Estimating the optimal dynamic treatment regime from longitudinal observational data

Liliana Orellana¹, Andrea Rotnitzky²,³ and James Robins²

¹Universidad de Buenos Aires, ²Harvard University, ³Universidad 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 ➔ 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 $< 200 \text{ cells/mm}^3$.
- *Unclear* if CD4 count $> 200 \text{ cells/mm}^3$.
  - Offer trx if $200 < \text{CD4} \leq 350 \text{ cells/mm}^3$.
  - Preferably defer trx if CD4 $> 350 \text{ cells/mm}^3$.

- A treatment strategy based on CD4 counts is an example of a *dynamic treatment regime*.
Dynamic treatment regimes

Data

\[ L_0, A_0, L_1, A_1, \ldots, L_K, A_K, L_{K+1} \]

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

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

\[
\left( \bar{L}_k, \bar{A}_{k-1} \right) \rightarrow d_k \left( \bar{L}_k, \bar{A}_{k-1} \right) \in A_k , k = 0, ..., K.
\]
Dynamic treatment regimes

Data

\[ L_0, A_0, L_1, A_1, \ldots, L_K, A_K, L_{K+1} \]

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

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

\[
(\overline{L}_k, \overline{A}_{k-1}) \rightarrow d_k (\overline{L}_k, \overline{A}_{k-1}) \in A_k , k = 0, \ldots, K.
\]

Optimal dynamic regime

- Maximizes the expectation of some utility function \( Y \equiv u(\overline{L}_{K+1}, \overline{A}_K) \) 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 ($\mu^{500}$) versus
2. a hypothetical world where regime $d^{200}$ was enforced ($\mu^{200}$).

It is difficult to conduct such a trial to compare many regimes. We must then rely on observational data.
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 ($\mu^{500}$) versus
2. a hypothetical world where regime $d^{200}$ was enforced ($\mu^{200}$).

We can estimate this contrast from our ideal clinical trial because

- randomization generates exchangeable groups and
- each subject can be assigned to any regime.
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 (\( \mu^{500} \)) versus
2. a hypothetical world where regime \( d^{200} \) was enforced (\( \mu^{200} \)).

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.

Naive analysis

- Define baseline as time when CD4 first falls below 500.
- Regard subject 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
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.
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 ⇒ 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 ⇒ 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

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

$$\mu^{200} = \text{expected utility in a world where all subjects followed regime } d^{200} \text{ (start HAART first time CD4 falls below 200).}$$
Inverse probability of censoring weighted estimation

Suppose we want to estimate

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

$\mu^{200}$ = 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 $k = 0, \ldots, K$. 
Inverse probability of censoring weighted estimation

Subject redistributed here among all “at risk” subjects at k+1 with the same past from 0 to k.

Subject failed to follow regime $d^{200}$ here.
Censored at occasion k, i.e. $C^{200} = k$
Inverse probability of censoring weighted estimation

Data recorded in the cohort study

\[ L_0, A_0, L_1, A_1, \ldots, L_K, A_K, L_{K+1} \]

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

Accumulated weight through occasion \( k \) for a subject is estimated as

\[
W_{200}^k = I(C_{200} > k) \prod_{j=1}^{c} \Pr(C_{200} > j | C_{200} > j - 1, A_j, L_j)
\]

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 \( L_j \) is a high dimensional vector, so a parametrical model is assumed for the censoring probabilities.
Inverse probability of censoring weighted estimation

Data recorded in the cohort study

\[ L_0, A_0, L_1, A_1, \ldots, L_K, A_K, L_{K+1} \]

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

Accumulated weight through occasion \( k \) for a subject is estimated as

\[
W_{k}^{200} = \frac{I (C^{200} > k)}{\prod_{j=1}^{k} \hat{\Pr} \left( C^{200} > j | C^{200} > j - 1, \overline{A}_{j-1}, \overline{L}_j \right)}
\]

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 $\mu^{200} \equiv E(Y^{200})$ with

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

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

- $\hat{\mu}^{200}$ is a weighted average of the outcomes of those patients who followed regime $d^{200}$ throughout.
Inverse probability of censoring weighted estimation

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

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

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

- $\hat{\mu}^{200}$ 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 \( d^x \equiv \text{start HAART first time CD4 falls below } x \), where \( x \in X = \{200, 201, \ldots, 500\} \).

- In principle, we can estimate each mean \( \mu^x \equiv E(Y^x) \) separately and then find \( \hat{x}_{opt} \) that maximizes \( \hat{\mu}^x \).

- However, estimates \( \hat{\mu}^x \) 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 \(d^x \equiv \text{start HAART first time CD4 falls below } x\), where \(x \in \mathcal{X} = \{200, 201, \ldots, 500\}\).

- In principle, we can estimate each mean \(\mu^x \equiv E(Y^x)\) separately and then find \(\hat{x}_{opt}\) that maximizes \(\hat{\mu}^x\).

- However, estimates \(\hat{\mu}^x\) 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

\[ \mu^x = E(Y^x) = h(x; \beta) \]  \hspace{1cm} (1)

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

For example,

\[ h(x; \beta) = \beta_0 + \beta_1 x + \beta_2 x^2 + \beta_3 x^3 + \beta_4 x^4 + \beta_5 x^5 \]
Estimating the optimal CD4 in a candidate set: Proposal

Assume that

\[ \mu^x = E(Y^x) = h(x; \beta) \]  \hspace{1cm} (1)

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

For example,

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

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

So, under model (1) the problem reduces to estimating \( \beta \).
Estimating the optimal CD4 in a candidate set: Proposal

- Let $\gamma = \text{number of regimes in the candidate set}$
  $\mathcal{X} = \{200, 201, \ldots, 500\}$.

- Create an artificial data set, with each subject contributing $\gamma$ observations $(W_i^{x_j}, Y_i, x_j)$, $j = 1, \ldots, \gamma$. 
Estimating the optimal CD4 in a candidate set: Proposal

- Let $\gamma = \text{number of regimes in the candidate set}$
  \[ \mathcal{X} = \{200, 201, \ldots, 500\} . \]

- Create an artificial data set, with each subject contributing $\gamma$ observations $\left( W_{i}^{x_j}, Y_i, x_j \right)$, $j = 1, \ldots, \gamma$.

- Find $\hat{\beta}$ 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 $\gamma = \text{number of regimes in the candidate set}$
  $\mathcal{X} = \{200, 201, \ldots, 500\}$.

- Create an artificial data set, with each subject contributing $\gamma$
  observations $(W_{ij}^x, Y_i, x_j)$, $j = 1, \ldots, \gamma$.

- Find $\hat{\beta}$ solving the weighted estimating equation

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

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

For instance,

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

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

For instance,

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

Can also consider more flexible **Semiparametric MSM models**

$$E (Y^x | Z = z) = h_{sem} (z, x; \beta) + q (z)$$
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 $(x_{1, opt}, x_{2, opt}, x_{3, opt})$. 
Assuming that treatment decisions are to be made at fixed times $t = 0, 1, \ldots, K$.

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 $\beta$ 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).

- Set of regimes $x_2 \in [100, 400]$.
- Proportion of patients following regime steadily decreased from 57% for regime $d_{100}$ to 27% for regime $d_{400}$.
- We assumed a Parametric MSM polynomial model in $x_5\,(5\text{th order})$ with no baseline covariates.
- We obtained $b_{x_{opt}} = 289$ cell counts/mm$^3$ with nominal 95% CI for $x_{opt} = (266; 312)$. 
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 $\hat{x}_{opt} = 289$ cell counts/mm$^3$ with nominal 95% CI for $x_{opt} = (266; 312)$. 

---

Liliana Orellana - IBS AR - NZ - Dec 2009

Optimal dynamic treatment regime
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:

