When To Start Antiretroviral Therapy

James Robins, Miguel Hernan, Lilliana Orellana, Andrea Rotnitzky

with collaboration of the investigators from

Multicenter AIDS Cohort Study (MACS, REF) and Women’s Interagency HIV Study (WIHS), French Hospital Database on HIV (FHDH)

particularly Steve Cole, Dominique Costagliola

Harvard School of Public Health
Motivation

2004 International Aids Society Recommendations in JAMA for asymptomatic patients

RCT only in those with CD4 counts $\leq 200$

Must rely on observational data for CD4 $>200$
Start

- CD4 above 350 only if HIV RNA is extremely high or the CD4 cell count has rapidly declined.

- when the CD4 falls below 350 if plasma HIV RNA exceeds 100,000 copies/ml or CD4 decline has exceeded 100/ul cells per year.

- At CD4 count 200 if HIV RNA is low and stable or CD4 decline is less than 50/ul cells per year.

- All patients with counts below 200 should receive therapy.

Trade Off: Protect immune system from declining versus more time to develop drug resistance and side effects
A Main Paper Leading to This Decision

Paella et al. Ann Internal Med 2003

Simplified Description

- Subjects who were already between 300 and 350 were followed from that measurement

- Those who ever started above 300 were regarded as group 1

- Those who started below 200 were regarded as group 2

- Those who never started were excluded
Analysis better than many since follow up for both groups started at same time.

Think if follow-up started at time that treatment began.

Then survival of those starting with CD4 >200 will be better than those with CD4 <200 even under null that CD4 when started makes no difference.
Nontheless Analysis has potential for severe bias

- Suppose prescribing trends increased with time.

First year noone above 300 treated.

Second year every one below 350 treated

- Start below 200 (group 2) implies CD4 cell count fell below 200 within a year, so subjects were very sick and possibly died before 1 year.

- Start above 300 (group 1) guaranteed to have at least a year survival and included healthier subjects whose CD4 count was stable for a year.
Under the null of no difference when you start (or even no treatment effect), recomendation based on this analysis: start above 300.

Reason: Treatment above 300 is a surrogate for slowly declining counts and for survival itself, so clearly it looks preferable.
Want to compare the 2 strategies: start HAART when CD4 falls below 350 \textbf{versus} start HAART when CD4 falls below 200.

- Take subjects with CD4 counts exceeding 350,

- when subject’s CD4 first falls below 350 call it start of followup (time0)

- randomize with p=1/2 to start immediately; p= 1/2 to start when their CD4 first fell below 200.

- Compare survival using Cox PH

Such a trial has not been conducted so we must rely on observational data
Samuel Beckett-like structure to talk

1. It can be done

2. It can’t be done

3. It can be done

4. It can’t be done

and so on ad infinitum

Because of time limitations we cover only point 1 and point 2,

leaving one in Beckettian despair.
Observational Analog: :

- Take subjects with CD4 counts exceeding 350.
- When a subject first falls below 350 he becomes on test at time 0.
- Regarded as group 1 (start below 350) if he initiates HAART at time zero.
- Otherwise in the group 2 (start below 200).

- The 2 groups are then compared using a standard COX PH model for time to AIDS or death that adjusts for the baseline (time 0) confounding factors such as HIV RNA, calendar date of entry, ethnicity, past rate of CD4 count decline, etc.
However any group 2 patient who begins HAART at time $t$ prior to falling below 200 or fails to start treatment at the time $t$ he falls below 200 has failed to follow group 2 ’protocol’

That group 2 patient should be considered censored at $t$,

but **dependently censored**.

No group 1 patient censored

Patient who has never started treatment and has CD4 count $>200$ at end of follow-up counts in every risk set as a non-failure in group 2 since never censored!! Compare Paella
Meaning of dependent censoring:

Consider all Group 2 subjects with CD4 cell count still greater than 200 cells/mL at time $t$ after entry.

- The subgroup who initiate treatment (and become censored) at $t$ will have, on average faster CD4 cell count decline form 0 to $t$ and lower HIV RNA at $t$ than those who do not initiate treatment (and remain uncensored) at time $t$.

- So individuals censored at $t$ will have a worse prognosis than those who remain uncensored.

In data, among those with a CD4 cell count less than 200 cells/mL for the first time at $t$, essentially all started treatment as supposed to do.
Conclusion: A standard Cox analysis (which assumes independent censoring) will show death rate of the group 2 subjects artificially lowered compared to what would have been seen in the above ideal randomized trial in which all group 2 subjects will have waited till 200 to start.

So if better to start at 350 standard, standard Cox may fail to detect this.
Method: Inverse probability of censoring weighting (IPCW) to adjust for dependent censoring.

Solution use of inverse probability of censoring weighting (IPCW) on each group subject in each risk set at time $u$.

Each group 2 subject uncensored in risk set at time $u$ divide by the probability of having remained uncensored to time $u$. This will fully adjust for dependent censoring due to measured time-dependent covariates, Here HIV RNA counts and rate of decline of $CD4$ cell count before $u$. 
Suppose a group 2 subject at risk at \( u \) with a CD4 count falling linearly from 350 to 250 from 0 to \( u \) has a probability of 1/4 that he would not have started treatment by \( u \).

Then he counts for 4 people: himself and the three other similar people who did start therapy.

That is his \( IPCW \) is 4 for the risk set at \( u \).

Subject CD4 falling 350-300 from 0 to \( u \) has probability of 1/2 uncensored so only counts for 2 persons
In contrast if one adds CD4 count as a time-dependent covariate in the Cox model (rather than using IPCW weights), bias due to dependent censoring remains because you have adjusted for a post randomization variable possibly affected by treatment.

It is ok to weight but not adjust for a post-randomization variable.
That is for each group 2 subject in risk set at time $u$ multiply by the inverse of the estimated probability of having remained uncensored to time $u$

\[
\hat{W} (u) = 1/ \prod_{j=1}^{u} \hat{pr} [C \neq j | C \neq j - 1, Z(j - 1)]
\]

\[
\log it \{pr [C \neq j | C \neq j - 1, Z(j - 1)]\} = \alpha^T Z(j - 1),
\]

\[
Z(j - 1) = (RCD4(j - 1), HIV(j - 1))
\]
IPCW valid if

\[ Z(j - 1) \] all joint risk factors for censoring and failure

\[ \text{mod } el \] for censoring correct

\[ \text{can be made doubly robust: ie either model for censoring or model for survival given } Z(j - 1) \text{ history (roughly)} \]
Actual Tentative Analysis:

2344 HIV-infected subjects included in the French Hospital Database on HIV (FHDH) who had their first CD4 cell count measurement below 500 cells/mL (not 350) between 1 January 1996 and 30 June 2004, and who had never received antiretroviral therapy before that measurement.

We followed these subjects from their first CD4 cell count measurement below 500 cells/mL (baseline) until a diagnosis of AIDS, death, or June 2004, whichever occurred earlier.

Data on HAART use, as well as on time-dependent covariates (e.g., CD4 cell count) were recorded throughout the follow-up.
There were 131 subjects in group 1,

2217 in the group 2.

655 subjects in group 2 were censored

when they either started HAART before their CD4 cell count dropped below 200 cells/mL or failed to start HAART first time below 200.
RR 0.9 (95% confidence interval: 0.4, 1.8) just straight Cox.

RR 0.5 (95% confidence interval: 0.2, 1.1) IPCW:

Appears better to start at 350 than 200
In practice want to answer:

- Find the optimal CD4 count $x$ at which to start.

- That is want to compare all $x$ in the candidate set $\{500, 499, \ldots, 200\}$ rather than just 2 dynamic regimes.

- Would like to use expected utilities $Y$ rather than Cox PH

- With $K$ end of FU time, need something like

$$Y = T \text{ if AIDS or death before } K$$

$$Y = K + 4\frac{CD4}{500} \text{ if survives to } K \text{ (or possibly other quality of life measures)}$$
Let $Y_x$ be utility if start HAART first time CD4 falls below $x$ a particular regime.

- Find $x_{opt}$ that maximizes $E \left[ Y_x \right]$.

- If randomized trial with full compliance where different subject randomized to different regimes $x$.

$\hat{E} \left[ Y_x \right]$ is average of $Y$ among subjects randomized to regime $x$ (even if never took HAART because CD4 always above $x$)
What if few subjects randomized to any one $x$.

- Fit ITT flexible polynomial (say 5th order) regression by least squares

$$Y = \beta_0 + \sum_{k=1}^{5} \beta_k x^k$$

(1)

to the $n$ study subjects.

- Use first year calculus to find the value $\hat{x}_{opt}$ of $x$ where the fitted polynomial $Y = \hat{\beta}_0 + \sum_{k=1}^{5} \hat{\beta}_k x^k$ obtains its maximum.
How to mimic in an observational study: (same method based on marginal structural models proposed by us and van der laan and petersen)

Consider a subject who started anti-retroviral therapy at a CD4 of 250 in week \( t \) whose lowest prior CD4 counts was 300. This subject followed all of the regimes \( x = 251, 252, \ldots, 300 \).

Consider a subject who never started therapy and whose lowest CD4 count was 225. This subject followed regimes \( x = 200, 201, \ldots, 225 \).

Consider a subject who started anti-retroviral therapy at a CD4 of 250 in week \( t \) whose lowest previous CD4 counts was less than 250. This subject failed to follow any regime. in the candidate set \( \{500, 499, \ldots, 200\} \)

- Consider the data set \((Y_i, x_{i1}), (Y_i, x_{i2}), \ldots, (Y_i, x_{i\Gamma_i})\), where the \( x_{ik}, k = 1, \ldots, \Gamma_i \) denote the \( \Gamma_i \) regimes followed by subject \( i, i = 1, \ldots, n \)
- Let $\widehat{W}_{ik}$ be the inverse of the probability that subject $i$ followed regime $x_{ik}$, without being censored (i.e without starting HAART at a value above $x_{ik}$).

- Fit the flexible polynomial (say 5th order) regression

$$Y = \beta_0 + \sum_{k=1}^{5} \beta_k x^k$$

by weighted least squares applied to the data set $(Y_i, x_{ik})$ with weights $\widehat{W}_{ik}, k = 1, \ldots, \Gamma_i, i = 1, \ldots, n$

- Use first year calculus to find the value $\hat{x}_{opt}$ of $x$ where the fitted polynomial $Y = \hat{\beta}_0 + \sum_{k=1}^{5} \hat{\beta}_k x^k$ obtains its maximum.
Applied this method to the MACS-WIHS data and obtained

\[ \hat{x}_{opt} = 289 \text{ with CI (266, 312)} \]

- Restricted to HIV-positive, AIDS-free participants who were antiretroviral therapy naïve by the time HAART was first available for use.
A more complex set of candidate regimes can be optimized.

- If current HIV RNA is greater than $z$, start HAART if the current CD4 count is less than $x$.

- If current HIV RNA is not greater than $z$, start HAART if the current CD4 count is less than $q$.

We can use the same methods to jointly find $(z_{opt}, x_{opt}, q_{opt})$. 
We can allow for the fact that the optimal treatment regime in our candidate set may differ depending on a subject’s measured pretreatment variables $V$.

- For instance we might replace the model $Y = \beta_0 + \sum_{k=1}^{5} \beta_k x^k$ with $Y = \beta_0 + \sum_{k=1}^{5} \beta_k x^k + \gamma^T V + \sum_{k=1}^{5} \gamma_k^T V x^k$

- Use first year calculus to find the value $\hat{x}_{opt}(V)$ of $x$ where the fitted polynomial $Y = \hat{\beta}_0 + \sum_{k=1}^{5} \hat{\beta}_k x^k + + \sum_{k=1}^{5} \hat{\gamma}_k^T V x^k$ obtains its maximum.

The method can be made robust by replacing the assumption of a linear main effect of $V$ by an arbitrary function $h(V)$. 
**Classic Problem:** Find the optimal regime given data \( L_0 A_0 L_1 A_2, \ldots, L_K A_K L_{K+1} \) where \( A_k \) is HAART at time \( k \). \( L_k \) covariates at time \( k \).

\[
\tilde{d}_{opt} = \left\{ d_{opt,k} \left( \overline{a}_{k-1}, \overline{l}_k \right), k = 0, 1, \ldots, K \right\}
\]

that maximizes the mean of \( Y_{\tilde{d}} \) over all regimes \( \tilde{d} \).

New statistical models for estimating \( \tilde{d}_{opt} \) from observational data, that are a clever empirical twist on dynamic programming (backward induction) by myself and Susan Murphy.

Can be very (too) complicated. But dynamic programming methods cannot simplify the class too much since must allow any possible past history of treatment and at least one covariate in \( \overline{l}_k \) say CD4 count.

If we want something simple like first time CD4 falls below \( x \) so easy to implement in many places.then previous methods: find optimal regime in a prespecified set
Problem with MACS WHIS

- Seen only every 6 months

*Treatment Decisions* made by treating physicians based on unknown CD4, HIV RNA clinical responses at time of decision (except for 6 month interval)

Solution: Continuous time records where all physician visits, labs, and pharmacy records are obtained eg HMO, so we know what the prescribing physician knew.

FHDH approximates this but many new problems arise, time to discuss just one
Patients come to hospital (ie HMO) at random times based on symptoms, routine follow-up appt, missed visit.

*Data on why a visit generally not available except perhaps in good HMO (even routine follow-up)*

Typical solution- censor at certain time from last visit but this is dependent censoring, usually nonignorable

If patients sick more (or less) likely to come back, we have association of treatment and risk

Only can control this confounding if data on health for people who return at time $t$ and on those who do not: Noone collects such data
· Thus only reasonable to assume, among subjects who have visit at time $t$ with same past history decision to treat or not treat is at random.

· But then $E[Y_x]$ means treat not first time CD4 falls below $x$, but rather first time CD4 falls below $x$ at a clinic visit.

· May then be impossible then to extrapolate to other populations with different reasons and frequency for coming to a clinic.

· Say data from US HMO HIV patients come every 3 weeks and subjects are compliant

· Consider extrapolation to another patient population where subjects are scheduled only every six months and few comply
In second pop, those who arrive at clinic at $t$ with CD4 below $x$ for the first time are a very different, often symptomatic, selected sample.

For example may have been below $x$ for many months so would have been treated months before if in HMO.
How large is this bias?

Try to quantify by comparing estimates that for subject who starts at CD4 $x_3$ at time $u_3$ on his third hosp visit who had earlier visits at times $u_1$ with CD4 $x_1$ and $u_2$ with CD4 $x_2$

- weights $W_{ik}$ of one over cond prob of not have started HAART at any of the previous visits but starting at this visit

versus

- weights $W_{ik}$ of one over prob of not have started HAART at any of the previous visits

but starting at this visit times one over the probability of having visits at exactly times $u_1, u_2, u_3$. 
· If these were very similar some evidence that the subjects with a given visit pattern are not a self-selected group whose risk differs from subjects with another visit pattern based on unmeasured factors.

· would allow one to extrapolate to another set of subjects for whom visits patterns are also not self-selected based on risk factors. So even if ok in no bias in US not extrapolable to the second group.
Problems of Extrapolation of 'When To Start results' Due to Medical Treatment Patterns in addition to the other issues.

- Africa has Tb Malaria different nutrition, health care etc

- Different Distributions of Time Since Infection if not in Seroconverter Group

Suggests only fairly important differences in results on when to start should be taken seriously