BIOMETRICS 000, 000–000 DOI: 000
000 0000

Web Appendix for "Using decision lists to construct interpretable and parsimonious treatment regimes"

Yichi Zhang[1,*], Eric B. Laber[1,], Anastasios Tsiatis[1,***], and Marie Davidian[1,****]**

[1]Department of Statistics NC State University, Raleigh, NC 27695, USA

*email: yzhang52@ncsu.edu

**email: laber@stat.ncsu.edu

***email: tsiatis@ncsu.edu

****email: davidian@ncsu.edu

This paper has been submitted for consideration for publication in *Biometrics*

1. An Illustrative Run Through the Algorithm for Finding an Optimal Decision List

In this section, we illustrate how the proposed algorithm for finding an optimal decision list works. For simplicity, the patient covariate is assumed to be two-dimensional.

- The algorithm starts at Step 1.
  - We choose $L_{\max} = 5$ and $\alpha = 0.05$.
  - We compute $\widetilde{a}_0 = \arg\max_{a_0 \in \mathcal{A}} \widehat{R}\,[\{a_0\}]$. Suppose the maximum found is $\widehat{R}\,[\{\widetilde{a}_0\}] = 10$. Figure 1 shows the decision list $\{\widetilde{a}_0\}$.
  - We set $\Pi_{\text{temp}} = \emptyset$ and $\Pi_{\text{final}} = \emptyset$.

[Figure 1 about here.]

- The algorithm proceeds to Step 2.
  - The goal is to estimate the first clause (c_1, a_1).
  - We compute $(\widetilde{c}_1, \widetilde{a}_1, \widetilde{a}'_1) = \arg\max_{(c_1, a_1, a'_1) \in \mathcal{C} \times \mathcal{A} \times \mathcal{A}} \widehat{R}\,[\{(c_1, a_1), a'_1\}]$. This is done, conceptually, by evaluating $\widehat{R}(\cdot)$ at each element in $\mathcal{C} \times \mathcal{A} \times \mathcal{A}$. Suppose the maximum found is $\widehat{R}\,[\{(\widetilde{c}_1, \widetilde{a}_1), \widetilde{a}'_1\}] = 15$ and the clause $\widetilde{c}_1$ has the form $x_1 \leqslant \tau_1$. Figure 2 shows the decision list $\{(\widetilde{c}_1, \widetilde{a}_1), \widetilde{a}'_1\}$.
  - We compute $\widehat{\Delta}_1 = \widehat{R}\,[\{(\widetilde{c}_1, \widetilde{a}_1), \widetilde{a}'_1\}] - \widehat{R}\,[\{\widetilde{a}_0\}]$ and compare $\widehat{\Delta}_1$ to $z_{1-\alpha}\big\{\widehat{\mathrm{Var}}\big(\widehat{\Delta}_1\big)\big\}^{1/2}$. In this case $\widehat{\Delta}_1 = 15 - 10 = 5$. Suppose we get $\widehat{\mathrm{Var}}\big(\widehat{\Delta}_1\big) = 4$ after calculations. Since $5 > z_{0.95} \times 4^{1/2}$, we add two decision lists, $\{(\widetilde{c}_1, \widetilde{a}_1), \widetilde{a}'_1\}$ and $\{(\widetilde{c}'_1, \widetilde{a}'_1), \widetilde{a}_1\}$, into the set Π_{temp} and proceed to estimate the second clause (c_2, a_2).
  - We make a remark on non-uniqueness here. The decision list $\{(\widetilde{c}_1, \widetilde{a}_1), \widetilde{a}'_1\}$ can be equivalently expressed as $\{(\widetilde{c}'_1, \widetilde{a}'_1), \widetilde{a}_1\}$, where $\widetilde{c}'_1$ is the negation of $\widetilde{c}_1$. Since these two decision lists provide the same treatment recommendation to every patient, we have $\widehat{R}[\{(\widetilde{c}'_1, \widetilde{a}'_1), \widetilde{a}_1\}] = \widehat{R}[\{(\widetilde{c}_1, \widetilde{a}_1), \widetilde{a}'_1\}] = 15$. However, their first clauses are different and may lead to consider-

ably different final decision lists. Currently it is impossible to determine whether $(\widetilde{c}_1, \widetilde{a}_1)$ or $(\widetilde{c}'_1, \widetilde{a}'_1)$ should be used in the first clause. Thus we add both decision lists into Π_{temp}, and move on to building the second clause while keeping in mind that there are two possibilities, $(\widetilde{c}_1, \widetilde{a}_1)$ and $(\widetilde{c}'_1, \widetilde{a}'_1)$, for the first clause. Figure 3 shows the decision list $\{(\widetilde{c}'_1, \widetilde{a}'_1), \widetilde{a}_1\}$. The diagram is the same as in Figure 2 while the description is different.

[Figure 2 about here.]

[Figure 3 about here.]

- The algorithm proceeds to Step 3.
  - We pick an element $\overline{\pi}$ from Π_{temp}. Currently Π_{temp} contains two decision lists: $\{(\widetilde{c}_1, \widetilde{a}_1), \widetilde{a}'_1\}$ and $\{(\widetilde{c}'_1, \widetilde{a}'_1), \widetilde{a}_1\}$. Suppose we get $\overline{\pi} = \{(\widetilde{c}_1, \widetilde{a}_1), \widetilde{a}'_1\}$. We remove $\overline{\pi}$ from Π_{temp}.
  - We compute $(\widetilde{c}_2, \widetilde{a}_2, \widetilde{a}'_2) = \arg\max_{(c_2, a_2, a'_2) \in \mathcal{C} \times \mathcal{A} \times \mathcal{A}} \widehat{R}\left[\{(\widetilde{c}_1, \widetilde{a}_1), (c_2, a_2), a'_2\}\right]$. During the maximization $(\widetilde{c}_1, \widetilde{a}_1)$ is held fixed. Intuitively, this is to partition $\mathcal{T}(\widetilde{c}_1)^c$ while keeping $\mathcal{T}(\widetilde{c}_1)$ fixed. Suppose the maximum found is $\widehat{R}[\{(\widetilde{c}_1, \widetilde{a}_1), (\widetilde{c}_2, \widetilde{a}_2), \widetilde{a}'_2\}] = 16$ and the clause $\widetilde{c}_2$ has the form $x_2 \leqslant \tau_{21}$. Figure 4 shows the decision list $\{(\widetilde{c}_1, \widetilde{a}_1), (\widetilde{c}_2, \widetilde{a}_2), \widetilde{a}'_2\}$.
  - We compute $\widehat{\Delta}_2 = \widehat{R}\left[\{(\widetilde{c}_1, \widetilde{a}_1), (\widetilde{c}_2, \widetilde{a}_2), \widetilde{a}'_2\}\right] - \widehat{R}\left[\{(\widetilde{c}_1, \widetilde{a}_1), \widetilde{a}'_1\}\right]$ and compare $\widehat{\Delta}_2$ to $z_{1-\alpha}\{\widehat{\text{Var}}(\widehat{\Delta}_2)\}^{1/2}$. In this case $\widehat{\Delta}_2 = 16 - 15 = 1$. Suppose we get $\widehat{\text{Var}}(\widehat{\Delta}_2) = 2.25$ after calculations. Since $\widehat{\Delta}_2 < z_{0.95}\{\widehat{\text{Var}}(\widehat{\Delta}_2)\}^{1/2}$, the simpler, more parsimonious decision list $\{(\widetilde{c}_1, \widetilde{a}_1), \widetilde{a}'_1\}$ is preferred and added to Π_{final}, while $\{(\widetilde{c}_1, \widetilde{a}_1), (\widetilde{c}_2, \widetilde{a}_2), \widetilde{a}'_2\}$ is discarded.

[Figure 4 about here.]

- The algorithm repeats Step 3.
  - Step 3 is repeated since Π_{temp} contains another element $\overline{\pi} = \{(\widetilde{c}'_1, \widetilde{a}'_1), \widetilde{a}_1\}$. We remove $\overline{\pi}$ from Π_{temp}.
  - We compute $(\widetilde{c}_2, \widetilde{a}_2, \widetilde{a}'_2) = \arg\max_{(c_2, a_2, a'_2) \in \mathcal{C} \times \mathcal{A} \times \mathcal{A}} \widehat{R}\left[\{(\widetilde{c}'_1, \widetilde{a}'_1), (c_2, a_2), a'_2\}\right]$. During the maximization $(\widetilde{c}'_1, \widetilde{a}'_1)$ is held fixed. Intuitively, this is to partition $\mathcal{T}(\widetilde{c}_1)$ while keeping

$\mathcal{T}(\widetilde{c}_1)^c$ fixed. Suppose the maximum found is $\widehat{R}[\{(\widetilde{c}_1', \widetilde{a}_1'), (\widetilde{c}_2, \widetilde{a}_2), \widetilde{a}_2'\})] = 18$ and the clause $\widetilde{c}_2$ has the form $x_2 \leqslant \tau_{22}$. Figure 5 shows the decision list $\{(\widetilde{c}_1', \widetilde{a}_1'), (\widetilde{c}_2, \widetilde{a}_2), \widetilde{a}_2'\}$.

– We compute $\widehat{\Delta}_2 = \widehat{R}\,[\{(\widetilde{c}_1', \widetilde{a}_1'), (\widetilde{c}_2, \widetilde{a}_2), \widetilde{a}_2'\}] - \widehat{R}\,[\{(\widetilde{c}_1', \widetilde{a}_1'), \widetilde{a}_1\}]$ and compare $\widehat{\Delta}_2$ to $z_{1-\alpha}\{\widehat{\mathrm{Var}}(\widehat{\Delta}_2)\}^{1/2}$. In this case $\widehat{\Delta}_2 = 18 - 15 = 3$. Suppose we get $\widehat{\mathrm{Var}}(\widehat{\Delta}_2) = 2$ after calculations. Then we have $\widehat{\Delta}_2 > z_{0.95}\{\widehat{\mathrm{Var}}(\widehat{\Delta}_2)\}^{1/2}$, which means that the second clause significantly improves the performance of the decision list. Thus we add decision lists $\{(\widetilde{c}_1', \widetilde{a}_1'), (\widetilde{c}_2, \widetilde{a}_2), \widetilde{a}_2'\}$ and $\{(\widetilde{c}_1', \widetilde{a}_1'), (\widetilde{c}_2', \widetilde{a}_2'), \widetilde{a}_2\}$ to Π_{temp}.

– Here the non-uniqueness comes into play again. Consequently, although the decision lists $\{(\widetilde{c}_1', \widetilde{a}_1'), (\widetilde{c}_2, \widetilde{a}_2), \widetilde{a}_2'\}$ and $\{(\widetilde{c}_1', \widetilde{a}_1'), (\widetilde{c}_2', \widetilde{a}_2'), \widetilde{a}_2\}$ are equivalent, it is important to have both of them added to Π_{temp}.

[Figure 5 about here.]

- The algorithm repeats Step 3.

– Now Π_{temp} contains two decision lists while Π_{final} contains one. Thus Step 3 is repeated. We first pick and remove an element $\overline{\pi}$ from Π_{temp}, say $\overline{\pi} = \{(\widetilde{c}_1', \widetilde{a}_1'), (\widetilde{c}_2, \widetilde{a}_2), \widetilde{a}_2'\}$.

– Next, we will build a decision list of length 3 and the first two clauses being $(\widetilde{c}_1', \widetilde{a}_1')$ and $(\widetilde{c}_2, \widetilde{a}_2)$. We compute $(\widetilde{c}_3, \widetilde{a}_3, \widetilde{a}_3') = \arg\max_{(c_3,a_3,a_3') \in \mathcal{C}\times\mathcal{A}\times\mathcal{A}} \widehat{R}\,[\{(\widetilde{c}_1', \widetilde{a}_1'), (\widetilde{c}_2, \widetilde{a}_2), (c_3, a_3), a_3'\}]$. During the maximization $(\widetilde{c}_1', \widetilde{a}_1')$ and $(\widetilde{c}_2, \widetilde{a}_2)$ are held fixed. Suppose the maximum found is $\widehat{R}[\{(\widetilde{c}_1', \widetilde{a}_1'), (\widetilde{c}_2, \widetilde{a}_2), (\widetilde{c}_3, \widetilde{a}_3), \widetilde{a}_3'\}] = 20$ and the clause $\widetilde{c}_3$ has the form $x_1 \leqslant \tau_{31}$. Figure 6 shows the decision list $\{(\widetilde{c}_1', \widetilde{a}_1'), (\widetilde{c}_2, \widetilde{a}_2), (\widetilde{c}_3, \widetilde{a}_3), \widetilde{a}_3'\}$.

– We then compute $\widehat{\Delta}_3 = \widehat{R}\,[\{(\widetilde{c}_1', \widetilde{a}_1'), (\widetilde{c}_2, \widetilde{a}_2), (\widetilde{c}_3, \widetilde{a}_3), \widetilde{a}_3'\}] - \widehat{R}\,[\{(\widetilde{c}_1', \widetilde{a}_1'), (\widetilde{c}_2, \widetilde{a}_2), \widetilde{a}_2'\}]$ and compare $\widehat{\Delta}_3$ to $z_{1-\alpha}\{\widehat{\mathrm{Var}}(\widehat{\Delta}_3)\}^{1/2}$. In this case $\widehat{\Delta}_3 = 20 - 18 = 2$. Suppose we get $\widehat{\mathrm{Var}}(\widehat{\Delta}_3) = 3$ after calculations. Then we have $\widehat{\Delta}_3 < z_{0.95}\{\widehat{\mathrm{Var}}(\widehat{\Delta}_3)\}^{1/2}$. Thus the simpler, more parsimonious, decision list $\{(\widetilde{c}_1', \widetilde{a}_1'), (\widetilde{c}_2, \widetilde{a}_2), \widetilde{a}_2'\}$ is preferred. So we add $\{(\widetilde{c}_1', \widetilde{a}_1'), (\widetilde{c}_2, \widetilde{a}_2), \widetilde{a}_2'\}$ to Π_{final} and drop $\{(\widetilde{c}_1', \widetilde{a}_1'), (\widetilde{c}_2, \widetilde{a}_2), (\widetilde{c}_3, \widetilde{a}_3), \widetilde{a}_3'\}$.

[Figure 6 about here.]

- The algorithm repeats Step 3.

  - Since Π_{temp} contains one element $\overline{\pi} = \{(\widetilde{c}'_1, \widetilde{a}'_1), (\widetilde{c}'_2, \widetilde{a}'_2), \widetilde{a}_2\}$, we repeat Step 3 once again. We remove $\overline{\pi}$ from Π_{temp}.
  - We compute $(\widetilde{c}_3, \widetilde{a}_3, \widetilde{a}'_3) = \arg\max_{(c_3, a_3, a'_3) \in \mathcal{C} \times \mathcal{A} \times \mathcal{A}} \widehat{R}\,[\{(\widetilde{c}'_1, \widetilde{a}'_1), (\widetilde{c}'_2, \widetilde{a}'_2), (c_3, a_3), a'_3\}]$ while keeping $(\widetilde{c}'_1, \widetilde{a}'_1)$ and $(\widetilde{c}'_2, \widetilde{a}'_2)$ fixed. Suppose $\widehat{R}(\{(\widetilde{c}'_1, \widetilde{a}'_1), (\widetilde{c}'_2, \widetilde{a}'_2), (\widetilde{c}_3, \widetilde{a}_3), \widetilde{a}'_3\}) = 20.5$ and the clause $\widetilde{c}_3$ has the form $x_1 \leqslant \tau_{32}$ and $x_2 > \tau_{33}$. Figure 7 shows the decision list $\{(\widetilde{c}'_1, \widetilde{a}'_1), (\widetilde{c}'_2, \widetilde{a}'_2), (\widetilde{c}_3, \widetilde{a}_3), \widetilde{a}'_3\}$.
  - We compute $\widehat{\Delta}_3 = \widehat{R}\,[\{(\widetilde{c}'_1, \widetilde{a}'_1), (\widetilde{c}'_2, \widetilde{a}'_2), (\widetilde{c}_3, \widetilde{a}_3), \widetilde{a}'_3\}] - \widehat{R}\,[\{(\widetilde{c}'_1, \widetilde{a}'_1), (\widetilde{c}'_2, \widetilde{a}'_2), \widetilde{a}_2\}]$ and compare $\widehat{\Delta}_3$ to $z_{1-\alpha}\big\{\widehat{\mathrm{Var}}\big(\widehat{\Delta}_3\big)\big\}^{1/2}$. In this case $\widehat{\Delta}_1 = 20.5 - 18 = 2.5$. Suppose we get $\widehat{\mathrm{Var}}\big(\widehat{\Delta}_1\big) = 2.56$ after calculations. Then we have $\widehat{\Delta}_3 < z_{0.95}\big\{\widehat{\mathrm{Var}}\big(\widehat{\Delta}_3\big)\big\}^{1/2}$. So the simpler, more parsimonious, decision list $\{(\widetilde{c}'_1, \widetilde{a}'_1), (\widetilde{c}'_2, \widetilde{a}'_2), \widetilde{a}_2\}$ is preferred. Consequently, we add $\{(\widetilde{c}'_1, \widetilde{a}'_1), (\widetilde{c}'_2, \widetilde{a}'_2), \widetilde{a}_2\}$ to Π_{final} and discard $\{(\widetilde{c}'_1, \widetilde{a}'_1), (\widetilde{c}'_2, \widetilde{a}'_2), (\widetilde{c}_3, \widetilde{a}_3), \widetilde{a}'_3\}$.

[Figure 7 about here.]

- The algorithm finishes Step 4, because Π_{temp} contains no element now.
- The algorithm proceeds to Step 5.

  - We would like to pick a decision list from Π_{final} that maximizes $\widehat{R}(\cdot)$.
  - In this example, we have three decision lists in Π_{final}: $\{(\widetilde{c}_1, \widetilde{a}_1), \widetilde{a}'_1\}$ with estimated value 15, $\{(\widetilde{c}'_1, \widetilde{a}'_1), (\widetilde{c}_2, \widetilde{a}_2), \widetilde{a}'_2\}$ with estimated value 18, and $\{(\widetilde{c}'_1, \widetilde{a}'_1), (\widetilde{c}'_2, \widetilde{a}'_2), \widetilde{a}_2\}$ with estimated value 18.
  - We then choose the one with the maximal estimated value (with ties broken using the first encountered). Therefore, the estimated optimal decision list $\widetilde{\pi}$ is described by $\{(\widetilde{c}'_1, \widetilde{a}'_1), (\widetilde{c}_2, \widetilde{a}_2), \widetilde{a}'_2\}$, as shown in Figure 5.

2. Asymptotic Properties of $\widehat{R}(\pi)$ for a Given π

We shall derive some asymptotic properties of the doubly robust estimator $\widehat{R}(\pi)$ introduced in Section 2.2 in the main paper. In the next section, we will use these properties to derive an estimator for $\mathrm{Var}\left\{\widehat{R}(\pi_1) - \widehat{R}(\pi_2)\right\}$, which is used by our proposed algorithm for finding an optimal decision list.

Hereafter denote the observed data for the ith subject by $O_i = (X_i^{\mathrm{T}}, A_i, Y_i)^{\mathrm{T}}$.

We first derive an i.i.d. representation of $\widehat{\gamma} = (\widehat{\gamma}_1^{\mathrm{T}}, \ldots, \widehat{\gamma}_{m-1}^{\mathrm{T}})^{\mathrm{T}}$, the maximum likelihood estimator of $\gamma = (\gamma_1^{\mathrm{T}}, \ldots, \gamma_{m-1}^{\mathrm{T}})^{\mathrm{T}}$ in the multinomial logistic regression model:

$$\mathbb{P}(A = a|X = x) = \exp\left(u^{\mathrm{T}}\gamma_a\right) \Big/ \left\{1 + \sum_{j=1}^{m-1} \exp\left(u^{\mathrm{T}}\gamma_j\right)\right\}.$$

If $u = u(x) \equiv 1$, then the maximum likelihood estimator of $\omega(x, a) = \mathbb{P}(A = a|X = x)$ reduces to $\mathbb{E}_n I(A = a)$. Thus the multinomial logistic regression model includes the sample proportion as its special case. The log-likelihood function is

$$\begin{aligned}
\ell_{\mathrm{t}}(\gamma) &= \frac{1}{n}\sum_{i=1}^{n}\left[\sum_{a=1}^{m-1} I(A_i = a)U_i^{\mathrm{T}}\gamma_a - \log\left\{1 + \sum_{a=1}^{m-1}\exp(U_i^{\mathrm{T}}\gamma_a)\right\}\right] \\
&= \frac{1}{n}\sum_{i=1}^{n}\left[\sum_{a=1}^{m-1} I(A_i = a)U_i^{\mathrm{T}}\Phi_a\gamma - \log\left\{1 + \sum_{a=1}^{m-1}\exp(U_i^{\mathrm{T}}\Phi_a\gamma)\right\}\right],
\end{aligned}$$

where $U_i = u(X_i)$, q is the dimension of U_i, and $\Phi_1 = \left(I_q \mid 0_{q\times(m-2)q}\right), \Phi_2 = \left(0_{q\times q} \mid I_q \mid 0_{q\times(m-3)q}\right), \ldots, \Phi_{m-1} = \left(0_{q\times(m-2)q} \mid I_q\right)$ are $(m-1)$ matrices of size $q \times (m-1)q$ satisfying $\Phi_a\gamma = \gamma_a$. Hence we have

$$\begin{aligned}
\frac{\partial \ell_{\mathrm{t}}(\gamma)}{\partial\gamma} &= \frac{1}{n}\sum_{i=1}^{n}\left\{\sum_{a=1}^{m-1} I(A_i = a)\Phi_a^{\mathrm{T}}U_i - \frac{\sum_{a=1}^{m-1}\exp(U_i^{\mathrm{T}}\Phi_a\gamma)\Phi_a^{\mathrm{T}}U_i}{1 + \sum_{a=1}^{m-1}\exp(U_i^{\mathrm{T}}\Phi_a\gamma)}\right\}, \\
\frac{\partial^2 \ell_{\mathrm{t}}(\gamma)}{\partial\gamma\partial\gamma^{\mathrm{T}}} &= -\frac{1}{n}\sum_{i=1}^{n}\frac{\sum_{a=1}^{m-1}\exp(U_i^{\mathrm{T}}\Phi_a\gamma)\Phi_a^{\mathrm{T}}U_iU_i^{\mathrm{T}}\Phi_a}{1 + \sum_{a=1}^{m-1}\exp(U_i^{\mathrm{T}}\Phi_a\gamma)} \\
&\quad + \frac{1}{n}\sum_{i=1}^{n}\frac{\{\sum_{a=1}^{m-1}\exp(U_i^{\mathrm{T}}\Phi_a\gamma)\Phi_a^{\mathrm{T}}U_i\}\{\sum_{a=1}^{m-1}\exp(U_i^{\mathrm{T}}\Phi_a\gamma)U_i^{\mathrm{T}}\Phi_a\}}{\{1 + \sum_{a=1}^{m-1}\exp(U_i^{\mathrm{T}}\Phi_a\gamma)\}^2}. \qquad (1)
\end{aligned}$$

Denote γ_0 as the maximizer of $\mathbb{E}\ell_{\mathrm{t}}(\gamma)$. By the likelihood theory, we have

$$\sqrt{n}(\widehat{\gamma} - \gamma_0) = -\sqrt{n}\left[\mathbb{E}\left\{\frac{\partial^2\ell_{\mathrm{t}}(\gamma_0)}{\partial\gamma\partial\gamma^{\mathrm{T}}}\right\}\right]^{-1}\left\{\frac{\partial\ell_{\mathrm{t}}(\gamma_0)}{\partial\gamma}\right\} + o_p(1),$$

where the partial derivatives are given in (1), and $o_p(1)$ denotes a random quantity that convergences to zero in probability. Define

$$\varphi_\gamma(O_i) = -\left\{\mathbb{E}\left(\frac{\partial^2 \ell_{\mathrm{t}}(\gamma_0)}{\partial\gamma\partial\gamma^{\mathrm{T}}}\right)\right\}^{-1}\left\{\sum_{a=1}^{m-1} I(A_i=a)\Phi_a^{\mathrm{T}}U_i - \frac{\sum_{a=1}^{m-1}\exp(U_i^{\mathrm{T}}\Phi_a\gamma_0)\Phi_a^{\mathrm{T}}U_i}{1+\sum_{a=1}^{m-1}\exp(U_i^{\mathrm{T}}\Phi_a\gamma_0)}\right\}.$$

Then we have

$$\sqrt{n}(\widehat{\gamma}-\gamma_0) = \frac{1}{\sqrt{n}}\sum_{i=1}^n \varphi_\gamma(O_i) + o_p(1).$$

Next we derive an i.i.d. representation of $\widehat{\beta} = (\widehat{\beta}_1^{\mathrm{T}},\ldots,\widehat{\beta}_m^{\mathrm{T}})^{\mathrm{T}}$, the maximum likelihood estimator of $\beta = (\beta_1^{\mathrm{T}},\ldots,\beta_m^{\mathrm{T}})^{\mathrm{T}}$ in the generalized linear model:

$$g\left\{\mathbb{E}(Y_i|X_i,A_i)\right\} = \sum_{a=1}^m I(A_i=a)Z_i^{\mathrm{T}}\beta_a.$$

We assume that Y_i given A_i and X_i has an distribution in the exponential family with density function

$$f_{Y_i}(y_i) = \exp\left\{\frac{y_i\theta_i - b(\theta_i)}{\phi} + c(y_i,\phi)\right\},$$

where θ_i and ϕ are parameters, and $b(\cdot)$ and $c(\cdot,\cdot)$ are known functions. Note that for normal distribution ϕ is known as the dispersion parameter while for Bernoulli distribution ϕ is always equal to one. For simplicity we assume $g(\cdot)$ is a canonical link function hereafter. Then we have $b'(\cdot) \equiv g^{-1}(\cdot)$ and $\theta_i = \sum_{a=1}^m I(A_i=a)Z_i^{\mathrm{T}}\beta_a$. The log-likelihood function is

$$\begin{aligned}\ell_{\mathrm{o}}(\beta,\phi) &= \frac{1}{n}\sum_{i=1}^n\left[\frac{Y_i\sum_{a=1}^m I(A_i=a)Z_i^{\mathrm{T}}\beta_a - b\left\{\sum_{a=1}^m I(A_i=a)Z_i^{\mathrm{T}}\beta_a\right\}}{\phi} + c(Y_i,\phi)\right]\\ &= \frac{1}{n}\sum_{i=1}^n\left[\frac{Y_i\sum_{a=1}^m I(A_i=a)Z_i^{\mathrm{T}}\Psi_a\beta - b\left\{\sum_{a=1}^m I(A_i=a)Z_i^{\mathrm{T}}\Psi_a\beta\right\}}{\phi} + c(Y_i,\phi)\right],\end{aligned}$$

where r is the dimension of Z_i, and $\Psi_1 = \left(I_q \mid 0_{q\times(m-1)q}\right)$, $\Psi_2 = \left(0_{q\times q} \mid I_q \mid 0_{q\times(m-2)q}\right),\ldots,\Psi_m = \left(0_{q\times(m-1)q} \mid I_q\right)$ are m matrices of size $q\times mq$ satisfying $\Psi_a\beta = \beta_a$. Then we have

$$\begin{aligned}\frac{\partial\ell_{\mathrm{o}}(\beta,\phi)}{\partial\beta} &= \frac{1}{n\phi}\sum_{i=1}^n\left[Y_i - b'\left\{\sum_{a=1}^m I(A_i=a)Z_i^{\mathrm{T}}\Psi_a\beta\right\}\right]\left\{\sum_{a=1}^m I(A_i=a)\Psi_a^{\mathrm{T}}Z_i\right\},\\ \frac{\partial^2\ell_{\mathrm{o}}(\beta,\phi)}{\partial\beta\partial\beta^{\mathrm{T}}} &= -\frac{1}{n\phi}\sum_{i=1}^n b''\left\{\sum_{a=1}^m I(A_i=a)Z_i^{\mathrm{T}}\Psi_a\beta\right\}\left\{\sum_{a=1}^m I(A_i=a)\Psi_a^{\mathrm{T}}Z_iZ_i^{\mathrm{T}}\Psi_a\right\}.\end{aligned}$$

By the property of the score function, we have

$$\mathbb{E}\left(\frac{\partial^2\ell_{\mathrm{o}}(\beta_0,\phi_0)}{\partial\beta\partial\phi}\right) = -\frac{1}{\phi}\mathbb{E}\left(\frac{\partial\ell_{\mathrm{o}}(\beta_0,\phi_0)}{\partial\beta}\right) = 0.$$

Therefore, by the likelihood theory and the property of block diagonal matrix, we conclude that

$$\sqrt{n}(\widehat{\beta} - \beta_0) = -\sqrt{n}\left[\mathbb{E}\left\{\frac{\partial^2 \ell_{\mathrm{o}}(\beta_0, \phi_0)}{\partial\beta\partial\beta^{\mathrm{T}}}\right\}\right]^{-1}\left\{\frac{\partial \ell_{\mathrm{o}}(\beta_0, \phi_0)}{\partial\beta}\right\} + o_p(1)$$
$$= \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\varphi_\beta(O_i) + o_p(1),$$

where

$$\varphi_\beta(O_i) = \left(\mathbb{E}\left[b''\left(\sum_{a=1}^{m} I(A_i = a)Z_i^{\mathrm{T}}\Psi_a\beta_0\right)\left\{\sum_{a=1}^{m} I(A_i = a)\Psi_a^{\mathrm{T}} Z_i Z_i^{\mathrm{T}}\Psi_a\right\}\right]\right)^{-1}$$
$$\cdot\left\{Y_i - b'\left(\sum_{a=1}^{m} I(A_i = a)Z_i^{\mathrm{T}}\Psi_a\beta_0\right)\right\}\left\{\sum_{a=1}^{m} I(A_i = a)\Psi_a^{\mathrm{T}} Z_i\right\}.$$

Finally we derive an i.i.d. representation of $\widehat{R}(\pi)$. To emphasize the dependence of $\omega(x, a)$ and $\mu(x, a)$ on the parameters γ and β, in the following we write $\omega(x, a)$ as $\omega(x, a, \gamma)$ and $\mu(x, a)$ as $\mu(x, a, \beta)$. Thus we have $\widehat{\omega}(x, a) = \omega(x, a, \widehat{\gamma})$ and $\widehat{\mu}(x, a) = \mu(x, a, \widehat{\beta})$. Note that

$$\omega(x, a, \gamma) = \frac{\exp(u^{\mathrm{T}}\Phi_a\gamma)}{\sum_{j=1}^{m}\exp(u^{\mathrm{T}}\Phi_j\gamma)},$$
$$\mu(x, a, \beta) = b'(z^{\mathrm{T}}\Psi_a\beta),$$

for $a = 1, \ldots, m$, where $\Phi_m = 0_{q\times(m-1)q}$. Hence we have

$$\frac{\partial\omega(x, a, \gamma)}{\partial\gamma} = \frac{\exp(u^{\mathrm{T}}\Phi_a\gamma)\left\{\sum_{j=1}^{m}\exp(u^{\mathrm{T}}\Phi_j\gamma)\cdot(\Phi_a^{\mathrm{T}} - \Phi_j^{\mathrm{T}})u\right\}}{\left\{\sum_{j=1}^{m}\exp(u^{\mathrm{T}}\Phi_j\gamma)\right\}^2},$$
$$\frac{\partial\mu(x, a, \beta)}{\partial\beta} = b''(z^{\mathrm{T}}\Psi_a\beta)\Psi_a^{\mathrm{T}} z. \tag{2}$$

By Taylor expansion, we have

$$\widehat{R}(\pi) = \frac{1}{n}\sum_{i=1}^{n}\sum_{a=1}^{m}\left[\frac{I(A_i = a)}{\omega(X_i, a, \widehat{\gamma})}\left\{Y_i - \mu(X_i, a, \widehat{\beta})\right\} + \mu(X_i, a, \widehat{\beta})\right] I\{\pi(X_i) = a\}$$
$$= \frac{1}{n}\sum_{i=1}^{n}\sum_{a=1}^{m}\left[\frac{I(A_i = a)}{\omega(X_i, a, \gamma_0)}\{Y_i - \mu(X_i, a, \beta_0)\} + \mu(X_i, a, \beta_0)\right] I\{\pi(X_i) = a\}$$
$$+ \frac{1}{n}\sum_{i=1}^{n}\sum_{a=1}^{m}\left[-\frac{I(A_i = a)}{\omega^2(X_i, a, \gamma_0)}\{Y_i - \mu(X_i, a, \beta_0)\} I\{\pi(X_i) = a\}\frac{\partial\omega(X_i, a, \gamma_0)}{\partial\gamma}\right]^{\mathrm{T}}(\widehat{\gamma} - \gamma_0)$$
$$+ \frac{1}{n}\sum_{i=1}^{n}\sum_{a=1}^{m}\left[\left\{-\frac{I(A_i = a)}{\omega(X_i, a, \gamma_0)} + 1\right\} I\{\pi(X_i) = a\}\frac{\partial\mu(X_i, a, \beta_0)}{\partial\beta}\right]^{\mathrm{T}}(\widehat{\beta} - \beta_0) + o_p(1)$$
$$= \frac{1}{n}\sum_{i=1}^{n}\sum_{a=1}^{m}\left[\frac{I(A_i = a)}{\omega(X_i, a, \gamma_0)}\{Y_i - \mu(X_i, a, \beta_0)\} + \mu(X_i, a, \beta_0)\right] I\{\pi(X_i) = a\}$$

$$+\mathbb{E}\left(\sum_{a=1}^{m}\left[-\frac{I(A_i=a)}{\omega^2(X_i,a,\gamma_0)}\left\{Y_i-\mu(X_i,a,\beta_0)\right\}I\{\pi(X_i)=a\}\frac{\partial\omega(X_i,a,\gamma_0)}{\partial\gamma}\right]\right)^{\mathrm{T}}(\widehat{\gamma}-\gamma_0)$$
$$+\mathbb{E}\left(\sum_{a=1}^{m}\left[\left\{-\frac{I(A_i=a)}{\omega(X_i,a,\gamma_0)}+1\right\}I\{\pi(X_i)=a\}\frac{\partial\mu(X_i,a,\beta_0)}{\partial\beta}\right]\right)^{\mathrm{T}}(\widehat{\beta}-\beta_0)+o_p(1).$$

Recall that

$$R(\pi)=\mathbb{E}\left(\sum_{a=1}^{m}\left[\frac{I(A_i=a)}{\omega(X_i,a,\gamma_0)}\left\{Y_i-\mu(X_i,a,\beta_0)\right\}+\mu(X_i,a,\beta_0)\right]I\{\pi(X_i)=a\}\right).$$

Define

$$\begin{aligned}\varphi_R(O_i)=&\sum_{a=1}^{m}\left[\frac{I(A_i=a)}{\omega(X_i,a,\gamma_0)}\left\{Y_i-\mu(X_i,a,\beta_0)\right\}+\mu(X_i,a,\beta_0)\right]I\{\pi(X_i)=a\}\\&-\mathbb{E}\left(\sum_{a=1}^{m}\left[\frac{I(A_i=a)}{\omega(X_i,a,\gamma_0)}\left\{Y_i-\mu(X_i,a,\beta_0)\right\}+\mu(X_i,a,\beta_0)\right]I\{\pi(X_i)=a\}\right)\\&+\mathbb{E}\left(\sum_{a=1}^{m}\left[-\frac{I(A_i=a)}{\omega^2(X_i,a,\gamma_0)}\left\{Y_i-\mu(X_i,a,\beta_0)\right\}I\{\pi(X_i)=a\}\frac{\partial\omega(X_i,a,\gamma_0)}{\partial\gamma}\right]\right)^{\mathrm{T}}\varphi_\gamma(O_i)\\&+\mathbb{E}\left(\sum_{a=1}^{m}\left[\left\{-\frac{I(A_i=a)}{\omega(X_i,a,\gamma_0)}+1\right\}I\{\pi(X_i)=a\}\frac{\partial\mu(X_i,a,\beta_0)}{\partial\beta}\right]\right)^{\mathrm{T}}\varphi_\beta(O_i),\end{aligned}\tag{3}$$

where $\partial\omega/\partial\gamma$ and $\partial\mu/\partial\beta$ are given in (2). Then we have

$$\sqrt{n}\left\{\widehat{R}(\pi)-R(\pi)\right\}=\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\varphi_R(O_i)+o_p(1).$$

Therefore, by the central limit theorem and the Slutsky's theorem, we conclude that

$$\sqrt{n}\left\{\widehat{R}(\pi)-R(\pi)\right\}\xrightarrow{d}N(0,\mathbb{E}\left\{\varphi_R^2(O_i)\right\}),\tag{4}$$

where $\xrightarrow{d}$ denotes convergence in distribution.

To estimate the asymptotic variance, we use the plug-in method. Namely, define $\widehat{\varphi}_R(O_i)$ as in (3) except that expectations are replaced with sample averages and true values are replaced with corresponding estimates. Then $\mathrm{Var}\left(\widehat{R}(\pi)\right)$ can be estimated by $\sum_{i=1}^{n}\widehat{\varphi}_R^2(O_i)/n^2$.

3. Asymptotic Properties of $\widehat{R}(\pi_1) - \widehat{R}(\pi_2)$

Define $\varphi_{R1}(O_i)$ as in (3) with π replaced by π_1. Define $\varphi_{R2}(O_i)$ as in (3) with π replaced by π_2. Define $\widehat{\varphi}_{R1}(O_i)$ and $\widehat{\varphi}_{R2}(O_i)$ similarly. Then we have

$$\sqrt{n}\left[\left\{\widehat{R}(\pi_1) - \widehat{R}(\pi_2)\right\} - \{R(\pi_1) - R(\pi_2)\}\right] = \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\{\varphi_{R1}(O_i) - \varphi_{R2}(O_i)\} + o_p(1)$$
$$\xrightarrow{d} N(0, \mathbb{E}\{\varphi_{R1}(O_i) - \varphi_{R2}(O_i)\}^2).$$

Therefore, we can estimate $\mathrm{Var}\left\{\widehat{R}(\pi_1) - \widehat{R}(\pi_2)\right\}$ by

$$\widehat{\mathrm{Var}}\left\{\widehat{R}(\pi_1) - \widehat{R}(\pi_2)\right\} = \frac{1}{n^2}\sum_{i=1}^{n}\{\widehat{\varphi}_{R1}(O_i) - \widehat{\varphi}_{R2}(O_i)\}^2. \tag{5}$$

The variance estimator $\widehat{\mathrm{Var}}(\Delta_j)$ used in the algorithm in Section 2.4.1 in the main paper can be obtained via (5) with $\pi_1 = \left\{(\bar{c}_1, \bar{a}_1), \ldots, (\bar{c}_{j-1}, \bar{a}_{j-1}), (\widetilde{c}_j, \widetilde{a}_j), \widetilde{a}'_j\right\}$ and $\pi_2 = \left\{(\bar{c}_1, \bar{a}_1), \ldots, (\bar{c}_{j-1}, \bar{a}_{j-1}), \bar{a}'_{j-1}\right\}$.

4. Implementation Details of Finding an Optimal Decision List

4.1 *Algorithm Description*

We give an equivalent version of the proposed algorithm for finding an optimal decision list. Compared to the algorithm presented in the main paper, this version makes use of recursive calls to avoid explicit constructions of sets Π_{temp} and Π_{final}, and facilitates the analysis of time complexity. The algorithm is as follows.

Input: $\widehat{R}(\cdot)$, $L_{\max}$, α

Output: a decision list $\widetilde{\pi}$ that maximize $\widehat{R}(\cdot)$

$\widetilde{a}_0 = \arg\max_{a_0 \in \mathcal{A}} \widehat{R}\left[\{a_0\}\right]$;

$\widetilde{\pi} = \texttt{FindList}(1, \{\}, \widetilde{a}_0)$;

The function `FindList` is defined below. When $j = 1$, we treat $(\bar{c}_1, \bar{a}_1), \ldots, (\bar{c}_{j-1}, \bar{a}_{j-1})$

as an empty array. Thus when $j = 1$, $\big\{(\bar{c}_1, \bar{a}_1), \ldots, (\bar{c}_{j-1}, \bar{a}_{j-1}), (\widetilde{c}_j, \widetilde{a}_j), \widetilde{a}'_j\big\}$ is the same as $\big\{(\widetilde{c}_1, \widetilde{a}_1), \widetilde{a}'_1\big\}$ and $\big\{(\bar{c}_1, \bar{a}_1), \ldots, (\bar{c}_{j-1}, \bar{a}_{j-1}), \bar{a}'_{j-1}\big\}$ is the same as $\{\bar{a}'_0\}$.

Function `FindList` $(j,\ \{(\bar{c}_1, \bar{a}_1), \ldots, (\bar{c}_{j-1}, \bar{a}_{j-1})\},\ \bar{a}'_{j-1})$

$(\widetilde{c}_j, \widetilde{a}_j, \widetilde{a}'_j) = \arg\max_{(c_j, a_j, a'_j) \in \mathcal{C} \times \mathcal{A} \times \mathcal{A}} \widehat{R}\left[\big\{(\bar{c}_1, \bar{a}_1), \ldots, (\bar{c}_{j-1}, \bar{a}_{j-1}), (c_j, a_j), a'_j\big\}\right]$;

$\widehat{\Delta}_j = \widehat{R}\left[\big\{(\bar{c}_1, \bar{a}_1), \ldots, (\bar{c}_{j-1}, \bar{a}_{j-1}), (\widetilde{c}_j, \widetilde{a}_j), \widetilde{a}'_j\big\}\right] - \widehat{R}\left[\big\{(\bar{c}_1, \bar{a}_1), \ldots, (\bar{c}_{j-1}, \bar{a}_{j-1}), \bar{a}'_{j-1}\big\}\right]$;

if $\widehat{\Delta}_j < z_{1-\alpha}\big\{\widehat{\mathrm{Var}}\big(\widehat{\Delta}_j\big)\big\}^{1/2}$ **then**

$\widetilde{\pi} = \big\{(\bar{c}_1, \bar{a}_1), \ldots, (\bar{c}_{j-1}, \bar{a}_{j-1}), \bar{a}'_{j-1}\big\}$;

else if $j = L_{\max}$ **then**

$\widetilde{\pi} = \big\{(\bar{c}_1, \bar{a}_1), \ldots, (\bar{c}_{j-1}, \bar{a}_{j-1}), (\widetilde{c}_j, \widetilde{a}_j), \widetilde{a}'_j\big\}$;

else

$\widetilde{\pi}_1 =$ `FindList`$(j+1,\ \{(\bar{c}_1, \bar{a}_1), \ldots, (\bar{c}_{j-1}, \bar{a}_{j-1}), (\widetilde{c}_j, \widetilde{a}_j)\},\ \widetilde{a}'_j)$;

$\widetilde{\pi}_2 =$ `FindList`$(j+1,\ \big\{(\bar{c}_1, \bar{a}_1), \ldots, (\bar{c}_{j-1}, \bar{a}_{j-1}), (\widetilde{c}'_j, \widetilde{a}'_j)\big\},\ \widetilde{a}_j)$,

where $\widetilde{c}'_j =$ negation of $\widetilde{c}_j$;

$\widetilde{\pi} = \arg\max_{\pi \in \{\widetilde{\pi}_1, \widetilde{\pi}_2\}} \widehat{R}(\pi)$;

end

return $\widetilde{\pi}$;

end

In the `FindList` function, a crucial step is to compute $(\widetilde{c}_j, \widetilde{a}_j, \widetilde{a}'_j)$. A straightforward implementation that involves a brute-force search over $\mathcal{C} \times \mathcal{A} \times \mathcal{A}$ can be time consuming. We provide an efficient implementation below.

We observe that some calculations can be performed only once at the beginning of the algorithm. First, define

$$\widehat{\xi}_{ia} = \frac{I(A_i = a)}{\omega(X_i, a, \widehat{\gamma})}\Big\{Y_i - \mu(X_i, a, \widehat{\beta})\Big\} + \mu(X_i, a, \widehat{\beta}).$$

Then we have

$$\widehat{R}(\pi) = \frac{1}{n}\sum_{i=1}^{n}\sum_{a=1}^{m}\widehat{\xi}_{ia}I\{\pi(X_i)=a\}.$$

Second, for the ith subject, denote x_{ij} as the observed value of his/her jth covariate. For the jth baseline covariate, there are $s_k = \#\mathcal{X}_j$ possible candidate cutoff values $\tau_{j1} \leqslant \cdots \leqslant \tau_{js_j}$, which divides the real line into $s_k + 1$ intervals:

$$(-\infty, \tau_{j1}],\ (\tau_{j1}, \tau_{j2}],\ \ldots,\ (\tau_{j(s_j-1)}, \tau_{js_j}],\ (\tau_{js_j}, \infty).$$

Then we code the observed values $x_{1j}, \ldots, x_{nj}$ into indices $b_{1j}, \ldots, b_{nj}$ according to which interval they fall.

In order to reduce the number of evaluations of $\widehat{R}(\cdot)$ when searching for the maximizer over $\mathcal{C}\times\mathcal{A}\times\mathcal{A}$, we organize the intermediate results as shown below. Let $\mathcal{I} = \{i : X_i \in \mathcal{T}(\bar{c}_\ell)^c \text{ for all } \ell < j\}$. Then $\mathcal{I}$ contains all the subjects that have not had treatment recommendations up to the jth clause. Since we have

$$n\widehat{R}(\pi) = \sum_{i\in\mathcal{I}}\sum_{a=1}^{m}\widehat{\xi}_{ia}I\{\pi(X_i)=a\} + \sum_{i\in\mathcal{I}^c}\sum_{a=1}^{m}\widehat{\xi}_{ia}I\{\pi(X_i)=a\}$$

and $\sum_{i\in\mathcal{I}^c}\sum_{a=1}^{m}\widehat{\xi}_{ia}I\{\pi(X_i)=a\}$ is constant during the maximization, we focus on maximizing $\sum_{i\in\mathcal{I}}\sum_{a=1}^{m}\widehat{\xi}_{ia}I\{\pi(X_i)=a\}$, which reduces to maximizing

$$\sum_{i\in\mathcal{I}}\sum_{a=1}^{m}\widehat{\xi}_{ia}I\{i\in\mathcal{T}(c_j), a=a_j\} + \sum_{i\in\mathcal{I}}\sum_{a=1}^{m}\widehat{\xi}_{ia}I\{i\notin\mathcal{T}(c_j), a=a'_j\}. \tag{6}$$

To identify the maximizer of (6), we first loop over all possible pairs of covariates. For each pair of covariates, say the kth and the ℓth covariates, define D, a three-dimensional array of size $m\times(s_k+1)\times(s_\ell+1)$, as $D_{auv} = \sum_{i\in\mathcal{I}}\widehat{\xi}_{ia}I(b_{ik}=u, b_{il}=v)$. Next, we loop over all possible cutoff values and construct the corresponding c_j. The values of a_j and a'_j that maximizes (6) for a given c_j can be easily obtained due to the additive structure. After enumerating all the possible conditions that c_j may take, we can find out $(\widetilde{c}_j, \widetilde{a}_j, \widetilde{a}'_j)$.

4.2 *Time Complexity Analysis*

Since computing $\widehat{\xi}_{ia}$s requires $O(nm)$ time and computing b_{ij}s requires $O(np)$ time. The calculations at the beginning of the algorithm take $O(nm+np)$ time in total.

The algorithm first computes $\widetilde{a}_0$, which requires $O(nm)$ time. Then it invokes a function call `FindList`$(1, \{\}, \widetilde{a}_0)$. Due to the recursive nature of the `FindList` function, we will compute the time complexity by establishing a recurrence relation between $T(j)$ and $T(j+1)$, where $T(j)$ is the time complexity of the function call `FindList` $(j,\ \{(\bar{c}_1, \bar{a}_1), \ldots, (\bar{c}_j, \bar{a}_j)\})$.

Suppose a call `FindList` $(j,\ \{(\bar{c}_1, \bar{a}_1), \ldots, (\bar{c}_j, \bar{a}_j)\})$ is invoked. The running time can be computed by going through the algorithm of the `FindList` function step-by-step as follows.

First, the function computes $(\widetilde{c}_j, \widetilde{a}_j, \widetilde{a}'_j)$. A naive implementation would involve looping over all the covariates, all the possible cutoff values and all the treatment options, whose running time is $O(nmp^2s^2)$, where $s = \max_j s_j$. However, the running time is greatly reduced if we use the efficient implementation described previously. For a given pair of covariates, we can compute D_{auv}s in $O(nm)$ time. Then we can find out the maximum of (6) in $O(ms^2)$ time by looping over all possible cutoff values. Therefore, the total time for computing $(\widetilde{c}_j, \widetilde{a}_j, \widetilde{a}'_j)$ is $O\{(n+s^2)mp^2\}$.

Second, the function computes $\widehat{\Delta}_j$, which takes $O(n)$ time.

Third, the function computes $\widehat{\mathrm{Var}}\big(\widehat{\Delta}_j\big)$, whose running time is $O(nmq+nmr)$, where q is the dimension of U_i and r is the dimension of Z_i. Since both U_i and Z_i are known feature vectors constructed from X_i, for most cases q and r are of the same order as p. So this step takes $O(nmp)$ time.

Fourth, the function executes the "if-then" statement. In the worst case, the function makes two recursive calls, taking $2T(j+1)$ time.

Combining these four steps, we have $T(j) = O\{(n+s^2)mp^2\} + 2T(j+1)$. The boundary condition is $T(L_{\max}) = O\{(n+s^2)mp^2\}$. Using backward induction, we get $T(0) = O\{2^{L_{\max}}(n+s^2)mp^2\}$. Recall that $s = \max_j \#\mathcal{X}_j$.

Combining $T(0)$ with the running time before invoking `FindList`$(1, \{\}, \widetilde{a}_0)$, we obtain that the time complexity of the entire algorithm is $O[2^{L_{\max}} mp^2\{n + (\max_j \#\mathcal{X}_j)^2\}]$.

5. Implementation Details of Finding an Equivalent Decision List with Minimal Cost

In this section we give an algorithmic description of the proposed method for finding an equivalent decision list with minimal cost. Recall that two decision lists are called equivalent if they give the same treatment recommendation for every patient in the population.

Input: a decision list $\bar{\pi}$

Output: an equivalent decision list $\pi_{\min}$ with minimal cost $N_{\min}$

Identify atoms in $\bar{\pi}$ as $d_1, \ldots, d_K$;

Compute $\mathcal{I}_a = \{i : \bar{\pi}(X_i) = a\}$ for each $a \in \mathcal{A}$;

Set $\pi_{\min} = \{\}$ and $N_{\min} = \infty$;

`FindMinCost` $(0, \{\}, \pi_{\min}, N_{\min})$;

The function `FindMinCost` is defined below.

Function `FindMinCost` *(j, $\{(c_1, a_1), \ldots, (c_j, a_j)\}$, π_{min}, N_{min})*

Compute a lower bound of the cost as $N_{\text{bd}} = \mathcal{N}_\ell \sum_{\ell=1}^{j} \mathbb{P}_n(X \in \mathcal{R}_\ell)$

$+ \mathcal{N}_j \mathbb{P}_n(X \in \cap_{\ell=1}^{j} \mathcal{R}_\ell^c)$, where $\mathbb{P}_n$ denotes the empirical probability measure;

if $N_{\text{bd}} \geqslant N_{\text{min}}$ **then return**;

$\mathcal{I} = \{i : X_i \in \mathcal{T}(c_\ell)^c$ for all $\ell \leqslant j\}$;

if $\mathcal{I} \subset \mathcal{I}_{a_0}$ *for some* a_0 **then**

  if $N[\{(c_1, a_1), \ldots, (c_j, a_j), a_0\}] < N_{\text{min}}$ **then**

    $\pi_{\text{min}} = \{(c_1, a_1), \ldots, (c_j, a_j), a_0\}$;

    $N_{\text{min}} = N(\pi_{\text{min}})$;

  end

else

  for $1 \leqslant k_1 < k_2 \leqslant K$ **do**

    Let $\mathcal{C}_{k_1,k_2}$ be the set consisting of all the logical clauses involving

    d_{k_1} or d_{k_2} or both using conjunction, disjunction, and/or negation;

    for $c_{j+1} \in \mathcal{C}_{k_1,k_2}$ **do**

      $\mathcal{J}_{j+1} = \{i \in \mathcal{I} : X_i \in \mathcal{T}(c_{j+1})\}$;

      if $\mathcal{J}_{j+1}$ *is non-empty and* $\mathcal{J}_{j+1} \subset \mathcal{I}_{a_{j+1}}$ *for some* $a_{j+1} \in \mathcal{A}$ **then**

        `FindMinCost` ($j+1$, $\{(c_1, a_1), \ldots, (c_j, a_j), (c_{j+1}, a_{j+1})\}$, π_{min}, N_{min}) ;

    end

  end

end

end

6. Point Estimate and Prediction Interval for $R(\widehat{\pi})$ with Bootstrap Bias Correction

In this section, we show how to estimate the value of the estimated treatment regime, $R(\widehat{\pi})$, and how to construct a prediction interval for it.

6.1 *Methodology*

To measure how well the estimated treatment regime $\widehat{\pi}$ performs, it is often of interest to construct an estimator of and a prediction interval for $R(\widehat{\pi})$. Since a natural candidate for estimating $R(\widehat{\pi})$ is $\widehat{R}(\widehat{\pi})$, it may be tempting to construct a prediction interval centering at $\widehat{R}(\widehat{\pi})$. However, $\widehat{R}(\widehat{\pi})$ is generally too optimistic to serve as an honest estimator of $R(\widehat{\pi})$. It has an upward bias due to the maximization process. As a remedy, we suggest using B bootstraps to correct this bias. Specifically, the perturbed version of $\widehat{R}(\pi)$ in the bth bootstrapping sample is

$$\widehat{R}_b^*(\pi) = \frac{1}{n}\sum_{i=1}^{n}\left(W_i \sum_{a=1}^{m}\left[\frac{I(A_i=a)}{\omega(X_i,a,\widehat{\gamma}^*)}\left\{Y_i-\mu(X_i,a,\widehat{\beta}^*)\right\}+\mu(X_i,a,\widehat{\beta}^*)\right] I\{\pi(X_i)=a\}\right),$$

where $W_1,\dots,W_n$ are identically and independently distributed with standard exponential distribution, $\widehat{\gamma}^*$ is the solution to

$$\sum_{i=1}^{n} W_i\left\{\sum_{a=1}^{m-1} I(A_i=a)\Phi_a^{\mathrm{T}}U_i - \frac{\sum_{a=1}^{m-1}\exp(U_i^{\mathrm{T}}\Phi_a\gamma)\Phi_a^{\mathrm{T}}U_i}{1+\sum_{a=1}^{m-1}\exp(U_i^{\mathrm{T}}\Phi_a\gamma)}\right\}=0,$$

and $\widehat{\beta}^*$ is the solution to

$$\sum_{i=1}^{n} W_i\left[Y_i - b'\left\{\sum_{a=1}^{m} I(A_i=a)Z_i^{\mathrm{T}}\Psi_a\beta\right\}\right]\left\{\sum_{a=1}^{m} I(A_i=a)\Psi_a^{\mathrm{T}}Z_i\right\}=0.$$

Let $\widehat{\pi}_b^*$ be the maximizer of $\widehat{R}_b^*(\pi)$ over Π. Then the actual bias $\widehat{R}(\widehat{\pi})-R(\widehat{\pi})$ can be estimated by the corresponding bias in the bootstrap world: $\widehat{\mathrm{Bias}}=\sum_{b=1}^{B}\{\widehat{R}_b^*(\widehat{\pi}_b^*)-\widehat{R}(\widehat{\pi}_b^*)\}/B$, where B is the number of bootstrap samples. The final estimator of $R(\widehat{\pi})$ is $\widehat{R}_{\mathrm{c}}(\widehat{\pi})=\widehat{R}(\widehat{\pi})-\widehat{\mathrm{Bias}}$.

To construct a prediction interval for $R(\widehat{\pi})$, we treat $\widehat{\pi}$ as a non-random quantity, and then utilize the asymptotic normality of $\widehat{R}(\widehat{\pi})$ given in (4). Let z_ρ be the 100ρ percentile of a standard normal distribution and $\widehat{\sigma}^2=\widehat{\mathrm{Var}}\big\{\widehat{R}(\widehat{\pi})\big\}$. Then a $(1-\alpha)\times 100\%$ prediction interval for $R(\widehat{\pi})$ is

$$\big[\widehat{R}_{\mathrm{c}}(\widehat{\pi}) + z_{\alpha/2}\widehat{\sigma},\ \widehat{R}_{\mathrm{c}}(\widehat{\pi}) + z_{1-\alpha/2}\widehat{\sigma}\big]. \tag{7}$$

A potential drawback of this interval is, though, that it ignores the variation introduced by $\widehat{\pi}$. Nevertheless, our numerical experiences suggest that this extra variation is generally small and the coverage probability is close to the nominal level. Taking into account the variability of $\widehat{\pi}$ has to deal with the associated non-regularity issue, which is beyond the scope of this paper.

As a final remark, for binary outcome we suggest to conduct the bias correction and construct the prediction interval based on $\mathrm{logit}\{\widehat{R}(\cdot)\}$ first and then transform back to the original scale, where $\mathrm{logit}(v) = \log\{v/(1-v)\}$.

6.2 *Simulations*

We present the point estimate and the coverage probabilities of the plain prediction interval and the prediction interval with bootstrap bias correction in Table 1. The setting used here is exactly the same as that in Section 3 in the main paper. We can see that the bias correction improves the coverage probability substantially in finite samples, especially as the number of covariates gets larger. Besides, the bootstrap prediction interval is prone to overcoverage for the binary response.

[Table 1 about here.]

7. Accuracy of Variable Selection

Consider the simulated experiments in the main paper. To quantify variable selection accuracy, we compute the true positive rate, the number of signal variables included in the decision list divided by the number of signal variables, and the false positive rate, the number of noise variables included in the decision list divided by the number of noise variables. A variable is called a signal variable if it appears in $\phi(x, a)$ and is a noise variable otherwise, irrespective of the actual functional form.

Table 2 presents the true positive rates and the false positive rates under different settings. The proposed method consistently identifies signal variables and screens out noise variables in most settings. The only exception is setting IV, where the optimal regime is far away from being well approximated by decision lists. Thus the proposed approach loses power in detecting useful covariates due to misspecifying the form of the regime.

[Table 2 about here.]

8. Impact of the Tuning Parameter in the Stopping Criterion

In the algorithm discussed in Section 2.4.1 in the main paper, we use a tuning parameter α to control the building process of the decision list and we suggest to fix α at 0.95. In the following we show that the final decision list is insensitive to the choice of α via simulation study. The setting used here is exactly the same as that in Section 3 in the main paper. We varied α among $\{0.9, 0.95, 0.99\}$.

Table 3 shows the impact of α on the value and the cost of the estimated regime. We can see that the value and the cost as well as the accuracy of variable selection, averaged over 1000 replications, are very stable across different choices of α. Table 4 shows the impact of α on the estimated regime. It is clear that α has little impact on the treatment recommendation made by the estimated regime.

[Table 3 about here.]

[Table 4 about here.]

9. Chronic Depression Data

In the application considered in Section 4.2 in the main paper, we applied the proposed method to construct an interpretable and parsimonious treatment regime. We follow Gunter et al. (2011) and Zhao et al. (2012), and use the following 50 covariates:

(1) Gender: 1 if female, 0 if male;

(2) Race: 1 if white, 0 otherwise;

(3) Marital status I: 1 if single, 0 otherwise;

(4) Marital status II: 1 if married or living with someone, 0 otherwise;

(5) Body mass index: continuous;

(6) Age at screening: continuous;

(7) Having difficulty in planning family activity: 1 if strongly agree, 2 if agree, 3 if disagree, 4 if strongly disagree;

(8) Supporting each other in the family: 1 if strongly agree, 2 if agree, 3 if disagree, 4 if strongly disagree;

(9) Having problems with primary support group: 1 if yes, 0 if no;

(10) Having problems related to the social environment: 1 if yes, 0 if no;

(11) Having occupational problems: 1 if yes, 0 if no;

(12) Having economic problems: 1 if yes, 0 if no;

(13) Receiving psychotherapy for current depression: 1 if yes, 0 if no or don't know;

(14) Receiving medication for current depression: 1 if yes, 0 if no or don't know;

(15) Having received psychotherapy for past depressions: 1 if yes, 0 if no or don't know;

(16) Having received medication for past depressions: 1 if yes, 0 if no or don't know;

(17) Number of major depressive disorder (MDD) episodes: 1 if one, 2 if two, 3 if at least three;

(18) Length of current MDD episode (in years): continuous;

(19) Age at MDD onset: continuous;

(20) MDD severity: 1 if mild, 2 if moderate, 3 if severe;

(21) MDD type I: 1 if melancholic, 0 otherwise;

(22) MDD type II: 1 if atypical, 0 otherwise;

(23) Number of dysthymia episodes: 0 if zero, 1 if one, 2 if at least two;

(24) Generalized anxiety: 0 if absent or inadequate information, 1 if subthreshold, 2 if threshold;

(25) Anxiety disorder (not otherwise specified): 0 if absent or inadequate information, 1 if subthreshold, 2 if threshold;

(26) Panic disorder: 0 if absent or inadequate information, 1 if subthreshold, 2 if threshold;

(27) Social phobia: 0 if absent or inadequate information, 1 if subthreshold, 2 if threshold;

(28) Specific phobia: 0 if absent or inadequate information, 1 if subthreshold, 2 if threshold;

(29) Obsessive compulsive: 0 if absent or inadequate information, 1 if subthreshold, 2 if threshold;

(30) Body dysmorphic disorder: 0 if absent or inadequate information, 1 if subthreshold, 2 if threshold;

(31) Post-traumatic stress disorder: 0 if absent or inadequate information, 1 if subthreshold, 2 if threshold;

(32) Anorexia nervosa: 0 if absent or inadequate information, 1 if subthreshold, 2 if threshold;

(33) Alcohol abuse: 0 if absent, 1 if abuse, 2 if dependent;

(34) Drug abuse (including cannabis, stimulant, opioid, cocaine, hallucinogen): 0 if absent, 1 if abuse, 2 if dependent;

(35) Global assessment of functioning: continuous;

(36) Chronic depression diagnosis I: 1 if no antecedent dysthymia and continuous full-syndrome type;

(37) Chronic depression diagnosis II: 1 if no antecedent dysthymia and incomplete recovery type;

(38) Chronic depression diagnosis III: 1 if superimposed on antecedent dysthymia;

(39) Chronic depression severity: integer between 1 (normal) and 7 (extremely ill);

(40) Hamilton anxiety rating scale (HAM-A) total score: continuous;

(41) HAM-A psychic anxiety score: continuous;

(42) HAM-A somatic anxiety score: continuous;

(43) Hamilton depression rating scale (HAM-D) total score: continuous;

(44) HAM-D anxiety/somatic score: continuous;

(45) HAM-D cognitive disturbance score: continuous;

(46) HAM-D retardation score: continuous;

(47) HAM-D sleep disturbance: continuous;

(48) Inventory of Depressive Symptoms - Self Report (IDS-SR) anxiety/arousal score: continuous;

(49) IDS-SR general/mood cognition score: continuous;

(50) IDS-SR sleep score: continuous.

10. Consistency of the decision list

Since the consistency of the decision list is difficult to analyze theoretically, we present some empirical evidence that the decision list tends to be consistent. We follow the simulated experiments considered in Section 4 in the main paper but increase the sample size. We consider settings I and V only as the optimal regime in other settings cannot be representable as a decision list.

The sample sizes considered and the associated results are presented in Table 5. For continuous response, the proposed method correctly identifies the form and the important covariates for treatment decision. As n increases, the loss in value decreases and the probability of recommending the best treatment increases. Also, the mean squared error of estimating the cutoff values decreases at the rate of n^{-1}. Results for binary response is qualitatively similar. Nevertheless, we may need a even larger sample size for the asymptotics to work. Therefore, the simulation results provides evidence that the proposed method is consistent.

[Table 5 about here.]

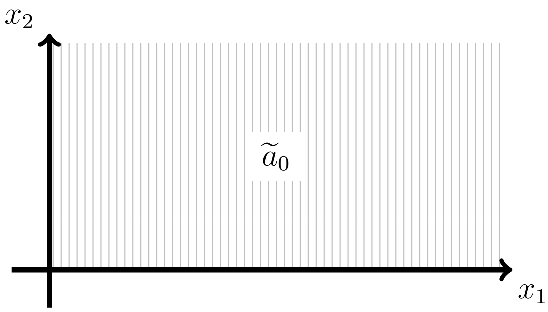
References

Gunter, L., Zhu, J., and Murphy, S. A. (2011). Variable selection for qualitative interactions. *Statistical Methodology* **8,** 42–55.

Zhao, Y., Zeng, D., Rush, A. J., and Kosorok, M. R. (2012). Estimating individualized treatment rules using outcome weighted learning. *Journal of the American Statistical Association* **107,** 1106–1118.

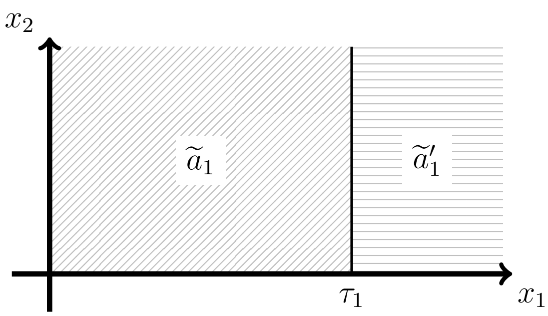
Received January 1900. *Revised February* 1900.

Accepted March 1900.



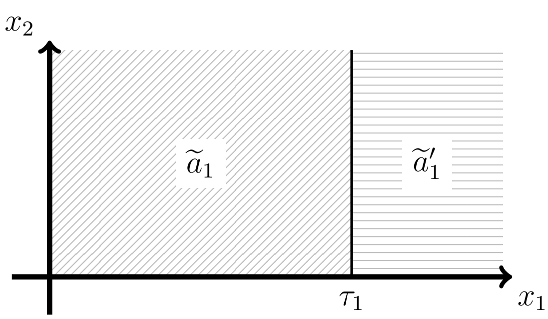
Everyone $\widetilde{a}_0$.

Figure 1. Diagram and description of the decision list $\{\widetilde{a}_0\}$.



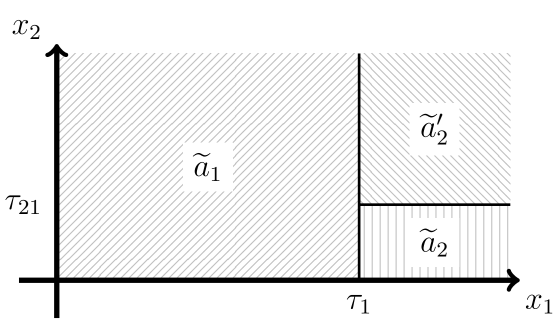
If $x_1 \leqslant \tau_1$ **then** $\widetilde{a}_1$;
else $\widetilde{a}'_1$.

Figure 2. Diagram and description of the decision list $\{(\widetilde{c}_1, \widetilde{a}_1), \widetilde{a}'_1\}$.



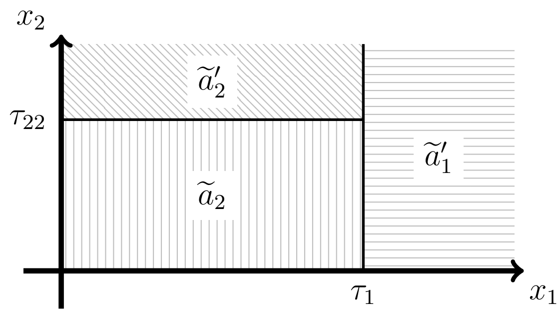
If $x_1 > \tau_1$ **then** $\widetilde{a}'_1$;
else $\widetilde{a}_1$.

Figure 3. Diagram and description of the decision list $\{(\widetilde{c}'_1, \widetilde{a}'_1), \widetilde{a}_1\}$.



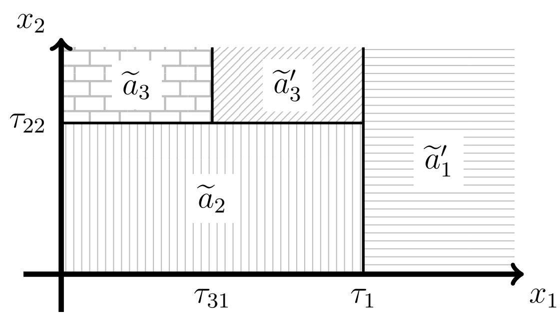
If $x_1 \leqslant \tau_1$ **then** $\widetilde{a}_1$;
else if $x_2 \leqslant \tau_{21}$ **then** $\widetilde{a}_2$;
else $\widetilde{a}'_2$.

Figure 4. Diagram and description of the decision list $\{(\widetilde{c}_1, \widetilde{a}_1), (\widetilde{c}_2, \widetilde{a}_2), \widetilde{a}'_2\}$. It is possible that $\widetilde{a}_2 = \widetilde{a}_1$ or $\widetilde{a}'_2 = \widetilde{a}_1$.



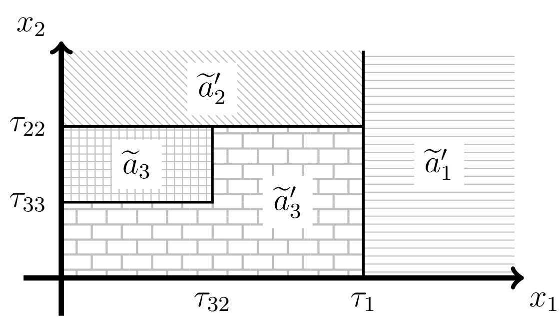
If $x_1 > \tau_1$ **then** $\widetilde{a}_1'$;
else if $x_2 \leqslant \tau_{22}$ **then** $\widetilde{a}_2$;
else $\widetilde{a}_2'$.

Figure 5. Diagram and description of the decision list $\{(\widetilde{c}_1', \widetilde{a}_1'), (\widetilde{c}_2, \widetilde{a}_2), \widetilde{a}_2'\}$. It is possible that $\widetilde{a}_2 = \widetilde{a}_1'$ or $\widetilde{a}_2' = \widetilde{a}_1'$.



If $x_1 > \tau_1$ **then** $\widetilde{a}_1'$;
else if $x_2 \leqslant \tau_{22}$ **then** $\widetilde{a}_2$;
else if $x_1 \leqslant \tau_{31}$ **then** $\widetilde{a}_3$;
else $\widetilde{a}_3'$.

Figure 6. Diagram and description of the decision list $\{(\widetilde{c}_1', \widetilde{a}_1'), (\widetilde{c}_2, \widetilde{a}_2), (\widetilde{c}_3, \widetilde{a}_3), \widetilde{a}_3'\}$. Some of the values of $\widetilde{a}_1', \widetilde{a}_2, \widetilde{a}_3, \widetilde{a}_3'$ can be equal.



If $x_1 > \tau_1$ **then** $\widetilde{a}_1'$;
else if $x_2 > \tau_{22}$ **then** $\widetilde{a}_2'$;
else if $x_1 \leqslant \tau_{32}$ and $x_2 > \tau_{33}$ **then** $\widetilde{a}_3$;
else $\widetilde{a}_3'$.

Figure 7. Diagram and description of the decision list $\{(\widetilde{c}_1', \widetilde{a}_1'), (\widetilde{c}_2', \widetilde{a}_2'), (\widetilde{c}_3, \widetilde{a}_3), \widetilde{a}_3'\}$. Some of the values of $\widetilde{a}_1', \widetilde{a}_2', \widetilde{a}_3, \widetilde{a}_3'$ can be equal.

Table 1

Point estimate and coverage probabilities of prediction intervals with and without bootstrap bias correction. Plain-PI refers to the coverage probability of the plain prediction interval, and Corrected-PI refers to the coverage probability of the bias-corrected prediction interval.

p	Setting	Continuous response				Binary response			
		$R(\widehat{\pi})$	$\widehat{R}_c(\widehat{\pi})$	Plain-PI	Corrected-PI	$R(\widehat{\pi})$	$\widehat{R}_c(\widehat{\pi})$	Plain-PI	Corrected-PI
10	I	2.78	2.78	0.95	0.94	0.77	0.76	0.97	0.96
	II	2.70	2.73	0.93	0.95	0.71	0.72	0.89	0.97
	III	2.59	2.61	0.95	0.95	0.73	0.74	0.88	0.96
	IV	2.89	2.98	0.88	0.94	0.71	0.72	0.63	0.98
	V	2.90	2.90	0.95	0.95	0.75	0.75	0.76	0.96
	VI	3.98	4.01	0.93	0.95	0.79	0.79	0.97	0.99
	VII	3.22	3.27	0.86	0.94	0.77	0.77	0.77	1.00
50	I	2.76	2.75	0.94	0.94	0.76	0.76	0.82	0.98
	II	2.70	2.72	0.93	0.94	0.71	0.71	0.80	0.96
	III	2.59	2.59	0.94	0.95	0.73	0.73	0.63	0.98
	IV	2.89	2.96	0.88	0.94	0.71	0.72	0.48	0.98
	V	2.87	2.87	0.93	0.94	0.74	0.74	0.33	0.96
	VI	3.95	3.99	0.91	0.94	0.78	0.79	0.89	0.99
	VII	3.21	3.27	0.88	0.94	0.76	0.77	0.63	0.99

Table 2

Accuracy of variable selection using decision list. TPR is the true positive rate and FPR is the false positive rate.

p	Setting	Continuous response		Binary response	
		TPR	FPR	TPR	FPR
10	I	1.00	0.00	1.00	0.07
	II	1.00	0.00	1.00	0.04
	III	1.00	0.00	1.00	0.11
	IV	0.93	0.00	0.79	0.07
	V	1.00	0.07	1.00	0.20
	VI	1.00	0.05	1.00	0.10
	VII	0.94	0.00	0.98	0.04
50	I	1.00	0.01	1.00	0.04
	II	1.00	0.00	1.00	0.02
	III	1.00	0.00	0.99	0.04
	IV	0.93	0.00	0.73	0.02
	V	1.00	0.02	0.99	0.06
	VI	1.00	0.02	1.00	0.03
	VII	0.94	0.00	0.97	0.02

Table 3

The impact of α on the value and the cost of the estimated regime. In the header, α is the tuning parameter in the stopping criterion; $R(\widehat{\pi})$ is the mean outcome under the estimated regime $\widehat{\pi}$, computed on a test set of 10^6 subjects; $N(\widehat{\pi})$ is the cost of implementing the estimated regime $\widehat{\pi}$, computed on the same test set; TPR is the true positive rate, namely, the number of signal variables involved in $\widehat{\pi}$ divided by the number of signal variables; FPR is the false positive rate, namely, the number of noise variables involved in $\widehat{\pi}$ divided by the number of noise variables. Recall that p is the dimension of patient covariates.

p	Setting	$\alpha = 0.9$				$\alpha = 0.95$				$\alpha = 0.99$			
		$R(\widehat{\pi})$	$N(\widehat{\pi})$	TPR	FPR	$R(\widehat{\pi})$	$N(\widehat{\pi})$	TPR	FPR	$R(\widehat{\pi})$	$N(\widehat{\pi})$	TPR	FPR
Continuous response													
10	I	2.78	1.65	1.00	0.01	2.78	1.65	1.00	0.00	2.78	1.65	1.00	0.00
	II	2.71	1.66	1.00	0.00	2.70	1.66	1.00	0.00	2.69	1.66	1.00	0.00
	III	2.59	1.69	1.00	0.00	2.59	1.69	1.00	0.00	2.59	1.69	1.00	0.00
	IV	2.89	2.51	0.93	0.00	2.89	2.51	0.93	0.00	2.89	2.51	0.92	0.00
	V	2.90	1.91	1.00	0.07	2.90	1.91	1.00	0.07	2.90	1.91	1.00	0.07
	VI	3.98	1.61	1.00	0.06	3.98	1.61	1.00	0.05	3.98	1.61	1.00	0.05
	VII	3.22	2.56	0.94	0.00	3.22	2.56	0.94	0.00	3.21	2.56	0.94	0.00
50	I	2.75	1.96	1.00	0.01	2.76	1.96	1.00	0.01	2.78	1.96	1.00	0.00
	II	2.70	1.66	1.00	0.00	2.70	1.66	1.00	0.00	2.69	1.66	1.00	0.00
	III	2.58	1.75	1.00	0.00	2.59	1.75	1.00	0.00	2.59	1.75	1.00	0.00
	IV	2.89	2.54	0.93	0.00	2.89	2.54	0.93	0.00	2.89	2.54	0.92	0.00
	V	2.87	2.19	1.00	0.03	2.87	2.19	1.00	0.02	2.88	2.19	1.00	0.02
	VI	3.95	1.70	1.00	0.02	3.95	1.70	1.00	0.02	3.95	1.70	1.00	0.02
	VII	3.22	2.56	0.94	0.00	3.21	2.56	0.94	0.00	3.21	2.56	0.93	0.00
Binary response													
10	I	0.76	2.16	1.00	0.12	0.77	2.16	1.00	0.07	0.77	2.16	1.00	0.02
	II	0.71	1.75	1.00	0.06	0.71	1.75	1.00	0.04	0.71	1.75	1.00	0.02
	III	0.73	2.24	1.00	0.15	0.73	2.24	1.00	0.11	0.74	2.24	0.99	0.04
	IV	0.71	2.48	0.81	0.08	0.71	2.48	0.79	0.07	0.71	2.48	0.70	0.06
	V	0.75	2.64	1.00	0.24	0.75	2.64	1.00	0.20	0.75	2.64	1.00	0.16
	VI	0.79	2.11	1.00	0.11	0.79	2.11	1.00	0.10	0.79	2.11	1.00	0.10
	VII	0.77	2.87	0.98	0.06	0.77	2.87	0.98	0.04	0.76	2.87	0.97	0.02
50	I	0.75	2.87	1.00	0.05	0.76	2.87	1.00	0.04	0.76	2.87	1.00	0.02
	II	0.71	1.93	1.00	0.02	0.71	1.93	1.00	0.02	0.71	1.93	0.99	0.01
	III	0.72	2.68	0.99	0.04	0.73	2.68	0.99	0.04	0.73	2.68	0.99	0.02
	IV	0.71	2.65	0.75	0.02	0.71	2.65	0.73	0.02	0.71	2.65	0.66	0.02
	V	0.73	3.32	0.99	0.07	0.74	3.32	0.99	0.06	0.74	3.32	0.99	0.05
	VI	0.78	2.47	1.00	0.03	0.78	2.47	1.00	0.03	0.78	2.47	1.00	0.02
	VII	0.76	3.04	0.97	0.02	0.76	3.04	0.97	0.02	0.76	3.04	0.95	0.01

Table 4

The impact of α on the estimated regime. In the header, α is the tuning parameter in the stopping criterion and $\widehat{\pi}_\alpha$ is the regime such obtained. For each pair of regimes $\widehat{\pi}_\alpha$ and $\widehat{\pi}_{\alpha'}$, we report the probability that they recommend the same treatment for a randomly selected patient in the population. Mathematically, this is to compute $\mathbb{P}\{\widehat{\pi}_\alpha(X) = \widehat{\pi}_{\alpha'}(X) | \widehat{\pi}_\alpha, \widehat{\pi}_{\alpha'}\}$ and then average over 1000 replications, where X is generated in the same way as in Section 3 in the main paper.

p	Setting	$\widehat{\pi}_{0.9}$ vs. $\widehat{\pi}_{0.95}$	$\widehat{\pi}_{0.95}$ vs. $\widehat{\pi}_{0.99}$	$\widehat{\pi}_{0.9}$ vs. $\widehat{\pi}_{0.99}$
Continuous response				
	I	0.998	0.998	0.996
	II	0.986	0.975	0.961
	III	0.993	0.992	0.986
10	IV	0.997	0.997	0.993
	V	0.999	1.000	0.998
	VI	0.998	0.997	0.995
	VII	0.991	0.988	0.979
	I	0.984	0.985	0.970
	II	0.986	0.977	0.962
	III	0.988	0.989	0.976
50	IV	0.998	0.996	0.994
	V	0.993	0.995	0.988
	VI	0.997	0.997	0.993
	VII	0.991	0.989	0.980
Binary response				
	I	0.971	0.969	0.941
	II	0.978	0.964	0.944
	III	0.971	0.952	0.926
10	IV	0.983	0.955	0.941
	V	0.973	0.969	0.944
	VI	0.992	0.993	0.985
	VII	0.976	0.968	0.944
	I	0.973	0.946	0.920
	II	0.980	0.958	0.942
	III	0.971	0.947	0.925
50	IV	0.985	0.962	0.947
	V	0.965	0.939	0.913
	VI	0.985	0.985	0.969
	VII	0.974	0.955	0.930

Table 5

Consistency of the decision list. In the header, n is the sample size; p is the number of predictors. Loss is $R(\pi^{\mathrm{opt}}) - R(\widehat{\pi})$, namely, the difference between the the value under the estimated regime and the value under the optimal regime. Pr(best) is $\mathbb{P}\{\widehat{\pi}(X) = \pi^{\mathrm{opt}}(X)|\widehat{\pi}\}$, namely, the probability that the treatment recommended by the estimate regime coincides with the treatment recommended by the optimal regime. Loss and Pr(best) are averaged over 1000 replications. Correct is the proportion of $\widehat{\pi}$ having the same form and covariates as π^{opt} among 1000 replications; MSE$_1$ is the mean squared error of the estimated cutoff for X_1; MSE$_2$ is the mean squared error of the estimated cutoff for X_2.

Setting	n	p	Loss	Pr(best)	Correct	MSE$_1$($\times n$)	MSE$_2$($\times n$)
Continuous response							
I	10^4	10	0.0023	0.9982	1.00	4.24	6.87
I	10^5	10	0.0006	0.9995	1.00	4.30	6.50
I	10^6	10	0.0002	0.9998	1.00	4.06	6.32
I	10^4	50	0.0022	0.9982	1.00	4.60	6.91
I	10^5	50	0.0006	0.9995	1.00	4.24	6.49
I	10^6	50	0.0002	0.9998	1.00	4.22	6.48
V	10^4	10	0.0039	0.9975	1.00	6.08	5.27
V	10^5	10	0.0010	0.9994	1.00	5.86	4.70
V	10^6	10	0.0003	0.9998	1.00	5.46	4.54
V	10^4	50	0.0036	0.9977	1.00	6.10	5.33
V	10^5	50	0.0010	0.9994	1.00	5.96	4.54
V	10^6	50	0.0003	0.9998	1.00	5.69	4.51
Binary response							
I	10^4	10	0.0007	0.9966	1.00	8.13	10.93
I	10^5	10	0.0001	0.9994	1.00	5.71	6.05
I	10^6	10	0.0000	0.9999	1.00	5.27	5.61
I	10^4	50	0.0007	0.9965	1.00	9.72	11.15
I	10^5	50	0.0001	0.9994	1.00	5.71	6.29
I	10^6	50	0.0000	0.9998	1.00	5.41	5.78
V	10^4	10	0.0094	0.9447	0.79	7.81	22.66
V	10^5	10	0.0081	0.9547	0.96	4.14	6.33
V	10^6	10	0.0079	0.9563	0.97	3.89	5.03
V	10^4	50	0.0099	0.9418	0.70	6.63	14.36
V	10^5	50	0.0078	0.9567	0.96	3.97	6.14
V	10^6	50	0.0081	0.9550	0.97	3.96	5.10