Considerations in Optimizing Personalized Treatments: Estimation and Evaluation in Light of Benefit and Risk

Yuanjia Wang, Ph.D.

Department of Biostatistics, Mailman School of Public Health & Division of Biostatistics, Department of Psychiatry
Columbia University
Motivation and Introduction
Personalized Medicine

President’s Council of Advisors on Science and Technology (PCAST) on Personalized Medicine refers to: "the tailoring of medical treatment to the individual characteristics of each patient”.

- Clinicians use individual characteristics to guide optimal treatment and provide best clinical care for a patient
  - Heterogeneity in responses:
    1. **Across patients**: what works for one may not work for another (MDD response rate 40%, Gaynes et al., 2009).
    2. **Within a patient**: what works now may not work later (MDD relapse rate 50%, APA 2000).

- Current practice: largely based on “trial-and-error”

- Research goal: evidence-based tailored treatment by distinguishing in advance those patients most likely to benefit from a given treatment
Individualized (dynamic) treatment rules (ITR/DTR, Lavori & Dawson 1998; Murphy 2005): tailoring by time-varying clinical characteristics, genetic profile, and intermediate outcomes.

Examples of ITR/DTR: Healing Emotion After Loss (HEAL, Shear et al. 2016):

- Administer grief-informed clinical management as the initial treatment; if a patient responds then continue; if a patient does not respond then offer an anti-depressant (Citalopram).

Adaptive Pharmacological Behavioral Treatments for Children with ADHD (Pelham 2002).

- Prescribe behavioral modification as the initial treatment; if a child responds then continue; if a child does not respond then switch to medication.
Rationale for Tailoring a Treatment

Three types of covariates:

- **prognostic variables** (associated with clinical response, no interaction with treatment $A$)
- **predictive variables** (quantitative interaction)
- **prescriptive variables/tailoring variables** (qualitative interaction)
Data Available for Personalized Medicine


- Demographics, Co-morbidity, Stage of disease, Genomic (oncology), Imaging (neurology, psychiatry), Mobile technology (neurology, psychiatry), Electronic health records

Yuanjia Wang, Department of Biostatistics, Columbia University
Data-Driven Approaches

Analytical challenges for discovering optimal ITR:

- Tailoring variables unknown
- High-dimensionality
- Correlation and structure among variables

Existing machine learning methods:

- Virtue twins (Foster et al. 2011)
- Q-learning (Qian and Murphy 2011; Nahum-Shani et al. 2012)
- O-learning (Zhao et al. 2012, 2014)
Q-learning and O-learning

Q-learning: Backwards induction. Single decision point, decompose expected outcome into two components and maximize over treatment domain:

$$E(Y|X, A; \psi, \beta) = G(X; \psi) + H(X, A; \beta)$$

- Treatment-free model (effect of patient history on outcome without treatment): $G(X; \psi)$
- Blip model (effect of treatment on the outcome): $H(X, A; \beta)$
- Pose models to maximize value function (Qian and Murphy 2012)

O-learning: Directly maximize the value function

$$\max_{D} E^D(Y)$$

Prior work on ITR focus on maximizing efficacy outcomes
Introduction to Our Work
Why considering safety outcomes?

Clinician: Complete picture of treatment decision making involves both efficacy and safety

- Most efficacious treatment for a patient may also lead to greater safety concern (escalating dosage of insulin may increase risk of hypoglycemia; Zhao et al. 2013)

- Patients with chronic disease (e.g., diabetes) and long duration treatment exposed to higher risk of adverse events (e.g., severe hypoglycemia, Wild et al., 2007)

Regulator/Industry: Important to characterize both the efficacy and risk profiles among patient populations (e.g., FDA guideline on evaluating cardiovascular risk for new antidiabetic therapies)
How to incorporate safety outcomes for estimating ITR?

- No treatment heterogeneity regarding safety outcomes
- Presence of heterogeneity:
  - Well known example: abundance of drug-metabolizing enzymes (cytochrome P45) varies across subjects, and thus adverse reactions to the same drug dosage
  - Risk of hypoglycemic events depends on patient characteristics and choice of treatment regimen (Sinclair et al., 2015)
Statistical Methodologies
Notation:

- $Y$: efficacy outcome (e.g., symptom reduction; change in HbA1c)
- $R$: risk outcome (hypoglycemia episodes)
- Two treatment arms $A \in \{-1, 1\}$
- Patient health history $X$
- Treatment rule $D(X)$: mapping from $X$ to $\{-1, 1\}$. 

Goal: estimate optimal $D^*$ while controlling for risk 

\[
\max_{D} \mathbb{E}_D(Y), \quad \text{s.t.} \quad \mathbb{E}_D(R) \leq \tau,
\]

\[
\mathbb{E}_D[\cdot]: \text{conditional expectation under probability measure } P_D \text{ for } (Y, R, A, X) \text{ given } A = D(X).
\]

$\tau$: pre-specified tolerance threshold of the risk
Framework for Personalized Benefit-Risk Analysis

Notation:
- $Y$: efficacy outcome (e.g., symptom reduction; change in HbA1c)
- $R$: risk outcome (hypoglycemia episodes)
- Two treatment arms $A \in \{-1, 1\}$
- Patient health history $X$
- Treatment rule $\mathcal{D}(X)$: mapping from $X$ to $\{-1, 1\}$.

Goal: estimate optimal ITR $\mathcal{D}^*$ while controlling for risk

$$\left\{ \begin{array}{l} \max_{\mathcal{D}} \mathbb{E}^{\mathcal{D}}(Y), \\
\text{s.t.} \quad \mathbb{E}^{\mathcal{D}}(R) \leq \tau, \end{array} \right.$$  

- $\mathbb{E}^{\mathcal{D}}[\cdot]$: conditional expectation under probability measure $\mathcal{P}^\mathcal{D}$ for $(Y, R, A, X)$ given $A = \mathcal{D}(X)$
- $\tau$: pre-specified tolerance threshold of the risk
Theoretical Optimal ITR Under Risk Constraint

Using data \((Y, R, A, X)\) collected from RCT, equivalent to (Qian and Murphy 2011):

\[
\begin{aligned}
\max_{D} & \quad E \left\{ \frac{I(A=D(X))}{p(A|X)} Y \right\} \\
\text{s.t.} & \quad E \left\{ \frac{I(A=D(X))}{p(A|X)} R \right\} \leq \tau
\end{aligned}
\]

Define \(D(X) = \text{sign}(f(X))\), the above is equivalent to

\[
\begin{aligned}
\max_{f} & \quad E \left\{ \delta_Y(X) I(f(X) > 0) \right\} \\
\text{s.t.} & \quad E[\delta_R(X) I(f(X) > 0)] \leq \alpha
\end{aligned}
\]

where

\[
\begin{aligned}
\delta_Y(X) &= E[Y|X, A = 1] - E[Y|X, A = -1], \\
\delta_R(X) &= E[R|X, A = 1] - E[R|X, A = -1],
\end{aligned}
\]

and \(\alpha = \tau - E[R|A = -1]\).
Theoretical Optimal ITR Under Risk Constraint

Key theoretical result: The optimal treatment rule under risk constraint is $D^*(X) = \text{sign}(f^*(X))$, where

$$f^*(X) = \begin{cases} 
\text{sign}(\delta Y(X)), & X \in A \\
\text{sign} (\delta Y(X) - \lambda^* \delta_R(X)), & X \in A^c 
\end{cases}$$

and $A = \{X : \delta_Y(X)\delta_R(X) \leq 0\}$. Here, $\lambda^* = 0$ if $E[\delta_R^+(X) | X \in A^c] \leq \alpha^*$; otherwise, $\lambda^*$ solves equation

$$E[\delta_R(X) I\{\delta_R(X) > 0, \delta_Y(X)/\delta_R(X) > \lambda\} | X \in A^c] + E[\delta_R(X) I\{\delta_R(X) < 0, \delta_Y(X)/\delta_R(X) < \lambda\} | X \in A^c] = \alpha^*$$

with $\alpha^* = \frac{\alpha - E[\delta_R(X) I(\delta_Y(X) > 0, X \in A)]}{P(X \in A^c)}$.

Remark 1. Solving for $D^*$ is analogous to finding the optimal rejection region at a given type I error rate as in the Neyman-Pearson lemma.

Remark 2. When no treatment heterogeneity on safety outcomes, apply with $\delta_R(X) = c$. 

Yuanjia Wang, Department of Biostatistics, Columbia University
Method 1: **BR-Q learning**

- Predictive modeling-based learning algorithm (reduces to Q-learning in the absence of $R$)
  - Step 1. Fit regression model for $Y$ given $(A, X)$, obtain
    \[
    \hat{\delta}_Y(X) = \hat{E}[Y|X, A = 1] - \hat{E}[Y|X, A = -1]
    \]
  - Step 2. Fit regression model for $R$ given $(A, X)$, obtain
    \[
    \hat{\delta}_R(X) = \hat{E}[R|X, A = 1] - \hat{E}[R|X, A = -1]
    \]
  - Step 3. Apply the theorem:
    \[
    \hat{f}(X) = \begin{cases} 
    \text{sign}(\hat{\delta}_Y(X)), & X \in \hat{A} \\
    \text{sign}\left(\hat{\delta}_Y(X) - \lambda\hat{\delta}_R(X)\right), & X \in \hat{A}^c.
    \end{cases}
    \]
Method 2: **BR-O learning**

- Directly estimate $D^*$ under risk constraint without posing a regression model (reduces to O-learning in the absence of $R$):

$$\left\{ \begin{array}{l}
\max_{D} E \left\{ \frac{I(A=D(X))}{p(A|X)} Y \right\}, \\
\text{s.t.} \ E \left\{ \frac{I(A=D(X))}{p(A|X)} R \right\} \leq \tau.
\end{array} \right.$$  

Maximizes empirical value function under constraint:

$$\left\{ \begin{array}{l}
\max_{f} n^{-1} \sum_{i=1}^{n} \frac{Y_i}{p(A_i|X_i)} I(A_i = \text{sign}(f(X_i))), \\
\text{s.t.} \ n^{-1} \sum_{i=1}^{n} \frac{R_i}{p(A_i|X_i)} I(A_i = \text{sign}(f(X_i))) \leq \tau.
\end{array} \right.$$
Implementation of BR-O learning

**Challenges:** constrained optimization with non-convex objective function and non-convex constraint.

**Solution:** approximate \( I(A_i \neq \text{sign}(f(X_i))) \) in objective function by a surrogate hinge loss, and approximate \( I(A_i = \text{sign}(f(X_i))) \) in the constraint by a shifted ramp loss (Huang et al. 2014) as upper bound

\[
\psi_\delta(u) = f^1_\delta(u) - f^0_\delta(u) \\
= \delta^{-1}(u + \delta)_+ - \delta^{-1}(u)_+.
\]
Implementation of BR-O learning

The optimization solved by difference of convex functions algorithm (DCA) (Tao and An 1998) and quadratic programming:

\[
\begin{align*}
\min_f & \quad C \sum_{i=1}^{n} \frac{Y^*}{p_i} \xi_i + \frac{1}{2} \beta^T_{(0)} K \beta_{(0)}, \\
\text{s.t.} & \quad \sum_{i=1}^{n} \frac{R_i}{p_i} \left[ \delta^{-1} \{ A_if(X_i) + \delta \} + - \delta^{-1} \{ A_if(X_i) \} \right] \leq n\tau, \\
& \quad \xi_i \geq 1 - A_i^* \{ \beta_0 + \sum_{j=1}^{n} \beta_j K(X_i, X_j) \}, \xi_i \geq 0 \quad \forall i.
\end{align*}
\]

Tuning parameters $C$ and $\delta$ selected by cross validation.
Numeric Results
Simulation Design

- 20 covariates as $X_1, \ldots, X_{20}$ i.i.d. $U(0, 1), n = 300$
- Efficacy responses are standard normal:
  \[ Y = 1 - 2X_1 + X_2 - X_3 + h_Y(X, A) + \epsilon_Y \]
  \[ h_Y = 2 \times (1 - X_1 - X_2) \times A. \]
- Safety responses are truncated normal (truncated at 1):
  \[ R = 2 + X_1 - 2X_2 - X_3 + h_R(X, A) + \epsilon_R \]
  \[ h_R = (1 + X_1 - X_2) \times A. \]
- Prognostic variables: $X_1, X_2, X_3$
- Prescriptive variables:
  - For $Y$ not considering $R$: optimal boundary $1 - X_1 - X_2$,
    positive indicates $A = 1$ is more efficacious
Figure: Regions of Optimal Treatments with and without Risk Constraint and Relationship with Average Benefit and Risk ($\tau = 1$): linear boundary
Theoretical Optimal Treatment Decision Boundaries

Figure: Regions of Optimal Treatments with and without Risk Constraint and Relationship with Average Benefit and Risk ($\tau = 1.75$): nonlinear boundary

Yuanjia Wang, Department of Biostatistics, Columbia University
Compare optimal rule computed using theoretical result, BR-Q and BR-O.

Table: Estimated average risk and optimal benefit†.

<table>
<thead>
<tr>
<th>τ</th>
<th>Safety outcome R</th>
<th>Efficacy outcome Y</th>
<th>% Correct</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Theo</td>
<td>BR-Q</td>
<td>BR-O‡</td>
</tr>
<tr>
<td>0.50</td>
<td>0.494</td>
<td>0.530</td>
<td>0.521</td>
</tr>
<tr>
<td>0.75</td>
<td>0.745</td>
<td>0.750</td>
<td>0.697</td>
</tr>
<tr>
<td>1.00</td>
<td>0.995</td>
<td>0.994</td>
<td>0.919</td>
</tr>
<tr>
<td>1.25</td>
<td>1.243</td>
<td>1.233</td>
<td>1.158</td>
</tr>
<tr>
<td>1.50</td>
<td>1.490</td>
<td>1.445</td>
<td>1.327</td>
</tr>
</tbody>
</table>

†: Average safety outcome is 1.503, and optimal value function without safety constraint is 0.662.
DURAbility of Basal Versus Lispro Mix 75/25 Insulin Efficacy (DURABLE) Trial (Buse et al., 2009):

- Randomized trial to compare the ability of two starter insulin regimens (once-daily basal insulin Glargin or twice-daily premixed insulin Lispro 75/25) to achieve glycemic control in patients with type 2 diabetes

- Insulin-naive patients with type 2 diabetes who did not achieve adequate control with oral antihyperglycemic drugs

- Efficacy outcome: glycemic control (change in HbA1C from baseline to end point)

- Safety outcomes: hypoglycemia (a plasma glucose value $\leq 70$ mg/dl or presence of typically associated symptoms)
Application to DURABLE Trial

Overall analyses results (Buse et al., 2009):

- **Efficacy**: Lispro 75/25 better control on glycemic than GL ($p = 0.005$)

- **Safety**: Lispro 75/25 higher hypoglycemia rate compared to GL ($p = 0.007$)

Application Data Description:

- Sample size: 965 Lispro Mix and 980 insulin Glargin.
- Efficacy endpoint: A1c change from baseline after 24 weeks treatment.
- Safety endpoint: Hypoglycemic event rate per day.
- Covariates: 18 baseline covariates (weight, BMI, blood pressure, heart rate, 7 points blood glucose values, fasting blood glucose, fasting insulin etc.).
Table. Average benefit, risk over 100 repetitions (300 patients as training set, the rest 1,645 patients as a testing dataset).

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>0.063</td>
<td>BR-Q</td>
<td>0.0639(0.005)</td>
<td>0.0689(0.004)</td>
<td>1.8668(0.142)</td>
<td>1.7201(0.049)</td>
</tr>
<tr>
<td></td>
<td>BR-O</td>
<td>0.0626(0.003)</td>
<td>0.0640(0.006)</td>
<td>1.7824(0.142)</td>
<td>1.6980(0.042)</td>
</tr>
<tr>
<td>0.064</td>
<td>BR-Q</td>
<td>0.0644(0.006)</td>
<td>0.0690(0.004)</td>
<td>1.8682(0.141)</td>
<td>1.7209(0.050)</td>
</tr>
<tr>
<td></td>
<td>BR-O</td>
<td>0.0632(0.003)</td>
<td>0.0650(0.006)</td>
<td>1.7905(0.135)</td>
<td>1.7004(0.050)</td>
</tr>
<tr>
<td>0.065</td>
<td>BR-Q</td>
<td>0.0652(0.006)</td>
<td>0.0692(0.004)</td>
<td>1.8736(0.142)</td>
<td>1.7228(0.050)</td>
</tr>
<tr>
<td></td>
<td>BR-O</td>
<td>0.0638(0.003)</td>
<td>0.0650(0.006)</td>
<td>1.7983(0.135)</td>
<td>1.7030(0.051)</td>
</tr>
<tr>
<td>0.066</td>
<td>BR-Q</td>
<td>0.0657(0.006)</td>
<td>0.0694(0.004)</td>
<td>1.8780(0.146)</td>
<td>1.7241(0.051)</td>
</tr>
<tr>
<td></td>
<td>BR-O</td>
<td>0.0644(0.003)</td>
<td>0.0655(0.006)</td>
<td>1.8048(0.135)</td>
<td>1.7021(0.046)</td>
</tr>
<tr>
<td>0.067</td>
<td>BR-Q</td>
<td>0.0667(0.006)</td>
<td>0.0696(0.004)</td>
<td>1.8827(0.148)</td>
<td>1.7250(0.052)</td>
</tr>
<tr>
<td></td>
<td>BR-O</td>
<td>0.0654(0.003)</td>
<td>0.0660(0.006)</td>
<td>1.8273(0.131)</td>
<td>1.7093(0.048)</td>
</tr>
<tr>
<td>∞</td>
<td>BR-Q</td>
<td>0.0756(0.010)</td>
<td>0.0712(0.003)</td>
<td>1.9392(0.153)</td>
<td>1.7378(0.048)</td>
</tr>
<tr>
<td></td>
<td>BR-O</td>
<td>0.0769(0.010)</td>
<td>0.0714(0.005)</td>
<td>1.9895(0.146)</td>
<td>1.7360(0.052)</td>
</tr>
</tbody>
</table>

**Conclusion:** BR-O controls risk below τ with similar benefit as BR-Q
**Ranks of Important Biomarkers**

Table: Ranking of Baseline Biomarkers Based on Average Standardized Effects over 100 Repetitions.

<table>
<thead>
<tr>
<th></th>
<th>$\tau = 0.063$</th>
<th>$0.064$</th>
<th>$0.065$</th>
<th>$0.066$</th>
<th>$0.067$</th>
<th>$\infty$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline A1C</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>BMI</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Fasting Blood Glucose</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>Height</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>Adiponectin</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>Duration of diabetes</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td>Body Weight</td>
<td>7</td>
<td>7</td>
<td>7</td>
<td>7</td>
<td>7</td>
<td>7</td>
</tr>
<tr>
<td>Diastolic blood pressure</td>
<td>8</td>
<td>8</td>
<td>8</td>
<td>8</td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td>Fasting Insulin</td>
<td>9</td>
<td>9</td>
<td>9</td>
<td>9</td>
<td>9</td>
<td>10</td>
</tr>
<tr>
<td>Heart rate</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>9</td>
</tr>
<tr>
<td>Systolic blood pressure</td>
<td>11</td>
<td>11</td>
<td>11</td>
<td>11</td>
<td>11</td>
<td>11</td>
</tr>
<tr>
<td>Glucose:Morning before meal</td>
<td>12</td>
<td>12</td>
<td>12</td>
<td>12</td>
<td>12</td>
<td>12</td>
</tr>
<tr>
<td>Glucose:3am at night</td>
<td>13</td>
<td>13</td>
<td>13</td>
<td>13</td>
<td>13</td>
<td>14</td>
</tr>
<tr>
<td>Glucose:Evening before meal</td>
<td>14</td>
<td>14</td>
<td>14</td>
<td>14</td>
<td>14</td>
<td>13</td>
</tr>
<tr>
<td>Glucose:Morning 2 hours after meal</td>
<td>15</td>
<td>15</td>
<td>16</td>
<td>15</td>
<td>15</td>
<td>16</td>
</tr>
<tr>
<td>Glucose:Evening after meal</td>
<td>16</td>
<td>16</td>
<td>15</td>
<td>16</td>
<td>16</td>
<td>15</td>
</tr>
<tr>
<td>Glucose:Noon before meal</td>
<td>17</td>
<td>17</td>
<td>18</td>
<td>17</td>
<td>17</td>
<td>18</td>
</tr>
<tr>
<td>Glucose:Noon 2 hours after meal</td>
<td>18</td>
<td>18</td>
<td>17</td>
<td>18</td>
<td>18</td>
<td>17</td>
</tr>
</tbody>
</table>

Yuanjia Wang, Department of Biostatistics, Columbia University
Figure: **Black** dashed line: O-learning without risk constraint. **Red** solid line: BR-O ($\tau = 0.065$). Patients above the lines recommended to take mix 75/25 and patients below recommended to take GL.
Conclusion and Extensions
We propose two methods (BR-Q and BR-O) for estimating optimal ITR while controlling for average risk.

- Both control theoretical risk adequately and approach theoretical optimal efficacy level.
- In the application, BR-O slightly conservative on the training, but controls risk better on testing.
- BR-O more computationally intensive

Extensions:

- Multiple efficacy and safety outcomes
- Multiple group-dependent thresholds
- Multi-stage trials (SMART, Lavori & Dawson 2000, 2004; Murphy 2005)
- Identify safest ITR while maintaining minimal benefit
Extension to Real World Setting

Electronic health records data: CUMC clinical data warehouse (20 years of health information for about 4.5 million patients with diverse ethnicity).

However, noisy, error prone, extract lab tests measurement patterns that reflect healthcare process.

Yuanjia Wang, Department of Biostatistics, Columbia University
Extension to Real World Setting

One type of most informative treatment distinguishing variables (lab tests recording patterns):

Simple frequencies are used. More sophisticated methods can be developed to construct decision tree.
Acknowledgements

- Collaborators and students:
  - Haoda Fu (Eli Lilly)
  - Christine Mauro (Biostatistics, Columbia University)
  - Xin Qiu (Biostatistics, Columbia University)
  - Katherine Shear (Department of Psychiatry, Columbia University)
  - Chunhua Weng (Biomedical Informatics, Columbia University)
  - Peng Wu (Biostatistics, Columbia University)
  - Donglin Zeng (Biostatistics, UNC Chapel Hill)

- Research support
  - NIH: NS073671, NS082062, MH60783
  - Research Foundation of Mental Hygiene (RFMH)

THANK YOU!