# Precision Medicine: Lecture 15
# Multiple Utilities

Michael R. Kosorok,
Nikki L. B. Freeman and Owen E. Leete

Department of Biostatistics
Gillings School of Global Public Health
University of North Carolina at Chapel Hill

Fall, 2021

# Outline

Incorporating patient preferences

Estimation and Optimization of Composite Outcomes

# Shared decision making

- So far we have focused on DTRs that tailor treatments to individual patient characteristics.
- Patient characteristics may include clinical information as well as patient preferences.
- Recall that an optimal DTR maximizes the mean of a pre-specified clinical outcome if it is applied to all patients in a population of interest.
- While this definition of optimality is mathematically convenient, it does not directly allow for shared decision making where patient preferences are integrated into the decision process.

# Why including patient preferences into DTR construction is important

- Including patient preferences in treatment selection in a mathematically rigorous way is important.
- First, it facilitates 'patient-centered care' in which patients play a key role in decision making and the evaluation of their own outcomes.
- Second, it offers a principled means for matching patient preferences to an optimal treatment based on potentially complex outcome profiles.

# Patient preference elicitation

- Eliciting patient preferences is not necessarily straightforward.
- For example, it would be convenient if patients could choose parameters indexing a composite outcome. However, without specialized training for the patients, this is not feasible.
- An alternative approach is to administer a questionnaire populated with items that are accessible to a patient in a domain context yet are informative about preferences in the outcome space.
- Butler, Laber, Davis, and Kosorok (2018) incorporate this latter approach to derive a preference-sensitive optimal DTR for each patient. We will explore their proposed methodology in detail.

# Setup and notation

- Observed data $\{\mathbf{W}_i, \mathbf{X}_i, A_i, Y_i, Z_i\}_{i=1}^n$ comprises $n$ independent and identically distributed tuples $(\mathbf{W}, \mathbf{X}, A, Y, Z)$ where
  - $\mathbf{W} \in \{0,1\}^p$ denotes answers to items in a preference questionnaire,
  - $\mathbf{X} \in \mathbb{R}^m$ denotes pre-treatment patient covariates,
  - $A \in \{-1, 1\}$ denotes the assigned treatment,
  - $Y, Z \in \mathbb{R}$ denote outcomes of interest, coded so that higher values are better.
- A DTR is denoted by $\pi : \text{dom } \mathbf{W} \times \text{dom } \mathbf{X} \to \text{dom } A$.

# Optimal DTR

To define an optimal DTR,

- Assume that each individual in the population has a latent preference, $H \in \mathbb{R}$, that indexes a utility function $U(y, z; h)$.

- Assume that the utility function induces an ordering on dom $Y \times$ dom $Z$ so that a patient with preference $H = h$ would prefer outcomes $(y, z)$ to $(y', z')$ if $U(y, z; h) \geq U(y', z'; h)$.

# Optimal DTR (cont.)

- Let $Y^*(a)$ and $Z^*(a)$ denote the potential outcomes under treatments $a \in \{-1, 1\}$ so that $U\{Y^*(a), Z^*(a)\}$ is the potential utility function under treatment $a$.
- Define the potential utility as

$$V_U(\pi) = \mathbb{E}\left[ \sum_{a \in \{-1,1\}} U\{Y^*(a), Z^*(a); H\} I(\pi(\mathbf{W}, \mathbf{X}) = a) \right]$$

  where the expectation is taken with respect to the joint distribution of $\{\mathbf{X}, \mathbf{W}, H, A, Y^*(0), Y^*(1), Z^*(0), Z^*(1)\}$.
- The optimal DTR, $\pi_U^{\text{opt}}$ satisfies

$$V_U(\pi_U^{\text{opt}}) \geq V_U(\pi) \quad \text{for all } \pi.$$

# Form of the utility

- Assume the utility has the form

$$U(y, z; h) = \Phi(h)y + \{1 - \Phi(h)\}z$$

  where $\Phi(\cdot)$ denotes the cumulative distribution function for a standard normal random variable.

- Interpretation: A patient with preference $h$ cares $\Phi(h)/(1 - \Phi(h))$ more about $Y$ than $Z$.

- Linear utility is a common assumption in multiobjective optimization, however the assumption of a constant gain in utility per unit increase in outcome may not be reasonable in some contexts.

# The optimal DTR for any rational utility can be expressed as the optimal DTR for some linear utility

- A rational utility will always prefer a treatment that is better on both outcomes to one that is worse on both outcomes.
- Butler et al. (2018) prove that the DTR for any rational utility may be expressed as the optimal DTR for some linear utility function.
- To state this formally, define the following

$$
\begin{aligned}
R_Z(\mathbf{w},\mathbf{x}) &= \{Z^*(1)|\mathbf{W}=\mathbf{w},\mathbf{X}=\mathbf{x}\} - \{Z^*(-1)|\mathbf{W}=\mathbf{w},\mathbf{X}=\mathbf{x}\} \\
R_Y(\mathbf{w},\mathbf{x}) &= \{Y^*(1)|\mathbf{W}=\mathbf{w},\mathbf{X}=\mathbf{x}\} - \{Y^*(-1)|\mathbf{W}=\mathbf{w},\mathbf{X}=\mathbf{x}\}, \\
R_U(\mathbf{w},\mathbf{x}) &= \{U\{Y^*(1),Z^*(1);H\}|\mathbf{W}=\mathbf{w},\mathbf{X}=\mathbf{x}\} \\
&\quad - \{U\{Y^*(-1),Z^*(-1);H\}|\mathbf{W}=\mathbf{w},\mathbf{X}=\mathbf{x}\}.
\end{aligned}
$$

- Note, it can be shown that $\pi_U^{\text{opt}}(\mathbf{w},\mathbf{x}) = \text{sign}\{R_U(\mathbf{w},\mathbf{x})\}$

# The optimal DTR for any rational utility is the optimal DTR for some linear utility (formally)

Lemma (2.1)

*Assume that* $\max\{R_U(\mathbf{w},\mathbf{x})R_Z(\mathbf{w},\mathbf{x}), R_U(\mathbf{w},\mathbf{x})R_Y(\mathbf{w},\mathbf{x})\} > 0$ *for all* $\mathbf{x}$, $\mathbf{w}$. *Then, there exists a real-valued random variable,* $H'$, *such that: (i)* $H' \perp\!\!\!\perp A, \{Z^*(a), Y^*(a) : a \in \{-1,1\}\}|\mathbf{X},\mathbf{W}$; *and (ii) the DTR*

$$\pi^{opt}_{CVX}(\mathbf{x},\mathbf{w}) = \underset{a\in\{-1,1\}}{\text{argmax}}\ E[\Phi(H')Y^*(a) + \{1-\Phi(H')\}Z^*(a)|\mathbf{X}=\mathbf{x},\mathbf{W}=\mathbf{w}]$$

*satisfies* $V_U(\pi^{opt}_{CVX}) = V_U(\pi^{opt}_{U})$.

# Identification

- The optimal DTR is defined in terms of potential outcomes. To identify the model, we assume
  - C1 Causal consistency, $(Y, Z) = \{Y^*(A), Z^*(A)\}$,
  - C2 Positivity, there exists $\epsilon > 0$ so that $P(A = a|\mathbf{X}, \mathbf{W}) \geq \epsilon$,
  - C3 Ignorability, $[\{Y^*(a), Z^*(a)\} : a \in \{-1, 1\}] \perp\!\!\!\perp A|\mathbf{X}, \mathbf{W}$, and
  - C4 $(A, Y, Z) \perp\!\!\!\perp H|\mathbf{X}, \mathbf{W}$.

- C1, C2, and C3 are standard assumptions.

- C4 holds if the assumptions of Lemma 2.1 hold.

- C4 can be weakened to $A \perp\!\!\!\perp H|\mathbf{X}, \mathbf{W}$ at the expense of postulating a model for the conditional mean of $\Phi(H)Y$ and $\Phi(H)Z$ given $(\mathbf{X}, \mathbf{W}, Z)$.

# Estimation strategy

- Under C1, C2, and C3, it can be shown that

$$\pi^{\text{opt}}(\mathbf{x}, \mathbf{w}) = \underset{a \in \{-1,1\}}{\text{argmax}}\, \mathbb{E}[\Phi(H)Y + \{1 - \Phi(H)\}Z | \mathbf{X} = \mathbf{x}, \mathbf{W} = \mathbf{w}, A = a]$$

which under C4 can be written as

$$\pi^{\text{opt}}(\mathbf{x}, \mathbf{w}) = \underset{a \in \{-1,1\}}{\text{argmax}}\, E[\Phi(H)|\mathbf{X} = \mathbf{x}, \mathbf{W} = \mathbf{w}]\mathbb{E}(Y|\mathbf{X} = \mathbf{x}, \mathbf{W} = \mathbf{w}, A = a)$$
$$+[1 - E\{\Phi(H)|\mathbf{X} = \mathbf{x}, \mathbf{W} = \mathbf{w}\}]\mathbb{E}(Z|\mathbf{X} = \mathbf{x}, \mathbf{W} = \mathbf{w}, A = a)].$$

- This suggests a strategy for estimating $\pi^{\text{opt}}$:
  - Construct estimators for $Q_Y(\mathbf{x}, \mathbf{w}, a) = \mathbb{E}(Y|\mathbf{X} = \mathbf{x}, \mathbf{W} = \mathbf{w}, A = a)$ and $Q_Z(\mathbf{x}, \mathbf{w}, a) = \mathbb{E}(Z|\mathbf{X} = \mathbf{x}, \mathbf{W} = \mathbf{w}, A = a)$
  - Postulate a latent preference model linking the unobservable preference $H$ with covariates $\mathbf{X}$ and preference questionnaire items $\mathbf{W}$ and use this model to estimate $\mu_H(\mathbf{x}, \mathbf{w}) = \mathbb{E}\{\Phi(H)|\mathbf{X} = \mathbf{x}, \mathbf{W} = \mathbf{w}\}$.
  - Plug in estimators to estimate $\pi^{\text{opt}}$.

# Estimation of the $Q$ functions

- To estimate $Q_Y$ and $Q_Z$, Butler et al. (2018) propose linear working models of the form

$$Q_Y(\mathbf{x}, \mathbf{w}, a; \psi_Y) = \mathbf{x}_{Y,0}^\intercal \psi_{Y,0} + \mathbf{w}_{Y,1}^\intercal \psi_{Y,1} + a\mathbf{x}_{Y,1}^\intercal \psi_{Y,2} + a\mathbf{w}_{Y,1}^\intercal \psi_{Y,3}$$
$$Q_Z(\mathbf{x}, \mathbf{w}, a; \psi_Z) = \mathbf{x}_{Z,0}^\intercal \psi_{Z,0} + \mathbf{w}_{Z,0}^\intercal \psi_{Z,1} + a\mathbf{x}_{Z,1}^\intercal \psi_{Z,2} + a\mathbf{w}_{Z,1}^\intercal \psi_{Z,3},$$

  where $\mathbf{x}_{\ell,j}$ and $\mathbf{w}_{\ell,j}$ for $\ell = Y, Z$ and $j = 0, 1$ are known feature vectors from $\mathbf{x}$ and $\mathbf{w}$ and $\psi_W$ and $\psi_Y$ are unknown parameter vectors.

- Let $\hat{\psi}_{Y,n} = \text{argmin}_{\psi_Y} \mathbb{P}_n\{Y - Q_Y(\mathbf{X}, \mathbf{W}, A; \psi_Y)\}^2$ and $\hat{\psi}_{Z,n} = \text{argmin}_{\psi_Z} \mathbb{P}_n\{Z - Q_Z(\mathbf{X}, \mathbf{W}, A; \psi_Z)\}^2$.

- Construct estimators $Q_Y(\mathbf{x}, \mathbf{w}, a; \hat{\psi}_Y)$ and $Q_Z(\mathbf{x}, \mathbf{w}, a; \hat{\psi}_Z)$ of $Q_Y(\mathbf{x}, \mathbf{w}, a)$ and $Q_Z(\mathbf{x}, \mathbf{w}, a)$, respectively.

# Specification of the latent preference model

- Assume that $H \perp\!\!\!\perp \mathbf{X}|\mathbf{W}$ for ease of exposition. (Note that $\mathbf{X}$ could be included in the latent patient preference model.)

- Assume that the latent patient preferences are connected to items on the questionnaire through a Rasch model of the form

$$\text{logit}\{P(W_j = 1|H = h)\} = \beta_{0,j} + \beta_{1,j}h, \quad j = 1, \ldots, p$$

  which is indexed by $\beta = (\beta_{0,1}, \beta_{1,1}, \ldots, \beta_{0,p}, \beta_{1,p})$.

- Let $\beta^*$ denote the true parameter value. The EM algorithm can be used to construct an estimator $\hat{\beta}_n$ of $\beta^*$.

# Estimation of $\mu_H$

- Given an estimator $\hat{\beta}_n$ and a postulated marginal distribution, $p_h$ for the latent preferences, the conditional distribution of $H$ given $\mathbf{W} = \mathbf{w}$ is proportional to $p(\mathbf{w}|h)p_h(h)$.
- This conditional distribution can be approximated using Metropolis Hastings.
- A computationally less burdensome approach is to apply a method of moments type estimator:
  - Let $\hat{h}(\mathbf{w})$ denote the solution to $\sum_{j=1}^{p} \hat{\beta}_{n,1,j} \text{expit}(\hat{\beta}_{n,j,0} + \hat{\beta}_{n,1,j} h) = \sum_{j=1}^{p} \hat{\beta}_{n,1,j} w_j$.
- Subsequently let $\hat{\mu}_{H,n}(\mathbf{x}, \mathbf{w}) = \Phi(\hat{h}_n(\mathbf{w}))$ denote our estimator of $\mu(\mathbf{w}, \mathbf{x})$.

# Estimation of $\pi^{\text{opt}}$

With estimates of $\mu_H$, $Q_Z$, and $Q_Y$ in hand, we can compute an estimate of the optimal DTR,

$$\hat{\pi}_n(\mathbf{x}, \mathbf{w}) = \underset{a \in \{-1,1\}}{\text{argmax}} [\hat{\mu}_{H,n}(\mathbf{x}, \mathbf{w}) \hat{Q}_{Y,n}(\mathbf{x}, \mathbf{w}, a) + \{1 - \hat{\mu}_{H,n}(\mathbf{x}, \mathbf{w})\} \hat{Q}_{Z,n}(\mathbf{x}, \mathbf{w}, a)].$$

# Assumptions for the theoretical results

Let $h^*(\mathbf{w})$ denote the solution to $\sum_{j=1}^{p} \beta_{1,j}^* \text{expit}(\beta_{j,0}^* + \beta_{j,1}^* h) = \sum_{j=1}^{p} \beta_{j,1}^* w_j$. Assume the following

(A1) The number of items satisfies $3 \leq p_n = o(e^n)$.

(A2) The estimator $\hat{h}_n(\mathbf{w})$ converges in probability to $h^*(\mathbf{w})$, pointwise for all $\mathbf{w}$.

(A3) The estimators $\hat{Q}_{Y,n}(\mathbf{x}, \mathbf{w}, a)$ and $\hat{Q}_{Z,n}(\mathbf{x}, \mathbf{w}, a)$ converge in probability to $Q_Y(\mathbf{x}, \mathbf{w}, a)$ and $Q_Z(\mathbf{x}, \mathbf{w}, a)$ pointwise for each $\mathbf{x}$, $\mathbf{w}$, and $a$.

# Theoretical results

The first theoretical result establishes consistency of the proposed estimator for the optimal DTR as the sample size diverges but the number of items remains fixed.

## Theorem (Thm 2.2 in Butler et al. (2018))

*Assume (A1) - (A3) and let the number of items, $p_n = p$, be fixed. Then $V_U(\pi^{opt}) - V_U(\hat{\pi}_n)$ converges to zero in probability as $n \to \infty$.*

The second theoretical result says that if the number of items is allowed to diverge with the sample size then the estimated DTR performs as well as an oracle that knows each patient's individual preference.

## Theorem (Thm 2.3 in Butler et al. (2018))

*Assume (A1) - (A3) and suppose $p_n \to \infty$ as $n \to \infty$. Then $V_U(\hat{\pi}_n) - V_U(\pi^{oracle})$ converges to zero in probability as $n \to \infty$.*

# Case Study: CATIE

- The Clinical Antipsychotic Trials of Intervention Effectiveness (CATIE) schizophrenia trial was designed to compare new antipsychotic drugs to conventional ones in a randomized, controlled, double-blind, multi-phase trial.
- The trial targeted patients already being treated for schizophrenia but who might benefit from a medicinal change.
- Patients received antipsychotic treatments and were offered psychosocial treatment with their families.

# First phase of CATIE

- The first phase of CATIE is suited for application of the method proposed in Butler et al. (2018).
- Patients were randomized to one of 5 medications,
  - 4 were atypical antipsychotics, and
  - 1 was a conventional antipsychotic (perphenazine).
- For ease of exposition, we will dichotomize the treatments into atypical antipsychotics and perphenazine.
- At baseline, patients answered a 10 question, binary response assessment, which can be used to measure patient preferences across two outcomes
  - Efficacy using the Positive and Negative Syndromes Scale (PANSS), and
  - Side effect burden measured as the sum of side effect and adverse event indicators.

# Patient preference questions

The patient preference information was collected using a 10 question Drug Attitude Inventory assessment. One question was excluded from analysis.

Patient preference questions

| Question | Coding |
|---|---|
| "For me, the good things about medication outweigh the bad" | Yes (1) No (0) |
| "I feel weird, like a 'zombie', on medication" | Yes (0) No (1) |
| "Medications make me feel more relaxed" | Yes (1) No (0) |
| "Medications make me feel tired and sluggish" | Yes (0) No (1) |
| "I take medications only when Im sick" | Yes (0) No (1) |
| "I feel more normal on medication" | Yes (1) No (0) |
| "It is unnatural for my mind and body to be controlled by medications" | Yes (0) No (1) |
| "My thoughts are clearer on medication" | Yes (1) No (0) |
| "By staying on medications I can prevent being sick" | Yes (1) No (0) |

Figure 1: Table 1 from Butler et al. (2018)

# Analysis

Tailoring covariates were selected based on clinical expertise and prior analyses.

Coefficients of the Q-Function for Y and Z: CATIE Case Study

| | $Q_Y$ | $Q_Z$ |
|---|---|---|
| Intercept | −5.51 | 13.04 |
| Treatment | 9.97 | −2.09 |
| Male (ref: Female) | −5.91 | 0.19 |
| Black (ref: Other) | 3.08 | −0.80 |
| White (ref: Other) | 4.02 | −0.14 |
| Age | −0.11 | −0.01 |
| Baseline Clinical Severity of Schizophrenia | 3.29 | −0.19 |
| Age at First Antipsychotic Medication | 0.24 | 0.02 |
| Olanzapine (at baseline, pre-randomization) | −3.77 | −0.08 |
| Quetiapine (at baseline, pre-randomization) | −5.91 | 0.24 |
| Risperidone (at baseline, pre-randomization) | −1.48 | 0.06 |
| Ziprasidone (at baseline, pre-randomization) | −0.94 | 0.60 |
| Perphenazine (at baseline, pre-randomization) | −3.01 | 1.17 |
| Haldol (at baseline) | −3.72 | 0.38 |
| Decanoate (at baseline) | −7.68 | 0.73 |
| BMI | −0.04 | −0.01 |
| Diastolic Blood Pressure | 0.05 | 0.01 |
| Systolic Blood Pressure | −0.10 | 0.0008 |
| Treatment*Male (ref: Female) | −5.05 | −0.21 |
| Treatment*Black (ref: Other) | −1.86 | −0.42 |
| Treatment*White (ref: Other) | −2.49 | 0.25 |
| Treatment*Age | 0.05 | 0.01 |
| Treatment*Baseline Clinical Severity of Schizophrenia | 0.21 | −0.01 |
| Treatment*Age at First Antipsychotic Medication | −0.04 | 0.02 |
| Treatment*Olanzapine (at baseline, pre-randomization) | −1.07 | 0.26 |
| Treatment*Quetiapine (at baseline, pre-randomization) | −5.22 | −0.08 |
| Treatment*Risperidone (at baseline, pre-randomization) | −2.78 | 0.22 |
| Treatment*Ziprasidone (at baseline, pre-randomization) | 3.97 | 0.62 |
| Treatment*Perphenazine (at baseline, pre-randomization) | −1.04 | 0.03 |
| Treatment*Haldol (at baseline, pre-randomization) | −1.38 | 0.20 |
| Treatment*Decanoate (at baseline, pre-randomization) | −0.13 | 1.60 |
| Treatment*BMI | 0.02 | 0.05 |
| Treatment*Diastolic Blood Pressure | −0.01 | 0.01 |
| Treatment*Systolic Blood Pressure | −0.05 | −0.02 |

Figure 2: Table 2 from Butler et al. (2018)

# Analysis (cont.)

- By examining the coefficients for the Q-functions, we see that the main effect of treatment has an opposite sign in the two Q-functions as well as several interactions involving treatment.

- This suggests a trade-off between the two outcomes that must be made in choosing a treatment and that this trade off varies across patient characteristics.

# Estimated optimal treatment allocation

While efficacy appears to favor atypical antipsychotics, side effect burden tends to favor perphenazine. The composite outcome occupies a middle ground between the two marginal outcomes.

Treatment recommendations

| | Perphenazine | Atypical Antipsychotics |
|---|---|---|
| $\hat{\pi}_Y$: Efficacy | 36% | 64% |
| $\hat{\pi}_Z$: Side Effect Burden | 67% | 33% |
| $\hat{\pi}_\Pi$: $\Phi(E)\,Y + \{1 - \Phi(E)\}Z$ | 41% | 59% |

Figure 3: Table 3 from Butler et al (2018)

# Percent agreement in treatment recommendations

This table shows the fraction of overlap between the proposed estimated optimal DTR and the optimal DTR based only on efficacy or side effect burden.

Percent of agreement in treatment recommendations

| | $\hat{\pi}_Y$ : Efficacy | $\hat{\pi}_Z$: Side Effect Burden |
|---|---|---|
| Both recommend perphenazine | 35% | 36% |
| Only $\hat{\pi}_n$ recommends perphenazine | 6% | 5% |
| Only $\hat{\pi}_n$ recommends atypical antipsychotic | 2% | 31% |
| Both recommend atypical antipsychotic | 57% | 28% |

Figure 4: Table 4 in Butler et al. (2018)

# Discussion

- Butler et al. (2018) propose a strategy for balancing multiple, possibly competing outcomes when estimating a dynamic treatment regime.
- A few other methods for incorporating multiple outcomes into precision medicine have been proposed for various scenerios under various assumptions. However, the literature in this area is relatively small.
- An interesting extension is the multi-decision setting. This is a challenging extension as patient preferences may change over time in response to treatment received and interim outcome experiences.

# Outline

# Introduction

- Almost all methods for estimating individualized treatment rules have been designed to optimize a scalar outcome
- Clinical decision making often requires balancing trade-offs between multiple outcomes
- Examples:
  - Bipolar disorder treatments must manage both depression and mania
  - Antidepressants can prevent depressive episodes but may also induce manic episodes
- Utility functions can be used to summarize multiple outcomes as a single scalar

# Setup and Notation

- We will assume two outcomes, but the method can be extended to more
- The available data are $(\mathbf{X}_i, A_i, Y_i, Z_i),\ i = 1, \ldots, n$
  - $\mathbf{X}_i \in \mathcal{X} \subseteq \mathbb{R}^p$ are patient covariates
  - $A \in \mathcal{A} = \{-1, 1\}$ is a binary treatment
  - $Y$ and $Z$ are two real-valued outcomes (higher is better)
- Let $Y^*(A)$ and $Z^*(A)$ be the potential outcomes under treatment $a$
- We will need the standard causal assumptions
  - Consistency: $Y = Y^*(A)$ and $Z = Z^*(A)$
  - Positivity: $\Pr(A = a|\mathbf{X} = \mathbf{x}) \geq c > 0$
  - Ignorability: $\{Y^*(-1), Y^*(1)\} \perp A \,|\, \mathbf{X}$

# Optimal Treatments for Individual Outcomes

- Define $$Q_Y(\mathbf{x}, a) = \mathbb{E}(Y \mid \mathbf{X} = \mathbf{x}, A = a)$$ $$Q_Z(\mathbf{x}, a) = \mathbb{E}(Z \mid \mathbf{X} = \mathbf{x}, A = a)$$
- Under the preceding assumptions we have
  - $d_Y^{\text{opt}}(\mathbf{x}) = \text{argmax}_{a \in \mathcal{A}} \, Q_Y(\mathbf{x}, a)$
  - $d_Z^{\text{opt}}(\mathbf{x}) = \text{argmax}_{a \in \mathcal{A}} \, Q_Z(\mathbf{x}, a)$
- In general, $d_Y^{\text{opt}}(\mathbf{x})$ need not equal $d_Z^{\text{opt}}(\mathbf{x})$
- If both $Y$ and $Z$ are clinically relevant, neither $d_Y^{\text{opt}}$ nor $d_Z^{\text{opt}}$ may be acceptable

# Utility Functions

- Define the composite outcome $U = u(Y, Z)$ for a utility function, $u$
  - $u : \mathbb{R}^2 \mapsto \mathbb{R}$ is the "goodness" of the outcome pair $(y, z)$
  - $u$ may be unknown and possibly depend on the covariates
- Define $Q_U(\mathbf{x}, a) = \mathbb{E}(U \mid \mathbf{X} = \mathbf{x}, A = a)$
- The optimal regime with respect to $U$ satisfies

$$\begin{aligned} d_U^{\text{opt}}(\mathbf{x}) &= \underset{a \in \mathcal{A}}{\text{argmax}}\, Q_U(\mathbf{x}, a) \\ &= \underset{a \in \mathcal{A}}{\text{argmax}}\, \mathbb{E}[u\{Y(a), Z(a)\} \mid \mathbf{x}] \end{aligned}$$

- Utility functions which are convex combinations of $Q_Y(\mathbf{x}, a)$ and $Q_Z(\mathbf{x}, a)$ are identifiable under the preceding assumptions

# Inverse Reinforcement Learning

- We assume that clinicians act with the goal of optimizing each patient's utility
- Inverse reinforcement learning uses decisions made by an expert to construct a utility function
- We assume that the clinicians are approximately (i.e., imperfectly) assigning treatment according to $d^{\text{opt}}(\mathbf{x})$
  - There would be no need to estimate the optimal treatment policy if the clinician were always able to correctly identify the optimal treatment
- We assume that the probability of making the correct treatment decision depends on individual patient characteristics

$$\Pr\{A = d_U^{\text{opt}}(\mathbf{X}) \,|\, \mathbf{X} = \mathbf{x}\} = \text{expit}(\mathbf{x}^{\mathsf{T}}\beta)$$

where $\beta$ is an unknown parameter

# Fixed Utility

- We begin by assuming the utility function is constant across patients
- Let the utility function be $u(y, z; \omega) = \omega y + (1 - \omega)z$ for some $\omega \in [0, 1]$
- The optimal ITR for a broad class of utility functions is equivalent to the optimal ITR for a utility function of this form (Butler 2018, Lemma 1)
- Define $Q_\omega(\mathbf{x}, a) = \omega Q_Y(\mathbf{x}, a) + (1 - \omega) Q_z(\mathbf{x}, a)$ and $d_\omega^{\text{opt}}(\mathbf{x}) = \text{argmax}_{a \in \mathcal{A}} Q_\omega(\mathbf{x}, a)$

# Fixed Utility, Cont.

- Let $\widehat{Q}_{Y,n}$ and $\widehat{Q}_{Z,n}$ be estimates based on regression models fit to the observed data
- For a fixed value of $\omega$, let

$$\widehat{Q}_{\omega,n}(\mathbf{x}, a) = \omega \widehat{Q}_{Y,n}(\mathbf{x}, a) + (1 - \omega)\widehat{Q}_{Z,n}(\mathbf{x}, a)$$

- Define $\widehat{d}_{\omega,n}(\mathbf{X}) = \text{argmax}_{a \in \mathcal{A}}\, \widehat{Q}_{\omega,n}(\mathbf{x}, a)$
- The joint distribution of $(\mathbf{X}, A, Y, Z)$ is

$$\begin{aligned} f(\mathbf{X}, A, Y, Z) &= f(Y, Z|\mathbf{X}, A) f(A|\mathbf{X}) f(\mathbf{X}) \\ &= f(Y, Z|\mathbf{X}, A) f(\mathbf{X}) \frac{\exp[\mathbf{X}^{\mathsf{T}} \beta 1\{A = d_{\omega}^{\text{opt}}(\mathbf{X})\}]}{1 + \exp(\mathbf{X}^{\mathsf{T}} \beta)} \end{aligned}$$

# Pseudo-likelihood Estimation of Utility Functions

- Assuming that $f(Y, Z|\mathbf{X}, A)$ and $f(\mathbf{X})$ do not depend on $\omega$ or $\beta$, the likelihood for $(\omega, \beta)$ is

$$\mathcal{L}_n(\omega, \beta) \propto \prod_{i=1}^{n} \frac{\exp[\mathbf{X}^\mathsf{T}\beta 1\{A = d_\omega^{\text{opt}}(\mathbf{X})\}]}{1 + \exp(\mathbf{X}^\mathsf{T}\beta)}$$

- Plugging in $\widehat{d}_{\omega,n}(\mathbf{X})$ for $d_\omega^{\text{opt}}(\mathbf{X})$ yields a pseudo-likelihood
- If we let $\widehat{\omega}_n$ and $\widehat{\beta}_n$ denote the maximum pseudo-likelihood estimators, an estimator of the utility function is $\widehat{u}_n(y, z) = \widehat{u}_n(y, z; \widehat{\omega}_n) = \widehat{\omega}_n y + (1 - \widehat{\omega}_n)z$ and $\text{expit}(\mathbf{X}^\mathsf{T}\widehat{\beta}_n)$ estimates the probability that a patient would be treated optimally

# Algorithm

- The pseudo-likelihood is non-smooth in $\omega$, so standard gradient-based optimization can't be used
- For a given value of $\omega$, it is straightforward to compute the profile estimator $\widehat{\beta}_n(\omega)$
- Compute the profile pseudo-likelihood over a grid for $\omega$ and select the value yielding the largest pseudo-likelihood
- Finding $\widehat{\beta}_n(\omega)$ can be accomplished using logistic regression

# Patient-specific Utility

- In some application domains outcome preferences can vary widely across patients
  - Schizophrenia
  - Pain management
  - etc.
- To accommodate this, we assume that the utility function takes the form $u(y, z; \mathbf{x}, \omega) = \omega(\mathbf{x})y + 1 - \omega(\mathbf{x})z$ where $\omega : \mathbf{X} \mapsto [0, 1]$ is a smooth function
  - e.g., Let $\omega(x; \theta) = \text{expit}(\mathbf{X}^{\mathsf{T}}\theta)$ where $\theta$ is an unknown parameter
- Define $Q_\theta(\mathbf{x}, a) = \omega(\mathbf{x}; \theta)Q_Y(\mathbf{x}, a) + (1 - \omega(\mathbf{x}; \theta))Q_Z(\mathbf{x}, a)$ and define $d_\theta^{\text{opt}}(\mathbf{x}) = \text{argmax}_{a \in \mathcal{A}}\, Q_\theta(\mathbf{x}, a)$

# Patient-specific Utility

- For $\widehat{Q}_{Y,n}$, $\widehat{Q}_{Z,n}$ and a fixed value of $\theta$, let

$$\widehat{Q}_{\theta,n}(\mathbf{x}, a) = \omega(\mathbf{x}; \theta)\widehat{Q}_{Y,n}(\mathbf{x}, a) + (1 - \omega(\mathbf{x}; \theta))\widehat{Q}_{Z,n}(\mathbf{x}, a)$$

  and $\widehat{d}^{\text{opt}}_{\theta,n}(\mathbf{x}) = \text{argmax}_{a \in \mathcal{A}}\, \widehat{Q}_{\theta,n}(\mathbf{x}, a)$

- We can compute the estimators $(\widehat{\theta}_n, \widehat{\beta}_n)$ by maximizing the pseudo-likelihood

$$\mathcal{L}_n(\theta, \beta) \propto \prod_{i=1}^{n} \frac{\exp[\mathbf{X}^{\mathsf{T}}\beta 1\{A = \widehat{d}_{\theta,n}(\mathbf{X})\}]}{1 + \exp(\mathbf{X}^{\mathsf{T}}\beta)}$$

- An estimator for the utility function is $\widehat{u}_n(y, z; \mathbf{x}) = \omega(\mathbf{x}; \widehat{\theta}_n)y + (1 - \omega(\mathbf{x}; \widehat{\theta}_n))z$
- An estimator for the optimal decision function is $\widehat{d}_{\widehat{\theta}_n,n}$

# Algorithm

- As before, the pseudo-likelihood is non-smooth in $\theta$
- It is again straightforward to compute the profile pseudo-likelihood estimator $\widehat{\beta}_n(\theta)$ for any $\theta \in \mathbb{R}^p$
- It is computationally infeasible to compute $\widehat{\beta}_n(\theta)$ over a grid for moderate $p$
- Instead we generate a random walk through the parameter space using the Metropolis algorithm (see next slide)
- After generating a chain $(\theta^1, \ldots, \theta^B)$, we select the $\theta^k$ that leads to the largest value of $\widetilde{L}_n(\theta^k)$ as the maximum pseudo-likelihood estimator

# Algorithm, Cont.

**Algorithm 2:** Pseudo-likelihood estimation of patient-dependent utility function

Set a chain length, $B$, fix $\sigma^2 > 0$, and initialize $\theta^1$ to a starting value in $\mathbb{R}^p$;

**for** $b = 2, \ldots, B$ **do**

  Generate $\mathbf{e} \sim N(0, \sigma^2 I)$;

  Set $\widetilde{\theta}^{b+1} = \theta^b + \mathbf{e}$;

  Compute $p = \min\{\widetilde{L}_n(\widetilde{\theta}^{b+1})/\widetilde{L}_n(\widetilde{\theta}^b), 1\}$;

  Generate $U \sim U(0,1)$; if $U \leq p$, set $\theta^{b+1} = \widetilde{\theta}^{b+1}$; otherwise, set $\theta^{b+1} = \theta^b$;

**end**

# Theoretical Results

- We assume that $\Pr\{A = d^{\text{opt}}(\mathbf{x})|\mathbf{X} = \mathbf{x}\} = \text{expit}(\mathbf{x}^{\mathsf{T}}\beta_0)$
- The true utility is $u(y, z; \mathbf{x}, \theta_0) = \omega(\mathbf{X}; \theta_0)y + \{1 - \omega(\mathbf{X}; \theta_0)\}z$ where $\omega(\mathbf{X}; \theta)$ has bounded continuous derivative on compact sets
- $d^{\text{opt}}_{\theta_0}(\mathbf{X}) = d^{\text{opt}}_{\theta}(\mathbf{X})$ almost surely implies $\theta = \theta_0$
- The main theoretical results rely on a number of assumptions
  - A rate of convergence for the estimated Q-functions
    - Automatically satisfied if the Q-functions are estimated using linear or generalized linear models
  - Positive probability of patients with $\mathbf{x}$ values near the boundary between where each treatment is optimal

# Asymptotic Inference

## Theorem

*Under regularity conditions, the pseudo-likelihood maximizers* $\hat{\beta}_n$ *and* $\hat{\theta}_n$ *satisfy*

$$\sqrt{n}\begin{pmatrix} \hat{\theta}_n - \theta_0 \\ \hat{\beta}_n - \beta_0 \end{pmatrix} \rightsquigarrow \begin{pmatrix} U \\ I_0^{-1}\left[Z_A - k_0(Z_Y, Z_Z, U)\right] \end{pmatrix} = \begin{pmatrix} U \\ B \end{pmatrix},$$

*where* $U = \text{argmin}_u\, \beta_0^{\mathsf{T}} k_0(Z_Y, Z_Z, u)$, *and*

$$\begin{pmatrix} Z_Y \\ Z_Z \\ Z_A \end{pmatrix} \sim N(0, \Sigma_0).$$

*A certain semiparametric bootstrap is also consistent in probability.*

# Asymptotic Inference

Main technical tools:

- The Argmax theorem
- The following for the bootstrap:

## Theorem

- *Let $H$ be compact with respect to a metric $d$ and $\mathcal{F} \subset C[H]$ be compact with respect to $\|\cdot\|_H$*
- *For each $f \in \mathcal{F}$, let $u(f) = \text{argmax}_{u \in H} f(u)$, where we arbitrarily choose a value if nonunique*
- *Suppose also that there exists an $\mathcal{F}_1 \subset \mathcal{F}$ such that each $f \in \mathcal{F}_1$ has a unique maximum*
- *Then*

$$\lim_{\delta \downarrow 0} \sup_{f \in \mathcal{F}_1} \sup_{g \in \mathcal{F}: \|f-g\|_H < \delta} d(u(f), u(g)) = 0$$

# Parametric Bootstrap

## Theorem

- *Assume* $\widehat{\Sigma}_n = \Sigma_0 + o_P(1)$
- *Let* $Z^* \sim N(0, I^{r\times r})$ *where* $r = p + q$, $\widetilde{Z}_n = \widehat{\Sigma}_n^{1/2} Z^* = (\widetilde{Z}_Y^\mathsf{T}, \widetilde{Z}_Z^\mathsf{T}, \widetilde{Z}_A^\mathsf{T})^\mathsf{T}$
- *Define* $\widetilde{U}_n = \text{argmin}_{u\in\mathbb{R}^d} \widehat{\beta}_n^\mathsf{T} \widetilde{k}_n(\widetilde{Z}_Y, \widetilde{Z}_Z, u)$ *and* $\widetilde{B}_n = I_n(\widehat{\beta}_n)^{-1}\{Z_A - \widetilde{k}_n(\widetilde{Z}_Y, \widetilde{Z}_Z, \widetilde{U}_n)\}$
- *Then*

$$\begin{pmatrix} \widetilde{U}_n \\ \widetilde{B}_n \end{pmatrix} \overset{P}{\underset{Z^*}{\rightsquigarrow}} \begin{pmatrix} U \\ B \end{pmatrix},$$

*where* $U$ *and* $B$ *are as defined on slide 43*

# Simulation Studies

- $\mathbf{X} = (X_1, ..., X_5)^\mathsf{T} \sim N(\mathbf{0}, \Sigma = 0.5^2 I)$
- $Y = A(4X_1 = 2X_2 + 2) + \epsilon_Y$
  $Z = A(2X_1 - 4X_2 - 2) + \epsilon_Z$
  where $\epsilon_Y \sim \epsilon_Z \sim N(\mathbf{0}, 0.5^2)$
- Setting 1:
  - $\Pr\{A = d^{\text{opt}}(\mathbf{x}) \,|\, \mathbf{X} = \mathbf{x}\} = \rho$
- Setting 2:
  - $\Pr\{A = d^{\text{opt}}(\mathbf{X})\} = \text{expit}(0.5 + X_1)$
- Setting 3:
  - $\Pr\{A = d^{\text{opt}}(\mathbf{X})\} = \text{expit}(0.5 + X_1)$
  - $\omega(\mathbf{X}; \theta) = \text{expit}(1 - 0.5X_1)$

# Simulation Results

- Value results for simulations where utility ($\omega$) and probability of optimal treatment ($\rho$) are fixed

| $n$ | $\omega$ | $\rho$ | Optimal | Estimated $\omega$ | Y only | Z only | Standard of care |
|---|---|---|---|---|---|---|---|
| 100 | 0.25 | 0.60 | 1.90 (0.07) | 1.75 (0.29) | 0.39 (0.12) | 1.77 (0.07) | 0.39 (0.23) |
| | | 0.80 | 1.90 (0.07) | 1.88 (0.07) | 0.39 (0.12) | 1.77 (0.07) | 1.14 (0.21) |
| | 0.75 | 0.60 | 1.89 (0.07) | 1.69 (0.40) | 1.76 (0.08) | 0.39 (0.12) | 0.40 (0.23) |
| | | 0.80 | 1.89 (0.07) | 1.89 (0.07) | 1.76 (0.08) | 0.39 (0.12) | 1.15 (0.21) |
| 200 | 0.25 | 0.60 | 1.90 (0.07) | 1.80 (0.25) | 0.39 (0.11) | 1.77 (0.07) | 0.38 (0.17) |
| | | 0.80 | 1.90 (0.07) | 1.89 (0.06) | 0.39 (0.11) | 1.77 (0.07) | 1.15 (0.15) |
| | 0.75 | 0.60 | 1.90 (0.07) | 1.79 (0.26) | 1.76 (0.07) | 0.38 (0.11) | 0.38 (0.17) |
| | | 0.80 | 1.90 (0.07) | 1.89 (0.06) | 1.76 (0.07) | 0.38 (0.11) | 1.16 (0.15) |
| 300 | 0.25 | 0.60 | 1.90 (0.07) | 1.86 (0.13) | 0.37 (0.11) | 1.76 (0.08) | 0.38 (0.13) |
| | | 0.80 | 1.90 (0.07) | 1.89 (0.07) | 0.37 (0.11) | 1.76 (0.08) | 1.14 (0.12) |
| | 0.75 | 0.60 | 1.90 (0.06) | 1.84 (0.19) | 1.76 (0.08) | 0.39 (0.11) | 0.39 (0.13) |
| | | 0.80 | 1.90 (0.06) | 1.90 (0.07) | 1.76 (0.08) | 0.39 (0.11) | 1.15 (0.12) |
| 500 | 0.25 | 0.60 | 1.90 (0.06) | 1.88 (0.08) | 0.38 (0.11) | 1.77 (0.07) | 0.37 (0.11) |
| | | 0.80 | 1.90 (0.06) | 1.90 (0.06) | 0.38 (0.11) | 1.77 (0.07) | 1.13 (0.09) |
| | 0.75 | 0.60 | 1.90 (0.07) | 1.88 (0.08) | 1.76 (0.08) | 0.39 (0.11) | 0.37 (0.10) |
| | | 0.80 | 1.90 (0.07) | 1.90 (0.07) | 1.76 (0.08) | 0.39 (0.11) | 1.13 (0.09) |

# Simulation Results

- Value results for simulations where utility ($\omega$) is fixed and probability of optimal treatment is variable

| n | $\omega$ | Optimal | Estimated $\omega$ | Y only | Z only | SoC |
|---|---|---|---|---|---|---|
| 100 | 0.25 | 1.90 (0.06) | 1.72 (0.41) | 0.40 (0.11) | 1.76 (0.07) | 0.33 (0.24) |
| | 0.75 | 1.90 (0.06) | 1.76 (0.29) | 1.76 (0.07) | 0.38 (0.12) | 0.58 (0.24) |
| 200 | 0.25 | 1.90 (0.06) | 1.84 (0.24) | 0.38 (0.11) | 1.75 (0.08) | 0.32 (0.16) |
| | 0.75 | 1.90 (0.06) | 1.84 (0.16) | 1.76 (0.07) | 0.38 (0.11) | 0.57 (0.16) |
| 300 | 0.25 | 1.89 (0.07) | 1.88 (0.14) | 0.38 (0.11) | 1.77 (0.07) | 0.32 (0.14) |
| | 0.75 | 1.90 (0.07) | 1.87 (0.09) | 1.76 (0.07) | 0.39 (0.12) | 0.56 (0.14) |
| 500 | 0.25 | 1.90 (0.07) | 1.90 (0.06) | 0.38 (0.11) | 1.77 (0.07) | 0.33 (0.10) |
| | 0.75 | 1.90 (0.07) | 1.89 (0.08) | 1.76 (0.07) | 0.39 (0.11) | 0.57 (0.10) |

# Simulation Results

- Value results for simulations where both utility and probability of optimal treatment are variable

| n | Optimal | Estimated $\omega$ | Y only | Z only | Standard of care |
|---|---|---|---|---|---|
| 100 | 1.74 (0.06) | 1.53 (0.19) | 1.59 (0.07) | 0.44 (0.11) | 0.51 (0.21) |
| 200 | 1.73 (0.06) | 1.61 (0.13) | 1.59 (0.07) | 0.44 (0.10) | 0.51 (0.15) |
| 300 | 1.74 (0.06) | 1.64 (0.12) | 1.59 (0.07) | 0.44 (0.10) | 0.50 (0.13) |
| 500 | 1.74 (0.06) | 1.68 (0.09) | 1.59 (0.07) | 0.43 (0.10) | 0.50 (0.09) |

# Misspecified Model for the Utility Function

- Let the true underlying utility function be $u(y, z; bx, \theta) = \omega \mathbf{x}; \theta)y + \{1 - \omega(\mathbf{x}; \theta)\}z$
- Where $\omega(\mathbf{x}; \theta) = \text{expit}(1 + x_1^2 + \mathbf{x}^T\theta_0)$
- Consider a misspecified model fit to estimate the utility function containing only an intercept, $X_1, X_2, X_3$, and $X_4$
  - i.e., one important covariate and a squared term are omitted from the model for the utility function

| n | Optimal | Correct | Misspecified | Standard of Care |
|---|---|---|---|---|
| 100 | 1.86 (0.07) | 1.61 (0.21) | 1.64 (0.20) | 0.59 (0.23) |
| 200 | 1.85 (0.07) | 1.68 (0.16) | 1.69 (0.17) | 0.57 (0.16) |
| 300 | 1.86 (0.07) | 1.72 (0.13) | 1.74 (0.13) | 0.57 (0.13) |
| 500 | 1.86 (0.07) | 1.77 (0.10) | 1.76 (0.11) | 0.58 (0.10) |

# Analysis of STEP-BD SCP Data

- Included an observational study with 1437 patients having bipolar disorder (Sachs et al, 2007, *NEJM*).
- Using the proposed method, we were able to estimate an improved decision rule which led to a 7% improvement (p-value $< 0.0001$).
- Both increased age and history of substance abuse were important factors leading to lower recommended use of antidepressants.
- If we selected the two outcomes to be depression and side effect burden, we obtain an improvement of 9% (p-value $< 0.001$).

# Conclusions

- We can estimate patient utilities if we assume that clinicians make treatment decisions with the goal of maximizing each patient's utility
- Accounting for patient specific utilities can improve outcomes over standard of care
- Early results suggest the method is robust to utility model misspecificaion, but more research is needed
- The approach could be extended to multiple decision times, more than two outcomes, and more than two possible treatments
- A Bayesian approach could be developed to handle the non-smooth pseudo-likelihood