

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

---

**Eric S. Rosenberg**

Massachusetts General Hospital and  
Harvard Medical School

**Marie Davidian** and **H.T. Banks**

North Carolina State University

# The project

---

- **Multidisciplinary collaboration** supported by a grant from NIAID
- **Big picture:** Use **mathematical-statistical modeling of disease progression** and **simulation** to design HIV treatment strategies and clinical trials to study them
- Develop better strategies for use of **existing antiretroviral therapies (ART)** to manage HIV infection **over time**
- 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**  $\implies$  more sophisticated strategies and trials

# HIV dynamic models and control

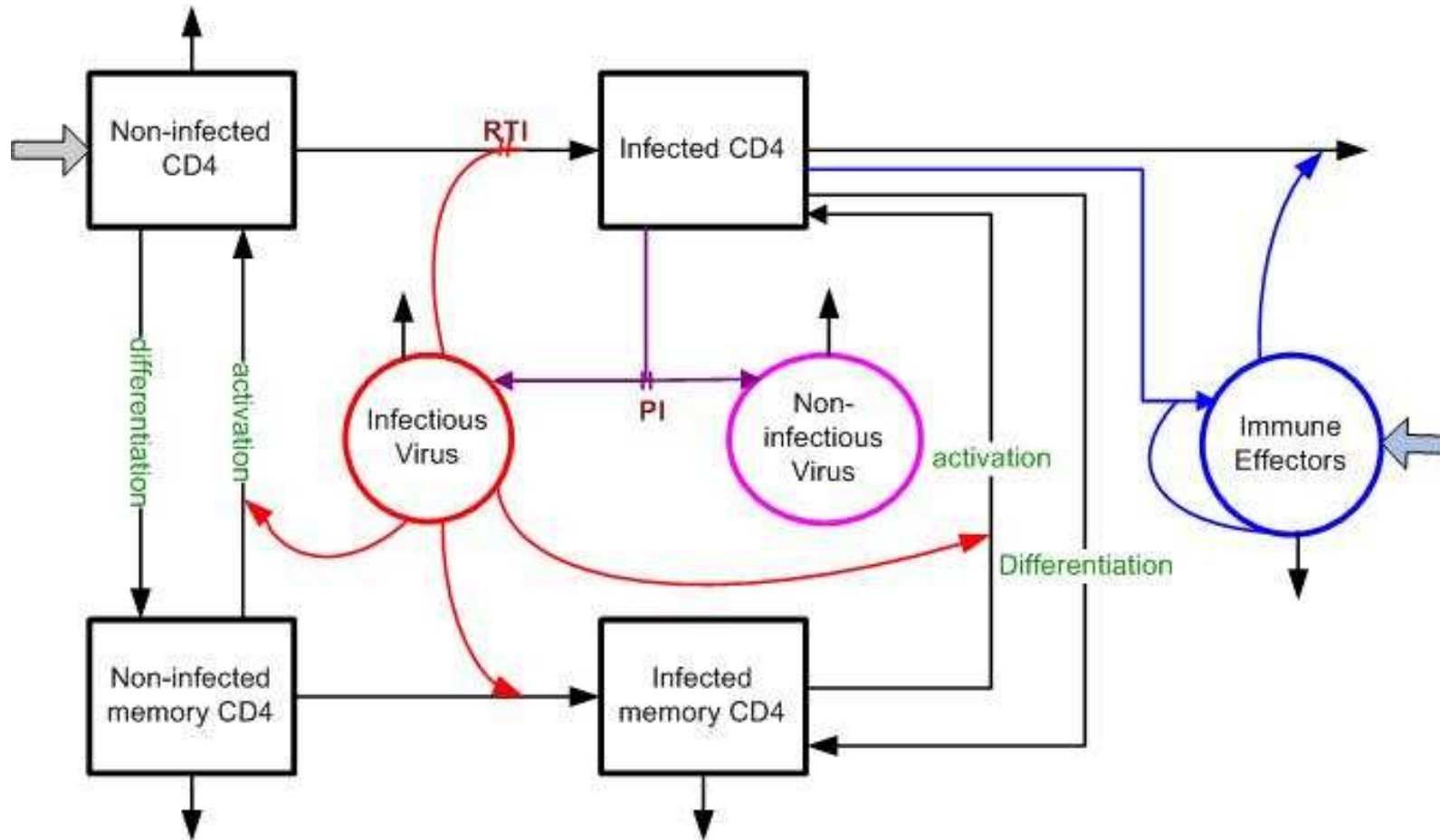
---

## HIV dynamic models:

- **Compartmental** representation of **mechanisms** involved in the **virus-immune system** interaction taking place over time **within a single subject**
- 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**
- $\implies$  **Viral load, CD4+ T cell count**, etc, at any time

# HIV dynamic models and control

Simple model for within-subject dynamics:



# HIV dynamic models and control

Model for within-subject dynamics:  $s = 7$  states

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

- $\theta = (\lambda_1, d_1, \epsilon_1, k_1, \dots)^T$  plus initial conditions
- Observable: **CD4 count** =  $T_1 + T_1^*$ , **viral load** =  $V_I + V_{NI}$
- $u(t)$  = ART 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

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

- 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

---

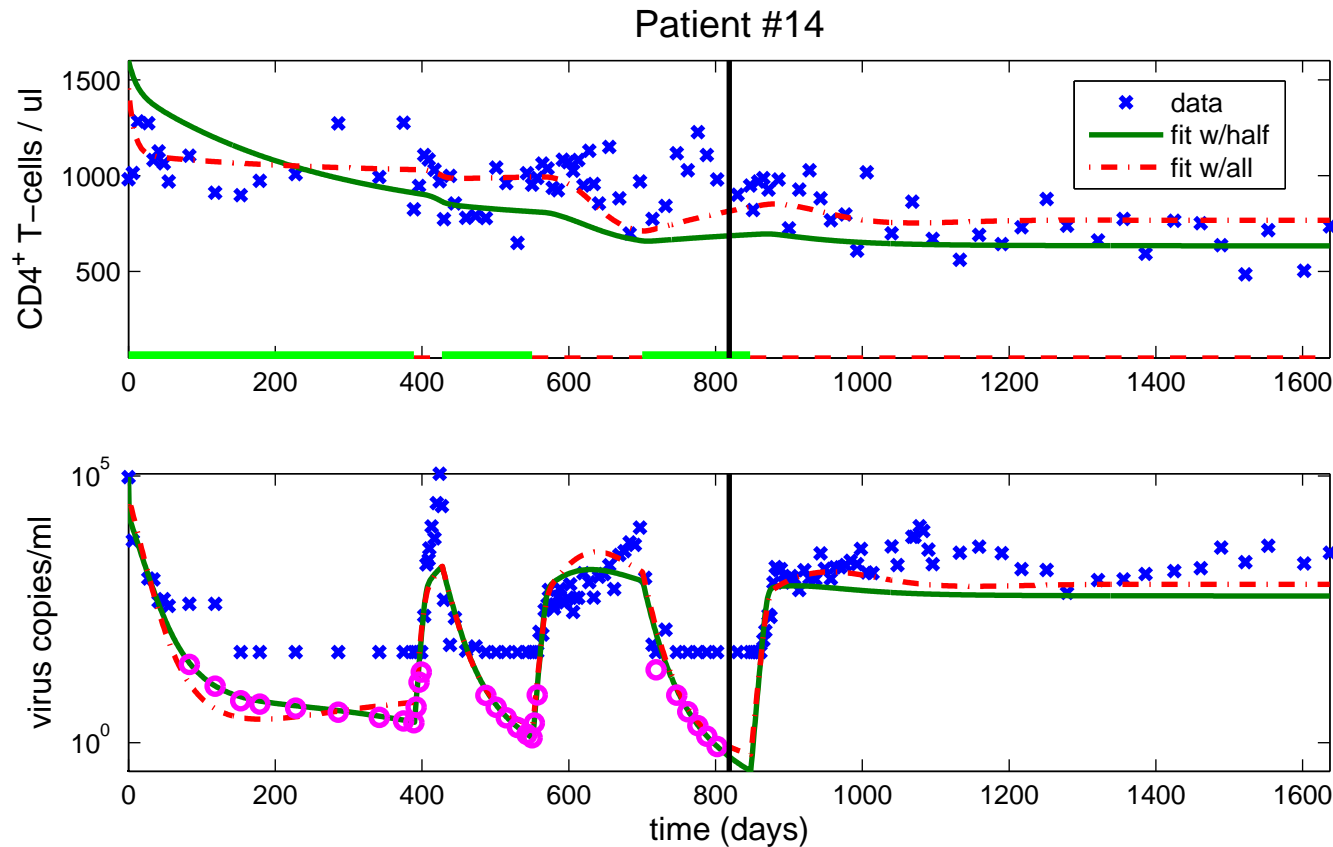
**Our ultimate goal:** Use **HIV dynamic models**, **control**, and **simulation** to **design treatment strategies**  $u(t)$  and **clinical trials** to study them

- Strategies that “**do well**” for individuals and for the population
- Need **evidence** supporting HIV dynamic model  $\implies$  **data** (e.g., **measured** CD4, VL, other stuff **over time** on lots of subjects)
- **Intra-subject variation** due to assay error, realization error
- 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**

# HIV dynamic models and control

**Data:** Eric has been collecting intensive **longitudinal** viral loads, CD4 counts on a **cohort** of  $\geq 150$  acutely-infected subjects for  $> 7$  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
- $\bar{x} = \mathcal{O}x$  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** of 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  $\implies$  effectively, knowledge of the **entire population**
- Basis for **simulation** of “**virtual**” subjects

**Needed:** Full characterization based on **data**

# Mathematical-statistical framework

---

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

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

**Nonlinear mixed effects model (bivariate response):** **Fit** to data

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

$$\theta_i \sim p(\theta_i; \theta^*, D), \quad i = 1, \dots, N$$

# Mathematical-statistical framework

---

## Challenges:

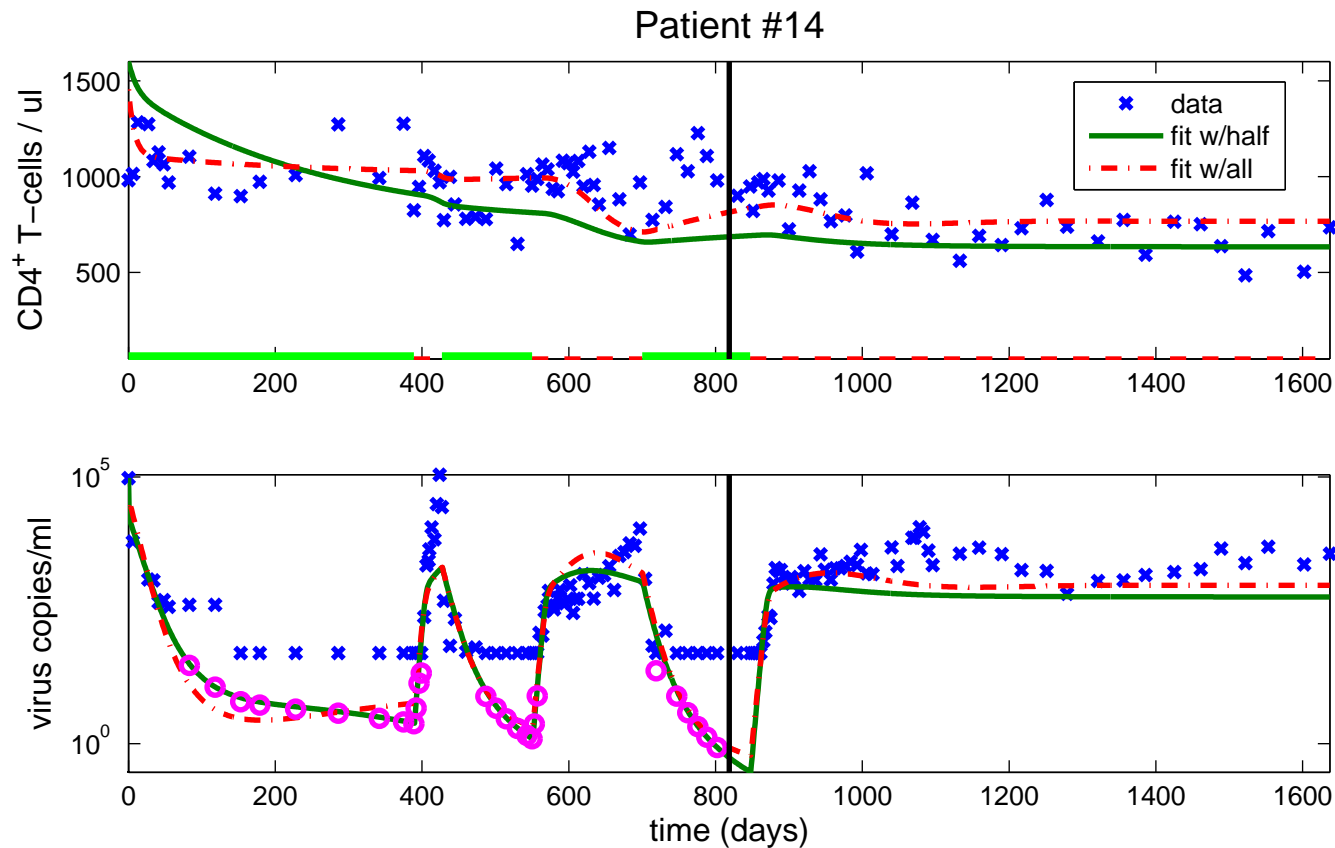
- **Left-censoring** of VL by **lower assay limit**
- $\dim(\theta) > 25$  and not all **identifiable** from CD4, VL only
- Components of  $\bar{x}$  only calculable **numerically** by **forward solution** of ODEs

## Two-stage approach:

- For each  $i$ , estimate  $\theta_i$  via EM algorithm to handle censoring incorporating **regularization** to address **identifiability** and **dimensionality**
- Use resulting  $\hat{\theta}_i$  as “data” to obtain  $\hat{\theta}^*$ ,  $\hat{D}$  using moment methods

# Mathematical-statistical framework

Predictive capability:



# Mathematical-statistical framework

---

## Simulation:

- Generate  $N_{sim}$  “**virtual subjects**” by generating  $\theta_i^*$ ,  $i = 1, \dots, N_{sim}$ , 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**”

# Design of a clinical trial

---

**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

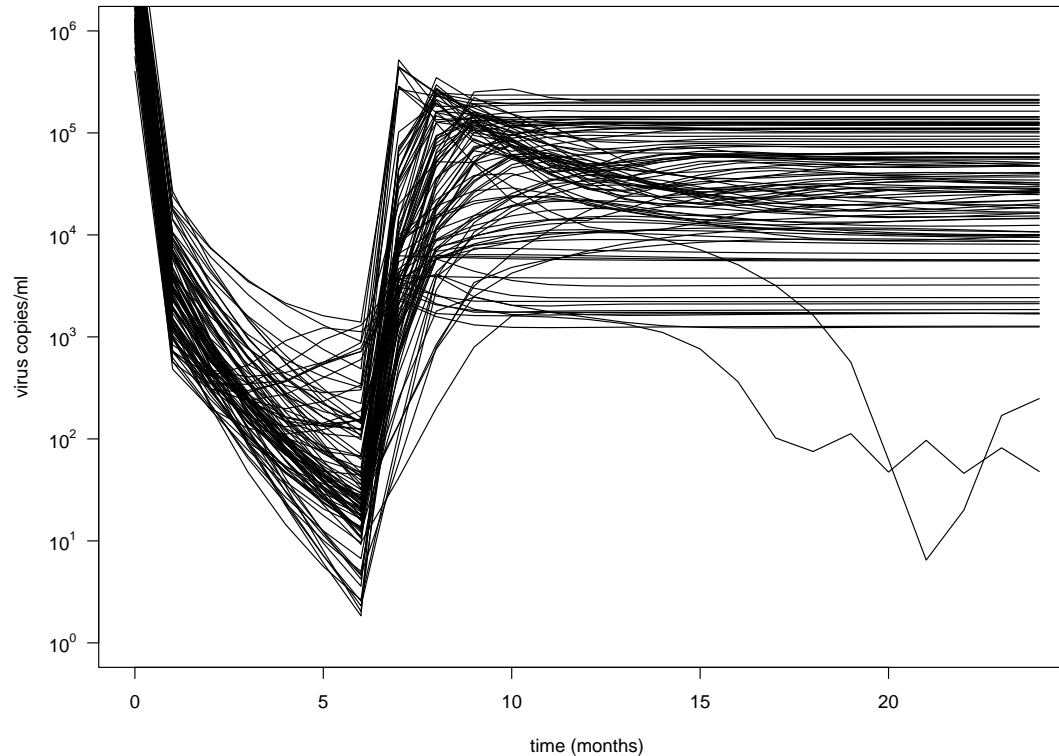
$$\begin{aligned}u(t) &= 1, 0 \leq t \leq \tau \\ &= 0, t > \tau\end{aligned}$$

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

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

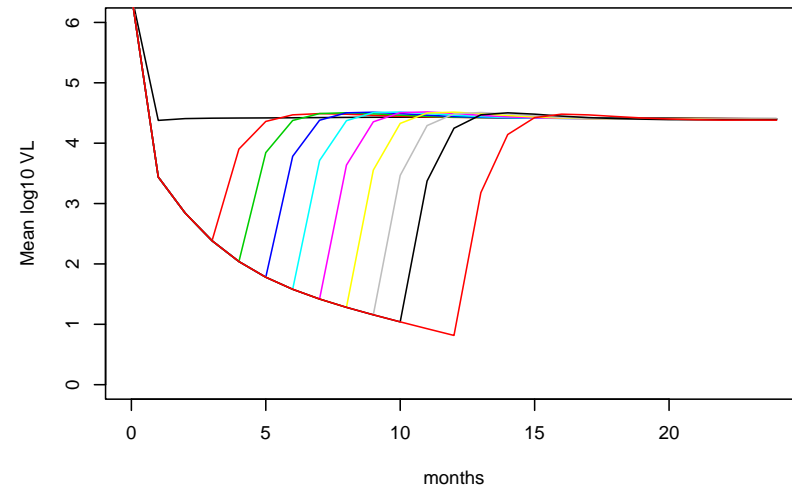
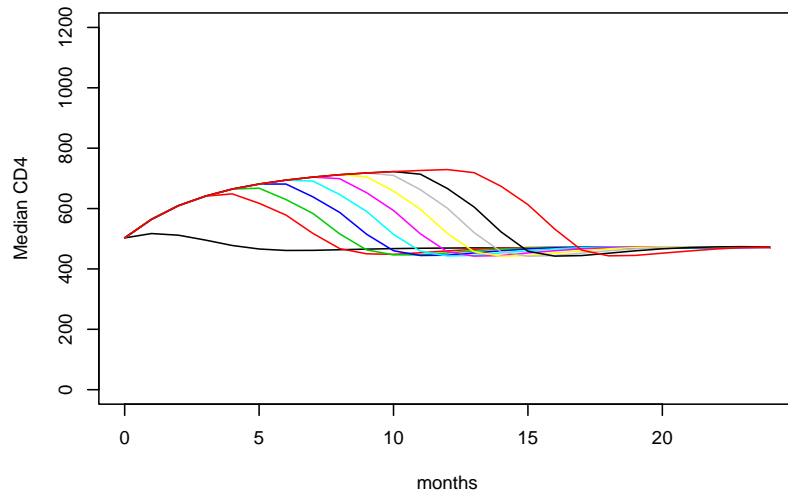
# Design of a clinical trial

**Strategy  $u(t)$  with  $\tau = 6$ :** 100 “virtual” “inherent” viral load trajectories with ART **terminated at 6 months**, i.e.,  $u(t) = 1, 0 \leq t \leq 6, u(t) = 0, t > 6$



# Design of a clinical trial

**Different termination times  $\tau$ :** Means of 15,000 “virtual” CD4 and viral load data profiles with  $u(t) = 1, 0 \leq t \leq \tau, u(t) = 0, t > \tau,$   
 $\tau = 0, 3, 4, \dots, 12$  months



# Design of a clinical trial

---

## Summary:

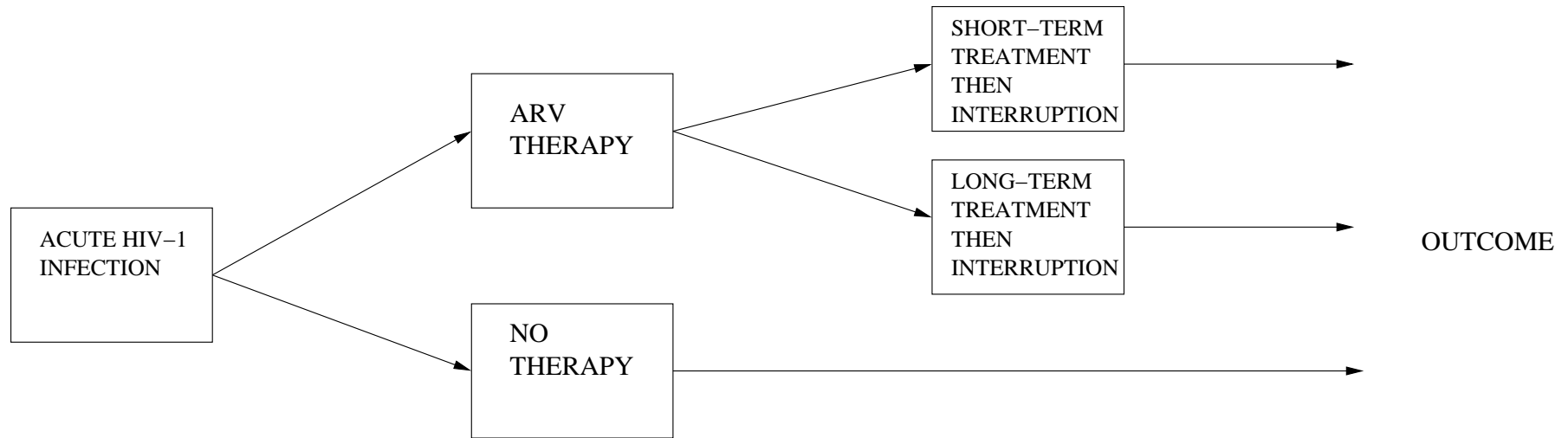
- Based on this (simple) HIV dynamic model, **no differences expected**
- Simple model does not represent adequately the **immune response**
- Simulations with a **refined model** showed larger **subpopulations** with **lowered VL set point** for larger  $\tau$  ...
- ... but are less reliable (very 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:** 18 pts randomized to ART, 18 pts to no ART



- **Primary comparison** – VL set point at 12 months
- **Intensive visit schedule** – CD4, VL, CTLs, viral fitness, etc
- Data collection more frequent when dynamics are anticipated to **be changing** (e.g., after **termination** of ART)

# 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 ...

# 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

# References

---

Adams, B.M., Banks, H.T., Davidian, M., and Rosenberg, E.S. (2007) Model fitting and prediction with HIV treatment interruption data. *Bulletin of Mathematical Biology* **69**, 563–584.

Rosenberg, E.S., Davidian, M., and Banks, H.T. (2007) Using mathematical modeling and control to develop structured treatment interruption strategies for HIV infection. *Drug and Alcohol Dependence* special supplement issue on “Customizing Treatment to the Patient: Adaptive Treatment Strategies” **88S**, S41-S51.