqrt{n}),$$

Then we have that

$$\hat{R}(\hat{\Psi}_k) - \hat{R}_0(\hat{\Psi}_k) = (\mathbb{P} - P_0)[D\{\hat{\bar{Q}}_k(P_0), Q_{W,0}, g_0, \psi_0\} + IC_g(P_0)] + o_P(1/\sqrt{n}),$$

for $D\{\bar{Q}, Q_W, g, \psi\} = L_{\bar{Q},g}(\psi) - Q_W L_{\bar{Q}}$ *the efficient influence function of* $R\{\psi, Q(P)\}$.

The proof of this theorem is presented in Appendix A. Next we will discuss the plausibility and implications of the assumptions of Theorem 1.

3.1 Discussion on the assumptions of result 1

1. This assumption is a natural assumption, equivalent to the positivity assumption for binary treatments, and needed to identify and also needed to estimate the risk using IPTW or A-IPTW estimators.

2. This is a very important assumption stating that $\hat{g}$ is a consistent estimator of g_0. It is required that the rate of convergence is $n^{-1/4}$ or faster. This condition is automatically true in randomized control trials (RCT), in which the treatment mechanism is known. It is also true if g is known to belong to a parametric model, and in semi-parametric models that assume enough smoothness of g_0. If g_0 is completely unknown, it is important to use aggressive data-adaptive estimation techniques such as the super learner [24] to find an estimator $\hat{g}$ that is more likely to satisfy this assumption.

3. This assumption states that the updated estimator $\hat{\bar{Q}}_1^*$ converges to some unspecified limit at a certain rate. It is worth to note that such limit is not assumed to be $\bar{Q}_{1,0}$, the only requirement is convergence to some value at a certain rate that depends on the rate of convergence of $\hat{g}$ to g_0. The desired rate of convergence can be achieved if, for example, $\hat{g}$ is $\sqrt{n}$-consistent (i.e. $\sqrt{n}\|\hat{g}(\mathbb{P}_T) - g_0\|_{0,S} = O_P(1)$) and $\hat{\bar{Q}}_1^*(\mathbb{P}_{T_s})$ converges to $\hat{\bar{Q}}_1^*(P_0)$ at any rate (i.e. $\|\hat{\bar{Q}}_1^*(\mathbb{P}_T) - \hat{\bar{Q}}_1^*(P_0)\|_{0,S} = o_P(1)$). The same is true for $\hat{\bar{Q}}_2^*$ and $\hat{\Psi}_k$.

4. In an RCT, in which g_0 is known, one could set $\hat{g}(\mathbb{P}) = g_0$ and this condition would be trivially satisfied. On the other hand, since cross-validation allows for the use of very aggressive techniques for estimation of $\bar{Q}_0$, we could have that $\hat{\bar{Q}}^*(P_0) = \bar{Q}_0$, and the condition would also be satisfied.

In other cases, this assumption seems to be conflicting with assumption 2. If the treatment mechanism is completely unknown, it is necessary to use very aggressive data-adaptive techniques to find estimators that satisfy assumption 2. The use of such estimators will usually lead to estimates of g_0 that do not provide the asymptotic linearity needed in 4. Likewise, the use of an inconsistent estimator that satisfies this condition (e.g. a parametric model) will violate assumption 2. In that case, it is necessary to rely on the consistency of $\hat{\bar{Q}}^*(\mathbb{P})$ in the sense that $\hat{\bar{Q}}^*(P_0) = \bar{Q}_0$, in which case assumption 4 will be trivially satisfied. This condition seems to suggest that the initial estimator $\hat{g}$ must also be fluctuated to target a smooth functional of g_0. This is a direction of future research beyond the scope of this article.

As opposed to the regular TMLE or A-IPTW, in which the Donsker conditions on D limit the use of very aggressive techniques for estimation of $\bar{Q}_0$, the use of cross-validation allows us to implement any type of algorithm, which in turn makes consistency of $\hat{\bar{Q}}^*(\mathbb{P})$ a very sensible assumption. We encourage the use of super learning for estimation of both $\bar{Q}_0$ and g_0. Super learner is a methodology that uses cross-validated risks to find an optimal estimator among a library defined by the convex hull of a user-supplied list of candidate estimators. One of its most important theoretical properties is that its solution converges to the oracle estimator (i.e. the candidate in the library that minimizes the loss function with respect to the true probability distribution). Proofs and simulations regarding these and other asymptotic properties of the super learner can be found in van der Laan et al. [20] and van der Laan and Dudoit [18].

4 Asymptotic optimality of the CDRC estimate selector based on CV-TMLE risk

If the objective is to choose the best candidate among a list of candidate estimators $\hat{\Psi}_k : k = 1, \ldots, K_n$, it suffices to construct a ranking based on the pseudo-risk

$$R^{\dagger}(\psi)(\bar{Q}_1, Q_W) = E_{Q_W} \int_{\mathscr{A}} \psi(a, Z)\{\psi(a, Z) - 2\bar{Q}_1(a, W)\}h(a, Z)d\mu(a).$$

which has the advantage that $\bar{Q}_{2,0}$ does not need to be estimated, providing additional robustness of the candidate selector. In an abuse of notation R^2 and $\bar{Q}_1$ will also be denoted by R and $\bar{Q}$ whenever the difference is clear from the context. Estimation of this pseudo-risk can be carried out in a similar fashion to estimation of the full risk presented in the previous section, with efficient influence function given by

$$
\begin{aligned}
D^{\dagger}(Q, g, \psi)(O) = {} & -2\frac{h(A, Z)\psi(A, Z)}{g(A, W)}\{Y - \bar{Q}_1(A, W)\} \\
& + \int_{\mathscr{A}} \psi(a, Z)\{\psi(a, Z) - 2\bar{Q}_1(a, W)\}h(a, Z)d\mu(a) - R(\psi)(\bar{Q}_1, Q_W),
\end{aligned}
\tag{14}
$$

which results in a CV-TMLE defined as

$$\hat{R}(\hat{\Psi}_k) \equiv E_S R\left\{\hat{\Psi}_k(\mathbb{P}_T), \hat{\bar{Q}}^*(\mathbb{P}_T), Q_W(\mathbb{P}_V)\right\}$$

with $\hat{\bar{Q}}^*(\mathbb{P}_{T_S})$ exactly as in eq. (12). We will discuss now asymptotic optimality of the selector based on the CV-TMLE. Assume that we have a list of candidate estimators for the CDRC given by $\hat{\Psi}_k; \ k = 1, \ldots K_n$. Each of these algorithms is viewed as a map $\hat{\Psi}_k : \mathscr{M} \to \mathscr{F}$, where $\mathscr{F}$ is the space of functions of a and Z. Define the CV-TMLE selector as

$$\hat{k} = \arg\min_{k=1,\ldots,K_n} \hat{R}(\hat{\Psi}_k),$$

and the oracle selector as

$$\tilde{k} = \arg\min_{k=1,\ldots,K_n} \hat{R}_0(\hat{\Psi}_k),$$

with $\hat{R}_0(\hat{\Psi}_k) = E_S R\{\hat{\Psi}_k(\mathbb{P}_T), Q_0\}$. The following theorem proves that these two selectors are asymptotically equivalent under certain consistency conditions of the initial estimator of g_0.

Theorem 2 (Oracle inequality). *For each k, define*

$$\hat{\varepsilon}_k = \arg\min_{\varepsilon \in B \subset \mathbb{R}} E_S \mathbb{P}_V L\{\hat{\bar{Q}}_k(\mathbb{P}_T)(\varepsilon)\}$$

where $|B| = n^c$ for finite c,

$$-L(\bar{Q})(O) = Y \log\{\bar{Q}(A, W)\} + (1 - Y)\log\{1 - \bar{Q}(A, W)\},$$

and

$$\operatorname{logit} \bar{Q}_k(\mathbb{P}_{T_S})(\varepsilon) = \operatorname{logit} \bar{Q}(\mathbb{P}_{T_S}) - 2\varepsilon \frac{h\hat{\Psi}_k(\mathbb{P}_{T_S})}{\hat{g}(\mathbb{P}_{T_S})}.$$

Let $\hat{\bar{Q}}_k^*(\mathbb{P}_{T_S}) = \hat{\bar{Q}}(\mathbb{P}_{T_S})(\hat{\varepsilon}_k)$ be the CV-TMLE targeted toward estimation of the true conditional risk

$$\hat{R}_0(\hat{\Psi}_k) = E_S R\{\hat{\Psi}_k(\mathbb{P}_T), Q_0\}.$$

Assume that $h/\hat{g}$, h/g_0, ψ_0, $\bar{Q}_0$, and $\hat{\bar{Q}}$ have supremum norm smaller than a constant $C < \infty$ with probability 1. Let M_n be the total number of possible points for (k, ε_k) across $k = 1, \ldots, K_n$, so that $M_n \leq n^c K_n$. Define $\hat{R}(\hat{\Psi}_k, \psi_0) \equiv \hat{R}(\hat{\Psi}_k) - \hat{R}(\psi_0)$ and $\hat{R}_0(\hat{\Psi}_k, \psi_0) \equiv \hat{R}_0(\hat{\Psi}_k) - R_0(\psi_0)$, where

$$\hat{R}(\hat{\Psi}_k) = E_S R\{\hat{\Psi}_k(\mathbb{P}_T), \hat{\bar{Q}}_k^*(\mathbb{P}_T), Q_W(\mathbb{P}_V)\}$$

is the TMLE of $\hat{R}_0(\hat{\Psi}_k)$. The expression $a_n \lesssim b_n$ means that $a_n \leq c b_n$ for a constant c. For a function $f(\mathbb{P}_{T_S})$ of O, define the norm $||f(\mathbb{P}_T)||_{0,S} = \sqrt{E_S P_0 f(\mathbb{P}_T)^2 h}$. We have for each $\delta > 0$, there exists a $c(M,\delta) < \infty$ so that

$$\sqrt{E\hat{R}_0(\hat{\Psi}_{\hat{k}}, \psi_0)} - \sqrt{(1+2\delta) E\hat{R}_0(\hat{\Psi}_{\hat{k}}, \psi_0)} \lesssim \sqrt{c(M,\delta)\frac{1+\log M_n}{n}}$$

$$+ \sqrt{(1+\delta) E||\hat{g}(\mathbb{P}_T) - g_0||_{0,S}\frac{1+\log M_n}{\sqrt{n}}}$$

$$+ (1+\delta) E||\hat{g}(\mathbb{P}_T) - g_0||_{0,S} E||\hat{\bar{Q}}_k^*(\mathbb{P}_T) - \bar{Q}_0||_{0,S}$$

$$+ (1+\delta) E||(\hat{g}(\mathbb{P}_T) - g_0)(\hat{\bar{Q}}_0^*(\mathbb{P}_T) - \bar{Q}_0)||_{0,S},$$

where $\hat{\bar{Q}}_0^*$ is the CV-TMLE of $\bar{Q}_0$ obtained when the target parameter is $R(\psi_0)$, and $\bar{k}$ is either $\hat{k}$ or $\tilde{k}$, whichever gives the worst bound.

A proof of this theorem is provided in Appendix B. The use of a grid of size n^c for constant c when estimating ε_k does not represent a limitation of the result of the theorem, since the result without the grid will be similar up to a term $O_P(1/\sqrt{n})$ that does not affect the asymptotic behavior of the CV-TMLE selector. However, a grid of size n^c allows the proof presented in Appendix B.

The following corollary provides the conditions under which the CV-TMLE selector is asymptotically equivalent to the oracle selector.

Corollary 1 (Asymptotic optimality). *In addition to the conditions of Theorem 2, assume that*

$$\frac{1+\log M_n}{n}\frac{1}{E\hat{R}_0(\hat{\Psi}_{\tilde{k}}, \psi_0)} \to 0 \quad as \quad n \to \infty$$

$$\frac{1+\log M_n}{\sqrt{n}}\frac{E||\hat{g}(\mathbb{P}_T) - g_0||_{0,S}}{E\hat{R}_0(\hat{\Psi}_{\tilde{k}}, \psi_0)} \to 0 \quad as \quad n \to \infty$$

$$\frac{E^2||\hat{g}(\mathbb{P}_T) - g_0||_{0,S} E^2||\hat{\bar{Q}}_k^*(\mathbb{P}_T) - \bar{Q}_0||_{0,S}}{E\hat{R}_0(\hat{\Psi}_{\tilde{k}}, \psi_0)} \to 0 \quad as \quad n \to \infty$$

$$\frac{E^2||(\hat{g}(\mathbb{P}_T) - g_0)(\hat{\bar{Q}}_0^*(\mathbb{P}_T) - \bar{Q}_0)||_{0,S}}{E\hat{R}_0(\hat{\Psi}_{\tilde{k}}, \psi_0)} \to 0 \quad as \quad n \to \infty.$$

then

$$\frac{E\hat{R}_0(\hat{\Psi}_{\hat{k}}, \psi_0)}{E\hat{R}_0(\Psi_{\tilde{k}}, \psi_0)} \to 1 \quad as \quad n \to \infty.$$

Since

$$E\hat{R}_0(\hat{\Psi}_{\tilde{k}}, \psi_0) = E\int\int (\hat{\Psi}_{\tilde{k}}(\mathbb{P}_T) - \psi_0)^2 d\mu \, dQ_{W,0} = E||(\hat{\Psi}_{\tilde{k}}(\mathbb{P}_T) - \psi_0)/\sqrt{g_0}||_{0,S}^2,$$

the convergences assumed in Corollary 1 are expected to hold, for example, if $\hat{g}$ converges to g_0 at a rate faster than $\hat{\Psi}$ converges to ψ_0.

In the following section, we will show the results of a simulation study in which the finite sample size properties of the CV-TMLE based selector of their risk are explored for a specific data generation process.

5 Simulation

In order to explore some of the finite sample size properties of the risk estimators and the selectors based on them, we performed a Monte Carlo simulation. We generated 500 samples of sizes 100, 500 and 1,000 from the following data generating process:

$$W_1 \sim U\{0, 1\}$$
$$W_2 \sim \text{Ber}\{0.7\}$$
$$W_3 \sim N\{W_1, 0.25 \times \exp(2W_1)\}$$
$$A \sim \text{Beta}\{v(W)\mu(W), v(W)[1 - \mu(W)]\}$$
$$Y \sim \text{Ber}\{Q_0(A, W)\},$$

where

$$v(W) = \exp\{1 + 2W_1\text{expit}(W_3)\}$$
$$\mu(W) = \text{expit}\{.03 - .8\log(1 + W_2) + .9\exp(W_1)W_2 - .4\arctan(W_3 + 2)W_2W_1\}$$
$$\bar{Q}_0(A, W) = \text{expit}\{-2 + 1.5A + 5A^3 - 2.5W_1 + .5AW_2 - \log(A)W_1W_2 + .5A^{3/4}W_1W_3\}.$$

Under this parameterization $E(A|W) = \mu(W)$. We considered four candidates algorithms given by marginal structural models (MSM) of the form logit $\Psi_p(a) = m_p(a, \beta)$, where m_p is a polynomial of degree $p = 1, \ldots, 4$ on a with coefficients $\beta_j : j = 0, \ldots, p$. The coefficients β_j were estimated with IPTW estimators as presented in Robins et al. [16] and Neugebauer and van der Laan [17]. The true value of $\psi_0(a) = E\{\bar{Q}_0(a, W)\}$ was computed from this data generating distribution by drawing a sample of size 100.000 and, for each a, computing the empirical mean of $\bar{Q}_0(a, W)$. All the simulations were performed assuming the true parametric model for the outcome and treatment mechanism was known. Figure 1 presents the true dose–response curve, as well as the expectation of the candidate estimators across the 500 samples. From this graph, we can see that among the candidates chosen, a polynomial of degree 2 seems to provide the closest approximation to the true dose–response curve with fewer parameters, therefore providing the best bias-variance trade-off. Table 1 shows the expectation of the random variable $\hat{R}(\Psi_p) - \hat{R}_0(\Psi_p)$, which from Theorem 1 should approach zero as the sample size increases. As we can see, that is not the case for the CV-A-IPTW estimator with sample size 100 due to the presence of empirical violations of the positivity assumption that cause very small treatment weights and therefore very unstable, non-regular estimates. However, that problem seems to be fixed asymptotically, since for large sample sizes empirical violations of the positivity assumption are less likely to occur.

Table 2 shows the proportion of estimates that fell outside the interval (–10, 10) or fell out of the parameter space. The interval (–10, 10) was chosen arbitrarily and represents an extreme of inadmissible bounds for an estimator of a parameter that ranges in the interval (0, 1). Since the TMLE is a substitution
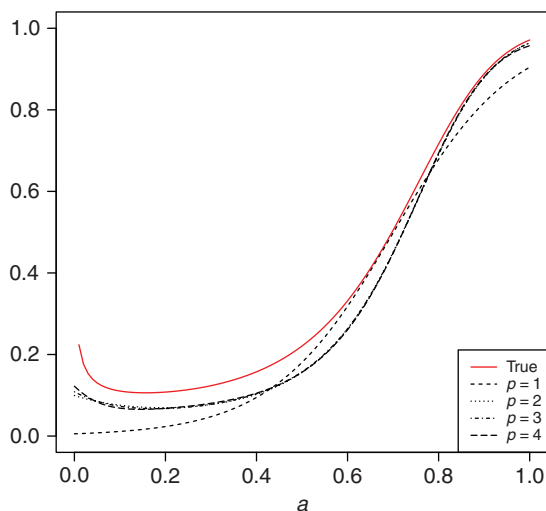
Figure 1 True $\psi_0(a)$ and expectations of the four candidate estimators of degree p

Table 1 Expectation of $\hat{R}(\Psi_p) - \hat{R}_0(\Psi_p)$ across 500 simulated samples

Risk estimator	Candidate degree p			n
		100	500	1,000
CV-A-IPTW	1	−1.1436	0.0067	0.0048
	2	−1.1716	−0.0061	0.0042
	3	−1.0788	−0.0054	0.0043
	4	−1.0194	−0.0053	0.0032
CV-TMLE	1	0.0063	0.0054	0.0046
	2	0.0085	0.0054	0.0041
	3	0.0091	0.0059	0.0043
	4	0.0094	0.0058	0.0042

Table 2 Proportion of estimates outliers (< -10 or > 10) and proportion of estimates out of bounds (< 0 or > 1)

Risk estimator	Candidate degree p							n
			Outliers				Out of bounds	
		100	500	1,000		100	500	1,000
CV-A-IPTW	1	0.0098	0.0000	0.0000		0.0547	0.0020	0.0020
	2	0.0078	0.0000	0.0000		0.0527	0.0020	0.0020
	3	0.0098	0.0000	0.0000		0.0488	0.0020	0.0020
	4	0.0098	0.0000	0.0000		0.0508	0.0020	0.0020

Table 3 Expectation of $\hat{R}(\Psi_p) - \hat{R}_0(\Psi_p)$ across 500 simulated samples after removing estimates out of bounds

Risk estimator	Candidate degree p			n
		100	500	1,000
CV-A-IPTW	1	0.0034	0.0064	0.0048
	2	−0.0076	−0.0062	0.0042
	3	−0.0083	−0.0054	0.0043
	4	−0.0083	−0.0055	0.0032
CV-TMLE	1	0.0059	0.0055	0.0045
	2	0.0082	0.0054	0.0039
	3	0.0089	0.0059	0.0041
	4	0.0093	0.0058	0.0042

estimator, all the estimates fell within the parameter space and are thus not presented. Due to practical violations of the positivity assumption previously mentioned, an important proportion (around 5%) of the A-IPTW estimates fell outside the parameter space for sample size 100.

Table 3 contains the expected values of $\hat{R}(\Psi_p) - \hat{R}_0(\Psi_p)$ across 500 simulated samples once the estimates that fell outside the interval (0, 1) were removed. In this case, the expectation of the A-IPTW–based estimator of the risk is much closer to what is expected theoretically and had already been achieved by the TML estimator.

Finally, Table 4 shows the proportion of times that a given candidate is chosen according to the A-IPTW, TMLE and the oracle selector. As we can see, both the A-IPTW- and the TMLE-based selectors perform similar to the oracle selector, particularly as the sample size increases, thus showing no apparent advantage (at least for this particular data generating mechanism) of either method when evaluated as a candidate selector procedure.

Table 4 Proportion of times that a given candidate is chosen according to each risk estimator

p									n
		100			500			1,000	
	TMLE	A-IPTW	Oracle	TMLE	A-IPTW	Oracle	TMLE	A-IPTW	Oracle
1	0.37	0.24	0.39	0.11	0.08	0.05	0.03	0.05	0.00
2	0.44	0.48	0.54	0.63	0.60	0.74	0.59	0.59	0.74
3	0.11	0.17	0.06	0.14	0.20	0.17	0.24	0.21	0.20
4	0.07	0.11	0.01	0.12	0.12	0.04	0.14	0.15	0.06

6 Discussion

In this article, we discuss estimation of the causal dose–response curve for cross-sectional studies and propose a procedure for selecting an estimator among a list of candidate algorithms. In particular, we introduced a cross-validated targeted minimum loss–based estimator of the risk of an estimator and its associated candidate selector. The approach we propose differs from commonly used approaches in that we aim to choose an algorithm among a list of candidates to fit the true dose–response curve in situations in which its functional form is unknown.

The use of augmented inverse probablity of treatment weigthed estimators of the risk has alread been discussed in van der Laan and Dudoit [18], in this article we provide a targeted minimum loss–based analogous estimator that has improved performance, as demonstrated by the simulations and analytical properties of the estimators. We provide a theorem establishing asymptotic linearity of the risk estimator, which together with its variance estimator can be used to construct hypothesis tests and confidence intervals. We also provide a theorem proving an oracle inequality bounding the expectation of the CV-TMLE risk estimator. The main corollary of this theorem is a result that establishes asymptotic equivalence of the CV-TMLE selector and the oracle selector.

The asymptotic optimality of cross-validation procedures such as the super learner [24] and our proposal is best exploited with very rich libraries of candidate estimators that contain algorithms that can provide an accurate description of the true underlying statistical (or causal) parameter. The aspects that must be taken into account when postulating candidate estimators for the dose–response curve are analogous to those in estimation of the regression function. For example, one can decide to add a candidate to the library because it represents prior knowledge about the nature of the phenomena under study, or because it is data-adaptive enough to accomodate various data generating processes. Since the extent to which the postulated algorithms approximate the true curve will often be unknown to the user, cross-validation selection will be optimal with libraries that contain a rich class of algorithms, some of which should be data-adaptive or non-parametric. Unfortunately, estimation of the causal dose–response curve has received very little attention from the machine learning research community and few data-adaptive estimators are available.

However, in a situation in which the library consists of a few parametric models (e.g. the MSMs in the simulations), our proposal provides an optimal model selection tool. Asymptotic equivalence to the oracle selector guarantees that our selector will asymptotically pick the parametric model that is closest to the true dose–response curve, among those in the library.

Appendix A

In this appendix, we will proof Theorem 1. This proof follows closely the proofs presented in Zheng and van der Laan [30] for a general CV-TMLE. Those proofs rely heavily on empirical process theory, of which van der Vaart and Wellner [31] provide a complete study.

Proof of Theorem 1. First of all note that

$$R(\psi, Q) - R(\psi, Q_0) = -P_0 D\{\bar{Q}, Q_W, g_0, \psi\},$$

which implies that

$$\hat{R}(\hat{\Psi}_k) - \hat{R}_0(\hat{\Psi}_k) = -E_S P_0 D\{\hat{\bar{Q}}^*(\mathbb{P}_T), Q_W(\mathbb{P}_V), g_0, \hat{\Psi}_k(\mathbb{P}_T)\}.$$

Note that

$$E_S \mathbb{P}_V D\{\hat{\bar{Q}}^*(\mathbb{P}_T), Q_W(\mathbb{P}_V), \hat{g}(\mathbb{P}_T), \hat{\Psi}_k(\mathbb{P}_T)\} = 0,$$

so that we can write

$$
\begin{aligned}
\hat{R}(\hat{\Psi}_k) - \hat{R}_0(\hat{\Psi}_k) &= E_S(\mathbb{P}_V - P_0) D\{\hat{\bar{Q}}^*(\mathbb{P}_T), Q_W(\mathbb{P}_V), \hat{g}(\mathbb{P}_T), \hat{\Psi}_k(\mathbb{P}_T)\} \\
&\quad + E_S P_0 [D\{\hat{\bar{Q}}^*(\mathbb{P}_T), Q_W(\mathbb{P}_V), \hat{g}(\mathbb{P}_T), \hat{\Psi}_k(\mathbb{P}_T)\} - D\{\hat{\bar{Q}}^*(\mathbb{P}_T), Q_W(\mathbb{P}_V), g_0, \hat{\Psi}_k(\mathbb{P}_T)\}] \quad (15) \\
&= E_S(\mathbb{P}_V - P_0) D\{\hat{\bar{Q}}^*(\mathbb{P}_T), Q_W(\mathbb{P}_V), \hat{g}(\mathbb{P}_T), \hat{\Psi}_k(\mathbb{P}_T)\} \\
&\quad + E_S P_0 \frac{g_0 - \hat{g}(\mathbb{P}_T)}{g_0 \hat{g}(\mathbb{P}_T)} h[\{\bar{Q}_{2,0} - \hat{\bar{Q}}_2^*(\mathbb{P}_T)\} - 2\hat{\Psi}_k(\mathbb{P}_T)\{\bar{Q}_{1,0} - \hat{\bar{Q}}_1^*(\mathbb{P}_T)\}]. \quad (16)
\end{aligned}
$$

We will first handle the term (16). By Cauchy–Schwarz, eq. (16) can be bounded by

$$
\begin{aligned}
&\sup_{(a,v)} |h(a,v)| \left\| \frac{g_0 - \hat{g}(\mathbb{P}_T)}{g_0 \hat{g}(\mathbb{P}_T)} \right\|_{0,S} \| \hat{\bar{Q}}_2^*(P_0) - \hat{\bar{Q}}_2^*(\mathbb{P}_T) \|_{0,S} \\
&\quad + E_S P_0 \frac{g_0 - \hat{g}(\mathbb{P}_T)}{g_0 \hat{g}(\mathbb{P}_T)} h\{\bar{Q}_{2,0} - \hat{\bar{Q}}_2^*(P_0)\} \\
&\quad - 2 \sup_{(s,a,v)} |h(a,v)\hat{\Psi}_k(\mathbb{P}_{T_s})(a,v)| \left\| \frac{g_0 - \hat{g}(\mathbb{P}_T)}{g_0 \hat{g}(\mathbb{P}_T)} \right\|_{0,S} \| \hat{\bar{Q}}_1^*(P_0) - \hat{\bar{Q}}_1^*(\mathbb{P}_T) \|_{0,S} \\
&\quad - 2 E_S P_0 \frac{g_0 - \hat{g}(\mathbb{P}_T)}{g_0 \hat{g}(\mathbb{P}_T)} h\hat{\Psi}_k(\mathbb{P}_T)\{\bar{Q}_{1,0} - \hat{\bar{Q}}_1^*(P_0)\}.
\end{aligned} \quad (17)
$$

Using a similar argument, the last term can be bounded (using assumption 1 and up to universal constants) by

$$\|\hat{g}(\mathbb{P}_T) - g_0\|_{0,S} \|\hat{\Psi}_k(\mathbb{P}_T) - \psi_0\|_{0,S} + E_S P_0 \frac{g_0 - \hat{g}(\mathbb{P}_T)}{g_0^2} h\psi_0\{\bar{Q}_{1,0} - \hat{\bar{Q}}_1^*(P_0)\} + \|\hat{g}(\mathbb{P}_T) - g_0\|_{0,S}^2,$$

whereas the second term in eq. (17) is bounded by

$$E_S P_0 \frac{g_0 - \hat{g}(\mathbb{P}_T)}{g_0^2} h\{\bar{Q}_{2,0} - \hat{\bar{Q}}_2^*(P_0)\} + \|\hat{g}(\mathbb{P}_T) - g_0\|_{0,S}^2.$$

Therefore

$$
\begin{aligned}
\hat{R}(\hat{\Psi}_k) - \hat{R}_0(\hat{\Psi}_k) &= E_S(\mathbb{P}_V - P_0) D\{\hat{\bar{Q}}^*(\mathbb{P}_T), Q_W(\mathbb{P}_V), \hat{g}(\mathbb{P}_T), \hat{\Psi}_k(\mathbb{P}_T)\} \\
&\quad + E_S P_0 \frac{g_0 - \hat{g}(\mathbb{P}_T)}{g_0^2} h[\{\bar{Q}_{2,0} - \hat{\bar{Q}}_2^*(P_0)\} - 2\psi_0\{\bar{Q}_{1,0} - \hat{\bar{Q}}_1^*(P_0)\}] \quad (18) \\
&\quad + \text{Rem}_n,
\end{aligned}
$$

where

$$
\begin{aligned}
\text{Rem}_n &\leq \|g_0 - \hat{g}(\mathbb{P}_T)\|_{0,S} \{ a\|\bar{Q}_2^*(P_0) - \bar{Q}_2^*(\mathbb{P}_T)\|_{0,S} - b\|\bar{Q}_1^*(P_0) - \hat{\bar{Q}}_1^*(\mathbb{P}_T)\|_{0,S} \\
&\quad + c\|\hat{\Psi}_k(\mathbb{P}_T) - \psi_0\|_{0,S} + d\|g_0 - \hat{g}(\mathbb{P}_T)\|_{0,S} \},
\end{aligned}
$$

for constants a, b, c, d.

On the other hand, since $E_S \mathbb{P}_V f = \mathbb{P}f$ when f does not depend on S, we may rewrite (15) as

$$E_S(\mathbb{P}_V - P_0)D\{\hat{\bar{Q}}^*(\mathbb{P}_T), Q_W(\mathbb{P}_V), \hat{g}(\mathbb{P}_T), \hat{\Psi}_k(\mathbb{P}_T)\}$$
$$= E_S(\mathbb{P}_V - P_0)[D\{\hat{\bar{Q}}^*(\mathbb{P}_T), Q_W(\mathbb{P}_V), \hat{g}(\mathbb{P}_T), \hat{\Psi}_k(\mathbb{P}_T)\} - D\{\bar{Q}^*(P_0), Q_{W,0}, g_0, \psi_0\}] \qquad (19)$$
$$+ (\mathbb{P} - P_0)D\{\bar{Q}^*(P_0), Q_{W,0}, g_0, \psi_0\}$$

Following similar arguments to those presented by Zheng and van der Laan [30], and using the assumptions of the theorem, it can be proven that (19) is $o_P(1/\sqrt{n})$, which implies

$$E_S(\mathbb{P}_V - P_0)D\{\hat{\bar{Q}}(\mathbb{P}_T)(\hat{\varepsilon}), Q_W(\mathbb{P}_V), \hat{g}(\mathbb{P}_T), \hat{\Psi}_k(\mathbb{P}_T)\}$$
$$= (\mathbb{P} - P_0)D\{Q(P_0)(\varepsilon_0), \hat{g}(P_0), \psi_0\} + o_P(1/\sqrt{n}).$$

This result, together with eq. (15) and assumptions 2, 3 and 4, yields

$$\hat{R}(\hat{\Psi}_k) - \hat{R}_0(\hat{\Psi}_k) = (\mathbb{P} - P_0)\big[D\{Q(P_0), Q_{W,0}, g_0, \psi_0 + IC_g(P_0)\big] + o_P(1/\sqrt{n}). \ \square$$

Appendix B

Before proceeding to prove Theorem 2, we will first present and prove the following useful theorem.

Theorem 3. *Define* $\hat{R}(\hat{\Psi}_k, \psi_0) \equiv \hat{R}(\hat{\Psi}_k) - \hat{R}(\psi_0)$ *and* $\hat{R}_0(\hat{\Psi}_k, \psi_0) \equiv \hat{R}_0(\hat{\Psi}_k) - R_0(\psi_0)$, *where*

$$\hat{R}(\hat{\Psi}_k) = E_S R\{\hat{\Psi}_k(\mathbb{P}_T), \hat{\bar{Q}}^*_k(\mathbb{P}_T), \mathbb{P}_V\}$$

is the TMLE of the true conditional risk

$$\hat{R}_0(\hat{\Psi}_k) = E_S R\{\hat{\Psi}_k(\mathbb{P}_T), Q_0\}.$$

Assume

$$(\hat{R} - \hat{R}_0)(\hat{\Psi}_k, \psi_0) = E_S(\mathbb{P}_V - P_0)D_k(\mathbb{P}_T, P_0, \hat{\varepsilon}_k) + \text{Rem}_k$$

for some function D_k *that depends on* $(\mathbb{P}_T, P_0, \hat{\varepsilon}_k)$ *such that*

$$E_S P_0 D_k(\mathbb{P}_T, P_0, \hat{\varepsilon}_k) = \hat{R}_0(\hat{\Psi}_k, \psi_0),$$

$||D_k(\mathbb{P}_T, P_0, \hat{\varepsilon}_k)||_\infty < M_1 < \infty$, *and* $P_0\{D_k(\mathbb{P}_T, P_0, \hat{\varepsilon}_k)\}^2 \le M_2 P_0 D_k(\mathbb{P}_T, P_0, \hat{\varepsilon}_k)$. *Assume also that* $\hat{\varepsilon}_k$ *falls in a finite set with maximally* n^c *points for some finite c and denote* $M_n = n^c K_n$. *Then,*

$$E\hat{R}_0(\hat{\Psi}_{\hat{k}}, \psi_0) \lesssim (1 + 2\delta)E\hat{R}_0(\hat{\Psi}_{\tilde{k}}, \psi_0) + c(M_1, M_2, \delta)\frac{1 + \log M_n}{n} + (1 + \delta)\{E\text{Rem}_{\tilde{k}} - E\text{Rem}_{\hat{k}}\},$$

where $c(M_1, M_2, \delta) = (1 + \delta)^2(M_1/3 + M_2/\delta)$

Proof of Theorem 3.

$$0 \le \hat{R}_0(\hat{\Psi}_{\hat{k}}, \psi_0)$$
$$= \hat{R}_0(\hat{\Psi}_{\hat{k}}, \psi_0) - (1 + \delta)\hat{R}(\hat{\Psi}_{\hat{k}}, \psi_0) + (1 + \delta)\hat{R}(\hat{\Psi}_{\hat{k}}, \psi_0)$$
$$\le \hat{R}_0(\hat{\Psi}_{\hat{k}}, \psi_0) - (1 + \delta)\hat{R}(\hat{\Psi}_{\hat{k}}, \psi_0) + (1 + \delta)\hat{R}(\hat{\Psi}_{\tilde{k}}, \psi_0)$$
$$= \hat{R}_0(\hat{\Psi}_{\hat{k}}, \psi_0) + (1 + \delta)\{\hat{R}(\Psi_{\tilde{k}}, \psi_0) - \hat{R}_0(\hat{\Psi}_{\tilde{k}}, \psi_0)\}$$
$$\quad - (1 + \delta)\{\hat{R}(\Psi_{\hat{k}}, \psi_0) - \hat{R}_0(\hat{\Psi}_{\hat{k}}, \psi_0)\}$$
$$\quad + (1 + \delta)\hat{R}_0(\hat{\Psi}_{\tilde{k}}, \psi_0)$$
$$\quad - (1 + \delta)\hat{R}_0(\hat{\Psi}_{\hat{k}}, \psi_0)$$
$$= (1 + 2\delta)\hat{R}_0(\hat{\Psi}_{\tilde{k}}, \psi_0) + H_{\hat{k}} + T_{\tilde{k}},$$

where

$$H_k = -(1+\delta)(\hat{R} - \hat{R}_0)(\hat{\Psi}_k, \psi_0) - \delta\hat{R}_0(\hat{\Psi}_k, \psi_0)$$

$$T_k = (1+\delta)(\hat{R} - \hat{R}_0)(\hat{\Psi}_k, \psi_0) - \delta\hat{R}_0(\hat{\Psi}_k, \psi_0).$$

By using the assumptions of the theorem we get

$$H_k = -(1+\delta)E_S(P_V - P_0)D_k(\mathbb{P}_T, P_0, \hat{\varepsilon}_k) - \delta E_S P_0 D_k(\mathbb{P}_T, P_0, \hat{\varepsilon}_k) - (1+\delta)\text{Rem}_k$$
$$\equiv H_k^* - (1+\delta)\text{Rem}_k.$$

Following arguments similar to those presented by Dudoit and van der Laan [21] and van der Vaart et al. [23] we have that

$$EH_k^* \lesssim c(M_1, M_2, \delta)\frac{1 + \log M_n}{n}.$$

The same bound applies to ET_k^*. As a consequence we obtain the desired result. □

Proof of Theorem 2. Recall from Theorem 1 that

$$\hat{R}(\hat{\Psi}_k) - \hat{R}_0(\hat{\Psi}_k) = E_S(\mathbb{P}_V - P_0)D\{\hat{\bar{Q}}_k^*(\mathbb{P}_T), Q_W(\mathbb{P}_V), \hat{g}(\mathbb{P}_T), \hat{\Psi}_k(\mathbb{P}_T)\}$$
$$+ 2E_S P_0 \frac{g_0 - \hat{g}(\mathbb{P}_T)}{g_0\hat{g}(\mathbb{P}_T)}h\hat{\Psi}_k(\mathbb{P}_T)\{\hat{\bar{Q}}_k^*(\mathbb{P}_T) - \bar{Q}_0\}, \tag{20}$$

where D is the efficient influence function

$$D(Q, Q_W, g, \psi)(O) = -2\frac{h(A, Z)\psi(A, Z)}{g(A, W)}\{Y - \bar{Q}(A, W)\}$$
$$+ \int_{\mathscr{A}} \psi(a, Z)\{\psi(a, Z) - 2\bar{Q}(a, W)\}h(a, Z)d\mu(a) - R(\psi)(\bar{Q}, Q_W).$$

Applying this same equality to the constant algorithm ψ_0 and subtracting it from eq. (20) yields

$$(\hat{R} - \hat{R}_0)(\hat{\Psi}_k, \psi_0) = E_S(\mathbb{P}_V - P_0)\Big[D\{\hat{\bar{Q}}_k^*(\mathbb{P}_T), Q_W(\mathbb{P}_V), \hat{g}(\mathbb{P}_T), \hat{\Psi}_k(\mathbb{P}_T)\} - D\{\hat{\bar{Q}}_0^*(\mathbb{P}_T), Q_W(\mathbb{P}_V), \hat{g}(\mathbb{P}_T), \psi_0\}\Big] \tag{21}$$

$$+ 2E_S P_0 \frac{g_0 - \hat{g}(\mathbb{P}_T)}{g_0\hat{g}(\mathbb{P}_T)}h\{\hat{\Psi}_k(\mathbb{P}_T) - \psi_0\}\{\hat{\bar{Q}}_0^*(\mathbb{P}_T) - \bar{Q}_0\} \tag{22}$$

$$+ 2E_S P_0 \frac{g_0 - \hat{g}(\mathbb{P}_T)}{g_0\hat{g}(\mathbb{P}_T)}h\hat{\Psi}_k(\mathbb{P}_T)\{\hat{\bar{Q}}_k^*(\mathbb{P}_T) - \hat{\bar{Q}}_0^*(\mathbb{P}_T)\} \tag{23}$$

$$= E_S(\mathbb{P}_V - P_0)\Big[D\{\hat{\bar{Q}}_k^*(\mathbb{P}_T), Q_W(\mathbb{P}_V), g_0, \hat{\Psi}_k(\mathbb{P}_T)\} - D\{\hat{\bar{Q}}_0^*(\mathbb{P}_T), Q_W(\mathbb{P}_V), g_0, \psi_0\}\Big] \tag{24}$$

$$+ E_S(\mathbb{P}_V - P_0)\Big(\Big[D\{\hat{\bar{Q}}_k^*(\mathbb{P}_T), Q_W(\mathbb{P}_V), \hat{g}(\mathbb{P}_T), \hat{\Psi}_k(\mathbb{P}_T)\} - D\{\hat{\bar{Q}}_k^*(\mathbb{P}_T), Q_W(\mathbb{P}_V), g_0, \hat{\Psi}_k(\mathbb{P}_T)\}\Big]$$
$$- \Big[D\{\hat{\bar{Q}}_0^*(\mathbb{P}_T), Q_W(\mathbb{P}_V), \hat{g}(\mathbb{P}_T), \psi_0\} - D\{\hat{\bar{Q}}_0^*(\mathbb{P}_T), Q_W(\mathbb{P}_V), g_0, \psi_0\}\Big]\Big) \tag{25}$$

$$+ \text{Rem}_{k,1} + \text{Rem}_{k,2}$$
$$\equiv E_S(\mathbb{P}_V - P_0)D_k(\mathbb{P}, P_0) + \text{Rem}_{k,1} + \text{Rem}_{k,2} + \text{Rem}_{k,3} \tag{26}$$

where $D_k(\mathbb{P}, P_0)$ denotes the function inside square brackets in eq. (24), and $\text{Rem}_{k,1}$, $\text{Rem}_{k,2}$ and $\text{Rem}_{k,3}$ denote eqs (22), (23) and (25), respectively. From the definition of the efficient influence function D, note that $D\{\bar{Q}, Q_W, g, \psi\} = L_{\bar{Q},g}(\psi) - R(\psi, \bar{Q}, Q_W)$, which implies

$$E_S(\mathbb{P}_V - P_0)D_k(\mathbb{P}, P_0) = E_S(\mathbb{P}_V - P_0)\Big[L_{\hat{\bar{Q}}_k^*(\mathbb{P}_T),g_0}(\hat{\Psi}_k(\mathbb{P}_T)) - L_{\hat{\bar{Q}}_0^*(\mathbb{P}_T),g_0}(\psi_0)\Big]$$
$$- E_S(\mathbb{P}_V - P_0)\Big[R\{\Psi(\mathbb{P}_T), \hat{\bar{Q}}_k^*(\mathbb{P}_T), Q_W(\mathbb{P}_V)\} - R\{\psi_0, \hat{\bar{Q}}_0^*(\mathbb{P}_T), Q_W(\mathbb{P}_V)\}\Big], \tag{27}$$

where the term inside square brackets in eq. (27) is a constant, and (27) equals zero. Note that $\hat{\bar{Q}}_k^*(\mathbb{P}_T) \equiv \hat{\bar{Q}}(\mathbb{P}_T)(\hat{\varepsilon}_k)$ depends on $\mathbb{P}_V$ only through $\hat{\varepsilon}_k$, thus allowing us to rewrite

$$(\hat{R} - \hat{R}_0)(\hat{\Psi}_k, \psi_0) = E_S(\mathbb{P}_V - P_0)D_k(\mathbb{P}_T, P_0, \hat{\varepsilon}_k) + \text{Rem}_{k,1} + \text{Rem}_{k,2} + \text{Rem}_{k,3},$$

for

$$D_k(\mathbb{P}_T, P_0, \hat{\varepsilon}_k) \equiv L_{\hat{\bar{Q}}_k^*(\mathbb{P}_T),g_0}(\hat{\Psi}_k(\mathbb{P}_T)) - L_{\hat{\bar{Q}}_0^*(\mathbb{P}_T),g_0}(\psi_0).$$

From the identity $P_0 L_{\bar{Q},g_0}(\psi) = P_0 L_{\bar{Q}_0}(\psi)$ ($L_{\bar{Q}}$ denotes the g-comp loss function) we have that

$$E_S P_0 D_k(\mathbb{P}_T, P_0, \hat{\varepsilon}_k) = \hat{R}_0(\hat{\Psi}_k, \psi_0).$$

This fact together with eq. (26) prove that D_k satisfies the conditions of Theorem 3, with $\text{Rem}_k = \text{Rem}_{1,k} + \text{Rem}_{2,k} + \text{Rem}_{3,k}$. By application of Theorem 3 we obtain

$$E\hat{R}_0(\hat{\Psi}_{\hat{k}}, \psi_0) \lesssim (1 + 2\delta)E\hat{R}_0(\hat{\Psi}_{\tilde{k}}, \psi_0)$$
$$+ c(M_1, M_2, \delta)\frac{1 + \log M_n}{n} + (1 + \delta)\{E\text{Rem}_{\tilde{k}} - E\text{Rem}_{\hat{k}}\}.$$

It remains to study $E\text{Rem}_k$. Let us first consider $E\text{Rem}_{1,k}$. Note that

$$||(\hat{\Psi}_k(\mathbb{P}_T) - \psi_0)/\sqrt{g_0}||_{0,S}^2 = \hat{R}_0(\hat{\Psi}_k, \psi_0)$$

Since g_0 and $\hat{g}(\mathbb{P}_T)$ are assumed bounded away from zero (positivity assumption), we can apply the Cauchy–Schwarz inequality to obtain

$$E\text{Rem}_{1,k} \lesssim E||(\hat{g}(\mathbb{P}_T) - g_0)(\hat{\bar{Q}}_0^*(\mathbb{P}_T) - \bar{Q}_0)||_{0,S}\sqrt{R_0(\hat{\Psi}_k, \psi_0)}.$$

We now consider $E\text{Rem}_{2,k}$. By applying the Cauchy–Schwarz inequality we obtain

$$\text{Rem}_{2,k} \lesssim ||\hat{g}(\mathbb{P}_T) - g_0||_{0,S}||\hat{\bar{Q}}_k^*(\mathbb{P}_T) - \hat{\bar{Q}}_0^*(\mathbb{P}_T)||_{0,S}. \tag{28}$$

From the definition of $\hat{\varepsilon}_k$ in the CV-TMLE of $R(\hat{\Psi}_k)$, note that $\hat{\bar{Q}}_k^*$ satisfies the equation

$$E_S \mathbb{P}_V \frac{h}{\hat{g}(\mathbb{P}_T)}\hat{\Psi}_k(\mathbb{P}_T)\{Y - \hat{\bar{Q}}_k^*(\mathbb{P}_T)\} = 0.$$

Applying the same equation for ψ_0 and $\hat{\bar{Q}}_0^*$, and subtracting it from the previous one yields

$$E_S \mathbb{P}_V \frac{h}{\hat{g}(\mathbb{P}_T)}\Big(\hat{\Psi}_k(\mathbb{P}_T)\{Y - \hat{\bar{Q}}_k^*(\mathbb{P}_T)\} - \psi_0\{Y - \hat{\bar{Q}}_0^*(\mathbb{P}_T)\}\Big) = 0,$$

which can be written as

$$E_S \mathbb{P}_V \frac{h}{\hat{g}(\mathbb{P}_T)}\{\hat{\Psi}_k(\mathbb{P}_T) - \psi_0\}\{Y - \hat{\bar{Q}}_k^*(\mathbb{P}_T)\} = E_S \mathbb{P}_V \frac{h}{\hat{g}(\mathbb{P}_T)}\psi_0\{\hat{\bar{Q}}_k^*(\mathbb{P}_T) - \hat{\bar{Q}}_0^*(\mathbb{P}_T)\},$$

which implies

$$E_S P_0 \frac{h}{\hat{g}(\mathbb{P}_T)}\psi_0\{\hat{\bar{Q}}_k^*(\mathbb{P}_T) - \hat{\bar{Q}}_0^*(\mathbb{P}_T)\} = E_S(\mathbb{P}_V - P_0)\frac{h}{\hat{g}(\mathbb{P}_T)}\{\hat{\Psi}_k(\mathbb{P}_T) - \psi_0\}\{Y - \hat{\bar{Q}}_k^*(\mathbb{P}_T)\}$$
$$- E_S(\mathbb{P}_V - P_0)\frac{h}{\hat{g}(\mathbb{P}_T)}\psi_0\{\hat{\bar{Q}}_k^*(\mathbb{P}_T) - \hat{\bar{Q}}_0^*(\mathbb{P}_T)\}$$
$$- E_S P_0 \frac{h}{\hat{g}(\mathbb{P}_T)}\{\hat{\Psi}_k(\mathbb{P}_T) - \psi_0\}\{\hat{\bar{Q}}_k^*(\mathbb{P}_T) - \bar{Q}_0\}.$$

By empirical process theory [31, theorem 2.14.1], noting that the first two terms are empirical processes applied to functions in a class of functions $\mathscr{F} = \{f(k, \varepsilon_k, \varepsilon_0) : k, \varepsilon_k, \varepsilon_0\}$, the expectations of the first two terms are bounded by $(1 + \log M_n)/\sqrt{n}$. The third term is bounded by $\sqrt{R_0(\hat{\Psi}_k, \psi_0)}\|\hat{\bar{Q}}_k^*(\mathbb{P}_T) - \bar{Q}_0\|_{0,S}$. These facts together with eq. (28) yield

$$E\mathrm{Rem}_{2,k} \lesssim E\|\hat{g}(\mathbb{P}_T) - g_0\|_{0,S}\frac{1 + \log M_n}{\sqrt{n}}$$
$$+ E\|(\hat{g}(\mathbb{P}_T) - g_0)\|_{0,S}\|(\hat{\bar{Q}}_k^*(\mathbb{P}_T) - \bar{Q}_0)\|_{0,S}\sqrt{R_0(\hat{\Psi}_k, \psi_0)},$$

Finally, consider the term $\mathrm{Rem}_{3,k}$. We can bound this term by $\max_{k,\varepsilon_k,\varepsilon_0} E_S(\mathbb{P}_V - P_0)D(\mathbb{P}_T, k, \varepsilon_k, \varepsilon_0)$, for

$$D(\mathbb{P}_T, k, \varepsilon_k, \varepsilon_0) = D\{\hat{\bar{Q}}_k^*(\mathbb{P}_T), Q_W(\mathbb{P}_V), \hat{g}(\mathbb{P}_T), \hat{\Psi}_k(\mathbb{P}_T)\} - D\{\hat{\bar{Q}}_k^*(\mathbb{P}_T), Q_W(\mathbb{P}_V), g_0, \hat{\Psi}_k(\mathbb{P}_T)\}$$
$$- D\{\hat{\bar{Q}}_0^*(\mathbb{P}_T), Q_W(\mathbb{P}_V), \hat{g}(\mathbb{P}_T), \psi_0\} + D\{\hat{\bar{Q}}_0^*(\mathbb{P}_T), Q_W(\mathbb{P}_V), g_0, \psi_0\}$$
$$= 2h\frac{\hat{g}(\mathbb{P}_T) - g_0}{\hat{g}(\mathbb{P}_T)g_0}\{\hat{\Psi}_k(\mathbb{P}_T)(Y - \hat{\bar{Q}}_k^*(\mathbb{P}_T)) - \psi_0(Y - \hat{\bar{Q}}_0^*(\mathbb{P}_T))\}$$

Let $F(\mathbb{P}_T)$ be the envelope of the class of functions $\mathscr{F}(\mathbb{P}_T) = \{D(\mathbb{P}_T, k, \varepsilon_k, \varepsilon_0) : k, \varepsilon_k, \varepsilon_0\}$, over which we take the maximum. We have $P_0F(\mathbb{P}_T)^2 \lesssim \|\hat{g}(\mathbb{P}_T) - g_0\|_{0,S}^2$. We will apply the following inequality for empirical processes indexed by a finite class of functions $\mathscr{F}$:

$$E\max_{f \in \mathscr{F}}|(\mathbb{P} - P_0)f| \lesssim \frac{1}{\sqrt{n}}\sqrt{\log(\#\mathscr{F})}\|F\|_2,$$

where F is an envelope of $\mathscr{F}$. Thus, given $\mathbb{P}_T$, we can bound the conditional expectation of $\mathrm{Rem}_{3,k}$ by $(1 + \log M_n)\|\hat{g}(\mathbb{P}_T) - g_0\|_{0,S}/\sqrt{n}$, which results in the following bound for the marginal expectation:

$$E\mathrm{Rem}_{3,k} \lesssim \frac{1 + \log M_n}{\sqrt{n}}E\|\hat{g}(\mathbb{P}_T) - g_0\|_{0,S}.$$

Accumulation of these bounds for the different components of $\mathrm{Rem}_{\hat{k}}$ and $\mathrm{Rem}_{\tilde{k}}$ yields the following inequality:

$$E\hat{R}_0(\hat{\Psi}_{\hat{k}}, \psi_0) \lesssim (1 + 2\delta)E\hat{R}_0(\hat{\Psi}_{\tilde{k}}, \psi_0)$$
$$+ c(M, \delta)\frac{1 + \log(K_n)}{n}$$
$$+ (1 + \delta)E\|(\hat{g}(\mathbb{P}_T) - g_0)(\hat{\bar{Q}}_0^*(\mathbb{P}_T) - \bar{Q}_0)\|_{0,S}\sqrt{E\hat{R}_0(\hat{\Psi}_{\hat{k}}, \psi_0)}$$
$$+ (1 + \delta)E\|\hat{g}(\mathbb{P}_T) - g_0\|_{0,S}E\|\hat{\bar{Q}}_k^*(\mathbb{P}_T) - \bar{Q}_0\|_{0,S}\sqrt{E\hat{R}_0(\hat{\Psi}_{\hat{k}}, \psi_0)}$$
$$+ (1 + \delta)E\|\hat{g}(\mathbb{P}_T) - g_0\|_{0,S}E\|\hat{\bar{Q}}_k^*(\mathbb{P}_T) - \bar{Q}_0\|_{0,S}\sqrt{E\hat{R}_0(\hat{\Psi}_{\tilde{k}}, \psi_0)}$$
$$+ (1 + \delta)\frac{1 + \log M_n}{\sqrt{n}}E\|\hat{g}(\mathbb{P}_T) - g_0\|_{0,S}$$
$$\lesssim (1 + 2\delta)E\hat{R}_0(\hat{\Psi}_{\tilde{k}}, \psi_0)$$
$$+ c(M, \delta)\frac{1 + \log(K_n)}{n}$$
$$+ (1 + \delta)E\|(\hat{g}(\mathbb{P}_T) - g_0)(\hat{\bar{Q}}_0^*(\mathbb{P}_T) - \bar{Q}_0)\|_{0,S}\sqrt{E\hat{R}_0(\hat{\Psi}_{\hat{k}}, \psi_0)}$$
$$+ (1 + \delta)E\|\hat{g}(\mathbb{P}_T) - g_0\|_{0,S}E\|\hat{\bar{Q}}_k^*(\mathbb{P}_T) - \bar{Q}_0\|_{0,S}\sqrt{E\hat{R}_0(\hat{\Psi}_{\hat{k}}, \psi_0)}$$
$$+ (1 + \delta)E\|\hat{g}(\mathbb{P}_T) - g_0\|_{0,S}\frac{1 + \log M(n)}{\sqrt{n}},$$

where $\bar{k}$ is either $\hat{k}$ or $\tilde{k}$, whichever gives the worst bound. This inequality can be written as $x^2 - bx \lesssim c$, for

$$b = (1+\delta)E||\hat{g}(\mathbb{P}_T) - g_0||_{0,S}||\hat{\bar{Q}}_{\bar{k}}^*(\mathbb{P}_T) - \bar{Q}_0||_{0,S}$$
$$+ (1+\delta)E||(\hat{g}(\mathbb{P}_T) - g_0)(\hat{\bar{Q}}_0^*(\mathbb{P}_T) - \bar{Q}_0)||_{0,S}$$

$$c = c(M,\delta)\frac{1 + \log M_n}{\sqrt{n}} + (1+\delta)E||\hat{g}(\mathbb{P}_T) - g_0||_{0,S}\frac{1 + \log M_n}{\sqrt{n}}$$

and can be solved using the quadratic formula as $x \le (b + \sqrt{b^2 + 4c})/2$, which in turn implies $x \le b + \sqrt{c}$, proving Theorem 2. $\square$

References

1. Pearl J. Causality: models, reasoning, and inference. Cambridge: Cambridge University Press, 2000.
2. Rubin DB. Estimating causal effects of treatments in randomized and nonrandomized studies. J Educ Psychol 1974;66 (5):688. Available at: http://www.eric.ed.gov/ERICWebPortal/detail?accno = EJ118470.
3. Mansson R, Sun W, Hennessy S, Joffe MM. On the estimation and use of propensity scores in case-control and case-cohort studies. Am J Epidemiol 2007;166:332–9.
4. Robins J. A new approach to causal inference in mortality studies with sustained exposure periods – application to control of the healthy worker survivor effect. Math Modell 1986;7:1393–512.
5. Rosenbaum P, Rubin D. The central role of the propensity score in observational studies for causal effects. Biometrika 1983;70:41–55.
6. Rubin D. Bayesian inference for causal effects: the role of randomization. Ann Stat 1978;6:34–58.
7. Rubin D. Matched sampling for causal effects. Cambridge, MA: Cambridge University Press, 2006.
8. van der Laan M, Robins J. Unified methods for censored longitudinal data and causality. New York: Springer, 2003.
9. van der Laan M, Rose S. Targeted learning: causal inference for observational and experimental data. New York: Springer, 2011.
10. Bickel P, Klaassen C, Ritov Y, Wellner J. Efficient and adaptive estimation for semiparametric models. Baltimore: Johns Hopkins University Press, 1997.
11. Díaz I, van der Laan M.. Assessing the causal effect of policies: an approach based on stochastic interventions. Technical report, 2012.
12. Díaz I, van der Laan M. Population intervention causal effects based on stochastic interventions. Biometrics 2012;68:541–9. Available at: http://dx.doi.org/10.1111/j.1541-0420.2011.01685.x.
13. Dawid AP, Didelez V. Identifying the consequences of dynamic treatment strategies: a decision-theoretic overview. CoRR 2010;abs/1010.3425.
14. Didelez V, Dawid AP, Geneletti S. Direct and indirect effects of sequential treatments. In Proceedings of the Twenty-Second Conference on Uncertainty in Artificial Intelligence, 138–146. UAI 2006.
15. Korb K, Hope L, Nicholson A, Axnick K. Varieties of causal intervention. In: Zhang C,Guesgen HW, Yeap W-K, editors. PRICAI 2004: trends in artificial intelligence. Lecture Notes in Computer Science, volume 3157. Berlin: Springer, 2004:322–31.
16. Robins JM, Hernan MA, Brumback B. Marginal structural models and causal inference in epidemiology. Epidemiology 2000;11:550–60.
17. Neugebauer R, van der Laan M. Nonparametric causal effects based on marginal structural models. J Stat Plann Inference 2007;137:419–34. Available at: http://www.sciencedirect.com/science/article/pii/S0378375806000334.
18. van der Laan M, Dudoit S.. Unified cross-validation methodology for selection among estimators and a general cross-validated adaptive epsilon-net estimator: finite sample oracle inequalities and examples. Technical report, Division of Biostatistics, University of California, Berkeley, CA, 2003.
19. van der Vaart A. Notes on cross-validation. Technical report, Department of Mathematics, Free University, Amsterdam, 2003.
20. van der Laan M, Dudoit S, Keles S. Asymptotic optimality of likelihood-based cross-validation. Stat Appl Genet Mol Biol 2004;3. Article 4.
21. Dudoit S, van der Laan M. Asymptotics of cross-validated risk estimation in estimator selection and performance assessment. Stat Methodol 2005;2:131–54.
22. van der Laan M, Dudoit S, van der Vaart A. The cross-validated adaptive epsilon-net estimator. Stat Decisions 2006;24:373–95.
23. van der Vaart A, Dudoit S, van der Laan M. Oracle inequalities for multi-fold cross-validation. Stat Decisions 2006;24:351–71.

24. van der Laan M, Polley E, Hubbard A. Super learner. Stat Appl Genet Mol Biol 2007;6(1):1–21.

25. van der Laan MJ, Petersen ML. Targeted learning. In: Zhang C,Ma Y, editors. Ensemble machine learning. New York: Springer, 2012:117–56.

26. Zheng W, van der Laan MJ. Cross-validated targeted minimum-loss-based estimation. In: van der Laan M,Rose S, editors. Targeted learning. New York: Springer Series in Statistics, Springer, 2011:459–74.

27. Wang Y, Bembom O, van der Laan M. Data adaptive estimation of the treatment specific mean. J Stat Plann Inference 2006;137(6):1871–1887.

28. van der Laan M, Rubin D. Targeted maximum likelihood learning. Int J Biostat 2006;2(1). Article 11.

29. van der Vaart A. Semiparameric statistics. In: Bolthausen, Erwin, Edwin Arend Perkins, and Aad W. van der Vaart. Lectures on Probability Theory and Statistics: Ecole D'Ete de Probabilites de Saint-Flour XXIX-1999. Ed. Pierre Bernard. Springer, 2002:331–457.

30. Zheng W, van der Laan MJ. Asymptotic theory for cross-validated targeted maximum likelihood estimation. Division of Biostatistics Working Paper Series, 2010.

31. van der Vaart AW, Wellner JA. Weak convergence and empirical processes. New York: Springer-Verlag, 1996.