Mathematical-Statistical Modeling to Inform the Design of HIV Treatment Strategies and Clinical Trials

2007 FDA/Industry Statistics Workshop

Marie Davidian*
Department of Statistics
North Carolina State University

http://www.stat.ncsu.edu/~davidian

* Joint work with H. Thomas Banks (NCSU) and Eric S. Rosenberg (MGH and Harvard Medical School)
Outline

- Objectives
- HIV therapy and structured treatment interruption (STI)
- HIV dynamic models and control
- Mathematical-statistical framework
- Design of a clinical trial in acute-infection
- Next steps - design of STI strategies
- Closing remarks
Objectives

• Describe a *multidisciplinary collaboration* supported by a 5-year grant from NIAID

• *Main players*: Applied mathematician/control theorist, immunologist/infectious disease clinician, and *statistician*

• Use *mathematical-statistical modeling of disease progression* and *simulation* to design HIV treatment strategies and clinical trials to study them

• Focus on *better use of existing antiretroviral therapies (ART)* to manage HIV infection *over time*

• Design and carry out a *clinical trial* assisted by modeling and simulation

• Collect *extensive data* to inform *refined modeling* $\implies$ more sophisticated strategies and trials
HIV therapy and STI

Potent ART for treatment of HIV-1 infection:

- Has led to *profound decrease* in morbidity and mortality from HIV and AIDS-related illnesses ...

- ...but *cannot* completely eradicate the virus \(\implies\) need for *life-long* treatment

- *Complications* – side effects, toxicities, adherence issues, cost, burden, development of drug resistance, life style issues,...

- \(\implies\) Continuous treatment *impossible* for most patients
HIV therapy and STI

Result: *Structured* (or *supervised*) *treatment interruption* (*STI*)

- ART for some period of time followed by *interruption* and *possible re-initiation*, perhaps through *several cycles*
- *Non-adaptive* (*non-dynamic*) strategies – planned *advance*, e.g., cycles of 8-weeks-on/8-weeks-off
- *Adaptive* (*dynamic*) strategies – decisions to interrupt and re-initiate based on *rules* taking patient information as input, e.g., stop or start based on *CD4+ T cell count* or *viral load*
- *Goals* – maintain *viral suppression*, preserve *immune response*, lower viral load “*set point*”
HIV therapy and STI

**Rationale:** Hypotheses

- **Chronic infection** – relieve *burden* and *side effects*; allow *wild-type virus* to *repopulate* in subjects with *drug resistance*

- **Acute infection** – preserve *HIV-specific immune response* and allow *discontinuation* of ART; on-off cycles serve as “*self-vaccination*” further stimulating immune response

**Studies so far:** *Mixed* results, e.g.,

- CPCRA *“Strategies for Management of Antiretroviral Therapy”* (*SMART*) trial (El-Sadr, Neaton, et al., 2006) in *chronically-infected* subjects

- Compared *continuous ART* to an *adaptive STI strategy* (“*drug conservation*”) – on-off ART dictated by *CD4+ T cell count*
HIV therapy and STI

Drug conservation strategy in the SMART study:

Result of SMART: Stopped early (∼ 5500 pts), drug conservation \(\Rightarrow\) 2x risk of primary endpoint (AIDS or death)
Our premise: Strategies so far may have been *unfortunately chosen*

- Based on *“educated guesses” expert opinion*, pieced-together *clinical evidence*
- E.g., CD4+ thresholds in SMART chosen after *much debate* among experts...
- ...and *decision rules* did not include *viral load* (or other info)
- \( \Rightarrow \) it is *premature* to dismiss STI and *adaptive treatment strategies*
- A formal, *evidence-based* approach combining *biological knowledge*, *data* in a *principled* way is needed to *design* strategies
HIV dynamic models:

- Represent \textit{mathematically} known and hypothesized \textit{mechanisms} involved in the \textit{virus-immune system} interaction taking place \textit{within a single subject}

- Series of "\textit{compartments}" characterizing different populations of virus and constituents of the immune system

- Interactions among compartments described by a system of (\textit{deterministic}) \textit{nonlinear ordinary differential equations}

- \(\implies\) In principle, \textit{viral load}, \textit{CD4+ T cell count}, etc, at any time
Possible model for within-subject dynamics:

- Non-infected CD4
- Infected CD4
- Infectious Virus
- Non-infectious Virus
- Immune Effectors
- Differentiation
- Activation
- RTI
- PI
HIV dynamic models and control

Model for within-subject dynamics: \( s = 7 \) “states”

\[
\begin{align*}
\dot{T}_1 & = \lambda_1 - d_1 T_1 - \{1 - \epsilon_1 u(t)\} k_1 V_I T_1 \\
\dot{T}_2 & = \lambda_2 - d_2 T_2 - \{1 - f \epsilon_1 u(t)\} k_2 V_I T_2 \\
\dot{T}_1^* & = \{1 - \epsilon_1 u(t)\} k_1 V_I T_1 - \delta T_1^* - m_2 E T_1^* \\
\dot{T}_2^* & = \{1 - f \epsilon_1 u(t)\} k_2 V_I T_2 - \delta T_2^* - m_2 E T_2^* \\
\dot{V}_I & = \{1 - \epsilon_2 u(t)\} 10^3 N_T \delta (T_1^* + T_2^*) - c V_I - \{1 - \epsilon_1 u(t)\} \rho_1 10^3 k_1 T_1 V_I \\
& \quad - \{1 - f \epsilon_1 u(t)\} \rho_2 10^3 k_2 T_2 V_I \\
\dot{V}_{NI} & = \epsilon_2 u(t) 10^3 N_T \delta (T_1^* + T_2^*) - c V_{NI} \\
\dot{E} & = \lambda_E + \frac{b_E (T_1^* + T_2^*)}{(T_1^* + T_2^*) + K_b} E - \frac{d_E (T_1^* + T_2^*)}{(T_1^* + T_2^*) + K_d} E - \delta_E E
\end{align*}
\]

- \( \theta = (\lambda_1, d_1, \epsilon_1, k_1, \ldots)^T \) plus initial conditions
- Observable: \textit{CD4 count} = \( T_1 + T_1^* \), \textit{viral load} = \( V_I + V_{NI} \)
- \( u(t) = \) ART input at \( t \) (\( 0 \leq u(t) \leq 1, 0 = \text{off}, 1 = \text{on} \))
HIV dynamic models and control

**In general:** HIV dynamic model with $s$ states

$$\dot{x}(t, \theta) = g\{t, x(t, \theta), \theta\}, \text{ solution } x(t, \theta) \ (s \times 1)$$

- Embodies *hypothesized mechanisms* through *model parameters* $\theta$

- $\theta$ includes cell and virus production, death, clearance rates; treatment efficacy parameters; etc

- $\theta$ dictates *pattern of progression* over time (*deterministic*) under any *treatment pattern* $u(t)$

**Control theory:** Mathematical theory and techniques for modifying (*controlling*) the behavior of such systems

- *Goal* – Optimize some *objective function*, e.g., drive viral load *set point* below a threshold while keeping “*cost of therapy*” low

- I.e., determine $u(t)$ to achieve this objective
HIV dynamic models and control

Our ultimate goal: Use HIV dynamic models and control along with simulation to design treatment strategies $u(t)$ for acute HIV infection and to design clinical trials to study them

- Find strategies that “do well” for individuals and for the population
- Need evidence supporting HIV dynamic model $\Rightarrow$ data (e.g., measured CD4, VL, other stuff over time on lots of subjects)
- Intra-subject variation due to assay error, realization error; left-censoring of viral loads due to assay lower limits of quantification
- Substantial inter-subject variation $\iff$ heterogeneity in mechanisms $\theta$ across the subject population

For both fitting to data and simulation: Must embed the (deterministic) mathematical model in a statistical framework that characterizes faithfully inter- and intra-subject variation
**Data:** Eric has been collecting intensive *longitudinal* viral loads, CD4 counts on a *cohort* of $\geq 150$ acutely-infected subjects for $> 7$ years.

![Graph showing CD4+ T-cells/ul and virus copies/ml over time for Patient #14.](image)
Mathematical-statistical framework

**Mathematical model:** \( \dot{x}(t, \theta) = g\{t, x(t, \theta), \theta\} \), solution \( x(t, \theta) \) \( (s \times 1) \)

- Observations *not available* on all \( s \) states
- \( \bar{x} = \mathcal{O}x \) for *observation operator* \( \mathcal{O} \)
- E.g., CD4 and VL only

**Statistical framework:** Embed \( \bar{x} \) in a *hierarchical statistical model*

- For each subject \( i \) in the population, conceive of a *bivariate* (CD4, VL) *subject-specific stochastic process* under \( u(t) \)

\[
Y_i\{t, u(t)\} = [Y_i^{CD4}\{t, u(t)\}, Y_i^{VL}\{t, u(t)\}]^T
\]

- Depends on treatment strategy \( u(t) \) up through time \( t \)
Mathematical-statistical framework

**Intra-subject model:** Decompose into

\[ Y_i\{t, u(t)\} = \bar{x}\{t, u(t), \theta_i\} + e_i\{t, u(t)\} \]

- \(e_i\{t, u(t)\}\) is the deviation process – realizations, assay errors
- \(e_i\{t, u(t)\}\) average out to zero over all possible realizations, assay errors (conditional on \(\theta_i\) and the strategy imposed)
- Interpret \(\bar{x}\{t, u(t), \theta_i\}\) as average trajectories over all possible realizations we could see on subject \(i\) under strategy \(u(t)\)
- *Also*, assumptions on (conditional) correlation (across \(t\) and among elements of \(e_i\{t, u(t)\}\)), variances, probability distribution

**Inter-subject model:** \(\theta_i\) is an “inherent characteristic” of subject \(i\)

- Probability distribution \(p(\theta_i; \theta^*, D)\), e.g., \(\theta_i \sim \mathcal{N}(\theta^*, D)\)
- Could also be conditional on subject characteristics
Mathematical-statistical framework

Result: Full description of the hypothesized data-generation process in continuous time

- For individual subjects (randomly-chosen from the population)
- And thus for samples of such subjects drawn at random from the population
- For large enough sample $\Rightarrow$ effectively, knowledge of the entire population
- Basis for simulation of “virtual” subjects

Needed: Full characterization based on data
Mathematical-statistical framework

Data: For subject $i$, $i = 1, \ldots, N$, observed at $n_i$ times $t_{i1}, \ldots, t_{in_i}$

- $U_i(t) = \text{actual ART pattern}$ over entire observation period ($\text{known}$)
- $Y_{ij} = (Y_{ij}^{CD4}, Y_{ij}^{VL})^T$ at time $t_{ij} \implies Y_i = (Y_{i1}, \ldots, Y_{in_i})^T$
- Conceive $Y_{ij} = Y_i\{t_{ij}, U_i(t_{ij})\}$ (similarly for $e_i$)
- $Eric\text{'s data} - N \approx 150, n_i \approx 30-60$
- $A_i = \text{possible subject characteristics}$

Nonlinear mixed model (bivariate response): $Fit$ to data

$$Y_{ij} = \bar{x}\{t_{ij}, U_i(t_{ij}), \theta_i\} + e_{ij}, \quad j = 1, \ldots, n_i$$

$$\theta_i \sim p(\theta_i; \theta^*, D), \quad i = 1, \ldots, N$$
Mathematical-statistical framework

Challenges:

• *Left-censoring* of VL by *lower assay limit*

• $\dim(\theta) > 25$ and not all *identifiable* from CD4, VL only

• Components of $\bar{x}$ only calculable *numerically* by *forward solution* of ODEs

Two-stage approach:

• For each $i$, estimate $\theta_i$ via EM algorithm to handle censoring incorporating *regularization* to address *identifiability/dimensionality*

• Use resulting $\hat{\theta}_i$ as “data” to obtain $\hat{\theta}^*, \hat{D}$ using moment methods
Mathematical-statistical framework

Predictive capability:

Patient #14

CD4⁺ T-cells/ul

0 200 400 600 800 1000 1200 1400 1600

0 200 400 600 800 1000 1200 1400 1600

virus copies/ml

0 10 5 10⁵

0 200 400 600 800 1000 1200 1400 1600

time (days)

data
fit w/half
fit w/all

Modeling for Design of HIV Treatment and Trials
Mathematical-statistical framework

Simulation:

- Generate $N_{\text{sim}}$ "virtual subjects" by generating $\theta^*_i$, $i = 1, \ldots, N_{\text{sim}}$, from $p(\theta_i; \hat{\theta}_*, \hat{D})$

- Generate "inherent trajectories" $x\{t, u(t), \theta^*_i\}$ under a given $u(t)$ (continuous time)

- Add within-subject deviations according to intra-subject model to obtain "virtual data"

- Suitable $p(\theta_i; \hat{\theta}_*, \hat{D})$ determined by comparing "virtual profile" distributions (VL, CD4) to those from Multicenter AIDS Cohort Study (MACS, $u(t) \equiv 0$) and Eric’s data (various $u(t)$)

- Mixture of normals
Armed with this framework: Use to *design treatment strategies* and *clinical trials*

Our first step: *Proof of principle* – can we use this capability to assist in addressing a question involving *non-adaptive treatment strategies*?

- **Unresolved** – Whether or not individuals *acutely-infected* with HIV should be treated with ART

- **More precisely** – Is it better to give ART for *some period* following acute infection ("*train*” the immune system, "*self-vaccinate*”) or is it better to give no treatment at all until later (delay *drug resistance*, etc)

- **Primary endpoint** – VL *set point* at 12 months
Design of a clinical trial

Which strategies to study? $u(t) \equiv 0$ vs. strategies of the form

\[
    u(t) = \begin{cases} 
        1, & 0 \leq t \leq \tau \\ 
        0, & t > \tau 
    \end{cases}
\]

for termination times $\tau = 3, 4, \ldots, 12$ months

**Approach:** Evaluate effects of candidate strategies on the (virtual) population by simulation

- Insight into which strategies to study based on their anticipated effects on the entire population
Strategy $u(t)$ with $\tau = 6$: 100 “virtual” “inherent” viral load trajectories with ART terminated at 6 months, i.e., $u(t) = 1$, $0 \leq t \leq 6$, $u(t) = 0$, $t > 6$
Different termination times $\tau$: Means of 15,000 “virtual” CD4 and viral load data profiles with $u(t) = 1, 0 \leq t \leq \tau$, $u(t) = 0, t > \tau$, $\tau = 0, 3, 4, \ldots , 12$ months
Design of a clinical trial

Summary:

• Based on this (simple) HIV dynamic model, *no differences expected*

• Simple model does not represent adequately the *immune response*

• Simulations with a *refined model* showed larger *subpopulations* with *lowered VL set point* for larger $\tau$ . . .

• . . . but are less reliable (very little data on immune response)

Result: Study ART under *more than one termination time*

• $\tau = 3$ ("*short-term*") and $\tau = 8$ months ("*long-term*")
Design of a clinical trial

**Trial schema:** 1/2 pts randomized to ART, 1/2 pts to no ART
Design of a clinical trial

**Plan:** 3 year accrual period (36 patients), 1 year follow-up

- Standard design considerations for primary VL comparison at 12 months
- *Intensive visit schedule* – collect CD4, VL, CTLs, viral fitness, etc
- Data collection more frequent when dynamics are anticipated to *be changing* (e.g., in the weeks *after ART termination*)
Design of STI strategies

Next step: Armed with more informative data (e.g., measurements reflecting aspects of immune response)

- Develop and validate more realistic HIV dynamic models...
- ...and refine the entire mathematical-statistical framework
- ...and use to develop and evaluate ("virtually") potential adaptive treatment strategies
- Receding horizon control methods
- And design the next trial to study the most promising strategies...
Design of STI strategies

Input data:
Clinical, immunological and virological data

Individual Model

Population Model

Clinical Trial Design

Model Simulation & Robust Control Design

Clinical Data

Clinical Practice

Modeling for Design of HIV Treatment and Trials
Closing remarks

- **Modeling and simulation** have a significant role to play in design of HIV treatment strategies

- *In principle* – could link dynamic models with models for PK, etc

- We envision cycles of smaller *“learning trials”* that provide richer information needed to develop more *refined adaptive strategies* that will then be evaluated in confirmatory trials

- We’ll see how this turns out!

**Slides at:**  http://www.stat.ncsu.edu/~davidian