Supplementary Materials For: Estimation of Personalized Effects Associated With Causal Pathways

Razieh Nabi
Computer Science Dept.
Johns Hopkins University
Baltimore, MD, 21218

Phyllis Kanki
Immunology and Infectious Diseases Dept.
Harvard T. H. Chan School of Public Health
Boston, MA, 02115

Ilya Shpitser
Computer Science Dept.
Johns Hopkins University
Baltimore, MD, 21218

Appendix

In Appendix A, we give a brief overview of semi-parametric statistical inference, which provides context for some of our subsequent results. Appendix B contains deferred proofs of our results. In Appendix C, we review the statistical modeling assumptions we made in our data analysis, provide figures which use decision tree classifiers to visualize policies we learned, and describe additional experimental results on policies that optimize effects not mediated by only adherence.

A: Statistical Inference In Semi-Parametric Models

Let $Z_1, \dots, Z_n$, be iid samples from a general class of probability densities $p(Z;\theta)$ parameterized by $\theta^T = (\beta^T, \eta^T)$, where $\beta \in \mathbb{R}^q$ denotes the set of target parameters, and η denotes a possibly infinite dimensional set of nuisance parameters. This type of model is termed semi-parametric, since it has both a parametric and a non-parametric component. The goal of statistical inference in semi-parametric models is to find "the best" estimator of β in the model, denoted by $\widehat{\beta}$. *Regular asymptotically linear (RAL)* estimators are considered in this setting, which are estimators of the form

$$\sqrt{n}(\hat{\beta} - \beta) = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \phi(Z_i) + o_p(1),$$

where $\phi \in \mathbb{R}^q$ with mean zero and finite variance, $o_p(1)$ denotes a term that approaches to zero in probability, and $\phi(Z_i)$ is the *influence function (IF)* of the ith observation for the parameter vector β. RAL estimators are consistent and asymptotically normal (CAN), with the variance of the estimator given by its IF:

$$\sqrt{n}(\hat{\beta} - \beta) \xrightarrow{\mathcal{D}} \mathcal{N}(0, \phi\phi^T).$$

Thus, there is a bijective correspondence between RAL estimators and IFs. In fact, IFs provide a geometric view of the behavior of RAL estimators. Consider a Hilbert space $\mathcal{H}$ of all mean-zero $q-$dimensional functions, equipped with an inner product, and define the inner product of two arbitrary elements of the Hilbert space, h_1 and h_2, to be equal to $\mathbb{E}[h_1^T h_2]$. Define a *parametric submodel* to be a subset of densities in the semi-parametric model parameterized by $\theta_\gamma^T = (\beta^T, \gamma^T)$, where $\gamma^T \in \mathbb{R}^r$, such that the subset contains the density $p(Z;\theta_0)$ in the semi-parametric model evaluated at the true parameter values θ_0. The *nuisance tangent space* Λ in the semi-parametric model is defined to be the mean square closure of elements of the nuisance tangent spaces $\Lambda_\gamma = \{B^{q\times r} S_\eta(Z;\theta)\}$ of every parametric submodel. The space Λ is important because it is known all influence functions lie in the orthogonal complement $\Lambda^\perp$ of Λ with respect to $\mathcal{H}$. For this reason, recovering $\Lambda^\perp$ is often the first step for constructing RAL estimators in semi-parametric models. Out of all IFs in $\Lambda^\perp$ there exists a unique one which lies in the tangent space, and which yields the most efficient RAL estimator by recovering the *semi-parametric efficiency bound*, see [5] for details.

B: Proofs

Here we give proofs of all claims in the main body of the paper.

Theorem 1 *Fix a causal model given by a complete DAG on variables $W_0, A_1, M_1, W_1, \dots, A_K, M_K, W_K$, listed in topological order, with a hidden common cause U of $W_0, \dots, W_K$. Let α be all directed edges out of $A_1, \dots, A_K$, and $\mathfrak{f}_\alpha$ which sets all edges $(A_i M_j)_\rightarrow$ to a_i, and all other edges in α to a policy $f_{A_i}(H_i)$. In this model, $p(W_K(\mathfrak{f}_\alpha))$ is identified as*

$$\sum_{H_K, M_K} p(W_0) \prod_{i=1}^{K} p(M_i | \bar{a}_i', \overline{W}_{i-1}, \overline{M}_{i-1}) p(W_i | \overline{M}_{i-1}, \overline{W}_{i-1}, f_{\underline{A}_i}(H_i)) \quad (1)$$

Proof: The causal model we describe is simply Pearl's functional model corresponding to the K stage version of the DAG in Fig. 2 (a). It is well known that in this model, given standard positivity assumptions, $p(W_0, \dots, W_K, M_1, \dots, M_k | \text{do}(a_1, \dots, a_K))$ is identified by the g-formula (2):

$$p(W_0) \prod_{i=1}^{K} p(M_i | \bar{a}_i', \overline{W}_{i-1}, \overline{M}_{i-1}) p(W_i | \overline{M}_{i-1}, \overline{W}_{i-1}, \bar{a}_i).$$

Since the recanting district criterion [3] does not hold, we have that $p(\{W_0,\ldots,W_K,M_1,\ldots,M_k\}(\mathfrak{a}'_\alpha))$ is identified by

$$p(W_0)\prod_{i=1}^{K} p(M_i|\bar{a}'_i,\overline{W}_{i-1},\overline{M}_{i-1})p(W_i|\overline{M}_{i-1},\overline{W}_{i-1},\bar{a}_i).$$

where α is all outgoing edges from $\mathbf{A}$, and $\mathfrak{a}'$ sets all edges of the form (A_iM_j) to a'_i, and all edges of the form (A_iW_j) to a_i.

Every $f_{A_i}(H_i)$ simply chooses a_i based on H_i, which is a subset of outcome variables in our distribution. Since the identifiability statement above holds regardless of how $a_1,\ldots,a_i$ are chosen, this implies $p(\{W_0,\ldots,W_K,M_1,\ldots,M_k\}(\mathfrak{f}_\alpha))$, where $\mathfrak{f}_\alpha$ is given in the Theorem statement is identified as

$$\sum_{H_K,M_K} p(W_0)\prod_{i=1}^{K} p(M_i|\bar{a}'_i,\overline{W}_{i-1},\overline{M}_{i-1})p(W_i|\overline{M}_{i-1},\overline{W}_{i-1},f_{\underline{A}_i}(H_i)) \quad (2)$$

which implies our result by a simple marginalization. □

Before proving Theorem 5, we show the following claim.

Theorem 1 *Within the model corresponding to Fig. 1 (a), the unique* efficient influence function $U(\beta)$ *of* $\beta = \mathbb{E}[Y(A = f(W), M(a'))]$ *is given by*

$$\frac{\widetilde{C}}{\widetilde{\pi}(W)}\frac{f(M|W,A=a')}{f(M|W,f(W))}\Big\{Y-\mathbb{E}[Y|f(W),M,W]\Big\}+$$
$$\frac{\mathbb{I}(A=a')}{\pi_{a'}(W)}\Big\{\mathbb{E}[Y|f(W),M,W]-\sum_M \mathbb{E}[Y|f(W),M,W]$$
$$p(M|W,A=a')\Big\}+\sum_M \mathbb{E}[Y|f(W),M,W]p(M|W,A=a')-\beta.$$

Proof: This proof follows as an extension of similar results on the influence function of the mediation functional, found in [4].

The model in Fig. 1 (a) imposes no restrictions on the observed data, and so is non-parametric saturated. As a result, the influence function $U(\beta)$ for any β is a unique solution to the following integral equation

$$\frac{\partial}{\partial t}\beta(F_t)\Big|_{t=0} = \mathbb{E}[S(W,A,M,Y)\phi(\beta)],$$

where F_t is the distribution function corresponding to a one dimensional regular parametric submodel of the non-parametric model on W,A,M,Y, indexed by a single parameter t, and S is the score. $\partial\beta(F_t)/\partial t$ is equal to

$$\frac{\partial}{\partial t}\sum_{w,m}\mathbb{E}[Y\mid a=f_A(w),m,w]p(m\mid a',w)p(w) =$$
$$\sum_{w,m,y} y\frac{\partial}{\partial t}(p(y\mid a=f_A(w),m,w)p(m\mid a',w)p(w)) =$$
$$\sum_{w,m,y} yS(y\mid a=f_A(w),m,w)p(y\mid a=f_A(w),m,w)\times$$
$$p(m\mid a',w)p(w)$$
$$+\sum_{w,m}\mathbb{E}[Y\mid a=f_A(w),m,w]S(m\mid a',w)p(m\mid a',w)p(w)$$
$$+\sum_{w,m}\mathbb{E}[Y\mid a=f_A(w),m,w]p(m\mid a',w)S(w)p(w),$$

where $S(.)$ represent appropriate conditional and marginal scores. By linearity of derivatives, we can solve this equation, term by term. We have, for the first term:

$$\sum_{w,m,y} yS(y\mid a=f_A(w),m,w)p(y\mid a=f_A(w),m,w)\times$$
$$p(m\mid a',w)p(w)$$
$$=\sum_{w,m,y,a''}\frac{\mathbb{I}(a''=f_A(w))p(m\mid a',w)}{p(a''=f_A(w)\mid w)p(m\mid a'',w)}$$
$$yS(y\mid a'',m,w)p(y\mid a'',m,w)p(m\mid a'',w)\times$$
$$p(a''=f_A(w)\mid w)p(w)$$
$$=\sum_{w,m,y,a''}\frac{\mathbb{I}(a''=f_A(w))p(m\mid a',w)}{p(a''\mid w)p(m\mid a'',w)}$$
$$\{y-\mathbb{E}[Y\mid a'',m,w]\}S(y,a'',m,w)p(y,a'',m,w)$$
$$=\mathbb{E}\Big[\frac{\mathbb{I}(A=f_A(W))p(M\mid a',W)}{p(A=f_A(W)\mid W)p(M\mid A,W)}\times$$
$$\{Y-\mathbb{E}[Y\mid A,M,W]\}S(Y,A,M,W)\Big].$$

So the first term contribution to $U(\beta)$ is $\frac{\mathbb{I}(A=f_A(W))p(M|a',W)}{p(A=f_A(W)|W)p(M|a,W)}\{Y-\mathbb{E}[Y\mid a,M,W]\}$.

For the second term, we have:

$$\sum_{w,m}\mathbb{E}[Y\mid a=f_A(w),m,w]S(m\mid a',w)\times$$
$$p(m\mid a',w)p(w)$$
$$=\sum_{w,m,a''}\frac{\mathbb{I}(a''=a')}{p(a''\mid w)}\mathbb{E}[Y\mid a=f_A(w),m,w]\times$$
$$S(m\mid a'',w)p(m,a'',w)$$
$$=\sum_{w,m,a''}\frac{\mathbb{I}(a''=a')}{p(a''\mid w)}\Big\{\mathbb{E}[Y\mid a=f_A(w),m,w]$$
$$-\sum_m \mathbb{E}[Y\mid a=f_A(w),m,w]p(m\mid a'',w)\Big\}\times$$
$$S(m,a'',w)p(m,a'',w)$$
$$=\sum_{w,m,a'',y}\frac{\mathbb{I}(a''=a')}{p(a''\mid w)}\Big\{\mathbb{E}[Y\mid a=f_A(w),m,w]$$
$$-\sum_m \mathbb{E}[Y\mid a=f_A(w),m,w]p(m\mid a'',w)\Big\}\times$$
$$S(y,m,a'',w)p(y,m,a'',w)$$
$$=\mathbb{E}\Big[S(Y,M,A,W)\frac{\mathbb{I}(A=a')}{p(a'\mid W)}\times$$
$$\big\{\mathbb{E}[Y|a=f_A(W),M,W]-\mathbb{E}_q[Y|a,a',W]\big\}\Big].$$

where $\mathbb{E}_q[Y\mid a=f_A(W),a',W]=\sum_m \mathbb{E}[Y\mid a=f_A(W),m,W]p(m\mid a',W)$.

So the second term contribution to $U(\beta)$ is

$$\frac{\mathbb{I}(A=a')}{p(a'\mid W)}\Big\{\mathbb{E}[Y\mid a=f_A(W),M,W]$$
$$-\mathbb{E}_q[Y\mid a=f_A(W),a',W]\Big\}.$$

For the third term, we have:

$$\sum_{w,m} \mathbb{E}[Y \mid a, m, w] p(m \mid a', w) S(w) p(w)$$
$$= \sum_{w} \left\{ \sum_{m} \mathbb{E}[Y \mid a, m, w] p(m \mid a', w) \right\} S(w) p(w)$$
$$= \sum_{w} \left\{ \sum_{m} \mathbb{E}[Y \mid a, m, w] p(m \mid a', w) \right.$$
$$\left. - \sum_{w,m} \mathbb{E}[Y \mid a, m, w] p(m \mid a', w) p(w) \right\} S(w) p(w)$$
$$= \sum_{w,a'',m,y} \left\{ \sum_{m} \mathbb{E}[Y \mid a, m, w] p(m \mid a', w) \right.$$
$$\left. - \sum_{w,m} \mathbb{E}[Y \mid a, m, w] p(m \mid a', w) p(w) \right\} \times$$
$$S(y, m, a'', w) p(y, m, a'', w)$$
$$= \mathbb{E}\left[\sum_{m} \mathbb{E}[Y \mid a = f_A(W), m, W] p(m \mid a', W) - \beta \right]$$

So the third term contribution to $U(\beta)$ is $\mathbb{E}_q[Y \mid a = f_A(W), a', W] - \beta$.

This establishes our result. □

Theorem 2 *Fix a causal model given by a complete DAG on variables $\mathbf{V} \equiv \{W_0, A_1, M_1, W_1, \ldots, A_K, M_K, W_K\}$, listed in topological order, with a hidden common cause U of $W_0, \ldots, W_K$. Let α be all directed edges present in the DAG out of $A_1, \ldots, A_K$ and into $M_1, \ldots M_K$, and $\mathfrak{a}_\alpha$ which sets all edges $(A_i M_j)_\rightarrow$ to a'_i. In this model, $p(\mathbf{V}(\mathfrak{a}_\alpha))$ is identified as*

$$p(\mathbf{V}(\mathfrak{a}_\alpha)) \equiv \tilde{p}(\tilde{W}_0, \tilde{A}_1, \tilde{M}_1, \tilde{W}_1, \ldots, \tilde{W}_K, \tilde{A}_K, \tilde{M}_K) =$$
$$p(W_0) \prod_{i=1}^{K} p(W_i | M_i, A_i, H_i) p(A_i | H_i) p(M_i | \overline{a}'_i, H_i \setminus \mathbf{A})$$

Proof: This is a corollary of Theorem 1, where we define each $f_{A_i}(H_i)$ to be the observed conditional distribution $p(A_i | H_i)$. □

Theorem 3 *Given that each $\tilde{Q}_i, i = 1, \ldots, K$ is specified correctly, the optimal treatment at stage i is equal to*

$$f^*_{A_i}(H_i) = \arg\max_{a_i} \tilde{Q}_i(H_i, a_i; \gamma_i).$$

Proof: This follows by the standard backwards induction argument giving the relationship between Q-functions and optimal policies, applied to $\tilde{p}$ and $\tilde{Q}_i$, and the definition of expected response corresponding to path-specific policies we have chosen.

The optimal policy set $\mathbf{f}^*_{\mathbf{A}}$ is defined as

$$\arg\max_{\{f^*_{A_i} \in \mathbf{f}^*_{\mathbf{A}}\}} \mathbb{E}[W_K(\mathfrak{f}_\alpha)]$$
$$= \arg\max_{\{f^*_{A_i} \in \mathbf{f}^*_{\mathbf{A}}\}} \int W_K \prod_{i=1}^{K} p(W_i | \overline{A}_i = f^*_{A_i}(H_i), H_i, M_i)$$
$$p(W_0) p(M_i | \overline{a}'_i, H_i) d\mathbf{V}$$
$$= \arg\max_{a_1} \int p(W_1 | a_1, H_1, M_1) p(M_1 | a'_1, H_1) dM_1, W_1$$
$$\ldots$$
$$\arg\max_{a_{K-1}} \int p(W_{K-1} | a_{K-1}, H_{K-1}, M_{K-1})$$
$$p(M_{K-1} | \overline{a}'_{K-1}, H_{K-1}) dM_{K-1}, W_{K-1}$$
$$\arg\max_{a_K} \int W_K p(W_K | a_K, H_K, M_K)$$
$$p(M_K | \overline{a}'_K, H_K) dM_K, W_K$$

It's immediately clear that the last line above yields $\tilde{Q}_K$, and given that line $i+1$ yields $\tilde{Q}_{i+1}$ assuming $a_{i+1}, \ldots, a_K$ were chosen optimally, line i yields $\tilde{Q}_i$. □

Theorem 4 *Assume models in the set $\{\tilde{Q}_i(\tilde{H}_i, \tilde{A}_i; \gamma_i), p(M_i | A_i, H_i; \phi) | \forall i\}$ are correctly specified. Then the estimation equations*

$$\mathbb{E}\left[\frac{\partial \tilde{Q}_K}{\partial \gamma_K} \{W_K - \tilde{Q}_K(A_K, H_K; \gamma_K)\} w_K(H_K; \widehat{\phi_K}) \right] = 0, \textit{ and}$$
$$\mathbb{E}\left[\frac{\partial \tilde{Q}_i}{\partial \gamma_i} \{V_{i+1}(H_{i+1}) - \tilde{Q}_i(H_i, A_i; \gamma_i)\} w_i(H_i; \widehat{\phi_i}) \right] = 0,$$

are consistent for γ_K and γ_i, where

$$w_i(H_i; \widehat{\phi_i}) \equiv \frac{p(M_i | \overline{A_i} = a', H_i; \widehat{\phi}_i)}{p(M_i | \overline{A_i}, H_i; \widehat{\phi_i})} \forall i = 1, \ldots K.$$

Proof: We show this inductively on the decision stage. For stage K, we have

$$\mathbb{E}\left[\frac{\partial \tilde{Q}_K; \gamma_K)}{\partial \gamma_K} \{W_K - \tilde{Q}_K(A_K, H_K; \gamma_K)\} w_K(H_K; \widehat{\phi_K}) \right]$$
$$= \int \frac{\partial \tilde{Q}_K}{\partial \gamma_K} \{W_K - \tilde{Q}_K(A_K, H_K; \gamma_K)\}$$
$$p(M_K | A_K = a'_k, H_K) p(H_K) dM_K dH_K$$
$$= \mathbb{E}\left[\frac{\partial \tilde{Q}_K}{\partial \gamma_K} \{\tilde{\mathbb{E}}[W_K | A_K, H_K] - \tilde{Q}_K(A_K, H_K; \gamma_K)\} \right],$$

where $\tilde{\mathbb{E}}$ is the expectation taken with respect to the appropriate conditional distribution derived from $\tilde{p}$. Consistency for γ_K then follows by standard results on regression estimators. Given consistency of the stage $i+1$ regression, we have a consistent estimator for $\tilde{V}_{i+1}$. This allows us to repeat the consistency argument for $\tilde{Q}_i$, as above. □

Theorem 5 *The estimator in (17) is consistent and asymptocally normal (CAN) if the models in the set $\{\pi(W; \psi), p(M | W, A; \phi)\}$ are correctly specified, and the estimator in (18) is CAN in the* union model, *where any two models in the set $\{\pi(W; \psi), \mathbb{E}[Y | A, M, W; \zeta], p(M | W, A; \phi)\}$ are correctly specified.*

Proof: This proof follows as an extension of consistency results derived for the triply robust estimator of the counterfactual expectation $\beta = \mathbb{E}[Y(a, M(a'))]$ associated with the natural direct effect in [4].

Assume the models in the set $\{\pi(W;\psi), p(M|W,A;\phi)\}$ are correctly specified. Let

$$g(W;\psi,\phi) \equiv \frac{p(M|A=a',W;\phi)}{p(M|A=f(W),W;\phi)\widetilde{\pi}(W;\psi)}$$

We have

$$\begin{aligned}\mathbb{E}\left[Y\widetilde{C}g(W;\psi,\phi)\right] &= \mathbb{E}\left[\mathbb{E}\left[Y\widetilde{C}g(W;\psi,\phi)\Big|W\right]\right]\\ &= \mathbb{E}\left[\mathbb{E}\left[Y|W\right]\widetilde{C}g(W;\psi,\phi)\right].\end{aligned}$$

This is equal to

$$\begin{aligned}&\int Yp(Y|M,A,W)p(M|A,W)p(A|W)\widetilde{C}g(A,W)dA,dM,dY\\ =&\int Yp(Y|M,A,W)p(M|a',W)p(A|W)\frac{\mathbb{I}(A=f(W))}{\widetilde{\pi}(W)}dA,dM,dY\\ =&\int Yp(Y|M,A=f(W),W)p(M|a',W)dM,dY.\end{aligned}$$

This is precisely β of interest.

The estimator $\widehat{\beta}_{triple}$ has the form

$$\mathbb{E}\left[\frac{\widetilde{C}}{\widetilde{\pi}(W;\psi)}\frac{f(M|W,A=a';\widehat{\phi})}{f(M|W,f(W);\widehat{\phi})}\Big\{Y-\mathbb{E}[Y|f(W),M,W;\widehat{\zeta}]\Big\}+ \frac{\mathbb{I}(A=a')}{\pi'_a(W;\widehat{\psi})}\Big\{\mathbb{E}[Y|f(W),M,W;\widehat{\zeta}] - \sum_M \mathbb{E}[Y|f(W),M,W;\widehat{\zeta}]. p(M|W,A=a';\widehat{\zeta})\Big\} + \sum_M \mathbb{E}[Y|f(W),M,W;\widehat{\zeta}]p(M|W,A=a';\widehat{\phi})\right],$$

Assume $\tilde{\pi}$ was specified incorrectly. The expectation in the estimator consists of three terms, where the last term is equal to true β if models for Y and M are correct. For the first term we have, by iterated expectation,

$$\begin{aligned}&\mathbb{E}\left[\widetilde{C}g(W;\phi,\psi)\Big\{Y-\mathbb{E}[Y|f(W),M,W;\widehat{\zeta}]\Big\}\right]=\\ &\mathbb{E}\left[\widetilde{C}g(W;\phi,\psi)\Big\{\mathbb{E}[Y|A,M,W]-\mathbb{E}[Y|f(W),M,W;\widehat{\zeta}]\Big\}\right]=\\ &\mathbb{E}\left[\widetilde{C}g(W;\phi,\psi)\Big\{\mathbb{E}[Y|f(W),M,W]-\mathbb{E}[Y|f(W),M,W;\widehat{\zeta}]\Big\}\right]=0,\end{aligned}$$

if the Y model is correct. For the second term we have, by iterated expectation,

$$\begin{aligned}&\mathbb{E}\left[\frac{\mathbb{I}(A=a')}{\pi'_a(W;\widehat{\psi})}\Big\{\mathbb{E}[Y|f(W),M,W;\widehat{\zeta}]-\sum_M\mathbb{E}[Y|f(W),M,W;\widehat{\zeta}] p(M|W,A=a';\widehat{\zeta})\Big\}\right]=\\ &\mathbb{E}\left[\mathbb{E}\left[\frac{\mathbb{I}(A=a')}{\pi'_a(W;\widehat{\psi})}\Big\{\mathbb{E}[Y|f(W),M,W;\widehat{\zeta}]-\sum_M\mathbb{E}[Y|f(W),M,W;\widehat{\zeta}] p(M|W,A=a';\widehat{\zeta})\Big\}\Big|W,A=a'\right]\right]=\\ &\mathbb{E}\left[\frac{\mathbb{I}(A=a')}{\pi'_a(W;\widehat{\psi})}\Big\{\mathbb{E}[\mathbb{E}[Y|f(W),M,W;\widehat{\zeta}]|A=a',W] - \sum_M\mathbb{E}[Y|f(W),M,W;\widehat{\zeta}]p(M|W,A=a';\widehat{\zeta})\Big\}\right]=0\end{aligned}$$

if the models for Y and M are correct.

Assume the model for M was specified incorrectly. The first term in the estimator is mean zero by above argument, since the Y model is still correct.

The second and last terms decompose into

$$\begin{aligned}&\mathbb{E}\left[\frac{\mathbb{I}(A=a')}{\pi'_a(W;\widehat{\psi})}\Big\{\mathbb{E}[\mathbb{E}[Y|f(W),M,W;\widehat{\zeta}]|A=a',W]\right]\\ &-\mathbb{E}\left[\frac{\mathbb{I}(A=a')}{\pi'_a(W;\widehat{\psi})}\sum_M\mathbb{E}[Y|f(W),M,W;\widehat{\zeta}]p(M|W,A=a';\widehat{\zeta})\Big\}\right]\\ &+\mathbb{E}\left[\sum_M\mathbb{E}[Y|f(W),M,W;\widehat{\zeta}]p(M|W,A=a';\widehat{\phi})\right]\\ =&\mathbb{E}\left[\frac{\mathbb{I}(A=a')}{\pi'_a(W;\widehat{\psi})}\Big\{\mathbb{E}[\mathbb{E}[Y|f(W),M,W;\widehat{\zeta}]|A=a',W]\right]\\ &+\mathbb{E}\left[\sum_M\mathbb{E}[Y|f(W),M,W;\widehat{\zeta}]p(M|W,A=a';\widehat{\zeta})\right.\\ &\left.\mathbb{E}\left[\left(1-\frac{\mathbb{I}(A=a')}{\pi'_a(W;\widehat{\psi})}\right)\Big|W\right]\right]=\\ &\mathbb{E}\left[\frac{\mathbb{I}(A=a')}{\pi'_a(W;\widehat{\psi})}\mathbb{E}[\mathbb{E}[Y|f(W),M,W;\widehat{\zeta}]|A=a',W]\right]\\ &+\mathbb{E}\left[\sum_M\mathbb{E}[Y|f(W),M,W;\widehat{\zeta}]p(M|W,A=a';\widehat{\zeta})\right.\\ &\left.\left(1-\frac{\pi'_a(W)}{\pi'_a(W;\widehat{\psi})}\right)\right]\\ =&\mathbb{E}\left[\frac{\mathbb{I}(A=a')}{\pi'_a(W;\widehat{\psi})}\mathbb{E}[\mathbb{E}[Y|f(W),M,W;\widehat{\zeta}]|A=a',W]\right]\\ =&\mathbb{E}\left[\mathbb{E}[\mathbb{E}[Y|f(W),M,W;\widehat{\zeta}]|A=a',W]\right]\end{aligned}$$

if the models for Y and A are correct. The remainder is precisely β.

Assume the model for Y was specified incorrectly. The terms in $\widehat{\beta}_{triple}$ then decompose into

$$\begin{aligned}&\mathbb{E}\left[\widetilde{C}g(W;\psi,\phi)Y\right]\\ &+\mathbb{E}\left[\left(\frac{\mathbb{I}(A=a)}{\pi_a(W;\psi)}-\widetilde{C}g(W;\psi,\phi)\right)\mathbb{E}[Y|f(W),M,W;\zeta]\right]\\ &+\mathbb{E}\left[\sum_M\mathbb{E}[Y|f(W),M,W;\widehat{\zeta}]p(M|W,A=a';\widehat{\zeta})\right.\\ &\left.\mathbb{E}\left[\left(1-\frac{1-\widetilde{C}}{1-\widetilde{\pi}(W;\psi)}\right)\Big|W\right]\right]\end{aligned}$$

The last term is mean zero if $\widetilde{\pi}$ is specified correctly. The second term is equal to

$$\begin{aligned}&\int\mathbb{E}[Y|f(W),M,W;\zeta]p(M|A=a',W)p(W)-\\ &\int\mathbb{E}[Y|f(W),M,W;\zeta]p(M|A=a',W)p(W)=0\end{aligned}$$

if the A and M models are specified correctly. The first term is equal to β by the argument above.

Both estimators are special cases of the RAL estimator for β based on the efficient influence function. As a result, standard regularity assumptions [2], and properties of maximum likelihood estimators imply both estimators are CAN. □

C: Experiments And Visualizations

C1. Models Used In Data Analysis

We used linear regression with interaction terms between treatment A_2 and history H_2 to model the outcome W_2: $\mathbb{E}[W_2|H_2, A_2, \mathbf{M_2}; \alpha] = \alpha_1(H_2, A_2, \mathbf{M_2}) + \alpha_2 A_2 H_2$, and logistic regression with interaction terms to model all dichotomous variables X with history H and immediate prior treatment A: $logit\{p(X = 1 \mid H, A; \beta)\} = \beta_1 H + \beta_2 AH$ for $X \in \{\mathbf{M_1}, \mathbf{M_2}, W_1\}$. We used the same form of linear regression with interaction terms to model Q-functions by excluding the mediators: $Q_2(H_2, A_2; \gamma^2) = \gamma_1^2 H_2 + \gamma_2^2 A_2 H_2$ and $Q_1(H_1, A_1; \gamma^1) = \gamma_1^1 H_1 + \gamma_2^1 A_1 H_1$. The parameters in all models were estimated by maximizing the likelihood.

For value search and G-estimation, we used log CD4 count at the end of sixth month as the outcome of interest, denoted by W_1. We used the same form of models, as described above, for W_1, $\mathbb{E}[W_1|H_1, A_1, \mathbf{M_1}; \alpha]$, and all the mediators, and used logistic regression with no interaction terms to model the treatment assignment: $logit\{p(A_1 = 1 \mid H_1; \beta)\} = \beta_0 + \beta_1 H_1$. We modeled the blip function in (19) as $\gamma(A_1, H_1; \psi) = \psi_1 A_1 + \psi_2 A_1 H_1$.

C2: Decision Tree Visualization Of Learned Policies

We derived the optimal policies using G-formula, Q-learning, G-estimation, and value search techniques. The value search method considered a simple class of policies based on a threshold, described in the main body of the paper. The optimal policies obtained from the first three methods were more complicated functions of prior history. To aid in interpretability of these policies, we approximated them by means of decision tree multi-label classifiers which treated history as a set of features, and treatment decision as the class label. The resulting decision tree classifiers are shown in in Fig. 1, 2, and 3. In these figures, the label "path policies" corresponds to policies that optimize the direct chemical effect of the drug where drug toxicity and adherence behave as if treatment was set to a reference level. In the following decision trees, the nodes *vl* and *adh* stand for viral load (log scale) and adherence level, repectively. *m00* and *m06* denote the measures at month zero (baseline) and the end of the first six months, and node *who* denotes the stage of disease (there is a total of 4 stages, with higher stages denoting progressively more severe disease). Since classifiers did not achieve perfect accuracy, these decisions trees should be viewed as easy to visualize approximations of the true learned policies.

Note that in Fig. 1, adherence level is relevant to the overall policy but is omitted in the path specific one. As mentioned above, the classification accuracies are not perfect and hence visualizations are not necessarily a good representation of the true policy. That said, finding the path-specific policy not via Ms corresponds to finding the overall policy in the world shown in Fig. 2 (b) in the main body. It is true that in this world, adherence at time one, $\tilde{M}_1$, influences A_2, and as a result it is in principle possible for adherence at time one to be informative in an interesting way for the decision at time two. However, one large source of variability in patient adherence is precisely due to the treatment we assign, and this source of variability is removed by construction in the world shown in Fig. 2(b) in the main body of the paper – the world where everyone adheres as if on a reference treatment. A low variability variable is certainly less likely to be relevant for decision making (consider what would happen in the limit where everyone had perfect adherence had they been on a reference treatment). Hence, we are certainly not surprised to find that a path-specific (not via adherence) policy did not include adherence as a relevant variable.

Table 1: Population log CD4 counts under different policies (under treatment assignments in the observed data, the value is 5.64 ± 0.01 in the 2-stage and 5.54 ± 0.01 in the 1-stage problem). G-formula and Q-learning are used with 2-stage decision points. Value search and G-estimation are used with 1-stage decision point.

	Path Policies (not through adherence)
G-formula	$6.78\ (5.65, 6.92)$
Q-learning	$7.00\ (4.82, 7.19)$
Value search	$5.56\ (5.44, 5.60)$
G-estimation	$5.56\ (5.55, 5.58)$

C3. Additional Experiments

To tie the results of this paper to earlier work [1], we ran additional experiments to find policies that optimize the chemical effect of the drug where only adherence behaves as if the treatments were set to a reference level. Expected outcomes under optimal policies we learned, along with 95% confidence intervals obtained by bootstrap, are shown in Table. 1. The results are consistent with the ones provided in the main body of the paper. For value search, under the same class of policies, $\mathbb{I}\{\text{CD4m00} < \alpha\}$, and the same modeling assumptions described above, the optimal path policy is chosen to be $\mathbb{I}\{\text{CD4m00} < 550 \text{ cells/mm}^3\}$.

References

[1] C. Miles, I. Shpitser, P. Kanki, S. Meloni, and E. T. Tchetgen. Quantifying an adherence path-specific effect of antiretroviral therapy in the nigeria pepfar program. *Journal of the American Statistical Society*, 2017.

[2] J. M. Robins, S. D. Mark, and W. K. Newey. Estimating exposure effects by modelling the expectation of exposure conditional on confounders. *Biometrics*, pages 479–495, 1992.

[3] I. Shpitser. Counterfactual graphical models for longitudinal mediation analysis with unobserved confounding. *Cognitive Science (Rumelhart special issue)*, 37:1011–1035, 2013.

[4] E. J. T. Tchetgen and I. Shpitser. Semiparametric theory for causal mediation analysis: efficiency bounds, multiple robustness, and sensitivity analysis. *Ann. Stat.*, 40(3):1816–1845, 2012.

[5] A. Tsiatis. *Semiparametric Theory and Missing Data*. Springer, 1 edition, 2006.

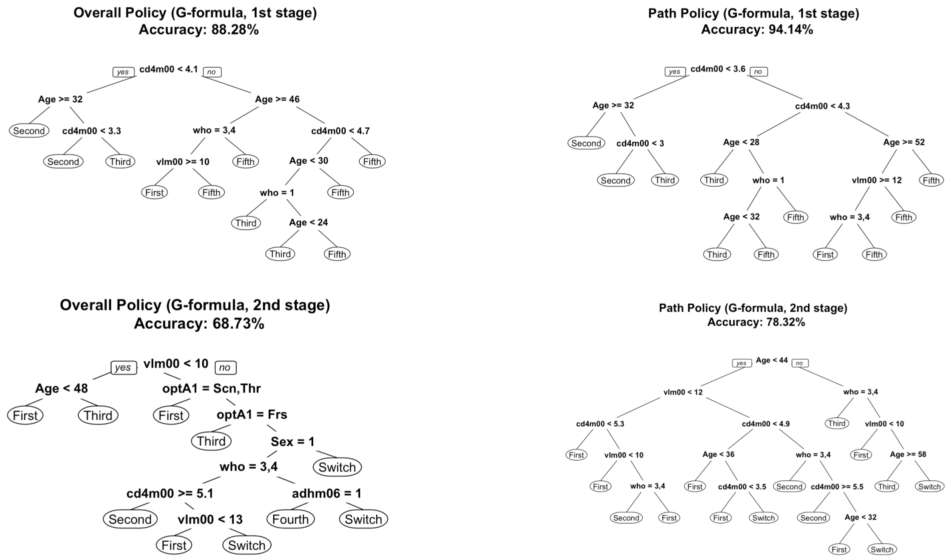

Figure 1: Decision trees for optimal policies obtained via G-formula.

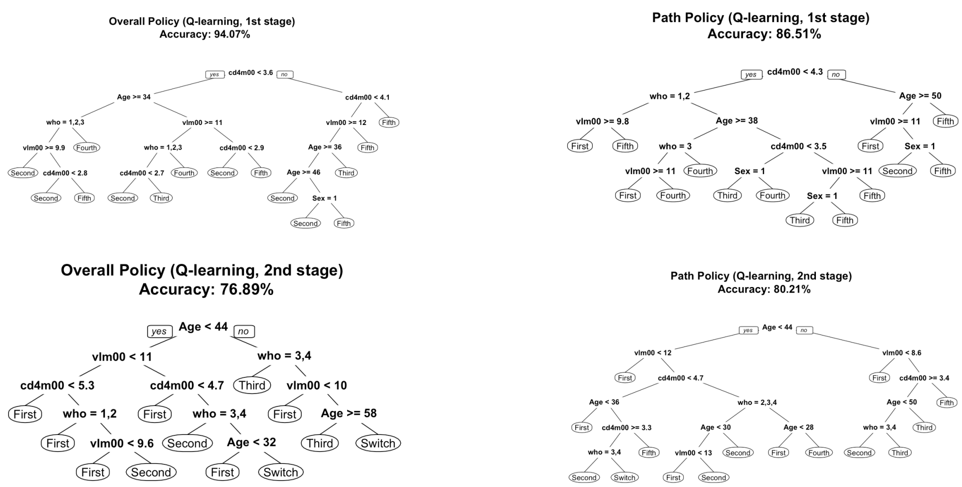

Figure 2: Decision trees for optimal policies obtained via Q-learning.

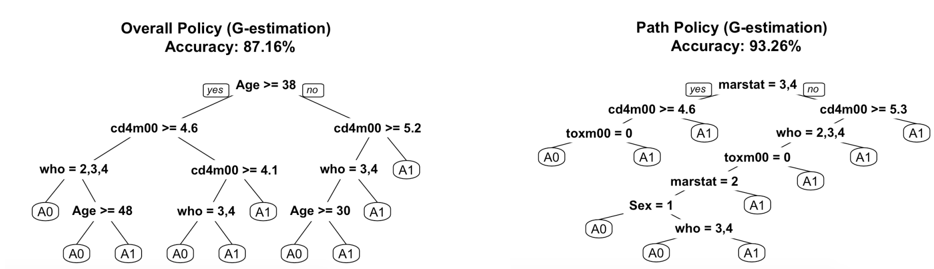

Figure 3: Decision trees for optimal policies obtained via G-estimation.