Estimating the optimal dynamic treatment regime from longitudinal observational data

Liliana Orellana[1], Andrea Rotnitzky[2,3] and James Robins[2]

[1]Universidad de Buenos Aires, [2]Harvard University, [3]Universidad Di Tella

IBS AR - NZ - Dec 2009

Motivation

- Highly active antiretroviral therapy (HAART) dramatically decreased morbidity and mortality due to infection with HIV.
- Eradication of HIV infection cannot be achieved with available antiretroviral regimens.
- Late initiation of HAART has both risks and benefits:
  - Risks: Irreversible damage of the immune system; AIDS.
  - Benefits: Avoid drug toxicity and side effects; delay drug resistance.

Motivation

- Highly active antiretroviral therapy (HAART) dramatically decreased morbidity and mortality due to infection with HIV.
- Eradication of HIV infection cannot be achieved with available antiretroviral regimens.
- Late initiation of HAART has both risks and benefits:
  - Risks: Irreversible damage of the immune system; AIDS.
  - Benefits: Avoid drug toxicity and side effects; delay drug resistance.

QUESTION $\Longrightarrow$ When to start HAART?

- Decision on "when to start" for *asymptomatic* HIV+ subjects is essentially based on CD4 cell count.

When to start HAART?

U.S. Treatment Guidelines for HIV-1 Infected Adults and Adolescents (October 2006)

Recommendations on when to start for asymptomatic HIV+ subjects:

- Definitively start if CD4 count < 200 cells/mm^3.
- *Unclear* if CD4 count > 200 cells/mm^3.
  - Offer trx if $200 <$ CD4 ≤ 350 cells/mm^3.
  - Preferably defer trx if CD4 > 350 cells/mm^3.
- A treatment strategy based on CD4 counts is an example of a *dynamic treatment regime*.

Dynamic treatment regimes

Data

$$L_0, A_0, L_1, A_1, ..., L_K, A_K, L_{K+1}$$

L_k = clinical and laboratory variables measured during the k^{th} clinic visit,
A_k = treatment prescription at visit k
$\overline{L}_k = (L_0, L_1, ..., L_k)$ and $\overline{A}_k = (A_0, A_1, ..., A_k)$

Dynamic treatment regimes

Data

$$L_0, A_0, L_1, A_1, ..., L_K, A_K, L_{K+1}$$

L_k = clinical and laboratory variables measured during the k^{th} clinic visit,
A_k = treatment prescription at visit k
$\overline{L}_k = (L_0, L_1, ..., L_k)$ and $\overline{A}_k = (A_0, A_1, ..., A_k)$

Dynamic treatment regime

- Sequential rule for determining, at each time k, the next treatment prescription A_k.
- Rule inputs the recorded health information up to time k and returns a treatment recommendation

$$\left(\overline{L}_k, \overline{A}_{k-1}\right) \rightarrow \; d_k\left(\overline{L}_k, \overline{A}_{k-1}\right) \in \mathcal{A}_k \;, k = 0, ..., K.$$

Dynamic treatment regimes

Data

$$L_0, A_0, L_1, A_1, ..., L_K, A_K, L_{K+1}$$

L_k = clinical and laboratory variables measured during the k^{th} clinic visit,
A_k = treatment prescription at visit k
$\overline{L}_k = (L_0, L_1, ..., L_k)$ and $\overline{A}_k = (A_0, A_1, ..., A_k)$

Dynamic treatment regime

- Sequential rule for determining, at each time k, the next treatment prescription A_k.
- Rule inputs the recorded health information up to time k and returns a treatment recommendation

$$\left(\overline{L}_k, \overline{A}_{k-1}\right) \rightarrow \; d_k\left(\overline{L}_k, \overline{A}_{k-1}\right) \in \mathcal{A}_k \;, k = 0, ..., K.$$

Optimal dynamic regime

- Maximizes the expectation of some utility function $Y \equiv u\left(\overline{L}_{K+1}, \overline{A}_K\right)$ among the set of candidate regimes.

An ideal randomized study to compare two regimes

Suppose we want to compare two dynamic regimes:

- start HAART when CD4 falls below 500 (d^{500})
- start HAART when CD4 falls below 200 (d^{200})

An ideal randomized study to compare two regimes

Suppose we want to compare two dynamic regimes:

- start HAART when CD4 falls below 500 (d^{500})
- start HAART when CD4 falls below 200 (d^{200})

DESIGN

- Follow patients periodically, say every 6 months, from HIV diagnosis
- When CD4 first falls below 500 randomize to
  - start immediately (say, $p = 1/2$)
  - start when CD4 first seen to fall below 200 (say, $p = 1/2$)
- Let Y be the outcome, a utility function of the health and treatment history (higher values are preferable)
- Compare outcome in the two groups after a number of years of follow-up (e.g., 5 years)

An ideal randomized study to compare two regimes

Estimation goal

We want to compare the expected utility in:

1. a hypothetical world where regime d^{500} was enforced $\left(\mu^{500}\right)$ versus
2. a hypothetical world where regime d^{200} was enforced $\left(\mu^{200}\right)$.

An ideal randomized study to compare two regimes

Estimation goal

We want to compare the expected utility in:

1. a hypothetical world where regime d^{500} was enforced $\left(\mu^{500}\right)$ versus
2. a hypothetical world where regime d^{200} was enforced $\left(\mu^{200}\right)$.

We can estimate this contrast from our ideal clinical trial because

- randomization generates exchangeable groups and
- each subject can be assigned to any regime.

An ideal randomized study to compare two regimes

Estimation goal

We want to compare the expected utility in:

1. a hypothetical world where regime d^{500} was enforced (μ^{500}) versus
2. a hypothetical world where regime d^{200} was enforced (μ^{200}).

We can estimate this contrast from our ideal clinical trial because

- randomization generates exchangeable groups and
- each subject can be assigned to any regime.

It is difficult to conduct such a trial to compare many regimes.
We must then rely on observational data

Observational analogue of a randomized trial

Interview HIV+ subjects periodically (say, every 6 months)

- record treatment modifications over the last time interval,
- measure biological and clinical markers at interview.

Observational analogue of a randomized trial

Interview HIV+ subjects periodically (say, every 6 months)

- record treatment modifications over the last time interval,
- measure biological and clinical markers at interview.

Naive analysis

- Define baseline as time when CD4 first falls below 500.
- Regard subject is in:
  - Group I: if he initiates HAART when first seen to fall below 500.
  - Group II: if he starts HAART when first seen to fall below 200.
- Because treatment was not randomized we compare groups after adjusting for baseline potential confounding factors.

Observational analogue of a randomized trial

Interview HIV+ subjects periodically (say, every 6 months)

- record treatment modifications over the last time interval,
- measure biological and clinical markers at interview.

Naive analysis

- Define baseline as time when CD4 first falls below 500.
- Regard subject is in:
  - Group I: if he initiates HAART when first seen to fall below 500.
  - Group II: if he starts HAART when first seen to fall below 200.
- Because treatment was not randomized we compare groups after adjusting for baseline potential confounding factors.

PROBLEM $\Rightarrow$ SELECTION BIAS

- Subjects not included in Group I or II can't be ignored.

Observational analogue of a randomized trial

Interview HIV+ subjects periodically (say, every 6 months)

- record treatment modifications over the last time interval,
- measure biological and clinical markers at interview.

Naive analysis

- Define baseline as time when CD4 first falls below 500.
- Regard subject is in:
  - Group I: if he initiates HAART when first seen to fall below 500.
  - Group II: if he starts HAART when first seen to fall below 200.
- Because treatment was not randomized we compare groups after adjusting for baseline potential confounding factors.

PROBLEM $\Rightarrow$ SELECTION BIAS

- Subjects not included in Group I or II can't be ignored.
- Selection bias can be corrected using Inverse Probability of Censoring Weighted (IPCW) methods.

Inverse probability of censoring weighted estimation

Suppose we want to estimate

$$\mu^{200} = E\left(Y^{200}\right)$$

μ^{200} = expected utility in a world where all subjects followed regime d^{200} (start HAART first time CD4 falls below 200).

Inverse probability of censoring weighted estimation

Suppose we want to estimate

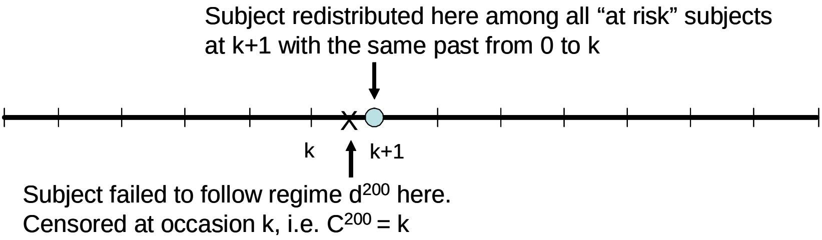
$$\mu^{200} = E\left(Y^{200}\right)$$

μ^{200} = expected utility in a world where all subjects followed regime d^{200} (start HAART first time CD4 falls below 200).

Strategy

- Censor a subject at occasion k if he/she:
  - started HAART at occasion k prior to falling below 200 or
  - failed to start HAART at occasion k when falling below 200.
- Redistribute the censored subject among those still **at risk** (following regime d^{200}) who have the same history up to k.
- The process is repeated for $k = 0, ..., K$.

Inverse probability of censoring weighted estimation



Inverse probability of censoring weighted estimation

Data recorded in the cohort study

$$L_0, A_0, L_1, A_1, ..., L_K, A_K, L_{K+1}$$

L_k = vector of covariates measured at time k,
A_k = HAART indicator.

Inverse probability of censoring weighted estimation

Data recorded in the cohort study

$$L_0, A_0, L_1, A_1, ..., L_K, A_K, L_{K+1}$$

L_k = vector of covariates measured at time k,
A_k = HAART indicator.

Accumulated weight through occasion k for a subject is estimated as

$$W_k^{200} = \frac{I\left(C^{200} > k\right)}{\prod_{j=1}^{k} \widehat{\Pr}\left(C^{200} > j \middle| C^{200} > j-1, \overline{A}_{j-1}, \overline{L}_j\right)}$$

where C^{200} = time to censoring under regime d^{200}.

- Numerator is the indicator of following regime d^{200} through k.
- Denominator estimates the probability a subject had his observed HAART history through k.
- Usually $\overline{L}_j$ is a high dimensional vector, so a parametrical model is assumed for the censoring probabilities.

Inverse probability of censoring weighted estimation

We estimate $\mu^{200} \equiv E\left(Y^{200}\right)$ with

$$\widehat{\mu}^{200} = \frac{\sum_{i=1}^{n} W_i^{200} Y_i}{\sum_{i=1}^{n} W_i^{200}}$$

where W^{200} is the accumulated weight at the end of study.

- $\widehat{\mu}^{200}$ is a weighted average of the outcomes of those patients who followed regime d^{200} throughout.

Inverse probability of censoring weighted estimation

We estimate $\mu^{200} \equiv E\left(Y^{200}\right)$ with

$$\widehat{\mu}^{200} = \frac{\sum_{i=1}^{n} W_i^{200} Y_i}{\sum_{i=1}^{n} W_i^{200}}$$

where W^{200} is the accumulated weight at the end of study.

- $\widehat{\mu}^{200}$ is a weighted average of the outcomes of those patients who followed regime d^{200} throughout.

The estimator is consistent and asymptotically normal if:

- Model for hazard of censoring is correctly specified.
- At each time k recorded data includes all covariates used by doctors to prescribe HAART.
  - Sequential Randomization or No Unmeasured Confounders Assumption.
  - Non-testable!

Estimating the optimal regime in a candidate set

- We want to compare regimes *$d^x \equiv$ start HAART first time CD4 falls below x*, where $x \in \mathcal{X} = \{200, 201, ..., 500\}$.
- In principle, we can estimate each mean $\mu^x \equiv E(Y^x)$ separately and then find $\widehat{x}_{opt}$ that maximizes $\widehat{\mu}^x$.
- However, estimates $\widehat{\mu}^x$ will have high variance because each regime will be followed by few subjects.
- Even in the ideal randomized trial we would also face this small cell problem.

Estimating the optimal regime in a candidate set

- We want to compare regimes $d^x \equiv$ *start HAART first time CD4 falls below* x, where $x \in \mathcal{X} = \{200, 201, ..., 500\}$.
- In principle, we can estimate each mean $\mu^x \equiv E(Y^x)$ separately and then find $\widehat{x}_{opt}$ that maximizes $\widehat{\mu}^x$.
- However, estimates $\widehat{\mu}^x$ will have high variance because each regime will be followed by few subjects.
- Even in the ideal randomized trial we would also face this small cell problem.

SOLUTION $\Rightarrow$ parametrically model $E(Y^x)$

Estimating the optimal CD4 in a candidate set: Proposal

Assume that

$$\mu^x = E\left(Y^x\right) = h\left(x;\beta\right) \tag{1}$$

where $h\left(x;\beta\right)$ is a known smooth function of a $p \times 1$ unknown parameter β.

For example,

$$h\left(x;\beta\right) = \beta_0 + \beta_1 x + \beta_2 x^2 + \beta_3 x^3 + \beta_4 x^4 + \beta_5 x^5$$

Estimating the optimal CD4 in a candidate set: Proposal

Assume that

$$\mu^x = E\left(Y^x\right) = h\left(x;\beta\right) \tag{1}$$

where $h\left(x;\beta\right)$ is a known smooth function of a $p \times 1$ unknown parameter β.

For example,

$$h\left(x;\beta\right) = \beta_0 + \beta_1 x + \beta_2 x^2 + \beta_3 x^3 + \beta_4 x^4 + \beta_5 x^5$$

Given an estimate $\widehat{\beta}$ of β we can find $\widehat{x}_{opt} = \arg\max\left(h\left(x;\widehat{\beta}\right)\right)$.

So, under model (1) the problem reduces to estimating β.

Estimating the optimal CD4 in a candidate set: Proposal

- Let $\gamma =$ number of regimes in the candidate set $\mathcal{X} = \{200, 201, ..., 500\}$.
- Create an artificial data set, with each subject contributing γ observations $\left(W_i^{x_j}, Y_i, x_j\right)$, $j = 1, ..., \gamma$.

Estimating the optimal CD4 in a candidate set: Proposal

- Let $\gamma =$ number of regimes in the candidate set $\mathcal{X} = \{200, 201, ..., 500\}$.
- Create an artificial data set, with each subject contributing γ observations $\left(W_i^{x_j}, Y_i, x_j\right)$, $j = 1, ..., \gamma$.
- Find $\widehat{\beta}$ solving the weighted estimating equation

$$\mathbb{P}_n \left\{ \sum_{x \in \mathcal{X}} \frac{\partial h(x;\beta)}{\partial \beta} W^x \left[Y - h(x;\beta) \right] \right\} = 0$$

Estimating the optimal CD4 in a candidate set: Proposal

- Let γ = number of regimes in the candidate set $\mathcal{X} = \{200, 201, ..., 500\}$.
- Create an artificial data set, with each subject contributing γ observations $\left(W_i^{x_j}, Y_i, x_j\right)$, $j = 1, ..., \gamma$.
- Find $\widehat{\beta}$ solving the weighted estimating equation

$$\mathbb{P}_n \left\{ \sum_{x \in \mathcal{X}} \frac{\partial h(x; \beta)}{\partial \beta} W^x \left[Y - h(x; \beta) \right] \right\} = 0$$

Model extensions (I)

Estimating equations can be modified to obtain estimators:

- doubly-robust
- locally efficient

Model extensions (I)

Estimating equations can be modified to obtain estimators:

- doubly-robust
- locally efficient

Can allow for the possibility that optimal CD4 count depends on baseline covariates Z by considering Parametric Marginal Structural Mean (MSM) models of the form

$$E\left(Y^{x} | Z = z\right) = h_{\text{par}}\left(z, x; \beta\right)$$

For instance,

$$h_{\text{par}}\left(z, x; \beta\right) = \beta_1 + \beta_2 z + \beta_3 x + \beta_4 xz + \beta_5 x^2 + \beta_6 x^2 z$$

Model extensions (I)

Estimating equations can be modified to obtain estimators:

- doubly-robust
- locally efficient

Can allow for the possibility that optimal CD4 count depends on baseline covariates Z by considering Parametric Marginal Structural Mean (MSM) models of the form

$$E\left(Y^x | Z = z\right) = h_{\text{par}}\left(z, x; \beta\right)$$

For instance,

$$h_{\text{par}}\left(z, x; \beta\right) = \underbrace{\beta_1 + \beta_2 z}_{q(z)} + \underbrace{\beta_3 x + \beta_4 xz + \beta_5 x^2 + \beta_6 x^2 z}_{h_{\text{sem}}(z,x;\beta)}$$

Can also consider more flexible Semiparametric MSM models

$$E\left(Y^x | Z = z\right) = h_{\text{sem}}\left(z, x; \beta\right) + q\left(z\right)$$

Model extensions (II)

The same approach can be used to optimize over a more complex set of candidates regimes where x is replaced by a vector $(x_1, ..., x_s)$.

Example:

- Start HAART the first time that
  - CD4 falls below x_1 or
  - CD4 falls in (x_1, x_2) and current HIV RNA is greater than x_3.
- Otherwise do not start.

The target of estimation in this approach is $(x_{1,opt}, x_{2,opt}, x_{3,opt})$.

General formulation (Summary)

- Assuming that treatment decisions are to be made at fixed times $t = 0, 1, ..., K$.
- We considered regimes indexed by a vector $x \in \mathcal{X}$, $\mathcal{X}$ possibly uncountable
- We developed estimators of the optimal treatment regime $x_{opt}(z)$ for subjects with baseline values $Z = z$ under:
  - Parametric Marginal Structural Mean Models for $E(Y^x|Z = z)$.
  - Semiparametric Marginal Structural Mean Models for $E(Y^x|Z = z)$.
- We established a set of assumptions for identification of $E(Y^x|Z = z)$ from the observed data distribution.
- We derived a class of consistent, doubly-robust and asymptotically normal estimators of β under each of the proposed models and the efficient estimator in the class.

Data analysis for illustrative purposes only

We applied this method to the publicly available MACS-WIHS data.

- Restricted to HIV-positive, AIDS-free participants who were antiretroviral therapy naïve by the time HAART was first available for use.
- Outcome of interest was the minimum of
  - time since baseline to death from any cause
  - time to first diagnosis of clinical AIDS
  - 7 years (five years follow-up).

Data analysis for illustrative purposes only

We applied this method to the publicly available MACS-WIHS data.

- Restricted to HIV-positive, AIDS-free participants who were antiretroviral therapy naïve by the time HAART was first available for use.
- Outcome of interest was the minimum of
  - time since baseline to death from any cause
  - time to first diagnosis of clinical AIDS
  - 7 years (five years follow-up).
- Set of regimes $x \in [100, 400]$.
- Proportion of patients following regime d^x steadily decreased from 57% for regime d^{100} to 27% for regime d^{400}.
- We assumed a Parametric MSM polynomial model in x (5^{th} order) with no baseline covariates.
- We obtained $\widehat{x}_{opt} = 289$ cell counts/mm^3 with nominal 95% CI for x_{opt} $= (266; 312)$.

Concluding remarks

- Dynamic MSM models have appealing properties
  - Easy to understand.
  - Easy to fit with standard software that allows for weighting.
  - It is possible to deal with missing outcomes (due to death for other causes or drop-out).

- We conducted simulation studies that confirmed the theoretical results.

Concluding remarks

- Dynamic MSM models have appealing properties
  - Easy to understand.
  - Easy to fit with standard software that allows for weighting.
  - It is possible to deal with missing outcomes (due to death for other causes or drop-out).
- We conducted simulation studies that confirmed the theoretical results.
- However... our proposal assumes that patients come to the clinic at fixed time intervals.
- This is not the realistic setting in the management of chronic diseases:
  - next visit date is decided based on patient health status and
  - patients are free to return earlier if they need to do so.

Main ideas of the talk based on:

- Orellana L.C. (2007) *Methodological challenges for the estimation of optimal dynamic treatment regimes from observational studies*. Harvard University, Dep of Biostatistics, Ph.D. Thesis.

- Robins J.M., Orellana, L., Rotnitzky A. (2008) Optimal treatment and testing strategies with possibly nonignorable observation processes. *Statistics in Medicine*, 27: 4678–4721.

- Orellana L., Rotnitzky A. and Robins J.M. Dynamic regime marginal structural mean models for estimation of optimal dynamic treatment regimes (to appear in *International Journal of Biostatistics*, 2009).