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

Marie Davidian and H.T. Banks
North Carolina State University

Eric S. Rosenberg
Massachusetts General Hospital and Harvard Medical School

Joint work with our Research Associates and Assistants at North Carolina State University
Outline

Overview
HIV therapy and acute infection
HIV dynamic models and control
Mathematical-statistical framework
Design of a clinical trial
Design of treatment strategies
Summary
References
Overview

- **Multidisciplinary collaboration** supported by a grant from the US National Institute of Allergy and Infectious Diseases (NIAID)
- **Main players**: Immunologist/infectious disease clinician (*Eric*, MGH), statistician (*Marie*, NCSU), applied mathematician/control theorist (*Tom*, NCSU)
- **Big picture**: Use *mathematical-statistical modeling* of disease progression and *simulation* to design *antiretroviral* (*ARV*) therapies to manage HIV infection and *clinical trials* to study them
- Design and carry out a *clinical trial* in subjects with *acute HIV infection* assisted by modeling and simulation
- Collect *extensive data* to inform *refined modeling* ⇒ more sophisticated strategies and trials
HIV therapy and acute infection

Eric’s practice at MGH: A 47 year old male presents to the ER

- 102.5 °F fever, headache nausea/vomiting, rash, ...
- MSM, recent unprotected sex, ...
- Tests for CMV, EBV, influenza negative
- HIV ELISA positive
- HIV RNA (viral load) > 750,000 copies/ml
- CD4+ T cell count = 432 cells/µl

Diagnosis: Acute HIV infection

- Within weeks of initial infection
HIV therapy and acute infection
HIV therapy and acute infection

**Question:** Should this individual be treated with *ARV therapy*?
### HIV therapy and acute infection

<table>
<thead>
<tr>
<th><strong>Disadvantages</strong></th>
<th><strong>Advantages</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>High cost, side effects, QoL</td>
<td>Delay of costs, side effects, risks</td>
</tr>
<tr>
<td>Unknown long term risks of ARV</td>
<td>Delay of drug resistance</td>
</tr>
<tr>
<td>Acquisition of drug resistance</td>
<td><em>Preservation of HIV-specific immune response</em></td>
</tr>
<tr>
<td>Limitation of future ARV options</td>
<td><em>Opportunity for treatment interruption</em></td>
</tr>
</tbody>
</table>

**Premise:** Cycles of treatment *interruption* and *re-initiation* may *augment* immune response and allow patient to maintain *viral control*

- Brief, *controlled* viral exposure may serve as a "*self-vaccination*"
- ???
HIV therapy and acute infection

The famous “Berlin patient”:

[Graph showing HIV RNA levels over time with key events marked: Epididymitis, day 15–22; Hepatitis A, day 121–137; Treatment stopped, day 176]
HIV therapy and acute infection

Current state of affairs:

- **Whether or not** and **how** to use ARV therapy during acute infection is **not known**
- **Treatment interruption** may be useful in acute infection, but the optimal approach is **not known**

Structured treatment interruption (STI):

- **Non-adaptive** (**non-dynamic**) strategies – planned in advance, e.g., cycles of 8-weeks-on/8-weeks-off, **terminal interruption**
- **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**
HIV therapy and acute infection

STI studies so far: Mixed results

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

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

- 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...
- ... and decision rules did not include viral load (or other info)
- ⇒ it is premature to dismiss treatment interruption and adaptive treatment strategies for managing HIV infection
- A formal, evidence-based approach combining biological knowledge, data in a principled way is needed to design and evaluate strategies
- In particular, can such an approach be used to determine the best way to manage patients from the time of acute infection?
HIV dynamic models and control

HIV dynamic models:

- Represent mathematically known and hypothesized mechanisms involved in the virus-immune system interaction taking place within a single subject
- Series of “compartments” characterizing different populations of virus and constituents of the immune system
- Interactions among compartments described by a system of (deterministic) nonlinear ordinary differential equations
- The solution to the system of ODEs yields a mechanistic model characterizing the joint behavior of the compartments over time (the “dynamics”)
- Viral load, CD4+ T cell count, etc, at any time
HIV dynamic models and control

Possible model for within-subject dynamics:

![Diagram of HIV dynamics model]
HIV dynamic models and control

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

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

- \( \theta = (\lambda_1, d_1, \epsilon_1, k_1, \ldots) \) plus initial conditions
- Observable: CD4 count = \( T_1 + T_1^* \), viral load = \( V_I + V_{NI} \)
- \( u(t) = \) ARV input at \( t \) (\( 0 \leq u(t) \leq 1 \), 0 = off, 1 = 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 the *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 (assay *lower limits of quantification*)
- Substantial *inter-subject variation* $\iff$ heterogeneity in mechanisms $\theta$ across the subject population

**To do this:** Must *embed* the (*deterministic*) mathematical model in a *statistical framework* that characterizes faithfully *intra-subject* and *inter-subject variation*
**Data:** Eric has been collecting intensive *longitudinal* viral loads, CD4 counts on a *cohort* of $\geq 270$ acutely-infected subjects for $\approx 12$ years.
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
- \( \overline{x} = \mathcal{O}x \) for observation operator \( \mathcal{O} \) – just the states observed
- E.g., CD4 and VL only

Statistical framework: A hierarchical statistical model that gives a hypothesized description of how observed (CD4, VL) arise

- For each subject \( i \) in the population, conceive of subject-specific, observed (CD4, VL) at any time \( t \) under treatment strategy \( u(t) \)

\[
Y_{i}\{t, u(t)\} = [Y_{i}^{CD4}\{t, u(t)\}, Y_{i}^{VL}\{t, u(t)\}]
\]

- If we could follow subject \( i \) continually, we could observe \( Y_{i}\{t, u(t)\} \) at all times \( t \)
Mathematical-statistical framework

**Intra-subject model:** How does $Y_i\{t, u(t)\}$ come about?

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

- $\bar{x}\{t, u(t), \theta_i\}$ is the *ideal smooth trajectory* from the math model
- $e_i\{t, u(t)\}$ represents how observed (CD4, VL) *deviate* from the smooth trajectory due to *realization* and *assay errors*
- Make *assumptions* on $e_i\{t, u(t)\} \implies$ probability distribution

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

- Characterizes the *ideal trajectory* of *disease progression* for subject $i$
- $\theta_i$ *vary across subjects* $\implies$ variation in trajectories
- Describe this variation by a *probability distribution*

$$p(\theta_i; \theta^*, D), \text{ e.g., } \theta_i \sim \mathcal{N}(\theta^*, D)$$
Mathematical-statistical framework

**Result:** Hypothesized *data-generation process* in *continuous time*

- For *individual* subjects (randomly chosen from the *population*)
- And thus for *samples of subjects* drawn from the population
- *Large enough* sample $\implies$ effective knowledge of the *entire population*
- Basis for *simulation* of “*virtual*” subjects

**To use this:** Need full characterization based on *data*, i.e.,

- *Estimate* the features of the *probability distributions*
- *Estimate* $\theta^* = \text{mean value}$ of the *model parameters* in the population and $D = \text{covariance matrix}$ describing how *subject-specific* $\theta_i$ vary about $\theta^*$
Mathematical-statistical framework

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

- $U_i(t) =$ actual *ARV pattern* over entire observation period (*known*)
- $Y_{ij} = (Y_{ij}^{CD4}, Y_{ij}^{VL})$ at time $t_{ij} \implies Y_i = (Y_{i1}, \ldots, Y_{in_i})$
- Conceive $Y_{ij} = Y_i\{t_{ij}, U_i(t_{ij})\}$ (similarly for $e_{ij}$)
- *Eric's data* – $N \approx 150$, $n_i \approx 30–60$

**Nonlinear mixed effects model:** *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, stability analysis*

• Use resulting $\hat{\theta}_i$ as “data” to obtain *estimates* $\hat{\theta}^*$, $\hat{D}$ using moment methods
Predictive capability:

Patient #14

CD4+ T-cells/ul

virus copies/ml

0 200 400 600 800 1000 1200 1400 1600

0 200 400 600 800 1000 1200 1400 1600

0 10 5

0 10^0 10^5

Modeling for HIV Treatment and Trials
Mathematical-statistical framework

Simulation:

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

- Generate "inherent trajectories" $\overline{x}\{t, u(t), \theta_{i}^{sim}\}$ 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 normal distributions
Design of a clinical trial

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 designing a clinical trial in acute HIV infection?

- Is it better to give ARV for some period following acute infection (“train” the immune system, “self-vaccinate”) followed by terminal interruption...
- A non-adaptive treatment strategy
- ...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 ARV therapy terminated at 6 months, i.e., $u(t) = 1$, $0 \leq t \leq 6$, $u(t) = 0$, $t > 6$
Design of a clinical trial

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
- Since the grant was awarded, we have developed a refined model
- Simulations with the refined model show larger subpopulations with lowered VL set point for larger $\tau$ . . .
- . . . but are less reliable (very little data on immune response)

Result: Study ARV 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 ARV, 1/2 pts to no ARV

Diagram:
- Acute HIV-1 Infection
  - ARV Therapy
    - Short-term Treatment then Interruption
  - No ARV Therapy
    - Long-term Treatment then Interruption
- Outcome
**Design of a clinical trial**

**Design:** 3 year accrual period, 1 year follow-up

- 36 subjects, 2:1:1 randomization to none, 3 months, 8 months
- Standard sample size 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 ARV termination*) based on the *math model*
Design of treatment strategies

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

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

- Modeling and simulation have a significant role to play in design of HIV treatment strategies and clinical trials to study them.

- In principle – could link HIV dynamic models with models for pharmacokinetics, 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.
References

