Modeling AIDS Clinical Trials and Antiviral Treatment Effects

Hulin Wu

Department of Biostatistics and Computational Biology
University of Rochester School of Medicine and Dentistry

hwu@bst.rochester.edu

Outline

- **Background and Objectives**
- **HIV Dynamic Models**
- **Models for Drug Exposure and Response**
- **Parameter Estimation and Model Fitting**
- **An AIDS Clinical Study**
- **Summary and Discussion**

Modeling Our Knowledge

- **Knowledge Sources for HIV/AIDS Treatments**
  - **Established mechanisms and biological theories**
  - **Prior information: published results and other studies**
  - **New/Current information: data at hand**

Modelers

- **Mathematicians:**
  - **Use established mechanisms for modeling/simulations.**
  - **Data from individual patients NOT efficiently used.**
- **Statisticians:**
  - **Focus on current information/data for statistical analysis.**
  - **The prior information/data and biological meachanisms/theories NOT efficiently used.**
- **Bayes Statisticians:**
  - **Use both current data and prior information for statistical inference.**
  - **Biological meachanisms/theories NOT efficiently used.**

How to Avoid the Problem?

How can we use the information from all different sources to achieve our goals?

- **Bridge the gap between mathematicians and statisticians**

Our Objectives

- **Develop mathematical models for the mechanisms of HIV infection and antiviral treatment effects**
  - **PK/PD models**
  - **Adherence models**
  - **Drug susceptibility**
- **Develop statistical methods for parameter identification, model fitting and prediction**
  - **Deal with the complexity of the models: Nonlinear differential equations**
  - **Deal with the unidentifiability issues**
  - **Deal with the intensive computations**
- **Apply the established models for AIDS clinical trial simulations and search for optimal treatment strategies**

A Mechanisms-Based Model for HIV Infection

- **A viral dynamic model: describe the population dynamics of HIV and its target cells in plasma**

$$\begin{aligned}
\tfrac{d}{dt}T &= \lambda - \rho T - [1-\gamma(t)]kTV \\
\tfrac{d}{dt}T^* &= [1-\gamma(t)]kTV - \delta T^* \\
\tfrac{d}{dt}V &= N\delta T^* - cV
\end{aligned} \tag{1}$$

  - T, T^*, V: **target uninfected cells, infected cells, virus**
  - $\gamma(t)$: **time-varying antiviral drug efficacy**
  - $(\lambda, \rho, k, \delta, N, C)$: **unknown parameters to be estimated**
  - **The equations (2): no closed-form solutions**

Selection of Mechanisms-Based Models

- Consider your objectives/goals to select the model
  - For prediction of clinical outcomes?
  - For understanding biological mechanisms?
  - For studing a new treatment strategy?
  - For modeling immunological responses or virological responses?
  - For modeling drug effects?
  - ??????

- **Consider the trade-off between the model accuracy and model complexity**
  - **Impossible to model everything in details**
  - **Important components missing: not accurate**
  - **Too many components included: too complex**
  - **What information/data do you have?**
    Do not use a model you cannot identify
  - **Try to use all information to identify more accurate (more complex too) model**
  - **Try to use a simpler model if your goal can be achieved**
- **Sensitivity analysis: dealing with some uncertainty of the model**

A Mechanisms-Based Model for HIV Infection

- **A viral dynamic model: describe the population dynamics of HIV and its target cells in plasma**

$$\begin{aligned} \frac{d}{dt}T &= \lambda - \rho T - [1-\gamma(t)]kTV \\ \frac{d}{dt}T^* &= [1-\gamma(t)]kTV - \delta T^* \\ \frac{d}{dt}V &= N\delta T^* - cV \end{aligned} \tag{2}$$

  - T, T^*, V: **target uninfected cells, infected cells, virus**
  - $\gamma(t)$: **time-varying antiviral drug efficacy**
  - $(\lambda, \rho, k, \delta, N, C)$: **unknown parameters to be estimated**
  - **The equations (2): no closed-form solutions**

Antiviral Drug Efficacy Model

- **A modified E_{max} model for drug efficacy:**

$$\gamma(t) = \frac{C(t)A(t)}{\phi IC_{50}(t) + C(t)A(t)} = \frac{IQ(t)A(t)}{\phi + IQ(t)A(t)}, \quad 0 \leq \gamma(t) \leq 1 \quad (3)$$

  - $C(t)$: **the plasma drug concentration**
  - $A(t)$: **drug adherence measurements**
  - IC_{50}: **in vitro phenotype drug resistance marker**
  - ϕ: **a conversion factor parameter**
  - $IQ = \frac{C(t)}{IC_{50}(t)}$: **the Inhibitory Quotient (IQ)**

- **If $\gamma(t) = 1$, the drug: 100% effective**
- **If $\gamma(t) = 0$, the drug: no effect**

Two or More Drug Regimens

$$\begin{aligned}\gamma(t) &= \frac{[C_1(t)A_1(t)/IC_{50}^1(t)] + [C_2(t)A_2(t)/IC_{50}^2(t)]}{\phi + [C_1(t)A_1(t)/IC_{50}^1(t)] + [C_2(t)A_2(t)/IC_{50}^2(t)]} \qquad (4)\\ &= \frac{IQ_1(t)A_1(t) + IQ_2(t)A_2(t)}{\phi + IQ_1(t)A_1(t) + IQ_2(t)A_2(t)} \qquad (5)\end{aligned}$$

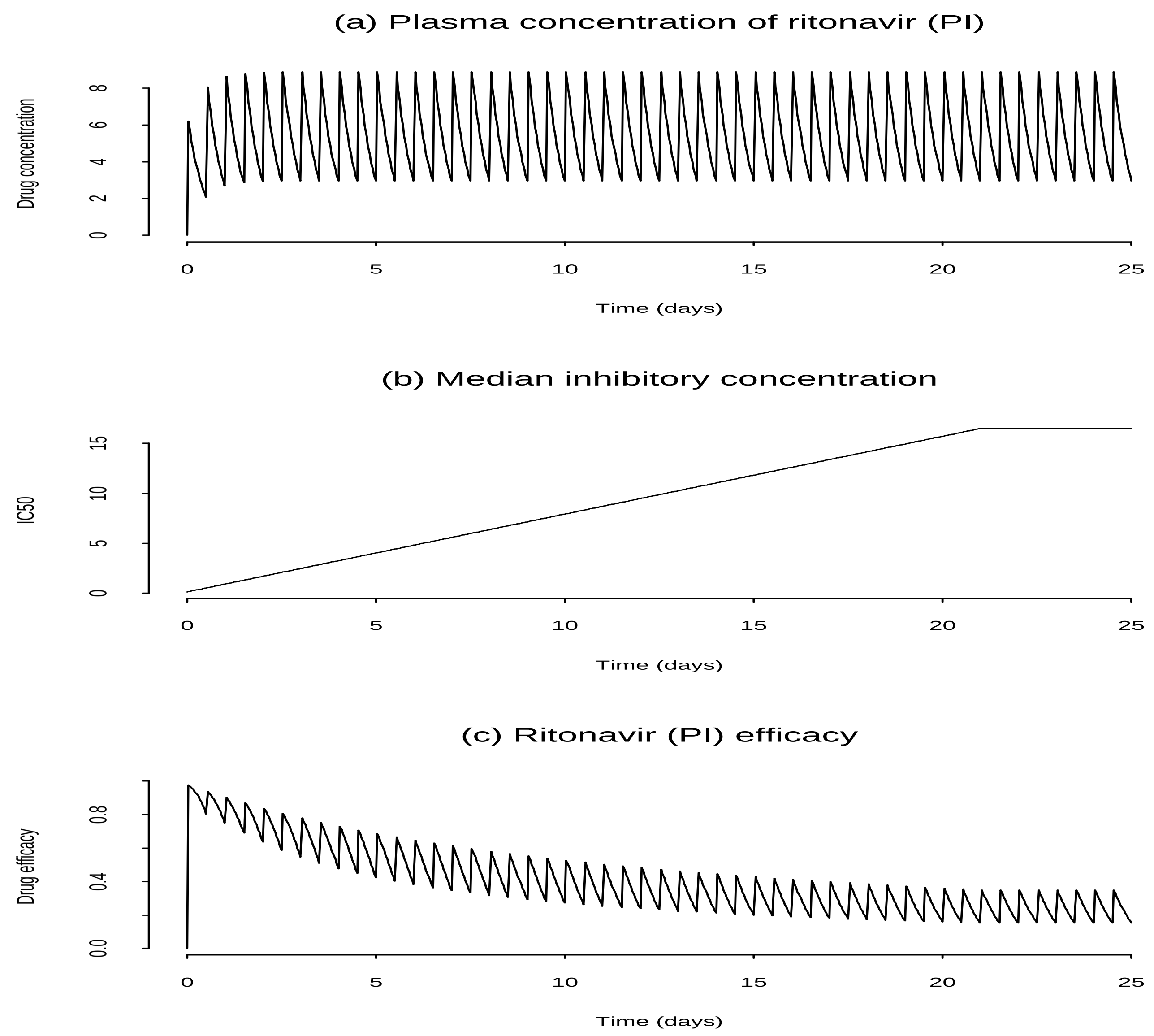
- $C_1(t)$ **and** $C_2(t)$ **: the plasma concentration for the two agents.**
- IC_{50}^1 **and** IC_{50}^2**: the median inhibitory concentration of the two agents.**
- $A_1(t)$ **and** $A_2(t)$**: the adherence rates of the two agents.**

Drug Susceptibility Model

- **Phenotype marker IC_{50} is used to quantify agent-specific drug sensitivity**
- **The function: to describe changes overtime in IC_{50}**

$$IC_{50}(t) = \begin{cases} I_0 + \frac{I_r - I_0}{t_r} t & \textbf{for } 0 < t < t_r, \\ I_r & \textbf{for } t \geq t_r, \end{cases} \tag{6}$$

  - **I_0 and I_r: respective values of $IC_{50}(t)$ at baseline and time point t_r at which drug resistant mutations appear**
  - **If $I_r = I_0$, no resistance mutation developed during treatment**


(a) Plasma concentration of ritonavir (PI)
Drug concentration
0
2
4
6
8
Time (days)
0
5
10
15
20
25
(b) Median inhibitory concentration
IC50
0
5
10
15
Time (days)
(c) Ritonavir (PI) efficacy
Drug efficacy
0.0
0.4
0.8
Time (days)

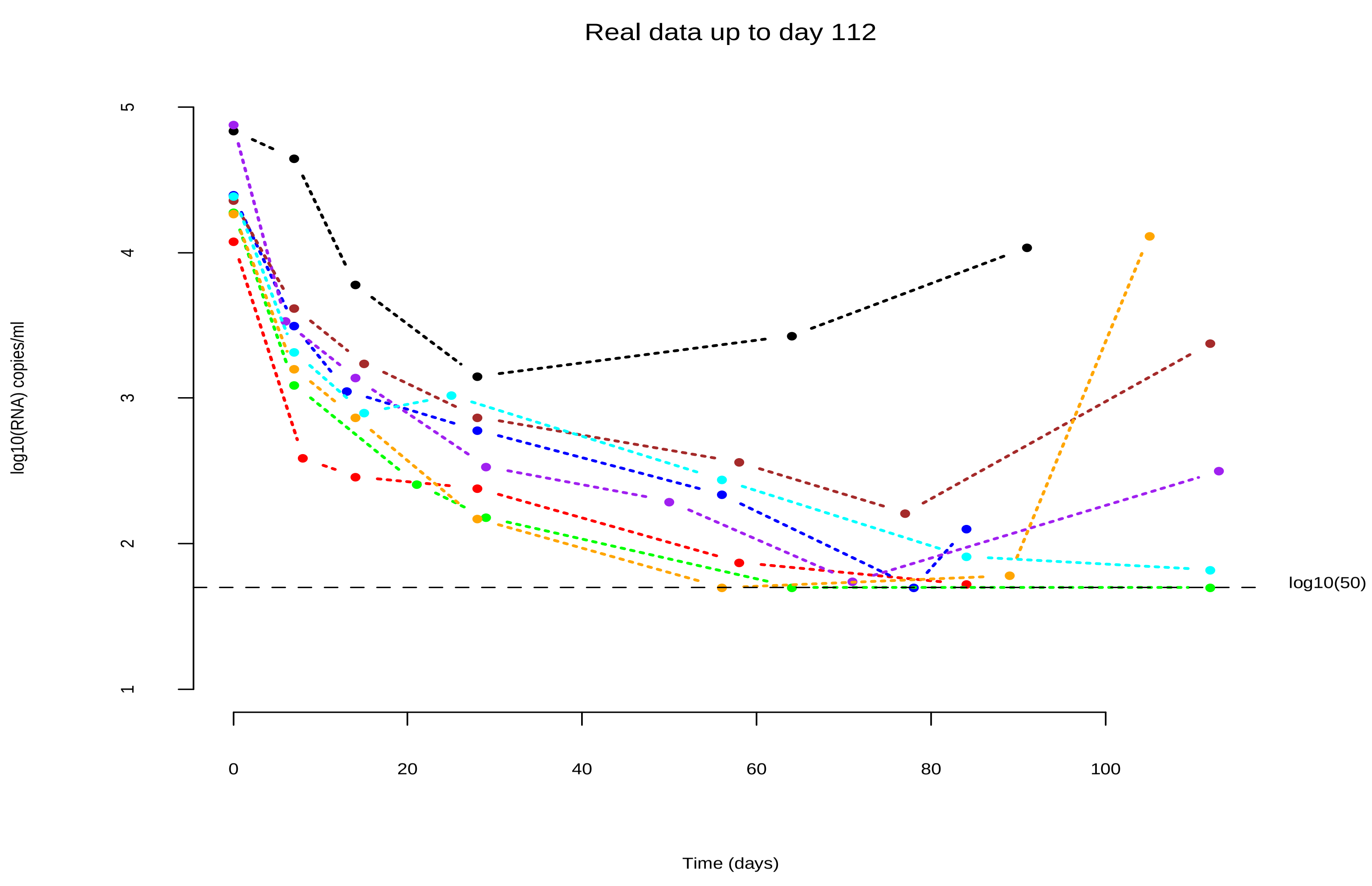
Properties of the HIV Dynamic Model

- Direct relationship between drug efficacy (drug exposure and drug sensitivity) and viral load
- A threshold of drug efficacy: $e_c = 1 - \frac{c\rho}{kN\lambda}$
  - if drug efficacy $\gamma(t) > e_c$, Model (2) converges to a stable uninfected steady-state
    * Virus will be eventually eradicated in theory
  - if $\gamma(t) < e_c$, the uninfected state is not stable and the endemically infected state exists
    * Viral load may rebound
  - The threshold e_c: may reflect the immune status of patients

A Challenging Problem

How to Estimate the Unknown Parameters in the Dynamic Model?

- **Difficulties:**
  - **Identifiability problem: Too many parameters,** $(\phi, \lambda, \rho, k, \delta, N, C)$
  - **Data from individuals: sparse**
  - **Different response patterns for different patients**
  - **Nonlinear differential equations model: no closed-form solutions**


Real data up to day 112
log10(RNA) copies/ml
1
2
3
4
5
0
20
40
60
80
100
Time (days)
log10(50)

Bayesian Hierarchical Model Approach

- Propose a three-stage hierarchical (mixed-effects) model
- Advantages of Bayesian hierarchical modeling approach
  - Naturally incorporate prior information
  - Deal with extremely complicated models such as nonlinear differential equation models
  - Use posterior distributions to easily answer inference questions
  - Estimate parameters for both population and individuals

Bayesian Modeling

- **A three-stage Bayesian hierarchical model**

Stage 1. Within-subject variation:

$$\mathbf{y}_i = \mathbf{f}_i(\boldsymbol{\theta}_i) + \mathbf{e}_i, \quad [\mathbf{e}_i|\sigma^2, \boldsymbol{\theta}_i] \sim \mathcal{N}(\mathbf{0}, \sigma^2 \mathbf{I}_{m_i})$$

- $\mathbf{f}_i(\boldsymbol{\theta}_i) = (f_{i1}(\boldsymbol{\theta}_i, t_1), \cdots, f_{im_i}(\boldsymbol{\theta}_i, t_{m_i}))^T$: **ODE solutions.**
- $\mathbf{y}_i = (y_{i1}(t_1), \cdots, y_{im_i}(t_{m_i}))^T$: **Data from Subject** i
- $\mathbf{e}_i = (e_i(t_1), \cdots, e_i(t_{m_i}))^T$: **Measurement error**

Stage 2. Between-subject variation:

$$\boldsymbol{\theta}_i = \boldsymbol{\mu} + \mathbf{b}_i, \quad [\mathbf{b}_i|\boldsymbol{\Sigma}] \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma})$$

Stage 3. Hyperprior distributions:

$$\sigma^{-2} \sim Ga(a, b), \quad \boldsymbol{\mu} \sim \mathcal{N}(\boldsymbol{\eta}, \boldsymbol{\Lambda}), \quad \boldsymbol{\Sigma}^{-1} \sim Wi(\boldsymbol{\Omega}, \nu)$$

- **Gamma (**Ga**), Normal (**$\mathcal{N}$**) and Wishart (**Wi**): independent distributions**
- **Hyper-parameters** $a, b, \boldsymbol{\eta}, \boldsymbol{\Lambda}, \boldsymbol{\Omega}$ **and** ν**: known**

Bayesian Estimation: Implementation

- Choose prior distributions
  - Informative prior and non-informative prior
  - Rule of thumb: choose non-informative prior distributions for parameters of interest
- Implement MCMC algorithm
  - Gibbs sampling step: closed form of conditional distributions for $\sigma^{-2}, \boldsymbol{\mu}, \boldsymbol{\Sigma}^{-1}$
  - Metropolis-Hastings step: no closed form of conditional distributions for θ_i
- Run a long chain: the number of iterations, initial "burn-in", every fifth simulation samples
- Obtain posterior distributions (posterior means or credible intervals) based on the final MCMC samples

A Clinical Study: A5055

- **A study of HIV-1 infected patients failing PI-containing therapies.**
- **Two salvage regimens:**
  - **Arm A: IDV 800 mg q12h+RTV 200mg q12h+two NRTIs**
  - **Arm B: IDV 400 mg q12h+RTV 400mg q12h+two NRTIs**
- **Plasma HIV-1 RNA (viral load) measured at days 0, 7, 14, 28, 56, 84, 112, 140 and 168 of follow-up**

Clinical Data –Results of population parameters

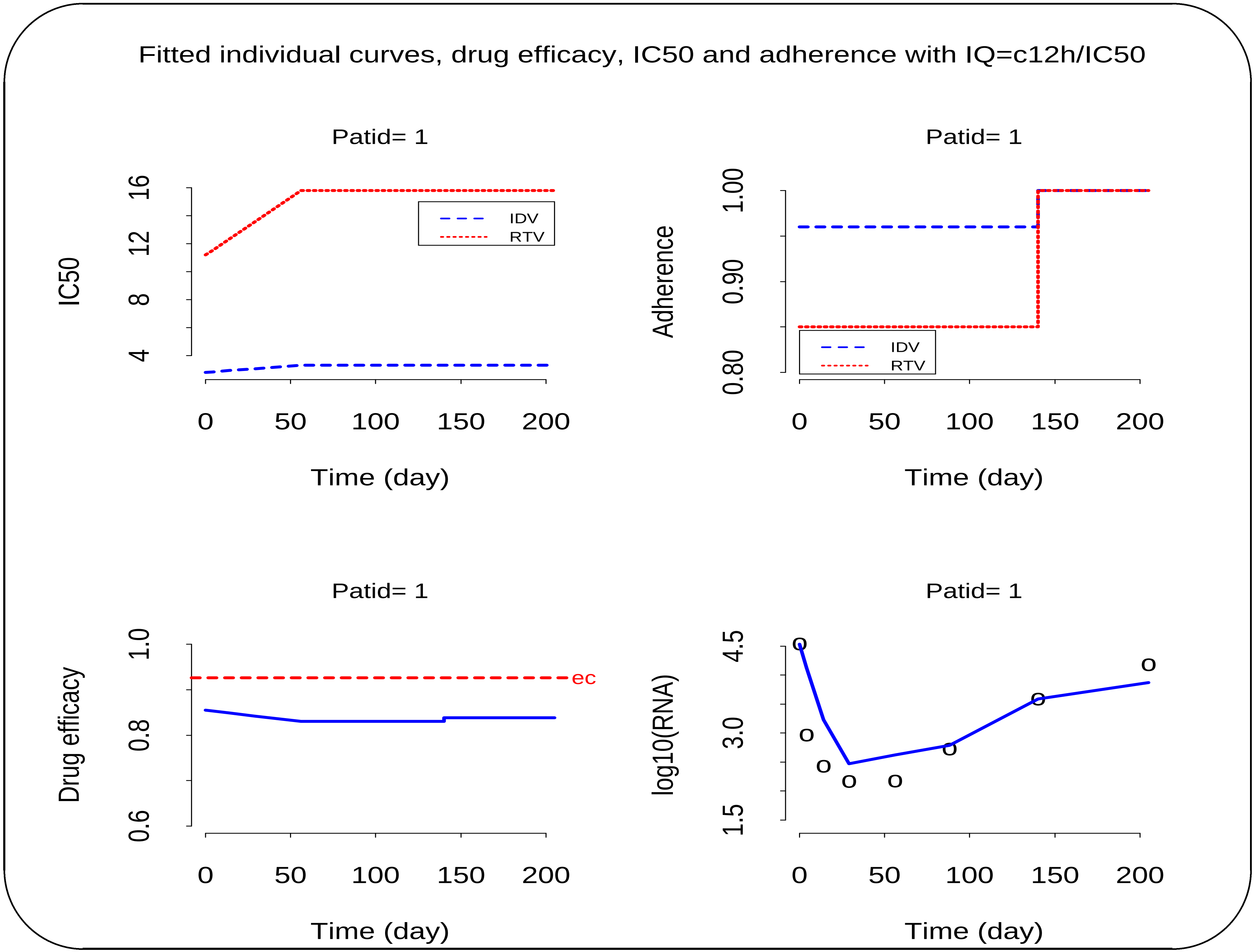
Parameter	PM	SD	95% CI
ϕ	**2.1091**	**0.6354**	**(1.2143, 3.6392)**
c	**2.9867**	**0.1466**	**(2.7139, 3.2881)**
δ	**0.3729**	**0.0184**	**(0.3387, 0.4105)**
λ	**100.645**	**4.9431**	**(91.497, 110.830)**
ρ	**0.0997**	**0.0049**	**(0.0905, 0.1099)**
N	**1004.988**	**49.795**	**(912.074, 1106.654)**
k	9.183×10^{-6}	0.290×10^{-6}	$(8.632 \times 10^{-6}, 9.774 \times 10^{-6})$

- **Posterior mean for the population parameter ϕ is 2.1091 with a SD of 0.6354 and the 95% CI of (1.2143, 3.6392)**
- **As ϕ plays a role of transforming the *in vitro* IC_{50} into *in vivo* IC_{50}, our estimate shows that there is about 2-fold difference between *in vitro* IC_{50} and *in vivo* IC_{50}**

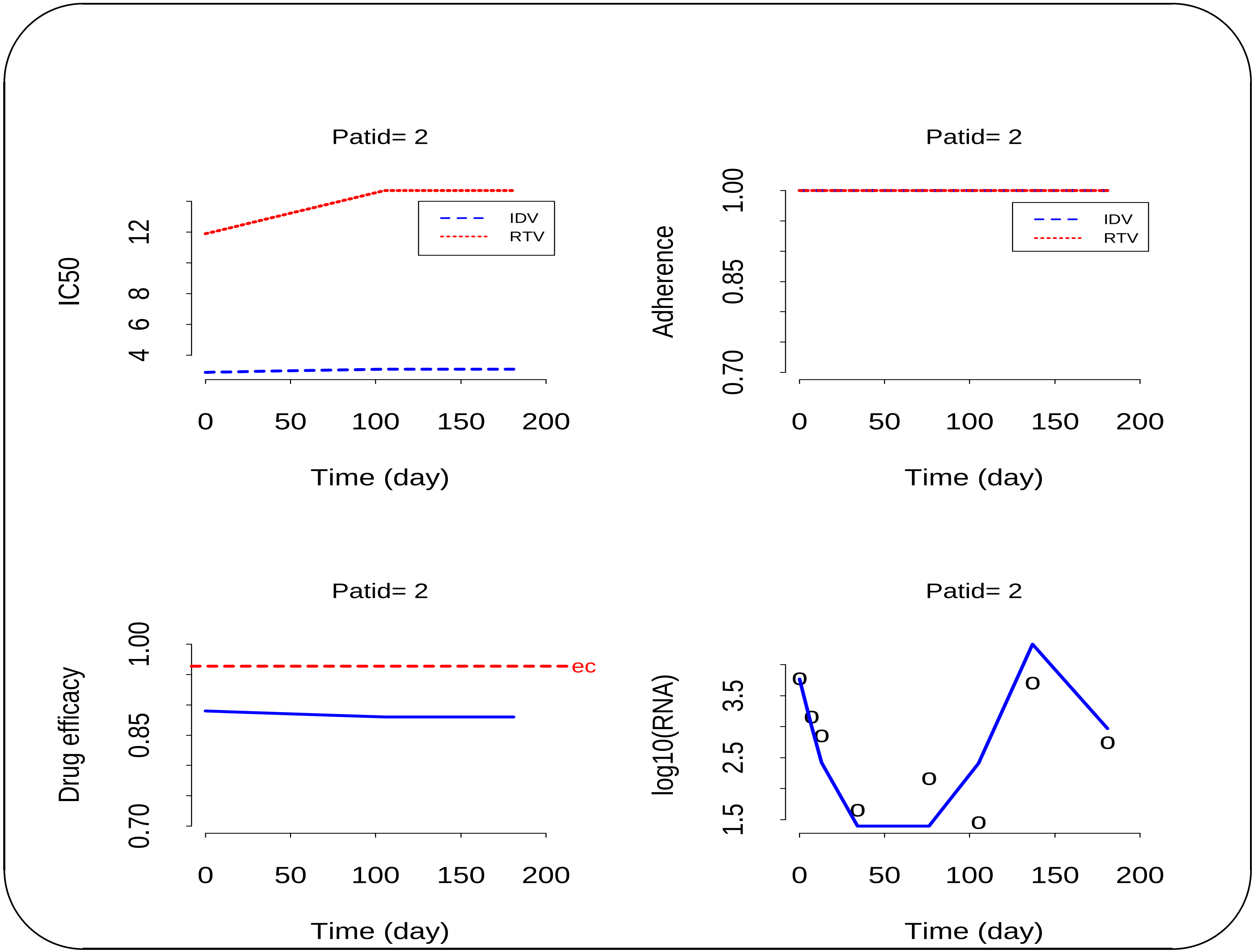
Clinical Data–Results of individual parameters

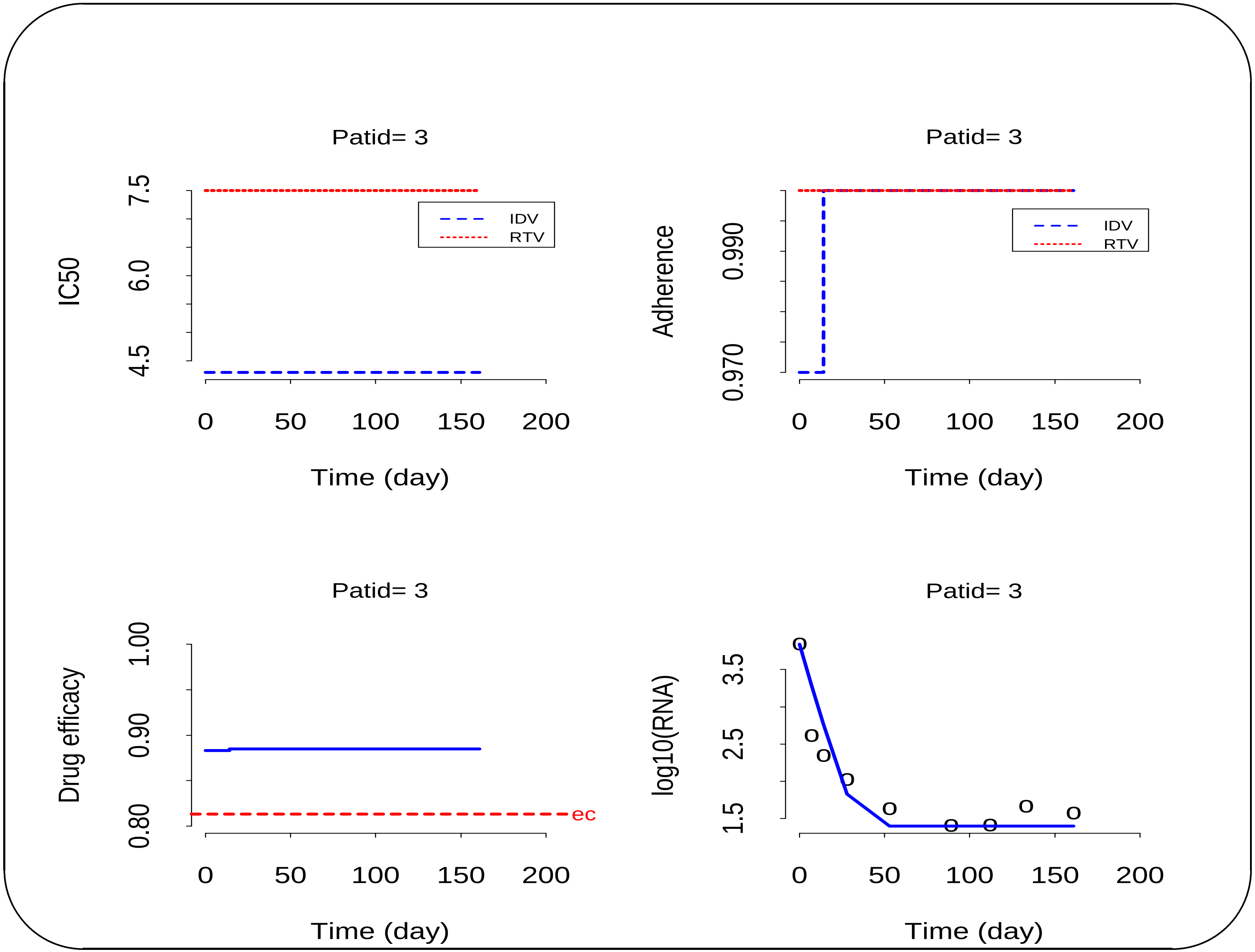
Patient	ϕ_i	c_i	δ_i	λ_i	ρ_i	N_i	k_i	e
1	0.447	2.254	0.270	410.462	0.024	456.757	8.33×10^{-6}	0.97
2	5.371	2.969	1.183	29.619	0.426	4795.813	10.84×10^{-6}	**0.17**
3	3.723	2.283	0.456	36.877	0.289	3258.347	8.66×10^{-6}	0.37
4	4.960	2.761	0.798	44.956	0.313	3051.988	9.09×10^{-6}	0.34
5	7.066	2.306	0.663	71.295	0.201	2735.239	6.54×10^{-6}	0.64
6	0.786	4.633	0.183	375.882	0.025	247.416	11.18×10^{-6}	0.89
7	0.091	7.008	0.299	4015.398	0.003	30.559	18.54×10^{-6}	0.98
8	8.484	2.280	0.663	32.722	0.416	4530.531	8.37×10^{-6}	**0.24**

- **The individual-specific parameter estimates suggest a large inter-subject variation**
- **The model provides a good fit to the clinical data**

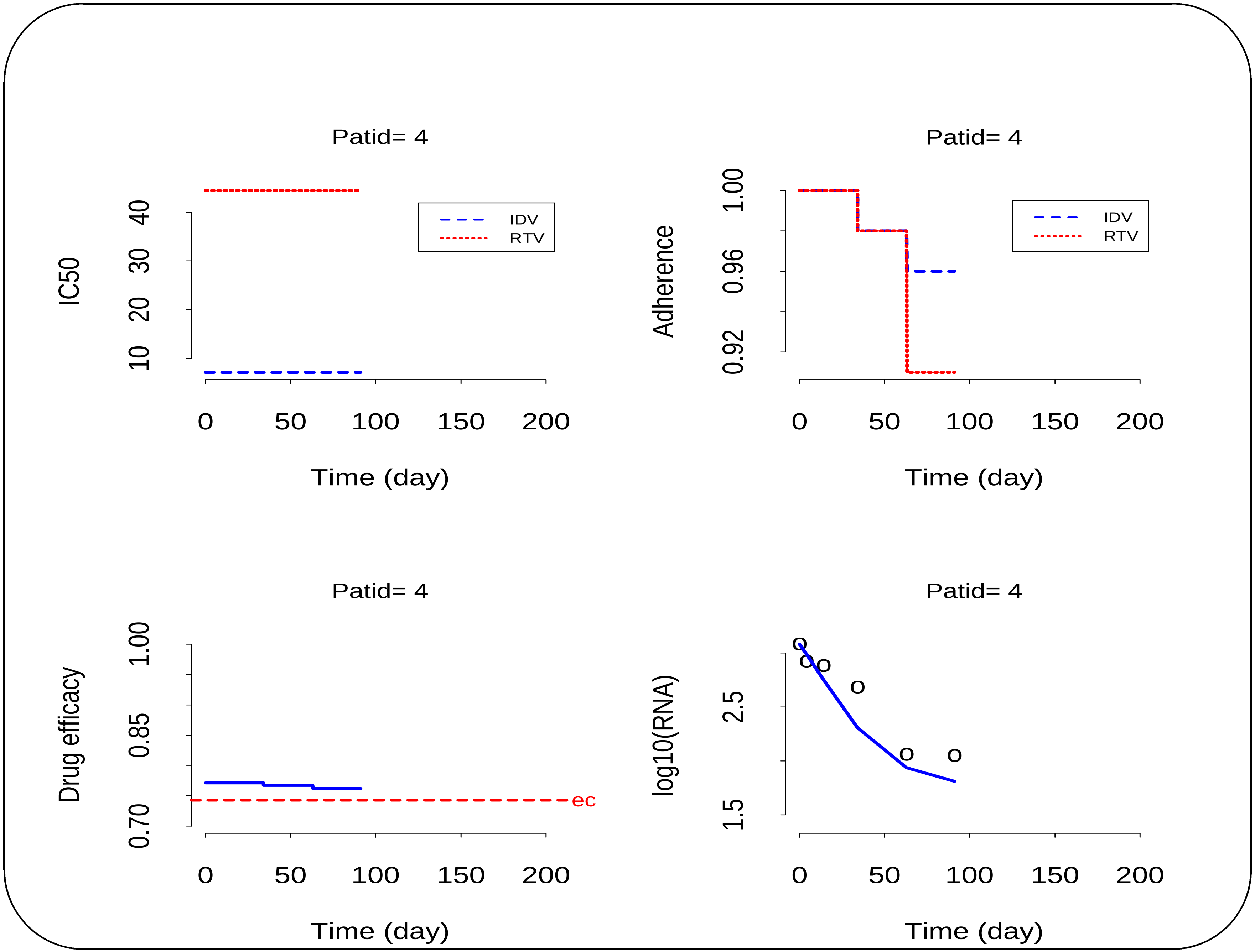


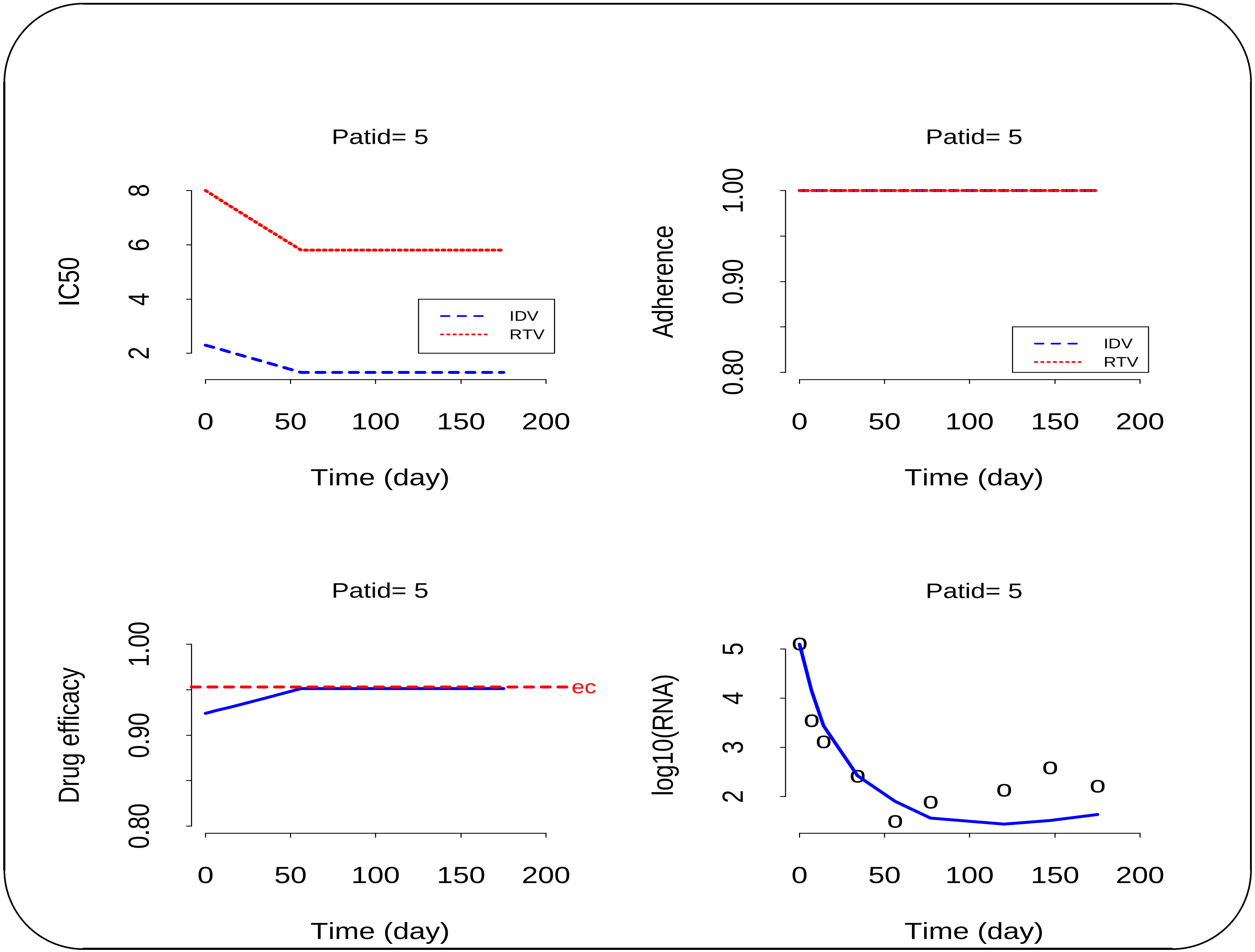
Fitted individual curves, drug efficacy, IC50 and adherence with IQ=c12h/IC50
Patid= 1
IC50
4
8
12
16
IDV
RTV
0
50
100
150
200
Time (day)
Patid= 1
Adherence
0.80
0.90
1.00
IDV
RTV
0
50
100
150
200
Time (day)
Patid= 1
Drug efficacy
0.6
0.8
1.0
ec
0
50
100
150
200
Time (day)
Patid= 1
log10(RNA)
1.5
3.0
4.5
0
50
100
150
200
Time (day)


Patid= 2
IC50
4 6 8 12
IDV
RTV
0 50 100 150 200
Time (day)
Patid= 2
Adherence
0.70 0.85 1.00
IDV
RTV
0 50 100 150 200
Time (day)
Patid= 2
Drug efficacy
0.70 0.85 1.00
ec
0 50 100 150 200
Time (day)
Patid= 2
log10(RNA)
1.5 2.5 3.5
0 50 100 150 200
Time (day)

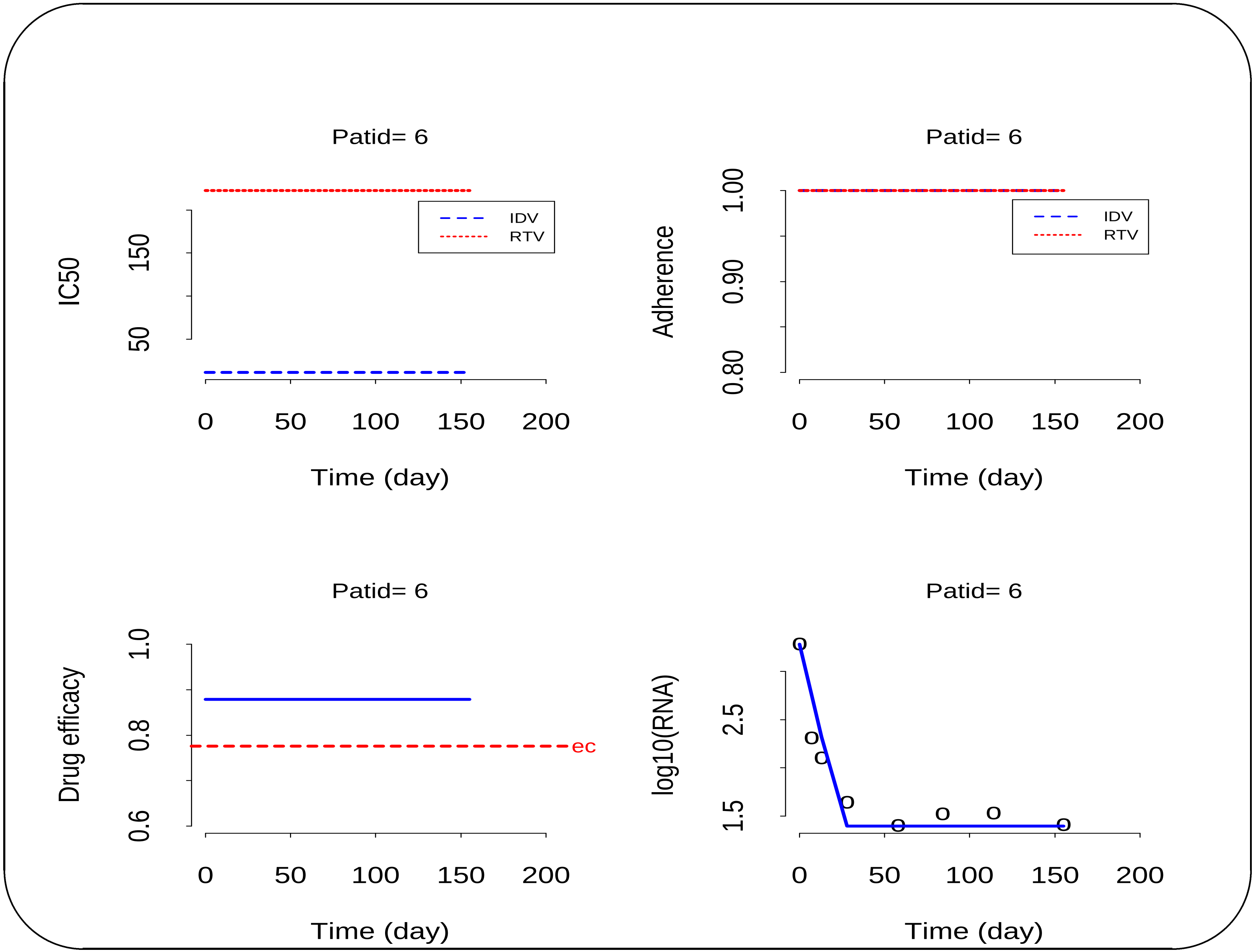


Patid= 3
IC50
7.5
6.0
4.5
IDV
RTV
0
50
100
150
200
Time (day)
Patid= 3
Adherence
0.990
0.970
IDV
RTV
0
50
100
150
200
Time (day)
Patid= 3
Drug efficacy
1.00
0.90
0.80
ec
0
50
100
150
200
Time (day)
Patid= 3
log10(RNA)
3.5
2.5
1.5
0
50
100
150
200
Time (day)


Patid= 4
IC50
10
20
30
40
IDV
RTV
0
50
100
150
200
Time (day)
Patid= 4
Adherence
0.92
0.96
1.00
IDV
RTV
0
50
100
150
200
Time (day)
Patid= 4
Drug efficacy
0.70
0.85
1.00
ec
0
50
100
150
200
Time (day)
Patid= 4
log10(RNA)
1.5
2.5
0
50
100
150
200
Time (day)



Patid= 5
IC50
IDV
RTV
Time (day)
Patid= 5
Adherence
IDV
RTV
Time (day)
Patid= 5
Drug efficacy
ec
Time (day)
Patid= 5
log10(RNA)
Time (day)



Patid= 6
IC50
50
150
IDV
RTV
0
50
100
150
200
Time (day)
Patid= 6
Adherence
0.80
0.90
1.00
IDV
RTV
0
50
100
150
200
Time (day)
Patid= 6
Drug efficacy
0.6
0.8
1.0
ec
0
50
100
150
200
Time (day)
Patid= 6
log10(RNA)
1.5
2.5
0
50
100
150
200
Time (day)

Questions

- **Model fitting looks good using the information:**
  - **PK: Trough-level drug concentration**
  - **Drug susceptibility: IC50**
  - **Adherence: Questionaire data**
- **Questions**
  - **Do all these factors contribute to good fitting of the model?**
  - **If not all, what are important factors?**
  - **Without using the complicated viral dynamic model, can we still see the effect of these factors on the response by simple regression analysis?**

Simple Regression or Correlation Analyses

- Diffculty: How to define the “response”?
- Viral load changes from baseline to week 4/week 24
- Simple regression or correlation: No effects

Figure 1: Simple Regression or Correlation Analyses

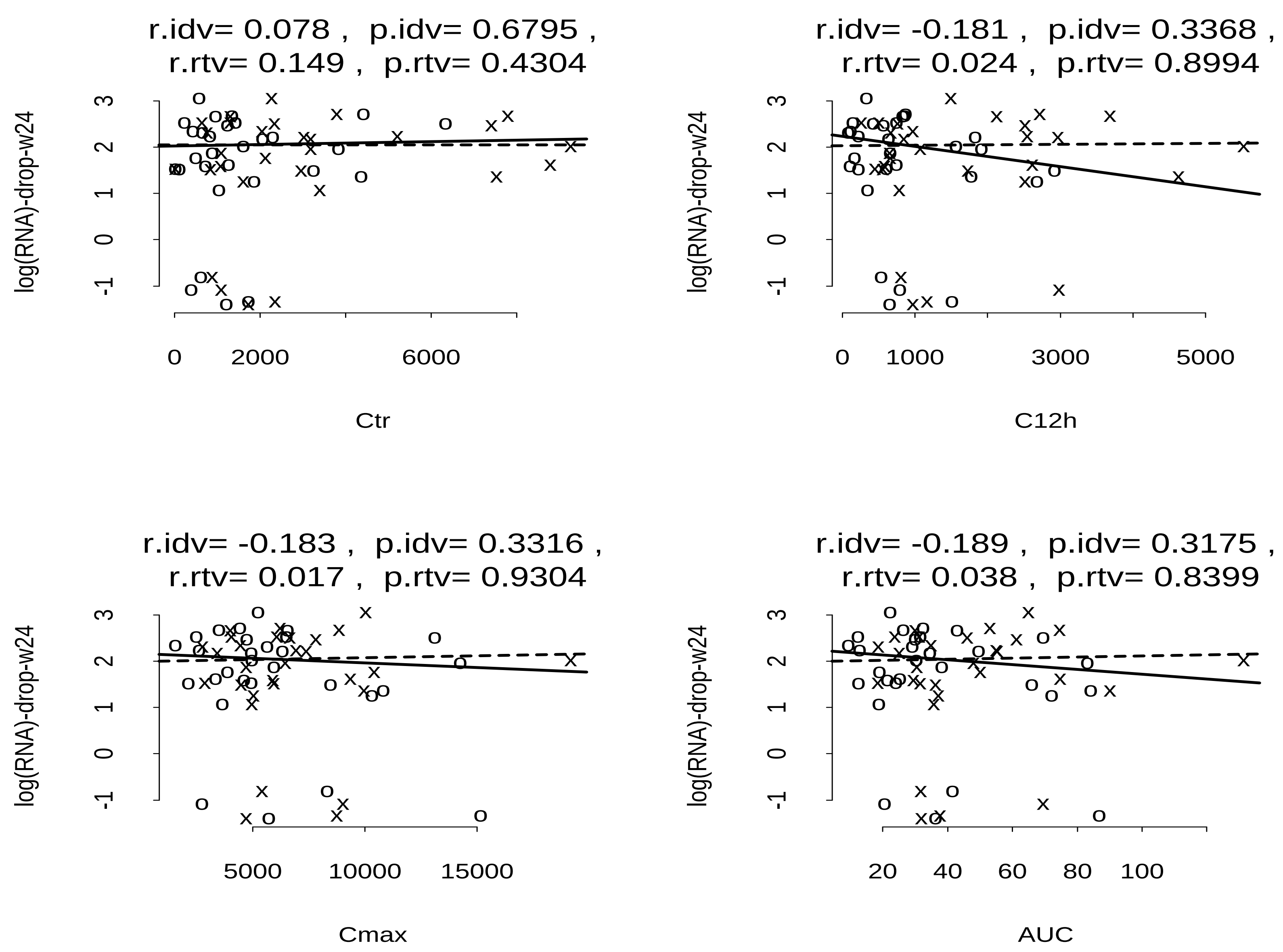
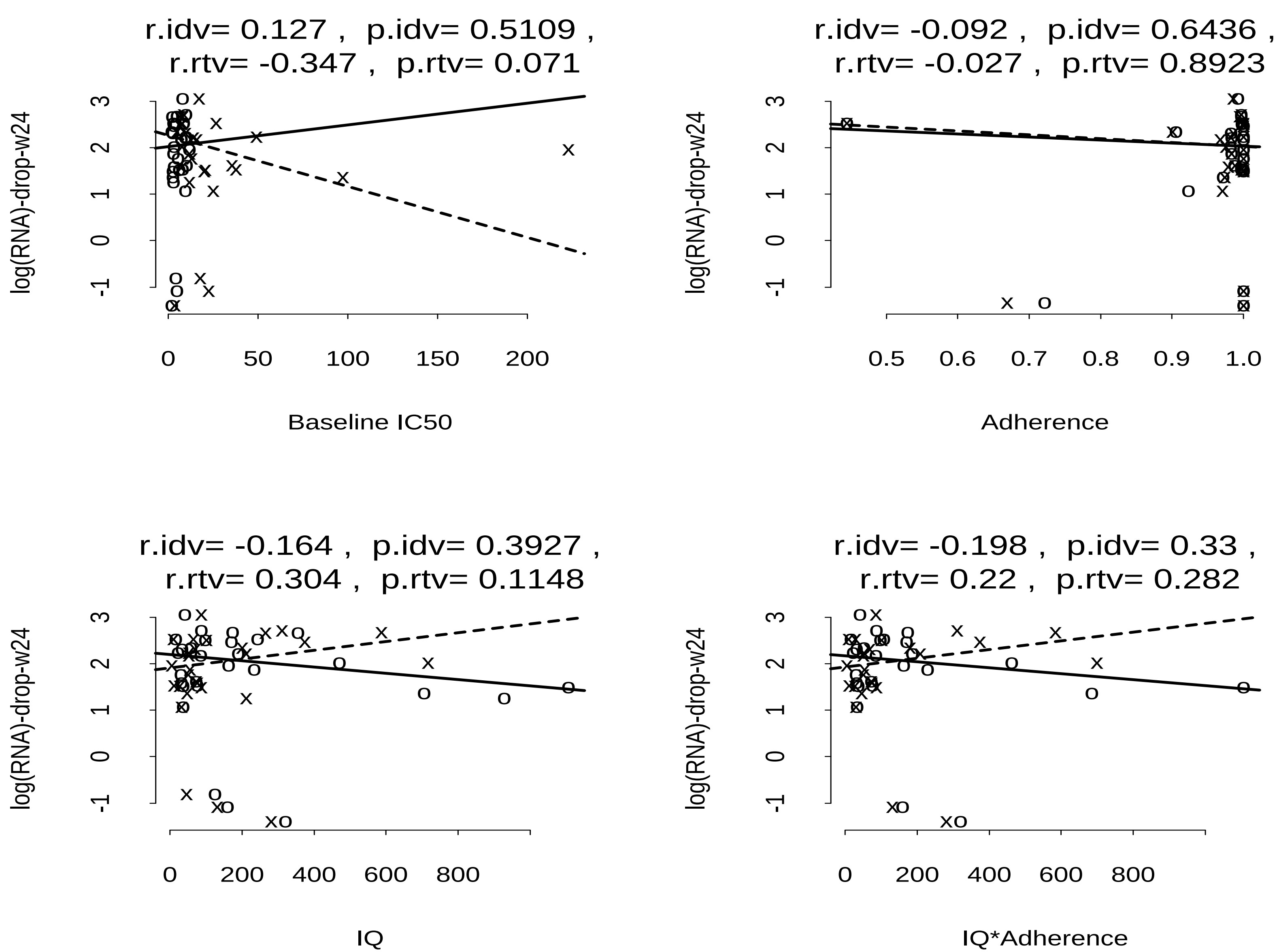


Figure 2: Simple Regression or Correlation Analyses



Mechanisms-Based Model Fitting

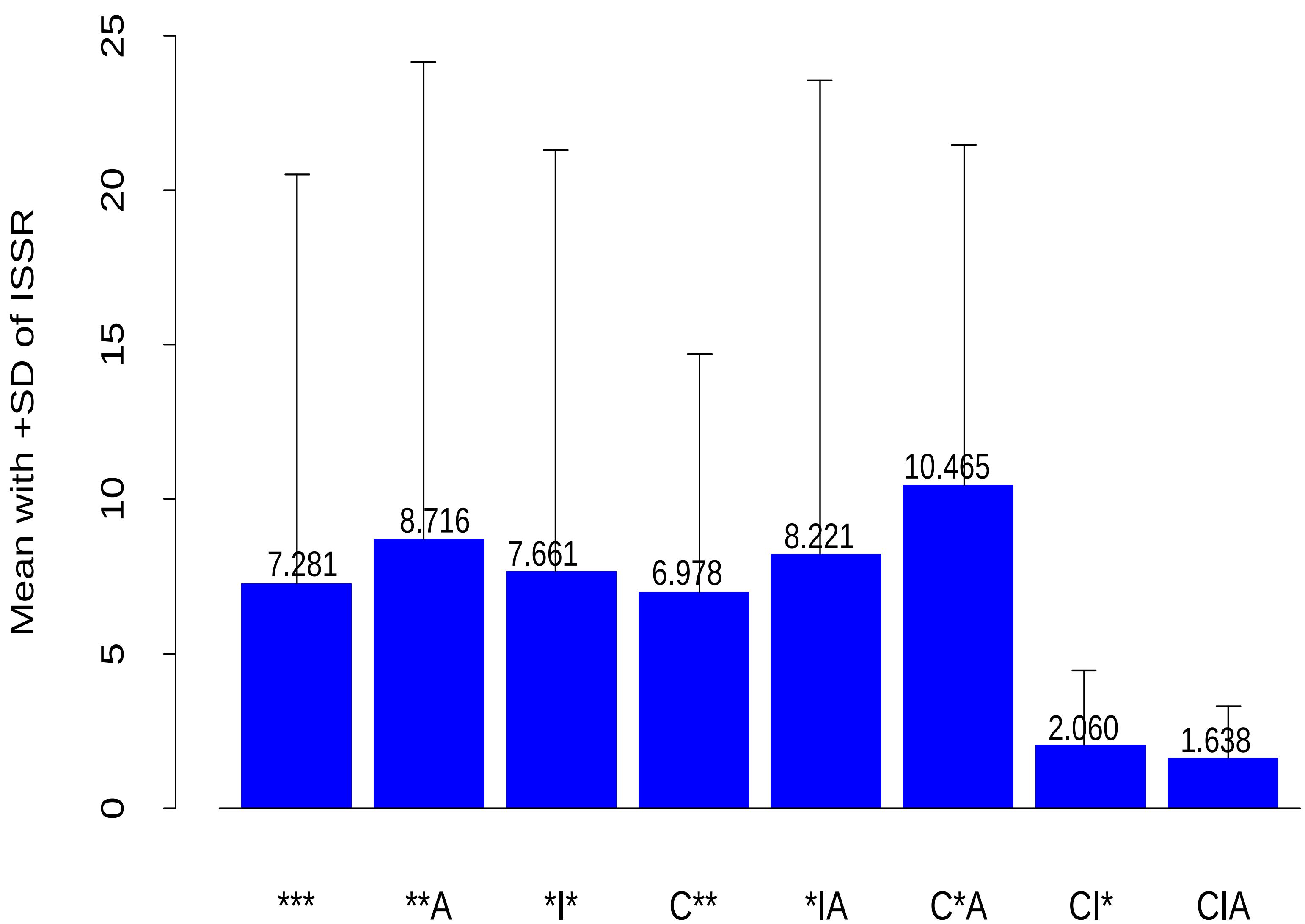
Three Factors: (1) PK, (2) adherence and (3) drug susceptibility

- **No factor considered**
- **Considering each of the three factors separately**
- **Considering all pairs of two-factor combinations**
- **Considering all three factors together**

Mechanisms-Based Model Fitting

- **Fit the data from all patients (Bayesian model)**
- **Get sum of squared residuals (SSR) from each patient**
- **Use the SSR from all inidividuals (ISSR) to compare model fittings**
- **The smallest ISSR is the best model**

Figure 3: **Mechanism Model Comparisons**



Mechanism Model Comparisons: p-values

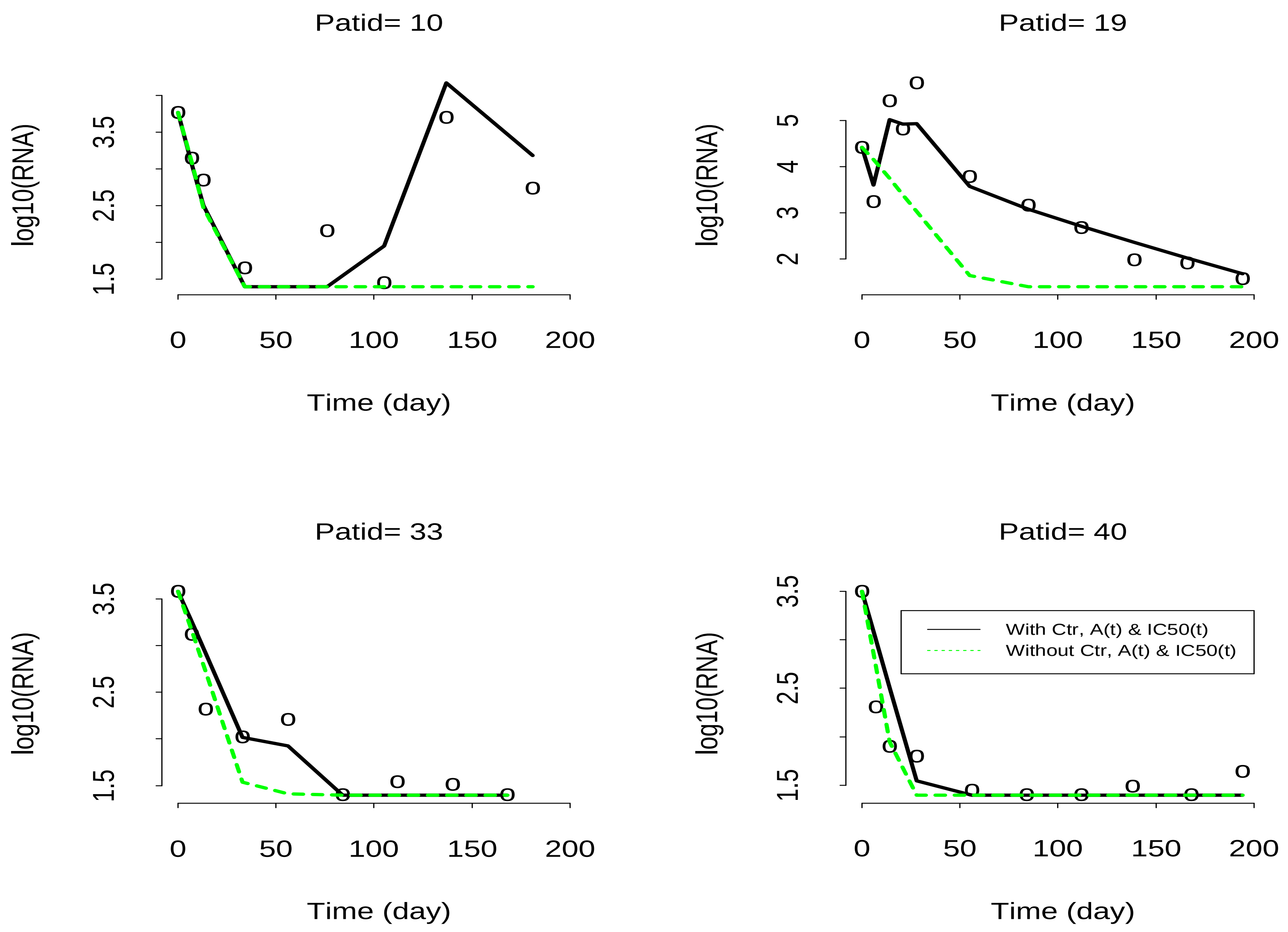
		***	**A	*I*	C**	*IA	C*A	CI*
	**A	0.1001						
	I	1.0000	0.0136					
	C**	0.7576	0.3545	0.758				
P	*IA	1.0000	0.0079	0.086	0.7576			
	C*A	0.0641	1.0000	0.12	0.0641	0.014		
	CI*	0.0136	<0.0001	0.019	0.0136	0.042	0.0007	
	CIA	0.0055	0.0001	0.031	0.0002	0.014	<0.0001	0.22

C: Drug Concetration (PK)

I: Drug susceptibility (IC50)

A: Adherence

Figure 4: Model Comparisons: No factor vs. 3 factors



Conclusions

C–PK, I–IC50, A–Adherence

- **PK and Drug susceptibility: Important**
  - **IA significantly better than A**
  - **I significantly better than A**
  - **CI: almost better than all others**
- **Adherence: No effect**
  - **CIA not significantly better than CI**
  - **IA's SSR larger than I**
  - **CA's SSR larger than C**
  - **A's SSR larger than that with no factor considered**
    Data quality problem? More noise or more signal?
- **IA significantly better than CA: I and A more independet**

Summary

Developed HIV Dynamic models by considering long-term antiviral treatment with the following factors

- **Drug efficacy**
- **Drug concentration**
- **Drug susceptibility**
- **Adherence**

Summary

HIV Dynamic models: Powerful to show a significant relationship between the above factors and response

- **Simple regression or correlation methods: failed to detect the effect**
- **Dynamic modeling method: more powerful because**
  - **More information used: biological mechanism theories, prior information and current data**
  - **The whole viral load trajectory used as the response**
  - **Complicated nonlinear relationship between the drug factors and antiviral response captured appropriately**
  - **Complicated nonlinear interactions among the factors captured appropriately**

Discussion and Open Problems

- **Data-Driven Parametric Models**
  - A model is selected after looking at the data
  - A linear or nonlinear functions available to fit the data
  - Good for predictions and interpretations
- **Data-Driven Nonparametric Models**
  - More flexible to fit complicated data patterns and robust against model assumptions
  - Not good for predictions and interpretations

Discussion and Open Problems

Mechanisms-Based Parametric Models

- **Advantages:**
  - The model can be determined before data collection
  - Biomedical mechanisms or physical laws: efficiently used
  - Great for predictions and interpretations
- **Drawbacks:**
  - Not robust to model assumptions
  - Well established biological theories and their mathematical representations required
- **More Work Needed:**
  - More statistical research needed for model identification
  - Apply the established models for AIDS clinical trial simulations and search for optimal treatment strategies

Acknowledgments

- **Dr. Yangxin Huang, U of R**
- **Drs. John G. Gerber and Edward P. Acosta: A5055 Co-Chairs**
- **A5055 Team Members**
- **ACTG DACS 210 Team Members**