The International Journal of Biostatistics

Volume 6, *Issue* 2 2010 *Article* 3

CAUSAL INFERENCE

Targeted Maximum Likelihood Based Causal Inference: Part II

Mark J. van der Laan, *University of California - Berkeley*

Recommended Citation:
van der Laan, Mark J. (2010) "Targeted Maximum Likelihood Based Causal Inference: Part II," *The International Journal of Biostatistics*: Vol. 6: Iss. 2, Article 3.
DOI: 10.2202/1557-4679.1241

Targeted Maximum Likelihood Based Causal Inference: Part II

Mark J. van der Laan

Abstract

In this article, we provide a template for the practical implementation of the targeted maximum likelihood estimator for analyzing causal effects of multiple time point interventions, for which the methodology was developed and presented in Part I. In addition, the application of this template is demonstrated in two important estimation problems: estimation of the effect of individualized treatment rules based on marginal structural models for treatment rules, and the effect of a baseline treatment on survival in a randomized clinical trial in which the time till event is subject to right censoring.

KEYWORDS: causal effect, causal graph, censored data, cross-validation, collaborative double robust, double robust, dynamic treatment regimens, efficient influence curve, estimating function, estimator selection, locally efficient, loss function, marginal structural models for dynamic treatments, maximum likelihood estimation, model selection, path-wise derivative, randomized controlled trials, sieve, super-learning, targeted maximum likelihood estimation

1 Introduction.

In the companion Part I article we developed targeted maximum likelihood estimation for estimation of causal effects of interventions on an outcome of interest for general longitudinal data structures. In this article in Section 2 we start out with translating/summarizing this into a general template for estimation of parameters of the intervention-specific counterfactual distributions, as identified by the so called G-computation formula. In Section 3, we demonstrate the application of this template with a complex longitudinal data structure one encounters in HIV research in order to compare mean clinical outcomes under different individualized rules for starting treatment. In Section 4 we apply the template to estimate a causal effect of a binary treatment on a clinical time till event outcome based on data generated by a randomized controlled trial. We end with a discussion in Section 5.

2 A general template for targeted MLE of parameters of the G-computation formula.

We present a road map for the computation of the targeted MLE and corresponding statistical inference for a target parameter of the distribution of the data

Code data: Represent the data on one unit as a time-ordered data structure

$$O = (L(0), A(0), L(1), A(1), \ldots, L(K), A(K), L(K+1)).$$

It is assumed that $L(t)$ occurs before $A(t)$, and we are interested in effect of interventions on the A-nodes of this graph.

Define target parameter: Let P_0 be the probability distribution of O, and let $\psi_0 = \Psi(P_0)$ be the target parameter of interest. The probability distribution of O factorizes as $p_0 = Q_0 g_0$, where $Q_0 = \prod_{t=0}^{K+1} Q_{0L(t)}$ and $g_0 = \prod_{t=0}^{K} g_{0A(t)}$, $Q_{0L(t)}$ is the conditional distribution of $L(t)$, given $Pa(L(t)) = \bar{L}(t-1), \bar{A}(t-1)$, and $g_{0A(t)}$ is the conditional distribution of $A(t)$, given $Pa(A(t)) = \bar{L}(t), \bar{A}(t-1)$. If it is known that the parent sets of these nodes are smaller, then these smaller parent sets need to be enforced. We will assume that each of these conditional distributions is unspecified. Typically we will have that $\psi_0 = \Psi^F(Q_0)$ is only a parameter of the Q-factor of density p_0 of O. In causal inference most target parameters can be defined as a parameter of a distribution obtained by

intervening on the A nodes in the complete system, which is thereby only a function of Q, i.e., the G-computation formula.

Determine efficient influence curve: In order to carry out the targeted MLE to estimate ψ_0 one will need to know the efficient influence curve $D^*(Q,g)$ for any (Q,g) identifying a distribution of the data. If ψ_0 only depends on P_0 through Q_0, then one can find an influence curve D_{IPCW} of Ψ in the model in which g_0 is known. Such influence curves can often be represented as so called inverse of probability of censoring weighted functions of a full data efficient influence curve (see van der Laan and Robins (2003) for a formal treatment of IPCW-estimating functions). In that case the efficient influence curve can be represented as a projection of such an IPCW-estimating function D_{IPCW} onto the tangent space of Q:

$$D^*(Q,g) = \Pi(D_{IPCW}(Q,g) \mid T_Q),$$

where T_Q is the tangent space of the Q-factor of the density p_0 of O. Our formulas of the clever covariates in the fluctuation function of the various factors in Q will be a direct function of this D_{IPCW}. Since for most target parameters the IPCW-estimating function is well known and easily constructed, this provides us with a straightforward way to obtain the right formulas for the clever covariates needed to define the fluctuation function of the targeted MLE step.

Determine binary factorization of likelihood: Consider a $L(t)$. Suppose $L(t) = (L(t,j) : j = 1,\ldots,n(t))$ consists of $n(t)$ components, which we denote with $L(t,j)$ for different j. Firstly, we determine a particular ordering, allowing us to model

$$Q_{L(t)} = \prod_{j=1}^{n(t)} Q_{L(t,j)},$$

where $Q_{L(t,j)}$ is the conditional probability distribution of $L(t,j)$, given $Pa(L(t,j)) = \bar{L}(t-1), L(t,1), \ldots, L(t,j-1), \bar{A}(t-1)$. It now remains to further factorize $Q_{L(t,j)}$. If $L(t,j)$ is binary, then we do not further factorize $Q_{L(t,j)}$. If $L(t,j)$ is a categorical variable with $n(t,j)$ categories, then assume an ordering of the categories $l = 1,\ldots,n(t,j)$, and factorize $Q_{L(t,j)}$ as

$$Q_{L(t,j)} = \prod_{l=1}^{n(t,j)-1} Q_{L(t,j,l)},$$

DOI: 10.2202/1557-4679.1241

where $Q_{L(t,j,l)}$ is the conditional distribution of the indicator $L(t,j,l)$ of $L(t,j) = l$, given $I(L(t,j) = m)$, $m = 1, \ldots, l-1$, and $Pa(L(t,j))$, where it is assumed that if one of these indicators $I(L(t,j) = m)$ for $m = 1, \ldots, l-1$ equals 1, then $Q_{L(t,j,l)}$ is degenerate at 0. Let $Pa(L(t,j,l))$ denote the parent set for this node $L(t,j,l)$.

If $L(t,j)$ is an ordered variable with $n(t,j)$ values, then we already have an ordering, and factorize $Q_{L(t,j)}$ as

$$Q_{L(t,j)} = \prod_{l=1}^{n(t,j)-1} Q_{L(t,j,l)},$$

where $Q_{L(t,j,l)}$ is the conditional distribuiton of the indicator $L(t,j,l)$ of $L(t,j) = l$, given $I(L(t,j) = m)$, $m = 1, \ldots, l-1$, and $Pa(L(t,j))$, where it is assumed that if one of these indicators $I(L(t,j) = m)$ for $m = 1, \ldots, l-1$ equals 1, then $Q_{L(t,j,l)}$ is degenerate at 0.

Note that the latter $Q_{L(t,j,l)}$ (conditional on the previous $l-1$ indicators all being zero) is identified by a so called hazard probability $Q_{L(t,j,l)}(1 \mid Pa(L(t,j,l))$, i.e., a probability of having the random variable fall at level l, conditional on being larger or equal than level l, and $Pa(L(t,j,l))$.

To conclude, we have the following factorization for $Q_{L(t)}$,

$$Q_{L(t)} = \prod_{j=1}^{n(t)} \prod_{l=1}^{n(t,j)-1} Q_{L(t,j,l)},$$

and thereby the factorization for the Q-factor of the density of O,

$$Q = \prod_{t} \prod_{j=1}^{n(t)} \prod_{l=1}^{n(t,j)-1} Q_{L(t,j,l)}$$

in terms of conditional distributions of binary variables $L(t,j,l)$, conditional on parent nodes $Pa(L(t,j.l))$.

Compute formulas for clever covariates: For each binary variable $L(t,j,k)$, $Pa(L(t,j,k))$ denotes the conditioning random variable:

$$Pa(L(t,j,k)) = \bar{L}(t-1), \bar{A}(t-1), L(t,1), \ldots, L(t,j-1), L(t,j,1), \ldots, L(t,j,k-1),$$

and the following formulas have been provided for general parent sets as well. Suppose that $D = D_{IPCW}$ can be represented as

$$D(Q) = \frac{D_1 + C_1(Q)}{g},\ C_1(Q) \text{ is only function of } O \text{ through } L(0), \bar{A}(K),$$

and D_1 does not depend on Q (it can depend on g).

The efficient influence curve can now be represented as $D^* = \Pi(D \mid T_Q) = D_0 + \sum_{t \geq 1, j, k} D_{tjk}$, where $D_0 = E(D^* \mid L(0))$, and for $t \geq 1$, $D_{tjk} = C_{tjk} \left\{ L(t,j,k) - Q_{L(t,j,k)}(1 \mid Pa(L(t,j,k))) \right\}$, with

$$C_{tjk}(Q,g) = \frac{1}{g(\bar{A}(t-1)|X)} \times \\ \{E_{tjk}(Q)(1, Pa(L(t,j,k))) - E_{tjk}(Q)(0, Pa(L(t,j,k)))\},$$

where we defined, for $\delta \in \{0,1\}$,

$$E_{tjk}(Q)(\delta, Pa(L(t,j,k))) = E_Q \left(\sum_{\bar{a}(t,K)} D_1 \mid L(t,j,k) = \delta, Pa(L(t,j,k)) \right).$$

Here we used short-hand notation for $\sum_{\bar{a}(t,K)} D_1(\bar{A}(t-1), \bar{a}(t,K), \bar{L}(K+1))$, and $\bar{a}(t,K) = (a(t), \ldots, a(K))$.

Given an estimator Q_n and g_n, $C_{tjk}(Q_n, g_n)$ denotes the clever covariate defining the fluctuation function of the estimator $Q_{L(t,j,k),n}$ obtained by adding $\epsilon C_{tjk}(Q_n, g_n)$ on the logit scale.

Define clusters of Q-probabilities that need to considered for pooling: We now need to define an initial estimator Q_n^0 of Q_0. One could estimate each $Q_{L(t,j,k)}$ separately with a machine learning algorithm for each t, j, k. Such an estimator can be considered as one particular candidate, but likelihood/loss function based cross-validation needs to be employed in order to evaluate this type of estimator relative to other estimators. Overall, smoothing in time t and/or category l is often very sensible and will improve the overall performance of the estimator: i.e. it will typically reduce the variance significantly at relatively minor loss in bias. Therefore, the user should define clusters of grid points in the (t,j,k)-grid for which the estimators of the corresponding $Q_{L(t,j,k)}$ need to be considered for being pooled: the particular machine learning algorithm applied to this cluster might still decide to not smooth in certain time-points or levels, but it will be guided by cross-validated risk using a specified loss function. Let $\tau_1, \ldots, \tau_L$ be such clusters. For example $\tau_1 = \{(t,1,k) : t = 1, \ldots, K, k = 1, \ldots, n(t,1) - 1\}$ might indicate that component indicated by $j = 1$ needs to be considered for smoothing in both time t across all time points and in its level k. So each cluster typically represents a particular ordered variable (e.g., CD4 count) and it might state that the estimation procedure needs to respect the fact

DOI: 10.2202/1557-4679.1241

that the hazard of this variable at different levels k and different time points might need to be smoothed across time and level. For a categorical variable, smoothing in the categories will make no sense and would thus be made clear by the definition of the cluster for that categorical variable, but that cluster might still suggest smoothing over time t.

Apply loss-based super learner to repeated measures data set to estimate each cluster of Q-probabilities:

For each cluster τ_l, we create a pooled data set which has for each unit as many rows as there are grid points in the cluster which will be used to fit all the conditional probabilities $Q_{L(t,j,k)}$ with $(t,j,k) \in \tau_l$. So we create a data set which has as columns a time stamp t, a j-stamp, and a level stamp l, an outcome column for the binary $L(t,j,k)$, various covariates extracted from $Pa(L(t,j,k))$, across all points $(t,j,k) \in \tau_l$, each grid point representing a line of data, thereby generating repeated measures type data for each unit of observation. Regarding extraction of covariates from $Pa(L(t,j,k))$, one needs to make sure that these covariates have the same meaning across these different grid points. So this step involves defining a list of extractions the histories $Pa(L(t,j,k))$, such as the most recent in time measurements on the particular variable indicated by j. If this is not possible for (say early time points), then that is an indications that some of these (t,j,k) should not have been included in the cluster and might thus need to be fitted separately.

Given this definition of a repeated measures data set corresponding with cluster τ_l, we can now apply the super learner or any other machine learning algorithm to fit the regression of the binary $L(t,j,k)$ onto (t,j,k) and these covariate extractions from $Pa(L(t,j,k))$, across (t,j,k). This requires a choice of loss function $L(Q)(O)$. One possibility is the log-likelihood loss function $-\sum_{(t,j,k)\in\tau_l} \log Q_{t,j,k}(O)$. We can use a potentially more targeted loss function given by the repeated measures squared error loss function $L(Q_{t,j,k} : (t,j,k) \in \tau_l)(O)$ defined as

$$\sum_{(t,j,k)\in\tau_l} w(t,j,k)R(t,j,k) \left\{L(t,j,k) - Q_{L(t,j,k)}(1 \mid Pa(L(t,j,k)))\right\}^2,$$

for some weight function $w(t,j,k)$. Here $R(t,j,k)$ denotes an indicator of $L(t,j,k)$ being at "risk" of changing value. If $R(t,j,k) = 0$, then $Q_{L(t,j,k)}(1 \mid Pa(L(t,j,k)))$ is either known to be zero or one, so that this loss function will only evaluate the conditional probability of $L(t,j,k)$, given its parents and given that it is at risk of changing.

As discussed previously, a particular weight function $w(t,j,k)$ that makes the risk of the loss function close to variance of efficient influence curve, and thereby targets the super learner fit towards the parameter of interest, is given by the square of the clever covariate:

$$w(t,j,k) = C(t,j,k)^2.$$

Since this weight depends on Q, g itself, this weighted super learner would require a two stage procedure, first an unweighted super learner (or other machine learning algorithm) to estimate Q_0, and a subsequent weighted super-learner using the first stage estimator to estimate the clever covariates, and thereby the weights.

Estimate treatment/censoring mechanism: The likelihood for g can also be factorized in binary conditional probability distributions, and will typically involve fewer binary conditional probability distributions. One can use log-likelihood based machine learning (e.g., super-learning) to estimate g_0. Regarding the choice of loss function for g_0, one needs to realize that the estimator of g_0 is only used to estimate the clever covariates, and that fact might guide the choice of loss function for g_0 so that the resulting estimator of g_0 is well suited for estimation of the clever covariates.

Targeted MLE algorithm at given initial estimator: Suppose that we are given an initial estimator Q_n of Q_0, described above, and an estimator g_n of g_0.

Define the fluctuation of the initial estimator Q_n:

$$\text{logit}Q_{L(t,j,k),n}(\epsilon) = \text{logit}Q_{L(t,j,k),n} + \epsilon C_{tjk}(Q_n, g_n).$$

We now have a variety of possibilities regarding estimation of ϵ depending on how much we want to smooth the estimator of ϵ across (t,j,k).

Firstly, we could create a single pooled repeated measures data set in which each unit contributes a line of data for each t,j,k. One now creates columns for the time stamp t, the variable indicator j, the category/level indicator k, the outcome $L(t,j,k)$, the offset $Q_{L(t,j,k),n}$, and the clever covariate C_{tjk}. One could now fit ϵ with logistic regression using the offset command, regressing the binary indicator $L(t,j,k)$ onto the clever covariate C_{tjk}, thereby obtaining a single estimator ϵ_n for each t,j,k. One now updates the initial estimator $Q_n^1 = Q_n^0(\epsilon_n)$, and this process

DOI: 10.2202/1557-4679.1241

is iterated till convergence. In this way, the targeted bias reduction is established with minimal extra fitting, and therefore this might be the preferred method relative to the alternatives considered below.

Alternatively, one could create an extra column that labels the cluster τ_l, and, for each l, one runs the same iterative algorithm as above but now only applied to the repeated measures data set corresponding with the $(t, j, k) \in \tau_l$, i.e., for which this cluster label column equals l. In this case, one obtains a separate estimator ϵ_{nl} for each cluster $l = 1, \ldots, L$, and this ϵ_{nl} is used for each $(t, j, k) \in \tau_l$. Note that in this approach one uses the same pooling to fit ϵ as was used in the initial estimator.

Finally, consider the ordering based on $L(t_1, j_1, k_1) < L(t_2, j_2, k_2)$ if and only if either $t_1 < t_2$, or, if $t_1 = t_2$, $j_1 < j_2$, or, if $t_1 = t_2, j_1 = j_2$, $k_1 < k_2$. Using this ordering, define disjoint and complementary clusters of points that represent an interval in the ordering: thus the intervals cover all points in the (t, j, k)-grid.

So, using this ordering any (t, j, k) can now be denoted with an integer $V(t, j, k) \in \{1, \ldots, N\}$, and a cluster now has to be of the form $[a, b]$ representing all integers between a and b, including a, b. A typical cluster will now only run over the integers corresponding with the k values for a particular variable indicated by t, j. For each interval one now runs the same iterative algorithm as above but now only applied to the repeated measures data set corresponding with the (t, j, k) with $V(t, j, k)$ in the interval. This iterative targeted ML algorithm thus only updates the $Q_{L(t,j,k),n}$ for $V(t, j, k)$ in the interval. However, we run these interval-specific targeted MLE algorithms *sequentially* starting with the last interval in the ordering. After having run the last interval algorithm thereby updating the $Q_{L(t,j,k),n}$ at the end of the ordering, we update these $Q_{L(t,j,k),n}$, and run the next interval (going backwards!) with the updated Q_n. In this second interval targeted ML algorithm one updates the $Q_{L(t,j,k),n}$ corresponding with the second interval while fixing the already obtained fits from the first interval. After having run this second interval algorithm and having updated the corresponding $Q_{L(t,j,k),n}$, one runs the targeted ML algorithm for the third interval (going backwards) updating the $Q_{L(t,j,k),n}$ corresponding with this third interval, while fixing the already obtained fits of the first and second interval. One iterates this updating process till one arrives at the first interval, at which time the algorithm is finished and the updated Q_n^* is complete.

This algorithm uses the fact that the clever covariates used in an interval

only depend on the $Q_{L(t,j,k),n}$ with $V(t,j,k)$ in the interval and to the right of that interval. As a consequence, an update of an interval does not affect the targeted ML algorithm for the intervals to the right of that interval. Thus, by first updating the $Q_{L(t,j,k),n}$ at the end of the ordering and moving backwards in the ordering we can finish the algorithm in one round. In other words, we exploit the monotonicity property of the clever covariates as presented in Theorem 3 in the companiion Part I article, that allowed a closed form backwards targeted MLE algorithm (converging in one step) if one uses a separate ϵ for each factor $Q_{L(t,j,k),n}$.

The Collaborative Targeted MLE: The targeted ML update of the initial super-learning fit Q_n is a function of the clever covariates and thereby depends on the choice of treatment/censoring mechanism estimator. Different choices of treatment/censoring mechanism estimator result in different clever covariates sets $(C_{tjk} : t, j, k)$ and thereby result in different increases in likelihood fits due to targeted MLE update.

As in van der Laan and Gruber (2009), we suggest to use log-likelihood-based cross-validation to select among a sequence of targeted maximum likelihood estimators using increasingly nonparametric estimators of the treatment mechanism to estimate the clever covariates, thereby fine-tuning the depth of bias-reduction pursued. For details about such procedures we refer to van der Laan and Gruber (2009), including the fact that the resulting targeted maximum likelihood estimator is now collaborative double robust. Collaborative double robustness means that the targeted maximum likelihood estimator is consistent if one uses an estimator g_n that converges to a true censoring mechanism that correctly adjusts the covariates that explain (i.e. increase the likelihood fit relative to Q_n) the residual bias $Q - Q_0$, where Q denotes the limit of the initial estimator Q_n. That is, covariates that are not helpful in explaining residual bias $Q_n - Q_0$ do not need to be adjusted for in the censoring mechanism inputted in the clever covariate/least favorable model. One particular collaborative targeted MLE algorithm presented in van der Laan and Gruber (2009) corresponds with using a greedy forward selection building of the treatment/censoring mechanism (adding one covariate at the time) based on the penalized log-likelihood of the corresponding targeted MLE, and choosing the size of the g_0-fit (i.e. number of steps in forward selection algorithm) with penalized log-likelihood based cross-validation, using a penalty to stabilize the procedure in sparse data situations in which clever covariates can reach large/outlier values.

DOI: 10.2202/1557-4679.1241

In order to save computer time one could decide to not cross-validate the initial estimator and also not cross-validate the treatment mechanism estimators in this cross-validated risk of this loss function. So the initial estimator is treated as a fixed off-set, and the clever covariates indexed by different treatment mechanism estimators are treated as given as well. The cross-validation thus concerns running the second stage targeted MLE algorithm described above on a training set, while fixing the treatment mechanism estimator (and thereby the clever covariate) and the offset-initial estimator at their fit based on the whole sample. In van der Laan and Gruber (2009) we point out that using a cross-validated initial estimator as off-set, in the sense that O_i is coupled with an initial estimator using the training sample excluding O_i, $i = 1, \ldots, n$, can be important to obtain the wished bias reduction with the targeted MLE in the case that the initial estimator is an overfit.

In addition, to save computer time, when carrying out this building and selection among different targeted maximum likelihood estimators we suggest that one might replace the candidate targeted maximum likelihood estimators with one-step (or few steps) targeted maximum likelihood estimators only carrying out one ϵ_n-updating step. Once a targeted maximum likelihood estimator is selected, it is fully iterated till convergence, and the latter true targeted maximum likelihood estimator is the reported estimator.

Evaluation of target parameter of targeted MLE: Above we defined a template for the targeted MLE $P_n \rightarrow Q_n^*(P_n)$. One now evaluates the target $\Psi^F(Q_n^*(P_n))$ to obtain the wished estimator of $\psi_0 = \Psi^F(Q_0)$.

Statistical Inference: Let's first consider the case in which we are willing to assume that the treatment/censoring mechanism estimator converges to the true g_0. Under this assumption, we can carry out influence curve based inference. In order to carry out statistical inference we can use the fact that

$$P_n D^*(Q_n^*(P_n), g_n(P_n)) = 0,$$

where $Q_n^*(P_n)$ is the targeted MLE. If D^* can be represented as an estimating function $D^*(\psi, Q, g)$ in ψ, then this corresponds with stating that $\psi_n^* = \Psi(Q_n^*)$ solves the estimating equation

$$P_n D^*(\psi_n^*, Q_n^*, g_n) = 0,$$

and statistical inference can now be based on the double robustness of the estimating function $P_0 D^*(\psi_0, Q, g) = 0$ if either $Q = Q_0$ or $g = g_0$.

In particular, under the assumption that g_n converges to g_0, asymptotically conservative first order statistical inference can be based on the influence curve $D^*(\psi, Q, g_0)$ and corresponding confidence intervals $^*_n \pm 1.96\sigma_n/\sqrt{n}$, where

$$\sigma_n^2 = \frac{1}{n}\sum_{i=1}^{n} D^*(\psi_n^*, Q_n^*, g_n)^2(O_i)$$

is an estimate of the variance of the influence curve. This follows from the fact that under regularity conditions, ψ_n^* is asymptotically linear with influence curve $D^*(\psi_0, Q, g_0)$ minus its projection on the tangent space of the model of g_n^*, where Q is the possibly miss-specified limit of Q_n^* (see van der Laan and Robins (2003)).

Collaborative double robust statistical inference: If we want to rely on double robustness, either due to high dimension of the treatment mechanism, or due to sparsity (w.r.t. target) so that the targeted maximum likelihood step is unstable, then we recommend the use of the collaborative targeted MLE mentioned above, which also involves the selection among targeted maximum likelihood estimators indexed by different candidate estimators of g_n and thereby different targeted MLE steps, where this choice is based on how much the clever covariates improve the fit of the log-likelihood (or loss-function specific risk) of the corresponding targeted maximum likelihood estimator. As shown in the Appendix of van der Laan and Gruber (2009), using this approach, one can still use the influence curve $D^*(Q, g)$ for statistical inference, with Q, g denoting the limits of Q_n^* and g_n, and one now relies on collaborative double robustness stating that g_n needs to converge to a true conditional distribution $g_0(Q)$, indexed by the limit Q of Q_n^*, that adjusts correctly for the residual bias due to misspecification of Q_n^* w.r.t Q_0. For details, we refer to van der Laan and Gruber (2009).

Off course, one can also use the bootstrap for statistical inference, which might provide better finite sample assessment of uncertainty.

3 Application to marginal structural model for realistic individualized rules.

Marginal structural models for a user supplied set of dynamic treatment regimens were developed and proposed in van der Laan (2006), van der Laan

DOI: 10.2202/1557-4679.1241

and Petersen (2007) and, simultaneously and independently, in Robins et al. (2008). van der Laan and Petersen (2007) also includes a data analysis application of these models to assess the mean outcome under a rule that switches treatment when CD4 count drops below a cut-off, and the optimal cut-off is estimated as well. Another practical illustration in sequentially randomized trials of these marginal structural models for realistic individualized treatment rules is presented in Bembom and van der Laan (2007). In this section we follow the approach of non-parametrically defining the causal parameters making use of marginal structural working model, as presented and advertised in R. Neugebauer (2007).

For the sake of concreteness, let's consider the "When to start treatment" question in HIV research. Let time 0 denote the time at which the patient enrolls in the study. At this time, various measurements are made, including baseline CD4 count and viral load. Subsequently each patient is monitored at various times on a possibly fine discrete time scale (e.g., the time unit might represent a week), and followed up till end of follow up or death. Certain time-dependent variables might be recorded at regularly spaced monitoring times such as a viral load and CD4 count, or the virus might be sequenced. Other indicator variables of particular events such as life-status indicator, heart-attack indicator, cancer occurrence indicator, infection indicator, might be observed at irregularly spaced monotoring times. In that case, if such an event occurs, it is recorded, and one also knows at each time, if it has already occurred (i.e., if it has not been recorded, then it did not happen). At such intermediate events, certain time-dependent variables might be recorded, beyond the variables recorded at the planned monitoring times. We will refer to any such time as a monitoring time, but one will have to specify different types of monitoring times. Different types of monitoring times might result in the recording of different variables. Beyond these different type of monitoring times and the corresponding data collection at these time points, there is a time till death and time till right-censoring, both marking the end of follow up. The right-censoring event could be indexed by different types, such as right-censoring by the end of study or by a medical doctor's decision. The indicator process which jumps at the start of treatment is also observed. We are concerned with using n such independently and identically distributed longitudinal data structures to compare the efficacy of different treatment strategies. For the sake of illustration we will focus on targeted maximum likelihood estimation of the effect of different rules for when to start a patient on antiretroviral therapy.

The organization of this section is as follows. To start with we will discuss the format of the data on one patient, and the formulation of the likelihood of such a longitudinal data structure factorized according to the statistical

graph defined by chronological time ordering. Subsequently, we will define various causal effects of "when to start rules" as parameters of interventions on this likelihood defined by the G-computation formula, including the unknown regression parameters in a marginal structural model for a family of realistic individualized "when to start treatment rules". We will then define a first stage super-learning maximum likelihood (or other loss-function based) estimator, and corresponding targeted maximum likelihood estimator of these unknown parameters.

3.1 The graph-factorized likelihood of the experimental data structure.

The data on a subject will involve various lines of data, each line corresponding with a monitoring time (e.g., corresponding with an intermediate event time such as an infection/heart attack etc), as indicated by a time-stamp column, and a final line corresponding with the follow up time till time at analysis or till death or another event marking end of follow up. At each monitoring time, updates on a collection of time-dependent variables will be recorded, while the columns coding time-independent covariates remain (obviously) constant. The time-dependent variables that are not measured at that monitoring time are either coded as missing or one imputes (e.g forward imputation) a value, and one creates an imputation indicator indicating if this measurement was imputed or an actual update. The final line of data provides the final time stamp, and either the censoring indicator column might jump to the value 1, and the type of censoring is coded, or, if this final time-stamp is time till death, then the life-status indicator jumps to the value 1.

Formally, we could code such a data set as follows. Let $W_j(t)$ be a time-dependent variable, defined as constant in between monitoring times, using forward imputation, and missing if no previous measurement is available, $j = 1, \ldots, J$. Let $\Delta_j(t)$ be an imputation indicator corresponding to $W_j(t)$, $j = 1, \ldots, J$. Let $N_j(t)$ be a counting process such as $R_j(t) = I(T_j \leq t)$ for time till event variables T_j, which are observed at all time points t, $j = 1, \ldots, J_1$. Updates of the variables $W(t)$ only occur at a time point for which at least one of the $dN_j(t) = 1$, i..e, at a time t at which one of the counting processes N_j jump. For example, if there is regular monitoring at fixed time points, beyond monitoring at random times, then one of the counting processes codes the regular monitoring times, while others code the random monitoring times. We can now define a process $O(t) = ((W_j(t), \Delta_j(t) : j), (N_j(t) : j))$, and we truncate this process at the minimum of end of follow up time and a maximal

DOI: 10.2202/1557-4679.1241

follow up time τ. We will suppress the missing indicators $\Delta_j(t)$ in notation below, and just refer to $W_j(t)$. The observed data structure for one unit is now given by $\bar{O} = (\bar{O}(t) : t \leq \tau)$. Note that in this file $\bar{O}$, a jump of N_j results in a new line of data.

If at a monitoring time stamp various measurements are made, then there might be additional time ordering within the time-stamp. For example, the time-stamp might correspond with a day, and it might be known that the medical doctor made a treatment decision that day based on a variety of newly recorded measures, and possibly additional measures were obtained that same day, after the treatment had been given. This kind of additional time ordering information is an important component of the causal graph necessary to obtain a valid G-computation formula respecting the time-ordering. In this case, for each t, $W(t)$ should be accompanied with a time-ordering vector making clear which groups of variables were measured at the same time and how these groups are ordered in time.

The factorization of likelihood respecting the time-ordering. The likelihood of this longitudinal data structure $\bar{O}$ could be written as a product over time t starting at time 0. One starts with drawing the baseline data at time 0. Subsequently, at each time point one draws from the conditional intensities of N_j, given past, including the planned monitoring intensities, time till event intensities, and end of follow up intensities such as time till death or time till right censoring intensities. As long as none of these events occur, one proceeds to the next time interval.

When one of these events occurs, then the possible additional random variables corresponding with that type of event are drawn, possibly sequentially according to an additional time ordering, thus always following the known time-ordering. One proceeds generating random variables like this, moving along in discrete time, till either death happens, another final event happens which marks the end of follow up, or the end of study is reached.

Such a likelihood $p_0(O) = \prod_t p_0(O(t) \mid \bar{O}(t-))$, can be factorized as

$$\begin{aligned} p_0(O) &= \prod_t \prod_j \lambda_j(t \mid Pa(dN_j(t)))^{dN_j(t)}(1 - \lambda_j(t \mid Pa(dN_j(t))))^{1-dN_j(t)} \\ & \quad P(W(t) \mid Pa(W(t))), \end{aligned}$$

where we used an ordering for the components $dN_j(t)$ so that $Pa(dN_j(t)) = \bar{O}(t-), dN_1(t), \ldots, dN_{j-1}(t)$, and $Pa(W(t)) = \bar{O}(t-), dN(t)$. If $dN(t) = 0$ (i.e., each $dN_j(t) = 0$), then no events and monitoring occurs at time t so that $P(W(t) \mid Pa(W(t)))$ is degenerate.

Let $W(t) = (W^-(t), A(t), W^+(t))$, where $A(t)$ is a treatment decision at time t, $W^-(t)$ are variables recorded before $A(t)$, and $W^+(t)$ are variables recorded after t. This particular time-ordering at time t can depend on the realization of $dN(t)$: i.e. for different types of events, different variables might be collected, and for each such group of variables that recorded, we need to know what variables are pre-treatment and post-treatment decision at time t.

So we have, respecting the time-ordering,

$$\begin{aligned} P(W(t) \mid Pa(W(t))) &= P(W^-(t) \mid Pa(W^-(t)))P(A(t) \mid Pa(A(t))) \\ & \quad P(W^+(t) \mid Pa(W^+(t))). \end{aligned}$$

The conditional probability distributions $P_{W^-(t)}$ and $P_{W^+(t)}$ can be factorized in terms of products of conditional densities of particular variables, and these conditional densities can be further factorized in terms of hazards of binary events. as we did in the previous sections.

We conclude that we have the following factorization of the likelihood in terms of conditional distribution of binary events, given the parent sets, factored according to the known time-ordering and user supplied orderings in the case that there is no time ordering provided,

$$p_0(O) = \prod_t \prod_j P_{dN_j(t)} \prod_{jl} P_{W^-_{jl}(t)} \prod_{jl} P_{W^+_{jl}(t)} \prod_t g_{(A(t)}.$$

Incorporation of causal graph knowledge beyond time-ordering. Causal graph knowledge can be incorporated by reducing the parent set of the nodes, and/or enforcing orderings of variables beyond the one implied by the time ordering. The time-ordering always has to be satisfied by any causal graph, so that any additional causal graph information provides additional ordering of all measured variables, and possible reductions of parent sets.

3.2 The likelihood factored in intervention mechanism and relevant factor.

We can refer to $\prod_t g_{A(t)}$, with $g_{A(t)}(A(t) \mid Pa(A(t)))$, as the treatment mechanism, and it provides us also with a likelihood criterion that can be used to generate maximum likelihood estimators of this treatment mechanism. A special case of a process $A(t)$ is of the form $A(t) = I(S \leq t)$, where S is the time at which an antiretroviral therapy is started. In this case, $A(t)$ only jumps once.

If $dN(t) = 0$, then no treatment decisions are made at time t so that $g_{A(t)}$ is degenerate at such a parent set realization. If $dN(t) \neq 0$, then the

DOI: 10.2202/1557-4679.1241

treatment assignment mechanism can still depend heavily on the type of event that occurred: i.e., which $dN_j(t) = 1$. In particular, it might be the case that only for one type of event, treatment decisions are made, while for any of the other events coded by the counting processes $N_j(t)$, $A(t)$ will not be assigned/changed, so that it will still follow a degenerate probability distribution. Note that this treatment mechanism product over all times t reduces to a product of treatment probabilities at the finite (but random) time points for which $dN(t) \neq 0$ and for which there is experimentation in $A(t)$. For example, if N_1 jumps at the time point at which doctors generate measurements and make treatment decisions, and all other events as coded by N_j, $j = 2, \ldots, J$ do not generate treatment decisions with probability 1, then we have

$$g = \prod_t g_{A(t)} = \prod_{t:dN_1(t)=1} g_{A(t)}(\cdot \mid Pa(A(t)), dN_1(t) = 1).$$

Interventions on treatment. Interventions on this treatment assignment mechanism define interesting causal effects. However, it needs to be understood that such an intervention does not control the actual time points at which these treatment decisions can be enforced: i.e.. the time points at which $A(t)$ changes value are kept uncontrolled under an intervention on this treatment mechanism, and, might be differential depending on the actual treatment intervention.

Interventions on treatment and timing of treatment changes. Therefore it is also of interest to intervene on both the treatment assignment mechanism as well as the monitoring mechanism that generates the monitoring times at which the treatment decisions can be enforced. For that purpose, suppose that N_1 is a counting process that jumps at times at which treatment decisions are made. It would now also be of interest to both intervene on the monitoring process N_1 as well as on the treatment decisions made at the monitoring times defined by N_1.

In that case the total mechanism defining the "treatment mechanism" is given by

$$g_1 \equiv \prod_t \lambda_{N_1(t)}(t)^{dN_1(t)}(1 - \lambda_{N_1(t)}(t))^{1-dN_1(t)} g_{A(t)}.$$

The first factor concerns the assignment of monitoring times of type 1 and the second factor concerns treatment decisions at such times, but possibly also at monitoring times of different types. As a special case, one might have that treatment decisions are only made at monitoring times of type 1, so that intervening on N_1 and the treatment assignment mechanism at these monitoring

times is an intervention on the complete treatment process. In this case, only $g_{A(t)}(A(t) \mid Pa(A(t)), dN_1(t) = 1)$ needs to be estimated, since conditioning on other realizations of $dN(t)$ makes $g_{A(t)}$ a degenerate distribution.

Right-censoring mechanism. We also define the factor that defines the likelihood for the right-censoring events. For example, if N_2 is the (only) counting process that codes right-censoring events that obstruct the complete observation of the outcome of interest Y, then we define the censoring mechanism as

$$g_2 = \prod_t \lambda_{dN_2(t)}(t)^{dN_2(t)}(1 - \lambda_{dN_2(t)}(t))^{1-dN_2(t)}.$$

Combined set of conditional distributions we intervene upon. We will denote the combined treatment and censoring mechanism with $g = g_1 * g_2$. Here g_1 can be either the treatment mechanism or it can be both the treatment and monitoring mechanism, depending on the scientific question we wish to address.

3.3 Target parameters of G-computation formula: Marginal structural working model for intervention rules.

We will consider causal effects of two interventions. Firstly, we discuss intervening on the treatment assignment process only, and subsequently, we discuss interventions on both treatment and monitoring.

Causal effect of intervention on treatment decisions at uncontrolled monitoring times. Suppose that, in words, we wish to assess the mean outcome Y measured at a fixed time K since baseline under a dynamic treatment rule d_θ that start treatment right after a measured CD4 count falls below θ. Here Y might be defined as the indicator of still being alive at time K, or an absolute level of CD4 count, or a combination of death and CD4 count such as an indicator of death or CD4 is below a critical value.

We now need to decide how we can formally define this parameter as an operation on the time-ordered/causal graph factorized likelihood of the data. Firstly, let's consider the case in which we do not intervene on the monitoring process N_1 that generates the monitoring times at which biomarker data (e.g., CD4 and viral load) is generated and treatment decisions are made. In this case we only intervene on the treatment assignment rule at the monitoring times generated by the true intensity λ_1 of N_1. Such a rule might be that

DOI: 10.2202/1557-4679.1241

if the most recent measured CD4 count is below θ, then the anti-retroviral therapy is started.

Since our outcome Y is subject to right-censoring by some of the other intensities and we are only interested in the effect of the treatment intervention on the uncensored outcome, we also need to intervene on the censoring mechanism. Let A_2 be the censoring process which jumps from zero to one at a time point t in which a subject is right-censored by an event coded by one or more of the counting processes. Let $g = g_1 g_2$ denote the factor of the likelihood that generates the treatment A and right-censoring A_2 events.

To generate the counterfactual data under such a rule d_θ, one would now generate the data according to the likelihood as described above, but at a monitoring time t with $dN_1(t) = 1$ at which a treatment decision is possible, we would now apply the dynamic rule to set the indicator $A(t)$ of starting treatment at time t, and at each time at which censoring can occur, we set the relevant $dN_j(t) = 0$ for all j that code right-censoring events. Note that this intervention would leave the monitoring process random as it is. So, if particular patients are badly monitored, even under a rule d_θ, they might start treatment at a much lower CD4 count than θ due to large time periods in which the patient's CD4 count is not observed.

Causal effect of intervention on monitoring times and treatment decisions. If one is concerned that the monitoring process heavily affects the clinical outcome and one is concerned with extrapolation of the results to a population in which monitoring times are differently distributed, then one might wish to assess the effect of rules that intervene on both the monitoring mechanism, as well as the treatment assignment mechanism. For example, a rule might be that one monitors a patient every θ_1 months and thereby measures the CD4 count and viral load at these times, and that one starts anti-retroviral treatment when the CD4 count measured at that time is below θ_2. To generate the data under such a rule d_{θ_1,θ_2}, one would now generate the data according to the likelihood as described above, except one would not generate monitoring times based on λ_1 till one reaches $t = \theta_1$, at which time one sets the monitoring time at $t = \theta_1$, one draws from the conditional distribution of time-dependent covariates at that time point, conditional on being monitored at that time, and one assigns the when to start treatment decision according to the rule indexed by threshold θ_2. One proceeds over time following the data generation according to the likelihood, the set monitoring times, the set treatment starting rule, and right-censoring set at infinity. It is also of interest to consider an intervention on monitoring that corresponds with randomly

drawing monitoring times from a user supplied monitoring mechanism, but we will not consider this case below.

Identifiability of target parameter. These rules have to be realistic rules in order to make the corresponding counterfactual probability distributions identifiable from the observed data. For example, the intensity λ_1 needs to have support on these regularly spaced monitoring times $k\theta_1$, $k = 1, 2, \ldots$, and the lower the support the harder it will be to reliably estimate the counterfactual distribution for that choice of θ_1. If in the actual study people were regularly followed, say every 3 months, and that variations on the monitoring times were at most one month off, then one would expect a good support for $\theta \in [3 - \delta, 3 + \delta]$ for an appropriately chosen δ. Similarly, the when to start treatment decision rule needs to be supported by the medical doctors that made these decisions in the actual study. For example, if the medical community supports the starting of the antiretroviral therapy at CD4 counts between 200 and 400, in the sense that there is experimentation across that range, then one should choose θ_2 in that range. Finally, the right-censoring intensity should never equal 1, whatever the history, up till the time point K at which the outcome can be measured. Thus, if K is selected too large, then the latter assumption might become practically violated. In addition, if there are events that imply right-censoring at the next time point with probability 1, then one might need to include such events in the definition of the outcome Y, so that the effect of treatment on that Y is still identifiable from the data, and such effects will then need to be honestly interpreted.

These counterfactuals Y_θ indexed by the rule d_θ for the treatment process $\bar{A}$, either only including the treatment decisions or also including the monitoring time process N_1, are now defined by its probability distribution Q_θ define above as an intervention on the the graph-factored likelihood, obtained by excluding the factors $g = g_1 g_2$ that are set by the rule, and setting the values of treatment $\bar{A}$ and right-censoring $\bar{A}_2$ according to the rule d_θ in any of the conditioning events of the other factors of the likelihood.

We can now define the parameter of interest as a projection of EY_θ onto a working model, thereby creating smoothing parameters of the complete response curve $\theta \to EY_\theta$ for which larger data support is available so that it can be estimated using semi-parametric model efficiency theory and methodology. That is, EY_θ is often not path-wise differentiable, while such a summary parameter will be path-wise differentiable.

Marginal structural models for realistic individualized treatment rules. Specifically, let $m_\beta(\theta, V)$ be a working model for the conditional mean of Y_θ, under rule d_θ (controlling treatment and censoring), given a baseline

DOI: 10.2202/1557-4679.1241

(for example) CD4 count V, such as

$$m_\beta(\theta, V) = \beta_0 + \beta_1\theta + \beta_2\theta^2 + \beta_3\theta V + \beta_4\theta^2 V.$$

We now define the target parameter as

$$\Psi(P_0) = \Psi^F(Q_0) = \arg\min_\beta E_{Q_0} \sum_\theta h(\theta, V)(Y_\theta - m_\beta(\theta, V))^2,$$

where we remind the reader that Q_0 denotes the factor of the probability distribution of O: $p_0 = Q_0 g$. Here $h(\theta, V)$ is a user supplied weight function. Thus Ψ is only a parameter of P_0 through its Q_0-factor. We refer to m_β as a marginal structural (working) model for realistic treatment rules (van der Laan and Petersen (2007)).

Evaluation of target parameter. Given such an estimator Q_n or its targeted MLE update Q_n^* defined below, the evaluation of $\Psi(Q_n)$ can be based on Monte-Carlo simulation. One first samples a larger number B of observations $Y_{\theta,b}, V_b$, $b = 1, \ldots, B$ for each possible θ, from the corresponding G-computation formula of the distribution of Y_θ. Then,

$$\Psi^F(Q_n) = \arg\min_\beta \sum_\theta \frac{1}{N} \sum_{b=1}^{B} h(\theta, V_b)(Y_{\theta,b} - m_\beta(\theta, V_b))^2.$$

In other words, we simply replace the expectation E_{Q_0} of a function $f_{\theta,\beta}(Y_\theta, V)$ of Y_θ, V in the definition of $\Psi^F(Q_0)$ by an expectation w.r.t. empirical distribution of the B draws $(Y_{\theta,b}, V_b)$, $b = 1, \ldots, B$.

3.4 Adaptive maximum likelihood estimation: Super learning.

Given an estimator Q_n of Q_0, one obtains the estimator $\Psi(Q_n)$ of $\psi_0 = \Psi^F(Q_0)$. Since the likelihood factor Q factorizes in a product over conditional distributions of binary factors, we can estimate this with loss-based super-learning methodology that can be implemented with standard software tools. Thus one applies super learners for binary predictions, possibly pooled across many of the binary predictions, pooling across time and or across different levels of ordered variables indexing the binary variables. The overall log-likelihood or pooled weighted squared error loss function for Q_0 could be employed for fine-tuning the choice and degree of pooling, only considering sensible pooling strategies. Super-learning could then be based on a library of algorithms

for estimation of the complete Q_0 based on this overall loss function for Q_0. Some of the candidate algorithms in the library of the super learner might involve super-learning itself of binary predictors possibly using different pooling strategies.

To be specific, let's consider modeling the conditional distribution of CD4 count at time t, given its parents. We will model this in terms of conditional binary distributions. Suppose that we can view CD4 as an ordered discrete variable with levels $l = 1, \ldots, L$, possibly defined by the L equally spaced quantiles of the marginal empirical distribution of CD4 counts. Let $Q_{CD4(t)}$ denote the conditional distribution of $CD4$ count at time t, conditional on the parent nodes and that the person is monitored at time t so that the CD4 count process is at risk of changing. We write this conditional probability distribution of $CD4(t)$ in terms of binary conditional distributions

$$Q_{CD4(t)}(CD4(t)) \;=\; \prod_{j=1}^{L} Q_{I(CD4(t)=j)}(1)^{CD4(t)=j} Q_{I(CD4(t)=j)}(0)^{CD4(t)\neq j}.$$

If we define the discrete hazard

$$\lambda_{CD4(t)}(j) \equiv P(CD4(t) = j \mid CD4(t) \geq j, Pa(CD4(t))$$

of CD4 count at time t, then it follows that

$$Q_{CD4(t)}(l) = \left\{ \prod_{j=1}^{l-1} (1 - \lambda_{CD4(t)}(j)) \right\} \lambda_{CD4(t)}(l).$$

The likelihood for this discrete hazard is thus given by

$$L(\lambda_{CD4(t)}) = \prod_{i=1, Mon_i(t)=1}^{n} \prod_{l=1}^{CD4_i(t)-1} (1 - \lambda_{CD4(t),i}(l)) \lambda_{CD4(t),i}(CD4_i(t)).$$

The likelihood of this discrete hazard $\lambda_{CD4} \equiv (\lambda_{CD4(t)}(j) : t, j)$ viewed as a function in both time t and the CD4 count level j is thus given by

$$L(\lambda_{CD4}) = \prod_{t} \prod_{i=1, Mon_i(t)=1}^{n} \left\{ \prod_{l=1}^{CD4_i(t)-1} (1 - \lambda_{CD4(t),i}(l)) \right\} \lambda_{CD4(t),i}(CD4_i(t)).$$

One can now carry out estimation of this nonparametric function λ_{CD4} based on this log-likelihood loss function. In particular, one can apply super-learning based on this loss function.

DOI: 10.2202/1557-4679.1241

We note that if monitoring is really random, then one will have few subjects that have a monitoring time at a given time point t within a fine grid of time points. As a consequence, in that case the t-specific likelihood for $\lambda_{CD4(t)}$ provides too little information for estimation of $\lambda_{CD4(t)}$. Thus, it will be essential to use the combined likelihood pooling across time the time points t provided above.

3.5 Calculation of least favorable model for targeted MLE step.

We now focus our attention on the definition of the fluctuation function required to carry out the targeted MLE step.

Inverse probability of censoring weighted function. Let X represent the collection of action specific counterfactuals controlling the intervention nodes defined by $\bar{A}_1, \bar{A}_2$. We will first define an IPCW-estimating function of O for the parameter EY_θ for a given θ, before presenting the IPCW-function of O for the MSM-parameter Ψ.

We have

$$EY_\theta = E\left(\frac{I(\bar{A}_1 = d_\theta(L), \bar{A}_2(K) = 0)}{g(\bar{A} \mid X)} Y\right),$$

where $g = g_1 g_2$ represents the product of conditional distributions of the intervention nodes $\bar{A}_1$ and $\bar{A}_2$. Thus an IPCW-estimating function of EY_θ is given by (see also van der Laan and Petersen (2007))

$$\frac{I(\bar{A}_1 = d_\theta(L)), \bar{A}_2(K) = 0)}{g(\bar{A} \mid X)} (Y - EY_\theta).$$

By a similar argument, it follows that the IPCW-estimating function of $\Psi(Q_0)$ is given by

$$D_{IPCW}(O) = \sum_\theta h(\theta, V) \frac{d}{d\psi_0} m_{\ 0}(\theta, V) \frac{I(\bar{A_1} = d_\theta(L), \bar{A_2}(K) = 0)}{g(\bar{A} \mid X)} (Y - m_{\psi_0}(\theta, V)).$$

For example, if we also intervene on the monitoring times at which treatment can be changed, then $g(\bar{A} \mid X)$ involves a product over time of the likelihood of monitoring events and a treatment event if monitoring occurred, and no censoring event, always conditioning on the parent sets implied by the graph implied by time ordering and possibly additional causal graph assumptions.

The efficient influence curve and corresponding clever covariates for binary factors. We first note that the IPCW-estimating function $D_{IPCW}(O)$ can be represented as

$\frac{1}{g(\bar{A}|X)} \sum_\theta h(\theta, V) \frac{d}{d\psi_0} m_{\psi_0}(\theta, V) I(\bar{A}_1 = d_\theta(O), \bar{A}_2(K) = 0)(Y - m_{\psi_0}(\theta, V))$
$\equiv \frac{D_1(O)}{g(\bar{A}|X)}$.

We have that the Q_0-factor of the density of the data is represented as

$$Q_0 \quad = \quad Q_{W(0)} \prod_t \prod_j Q_{dN_j(t)} \prod_{jl} Q_{W_{jl}^-(t)} \prod_{jl} Q_{W_{jl}^+(t)}.$$

Here we exclude the conditional distributions of the counting processes corresponding with right-censoring and or monitoring events, depending on how $\bar{A}$ is defined. This Q_0-factor identifies the G-computation formula for the distribution of the data under the individualized interventions d_θ.

Let T_Q be the tangent space of Q_0 at $P_{Q,g}$. This tangent space T_Q can be decomposed orthogonally as:

$$T_Q = T_{Q_{W(0)}} + \sum_{tj} T_{\lambda_j,t} + \sum_{tjl} T_{Q_{tjl}^-} + \sum_{tjl} T_{Q_{tjl}^+},$$

where $T_{Q_{W(0)}}$ is the tangent space of the marginal probability distribution of $W(0)$, $T_{\lambda_j,t}$ is the tangent space of the j-th intensity $\lambda_j(t)$, $T_{Q_{tjl}^-}$ is the tangent space of conditional distribution of $W_{jl}^-(t)$, and $T_{Q_{tjl}^+}$ is the tangent space of conditional distribution of $W_{jl}^+(t)$.

By our general Theorem, we have that the efficient influence curve $D^* = \Pi(D \mid T_Q)$ can be represented as

$$D^* = \Pi(D \mid T_Q) = D_0 + \sum_{tj} D_{t,\lambda_j} + \sum_{tjl} D_{tjl}^- + \sum_{tjl} D_{tjl}^+,$$

where $D_0 = E(D^* \mid W(0))$, and for each other factor, we can represent it as $D_{\lambda_j,t} = C_{tj}(dN_j(t) - \lambda_j(t))$, $D_{tjl}^- = C_{tjl}^-(W_{jl}^-(t) - Q_{tjl}^-(1))$, $D_{tjl}^+ = C_{tjl}^+(W_{jl}^+(t) - Q_{tjl}^+(1))$, where these (clever covariates) $C_{tj}, C_{tjl}^+, C_{tjl}^-$ are defined below. These functions D_{t,λ_j}, D_{tjl}^+ and D_{tjl}^- are zero at corresponding parent histories which deterministically predict the value of the corresponding binary variable $dN_j(t)$, $W_{jl}^+(t)$, $W_{jl}^-(t)$, respectively. For example, if $W_{jl}^-(t)$ is only generated at a monitoring time t generated by N_1, then $\sum_t D_{tjl}^- = \sum_{t:dN_1(t)=1} D_{tjl}^-$ reduces to a sum at the random monitoring times at which $dN_1(t) = 1$. On the other hand, if N_1 is at risk of jumping at any of the time points t, then $\sum_t D_{t,\lambda_1}$ remains a sum over all time points t.

DOI: 10.2202/1557-4679.1241

For each binary variable $dN_j(t)$, $W_{jl}^+(t)$, $W_{jl}^-(t)$ we define $\bar{A}(t,j)$, $\bar{A}^+(t,j,l)$, $\bar{A}^-(t,j,l)$, respectively, as the A-nodes that are included in the parent set of that binary variable. For each binary variable $dN_j(t)$, $W_{jl}^+(t)$, $W_{jl}^-(t)$ we define $\overline{a}(t,j)$, $\overline{a}^+(t,j,l)$, $\overline{a}^-(t,j,l)$ as the future path, corresponding with a complete path $\bar{a}$, starting right after where $\bar{A}(t,j)$, $\bar{A}^+(t,j,l)$, $\bar{A}^-(t,j,l)$ stops.

Suppose that the process A we control includes the monitoring process N_1. In that case, $\bar{A}(t,j) = \bar{A}(t-), dN_1(t), dA_2(t))$ includes all actions up till and including previous time point and the censoring and monitoring event at time t, $\bar{A}^-(t,j,l)$ equals $\bar{A}(t-), dN_1(t), dA_2(t)$ as well, and $\bar{A}^+(t,j,l)$ equals $\bar{A}(t-), dN_1(t), dA_1(t), dA_2(t)$ including now also the treatment decision at time t. If our process A does not control monitoring N_1, then exclude N_1 from the statement in the previous sentence.

The formulas for C_{tj}, C_{tjl}^+, and C_{tjl}^- can now be defined as

$$\begin{aligned}
C_{tj} &= \tfrac{1}{g(\bar{A}(t,j)|X)} \times \{C_{tj}(1, Pa(dN_j(t))) - C_{tj}(0, Pa(dN_j(t)))\} \\
C_{tjl}^+ &= \tfrac{1}{g(\bar{A}(t,j,l)^+|X)} \times \{C_{tj}^+(1, Pa(W_{jl}^+(t))) - C_{tj}^+(0, Pa(W_{jl}^+(t)))\} \\
C_{tjl}^- &= \tfrac{1}{g(\bar{A}(t,j,l)^-|X)} \times \{C_{tj}^-(1, Pa(W_{jl}^-(t))) - C_{tj}^-(0, Pa(W_{jl}^-(t)))\},
\end{aligned}$$

where, for $\delta \in \{0,1\}$,

$$C_{tj}(\delta, Pa(dN_j(t))) = E_Q\left(\sum_{\overline{a}(t,j)} D_1 \mid dN_j(t) = \delta, Pa(dN_j(t))\right),$$

and, similarly, we define the other terms $C_{tj}^+(\delta, Pa(W_{jl}^+(t)))$ and $C_{tj}^-(\delta, Pa(W_{jl}^-(t)))$.

Here we used short-hand notation for $\sum_{\overline{a}(t,j)} D_1(O_{\bar{A}(t,j),\overline{a}(t,j)})$, and similarly for the other two terms. If the parent set implies that there is no experimentation in the node, then this clever covariate is not defined, and is also never needed, as stated above. If one wants to extend the definition, then one could simply define the clever covariate as zero for any parent set in which there is no experimentation in the binary node.

For each binary node, the clever covariate can also be represented as the difference of the conditional expectation of $D_{IPCW} = D_1/g(\bar{A} \mid X)$ given the binary node equals 1, and the conditional expectation of D_{IPCW} given the binary node equals zero, and in both cases one also conditions on the parent set of the binary node. In other words, it is a choice to either integrate out over the future sample paths $\overline{a}$ so that the clever covariate factors in a g and Q-factor, or not. One can evaluate these prediction representations of the clever covariate by Monte-Carlo simulation which involves drawing from all future factors of the density (either from Q only, or, from both Q and g, depending on the representation), starting at which the parent set left-off.

These clever covariates provides us with the least favorable model through Q_n (at $\epsilon = 0$) with fluctuation parameter ϵ whose score at $\epsilon = 0$ equals the efficient influence curve $D^*(Q_n, g_n)$, and defines the corresponding targeted MLE Q_n^* and $\Psi(Q_n^*)$. We will not repeat this definition of the targeted MLE, since it was presented in the previous section and our Part I companion paper.

3.6 Targeted maximum likelihood estimation at degenerate initial estimator of intermediate conditional distribution.

For simplicity, let's consider the case that our initial estimator provides deterministic predictions for any of the intermediate time-dependent covariates, so that the clever covariates for all intermediate factors equals zero. As a consequence, the targeted MLE only involves updating the conditional distribution of the final node Y, given its parents.

Consider the IPCW-estimating function:

$$D_{IPCW}(O) = \sum_{\theta} h(\theta, V) \frac{d}{d\psi_0} m_{\ 0}(\theta, V) \frac{I(A_1 = d_\theta(O), A_2(K) = 0)}{g(\bar{A} \mid X)} (Y - m_{\psi_0}(\theta, V)).$$

Under a degenerate distribution for all intermediate variables, we only have to project this onto the tangent space of the distribution of $L(0)$ and the conditional distribution of Y, given $\bar{L}(K), \bar{A}(K)$. We will now present this efficient influence curve at such a Q, and, for the sake of illustration, we will show that it represents an unbiased estimating function of ψ at a correctly specified g_0, and arbitrarily misspecified Q. The efficient influence curve has now only two components we will denote with $D_1^*(Q, g_0)$ and $D_2^*(Q, g_0)$ respectively. We have

$$D_1(Q, \psi_0) = \sum_{\theta} h(\theta, V) \frac{d}{d\psi_0} m_{\ 0}(\theta, V) (E_Q(Y_\theta \mid L(0)) - m_{\psi_0}(\theta, V))$$

and, using $d_{\theta,0}$ to denote both the treatment rule and the no-censoring intervention,

$$D_2(Q, g_0) = \sum_{\theta} h(\theta, V) \frac{d}{d\psi_0} m_{\ 0}(\theta, V) \frac{I(\bar{A} = d_{\theta,0}(O))}{g(\bar{A} \mid X)} (Y - E_Q(Y \mid \bar{A}(K), \bar{L}(K))).$$

Thus the clever covariate we add to an initial estimator of the conditional distribution of Y, given $\bar{A}(K), \bar{L}(K)$, is given by

$$C(g_0) = \sum_{\theta} h(\theta, V) \frac{d}{d\psi_0} m_{\ 0}(\theta, V) \frac{I(\bar{A} = d_{\theta,0}(O))}{g_0(\bar{A} \mid X)}.$$

DOI: 10.2202/1557-4679.1241

Let Q_n^* be the targeted MLE based on an initial estimator Q_n defined by the empirical distribution $Q_{n,L(0)}$ of $L(0)$ and an initial estimator $Q_{n,Y}$ of Y, given $\bar{A}(K), \bar{L}(K)$, and degenerate distributions $Q_{n,d}$ for all the conditional distributions of intermediate variables/time-dependent confounders. The targeted maximum likelihood estimator solves the efficient influence curve equation $P_n D^*(Q_n^*, g_n, \Psi(Q_n^*)) = 0$.

We now show that indeed, as predicted by the double robustness of the efficient influence curve, $P_0 D^*(Q, g_0, \Psi(Q)) = 0$ implies $\Psi(Q) = \Psi(Q_0)$, showing that the targeted MLE $\Psi(Q_n^*)$ solves an unbiased estimating function in ψ at a correctly specified g_0. Note, by first conditioning on X in $P_0 D_2$,

$$\begin{aligned} P_0 D^*(Q, g_0, \Psi(Q)) &= P_0 D_1(Q, \Psi(Q)) + P_0 D_2(Q, g_0) \\ &= E_{Q_0} \sum_\theta h(\theta, V) \frac{d}{d\psi} m \ (\theta, V)(E_Q(Y_\theta \mid L(0)) - m_\psi(\theta, V)) \\ + E_{Q_0} \sum_\theta h(\theta, V) \frac{d}{d\psi} m \ (\theta, V)(Y_\theta - E_Q(Y \mid \bar{A}(K) = d_\theta(\bar{L}(K), \bar{L}(K))). \end{aligned}$$

Now, note that, under the degenerate distribution Q, we have

$$E_Q(Y \mid \bar{A}(K) = d_\theta(\bar{L}(K), \bar{L}(K))) = E_Q(Y_\theta \mid L(0)).$$

Thus, we have

$$P_0 D^*(Q, g_0, \Psi(Q)) = E_{Q_0} \sum_\theta h(\theta, V) \frac{d}{d\psi} m \ (\theta, V)(Y_\theta - m_{\Psi(Q)}(\theta, V)).$$

By definition of $\Psi(Q_0)$ we have that the latter equation equals zero at ψ_0. Thus, we can conclude that indeed, under a weak identifiability condition on the working model m_ψ, $P_0 D^*(Q, g_0, \Psi(Q)) = 0$ implies $\Psi(Q) = \psi_0 = \Psi(Q_0)$.

This targeted MLE only involves adding a clever covariate to $Q_{n,Y}$ and doing a single step update. This updated distribution, only updating $Q_{n,Y}$, equals the targeted MLE at such an initial Q_n.

Discussion on using degenerate fits to simplify targeted MLE. Even though the degeneracy of the initial estimator results in a simple to compute T-MLE, it is questionable till what degree this should be an issue to consider. Given available software, the actual practical performance will be the driving force in such a decision over time. We can simplify the T-MLE in less dramatic ways, by enforcing the degeneracy for most time-dependent variables, but truly modelling the conditional probability distribution for the most important time-dependent confounders. This might result in a highly efficient T-MLE, while

it still only involves updating the non-degenerate conditional distributions. In particular, such a T-MLE can still be obtained in closed form by using our backwards solving algorithm and using a separate updating step (i.e. variation independent fluctuation parameters for the conditional distributions) for these non-degenerate conditional distributions.

Specifically, in the when to start treatment application, it is well known that CD4 count and viral load are the most important time-dependent confounders. In addition, we consider dynamic rules for when to start the treatment responding to these time-dependent confounders. Thus, in this case, it seems particularly appropriate to estimate the actual conditional distributions of CD4 and viral load at the intermediate monitoring times.

The clever covariates of the CD4 count or viral load involve the evaluation of conditional expectations of the future outcome given the parent nodes. Since these conditional expectations are calculated under a Q-fit that uses a deterministic system for all variables except viral load, CD4, and the final outcome Y, this only involves random generation of the future CD4 counts, viral loads, and final Y. All other nodes are generated deterministically according to a fitted prediction function.

We also note that the degenerate conditional distributions do not need to be factored in terms of binary conditional distributions since these will not be updated anyway. Instead, one could simply predict the mean outcome for any continuous or ordered categorical variable from its parent nodes, and put probability 1 on that predicted value.

3.7 Dimension reduction.

We refer to our subsection on dimension reduction in the Part I article, which shows that one can reduce the dimension of the time-dependent process $L(t)$ to a few univariate time dependent processes, beyond the time-dependent covariates used in the rule d_θ, at little loss of information, while still using the same fit of the g-factor adjusting for all relevant variables.

DOI: 10.2202/1557-4679.1241

4 Application to causal effect of point treatment, allowing for right-censoring and utilizing time-dependent covariates.

In this section, we consider a simplified version of the general data structure covered in previous section. Suppose we observe as unit-specific data structure

$$O = (W, A, A_2(1), L(1), \ldots, A_2(K), L(K), A_2(K+1), Y),$$

where W is baseline covariates, A is a treatment assigned at baseline, $A_2(j) = I(C \leq j)$ is the indicator of being right-censored at time j, C is right-censoring time, $L(j)$ is the biomarker (e.g., CD4 count) measured at time j, using forward imputation if $C \leq j$, and Y is the final outcome of interest, but affected by right-censoring. For example, $Y = I(T \leq K+1)(1 - A_2(K+1))$ is the indicator of a time T till failure at time $K+1$ and not being censored. We assume $L(j)$ is ordered and discrete values with values $m = 1, \ldots, M$. The treatment A could be randomized as in a randomized controlled trial.

Suppose that the model for the distribution P_0 of O is nonparametric, and let $\psi_0 = EY(1) - Y(0)$ be the additive causal effect of the binary treatment, where $Y(a)$ is defined as the random variable with probability distribution defined by the G-computation formula under the intervention $A = a$ and $C = \infty$ (i.e., no censoring), $a \in \{0, 1\}$.

The G-computation formula for this target parameter requires estimation of the marginal distribution of W, the conditional distribution of $L(j)$, given past and not being right-censored, $j = 1, \ldots, K$, and the conditional distribution of Y, given past and not being right-censored. The marginal distribution is estimated with the empirical distribution of W. If Y is binary, the conditional distribution of Y, given past and not being right-censored, is estimated with loss-based super-learning based on the log-likelihood loss function, and, if Y is continuous, we estimate the conditional mean of Y with loss-based super-learning based on the squared error loss function: the efficient influence curve of ψ_0 only depends on the conditional distribution of Y through its conditional mean.

The conditional distribution of $L(j)$, given past, and not being right-censored, is also estimated with loss-based (super) learning using the log-likelihood loss function, but we will first factorize the conditional density as

$$P(L(j) = l \mid \cdot) = \prod_{m \leq l-1} (1 - P(L(j) = m \mid L(j) \geq m, \cdot)) P(L(j) = l \mid L(j) \geq l, \cdot).$$

In other words, we code $L(j)$ as a a vector of binaries $I(L(j) = 1), I(L(j) = 2), \ldots, I(L(j) = M)$, and, factor the likelihood of $L(j)$ accordingly. Thus, it remains to estimate the conditional hazard $P(L(j) = m \mid L(j) \geq m, Pa(L(j)))$. We could estimate this with super-learning smoothing in both time j and level m.

The G-computation formula for the counterfactual distribution of $W, \bar{L}, Y$ under intervention $A = a$ and no-censoring $\bar{A}_2 = 0$ is given by

$$\begin{array}{l} P_a(W, L(1), \ldots, L(K), Y) = \\ Q_W(W) \prod_{j=1}^{K} \prod_m Q_{L(j,m)}(L(j,m) \mid Pa(L(j,m), A = a, \bar{A}_2(j) = 0) \\ Q_Y(Y \mid Pa(Y), A = a, \bar{A}_2(K) = 0), \end{array}$$

where the conditional distribution of $L(j,m) = I(L(j) = m)$, given its parents $(L(j,1), \ldots, L(j,m-1)), Pa(L(j))$ is degenerate if one of the indicators $L(j,l)$ with $l \leq m-1$ is already equal to 1.

The targeted maximum likelihood step now involves adding clever covariates (on the logit-scale) to the logistic regression fits of the conditional distributions of $L(j,m)$, $j = 1, \ldots, K$, $m = 1, \ldots, M$, and the logistic or normal error (i.e., least squares, if Y is continuous) regression fit of Y. These clever covariates for the conditional distribution of the binary $L(j,m)$ are

$$\begin{array}{rcl} C_{jm}(Q,g) & = & E_{Q,g}(D_{IPCW} \mid L(j,m) = 1, Pa(L(j,m)), \bar{A}_2(j) = 0) \\ & & -E_{Q,g}(D_{IPCW} \mid L(j,m) = 0, Pa(L(j,m)), \bar{A}_2(j) = 0), \end{array}$$

where

$$D_{IPCW}(O) = Y \left\{ \frac{I(A = 1, \bar{A}_2 = 0)}{g(A, \bar{A}_2 \mid X)} - \frac{I(A = 0, \bar{A}_2 = 0)}{g(A, \bar{A}_2 \mid X)} \right\}.$$

So, calculation of the clever covariates requires, for each subject i, for each time j with $C_i > j$, and each m with $L_i(j) \geq m$, Monte-Carlo simulation to evaluate the conditional mean of D_{IPCW}, conditional on $Pa_i(L_i(j,m))$ for which $L_i(j) \geq m$ and $C_i > j$. This corresponds with imputing a $Y(1)$ and $Y(0)$, and thereby a $Y(1) - Y(0)$, for each subject, based on history of that subject at time j, across j.

The clever covariate to fluctuate the conditional distribution of Y is given by

$$C_Y = \left\{ \frac{I(A = 1, \bar{A}_2 = 0)}{g(A, \bar{A}_2 \mid X)} - \frac{I(A = 0, \bar{A}_2 = 0)}{g(A, \bar{A}_2 \mid X)} \right\}.$$

If Y is binary, the ϵC_Y is added on the logistic scale, and if Y is continuous, one adds ϵC_Y to the fitted conditional mean of Y. The iterative targeted

DOI: 10.2202/1557-4679.1241

maximum likelihood algorithm can now be applied to obtain the targeted maximum likelihood estimator Q_n^* and corresponding $\Psi(Q_n^*)$.

If one assumes that g_n converges to g_0, faster than Q_n^* converges to Q_0 (e.g Q_n^* is inconsistent, or converges at slow rate), then it makes sense to use as loss function for Q $L(Q) = D^*(Q, g_n)^2$. Thus, one would use loss-function based cross-validation to select among different targeted maximum likelihood estimators Q_n^* indexed by different initial estimators. In this manner, one is guaranteed to asymptotically select the targeted maximum likelihood estimator that results in the most efficient estimator of ψ_0.

Statistical inference can be based on the norrmal distribuition approximation of $\sqrt{n}(\psi_n^* - \psi_0)$, given by $N(0, \sigma_n^2)$, where σ_n^2 is an estimate of the variance of $D^*(Q_n^*, g_n)$.

5 Discussion.

As mentioned in our template, in van der Laan and Gruber (2009) we present a variety of proposals for generating a sequence of targeted MLE's Q_j^* coupled with a treatment mechanism estimator g_j, each defined as the result of the targeted MLE algorithm that maps an initial estimator Q_j and treatment mechanism estimator g_j into an targeted MLE update Q_j^*, indexed by treatment/censoring mechanism estimators g_j that are increasingly nonparametric in j, and "initial" Q_j that themselves might represent a targeted MLE update of a previous initial estimator.

Such a sequence of candidate targeted MLE's is constructed to be increasing in the empirical risk of a loss function for Q_0 (e.g., log-likelihood), and corresponds with increasing levels of targeted bias reduction (since the later ones use an estimator of the treatment mechanism g_0 that is more nonparametrically estimating g_0 than estimator used in previous, and using g_0 results in the full bias reduction for the target parameter). Given such a constructed sequence of candidate targeted MLE's, one now selects the index of this sequence with the minimizer of the cross-validated risk of the loss function (see van der Laan and Gruber (2009)).

The main idea of collaborative targeted MLE is that targeted maximum likelihood estimators of ψ_0 are defined by an estimator Q_n^* of Q_0, so that a loss function (i.e. empirical criterion) can be used to evaluate different targeted maximum likelihood estimators that only differ in different degrees of targeted bias reduction. In this manner we can fine tune the bias reduction. For example, this adaptive selection guarantees that the targeted maximum likelihood step (i.e., the choice of g_j) is actually improving the fit of Q_n^* w.r.t.

the loss function, thereby dealing with the possible problem that the choice of g_j actually deteriorates the estimator relative to the initial estimator.

This selection approach for selection among candidate estimators g_j is not only theoretically grounded by oracle properties of the cross-validation selector, but also by the collaborative double robustness of the efficient influence curve as proved in van der Laan and Gruber (2009). This collaborative double robustness shows that the bias reduction for ψ_0 is achieved by an estimator g_n that correctly adjusts for the covariates that are still helpful in improving the fit of Q_n (i.e., the covariates that deal with the residual bias taking into account the initial estimator), but that covariates that are not needed to remove bias w.r.t. ψ_0 can be ignored. The fine-tuning of the bias reduction of the targeted maximum likelihood estimator through the collaborative targeted maximum likelihood estimator can, and typically should, be applied to further improve the finite sample mean squared error of the resulting estimator of ψ_0.

References

Oliver Bembom and Mark van der Laan. Statistical methods for analyzing sequentially randomized trials, commentary on jnci article adaptive therapy for androgen independent prostate cancer: A randomized selection trial including four regimens, by peter f. thall, c. logothetis, c. pagliaro, s. wen, m.a. brown, d. williams, r. millikan (2007). *Journal of the National Cancer Institute*, 99(21):1577–1582, 2007.

M. J. van der Laan R. Neugebauer. Nonparametric causal e?ects based on marginal structural models. *Journal of Statistical Planning and Inference*, 137:419434, 2007.

J. Robins, L. Orallana, and A. Rotnitzky. Estimaton and extrapolation of optimal treatment and testing strategies. *Statistics in Medicine*, 27(23): 4678–4721, 2008.

M.J. van der Laan. Causal effect models for intention to treat and realistic individualized treatment rules. Technical report 203, Division of Biostatistics, University of California, Berkeley, 2006.

M.J. van der Laan and S. Gruber. Collaborative double robust penalized targeted maximum likelihood estimation. *The International Journal of Biostatistics*, 2009.

DOI: 10.2202/1557-4679.1241

M.J. van der Laan and M.L. Petersen. Causal effect models for realistic individualized treatment and intention to treat rules. *International Journal of Biostatistics*, 3(1), 2007.

M.J. van der Laan and J.M. Robins. *Unified methods for censored longitudinal data and causality*. Springer, New York, 2003.