Dose Individualization for Cancer Chemotherapy

Gary Rosner

Department of Biostatistics

U.T. M.D. Anderson Cancer Ctr

PODE Workshop June 2009

The Motivating Problem

- **Dr. Borje Andersson @ MDACC**
  - **Ultra high-dose therapy to treat leukemia**
    - **Dose too high?**
      - Fatal toxicity
    - **Dose too low?**
      - Less chance for benefit
    - **Solution**
      - Choose “optimal” dose

Bayesian Modeling

- Incorporates prior info
  - Previous studies
  - "Expert" opinion
- Predict & estimate precision
  - Based on current knowledge
    - Can incorporate information from outside of current study

Optimal Individualized Dose

- Jelliffe et al.
- D'Argenio & Rodman
- Wakefield
  - Constant & quadratic loss functions
- Sandstrom, Karlsson, Ljungman, et al.
  - Individualization of oral busulfan

Back to Transplant Study

- **Have previous studies**
  - 1/day dosing & 2/day dosing
- **Current study**
  - 12 mg i.v. test dose
  - Uniform high dose (mg/m2)
- **Future study**
  - Test dose ⇒ PK ⇒ “Optimal”

What We Would Like To Do

- **Combine**
  - **Meta-analysis of PK from earlier studies**
  - **PK of test dose for new patient**
- **Predict PK for new patient**
  - **Determine optimal dose**
    - **Bayesian Decision Theory**

Data

Study	Pt. #	Dose	Time	Conc	Covar
1	1	HIGH	$t_{1,1,1},\ldots$	$y_{1,1,1},\ldots$	$x_{1,1},\ldots$
…	…	…			
1	i	HIGH	$t_{1,i,j},\ldots$	$y_{1,i,j},\ldots$	$x_{1,i},\ldots$
…	…	…			
1	n_1	HIGH	$t_{1,n1,j},\ldots$	$y_{1,n1,1},\ldots$	$x_{1n1},\ldots$
2	1	low	$t_{2,1,1},\ldots$	$y^*_{2,1,1},\ldots$	$x_{2,1},\ldots$
2	1	HIGH	$t_{2,1,1},\ldots$	$y_{2,1,1},\ldots$	$x_{2,1},\ldots$
…	…	…			
2	i	low	$t_{2,i,1},\ldots$	$y^*_{2,i,1},\ldots$	$x_{2,i},\ldots$
2	i	HIGH	$t_{2,i,1},\ldots$	$y_{2,i,1},\ldots$	$x_{2,i},\ldots$
…	…	…			
2	n_2	low	$t_{2,n2,1},\ldots$	$y^*_{2,n2,1},\ldots$	$x_{2,i,n2},\ldots$
2	n_2	HIGH	$t_{2,n2,1},\ldots$	$y_{2,n2,j},\ldots$	$x_{2,n2},\ldots$
3	1	low	$t_{3,1,1},\ldots$	$y^*_{3,1,1},\ldots$	$x_{3,1},\ldots$

Sampling Distribution
Nonlinear Mean Function

- **For pt. i, in study s, at time t_{sij}**

$$\begin{aligned} y_{s,i,j} &= \log\left[\mathrm{Conc}(t_{s,i,j})\right] \\ &= \log\left[f(t_{s,i,j})\right] + error \end{aligned}$$

Population Model as Hierarchical Model

- **Observation**

$$p\left(y_{s,i,j} \mid \theta_i\right), s = 1, \ldots, S; j = 1, \ldots, n_i$$

- **Subject-specific parameters**

$$p\left(\theta_i \mid \phi\right), i = 1, \ldots, I$$

- **Population**

$$p(\phi)$$

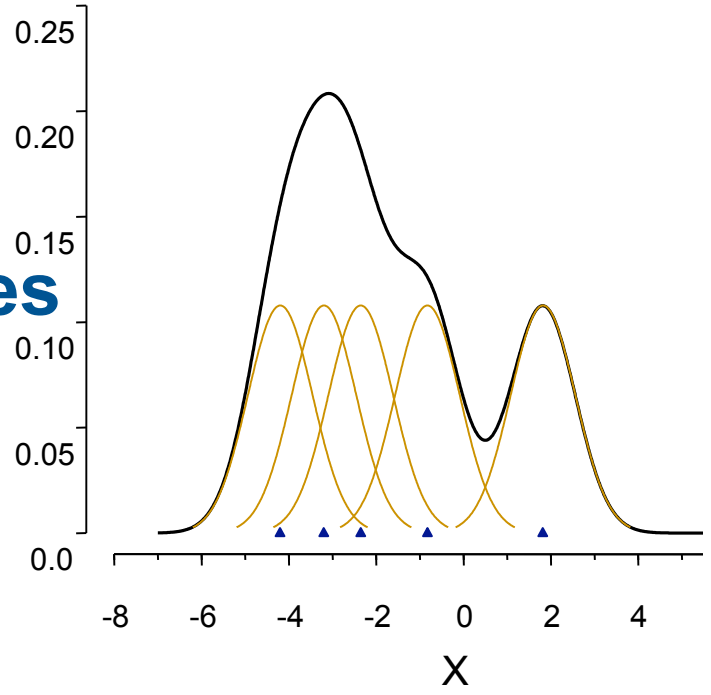
Mixture Models for Population Distribution

- **Distribution of subject-specific parameters:** $p(\theta_i \mid \phi), i = 1, \ldots, I$
  - **Characterize as a mixture of simpler distributions**
    - **Weighted sum of densities**

$$p(\theta_i \mid \phi) = \text{Mixture of MV Normals} = \sum_k w_k N(\mu_k, V)$$

Why Mixture?

- **Flexible: Depict Many Shapes**
  - **Heterogeneity:**
    - **patient characteristics**
    - **genetic diversity**
- **Example:**
  - **Mixtures of normal densities (kernels)**



Dirichlet Process (DP)
Mixture of Normals

- **Dirichlet Process is distribution on space of distributions**
- **DP has 2 parameters**
- G_0 **is base measure**
- M **is total mass parameter**

$$\mu_i \sim G$$

$$G \sim DP(G_0, M) \quad i.e., G = \sum_{h=1}^{\infty} w_h \delta(\mu_h)$$

$$G_0(\mu) = N(b, B)$$

Hierarchical Model: Within Study

$$\log\left(y_{ij}\right) = \log\left[f\left(\theta_i, t_{ij}\right)\right] + \varepsilon_{ij}$$

$$\varepsilon_{ij} \sim N\left(0, \tau^{-1}\right), \quad \tau \sim Ga\left(\alpha_\tau, \beta_\tau\right)$$

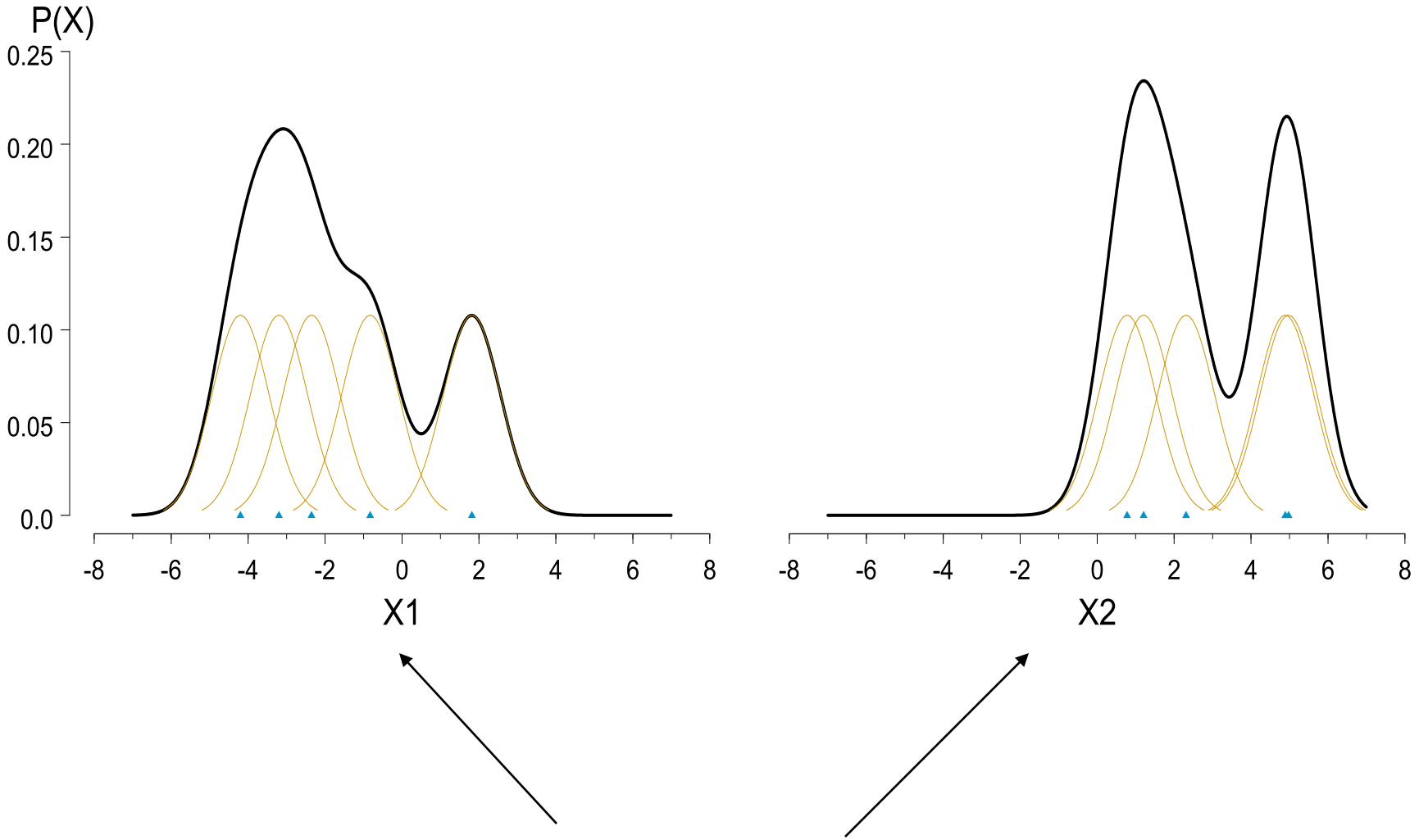
$$\left.\begin{array}{l} \theta_i \sim N\left(\mu_i, S\right) \\ \mu_i \sim G \end{array}\right\} \Rightarrow \theta_i \sim \sum_{h=1}^{\infty} w_h N\left(\mu_h, S\right)$$

$$G \sim DP\left(G_0, m\right) \quad i.e., G = \sum_{h=1}^{\infty} w_h \delta\left(\mu_h\right)$$

Multiple Studies
Hyperprior for Meta-Analysis

- **Straightforward if parametric hyperprior (random effects)**
  - **Parametric model provides structure**

- **Complex if mixture hyperprior**
  - **Link studies at what level?**
  - **How enforce linkage?**



How Model Commonality?

Want to Share More Info From Prior Studies

- **Dependent Dirichlet Process**
  - **ANOVA-like structure for means**
    - **Categorical covariates:** $x = (v, w)$
      - Locations linked across factor levels of x via ANOVA model

$$G_x(\phi) = \sum w_h \delta(m_{xh}),$$

$$m_{xh} = \mu_h + A_{v,h} + B_{w,h}$$

Dependence Structure
DP Mixture of ANOVA (categorical)

- **Across levels of x (C levels)**

$$x = x_1, x_2, \ldots, x_C$$

$$x = x_1: \quad G_{x_1} = \omega_1 \delta(m_{11}) + \omega_2 \delta(m_{12}) + \omega_3 \delta(m_{13}) + \ldots$$

$$x = x_2: \quad G_{x_2} = \omega_1 \delta(m_{21}) + \omega_2 \delta(m_{22}) + \omega_3 \delta(m_{23}) + \ldots$$

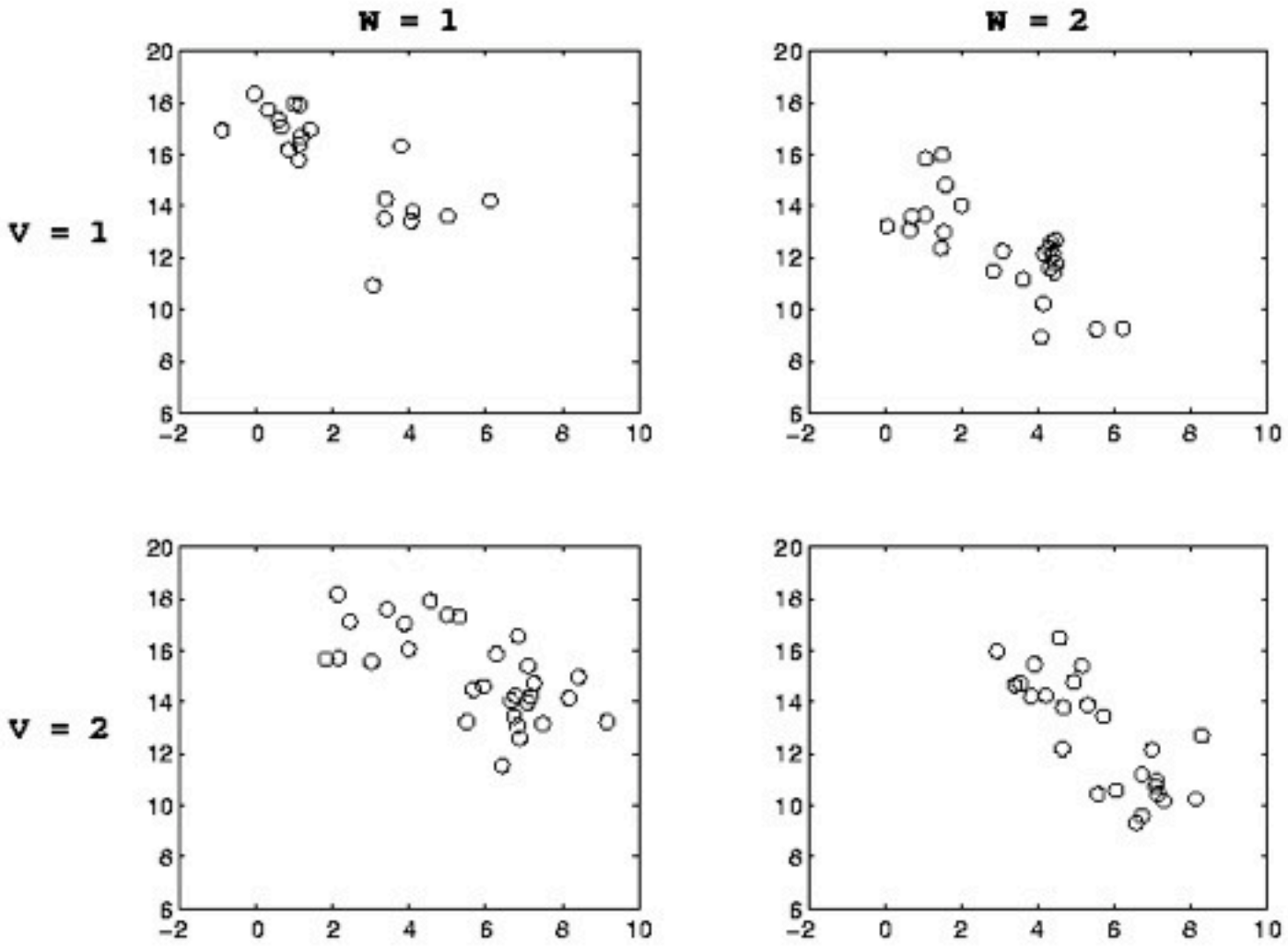
$$x = x_3: \quad G_{x_3} = \omega_1 \delta(m_{31}) + \omega_2 \delta(m_{32}) + \omega_3 \delta(m_{33}) + \ldots$$

$\ldots$

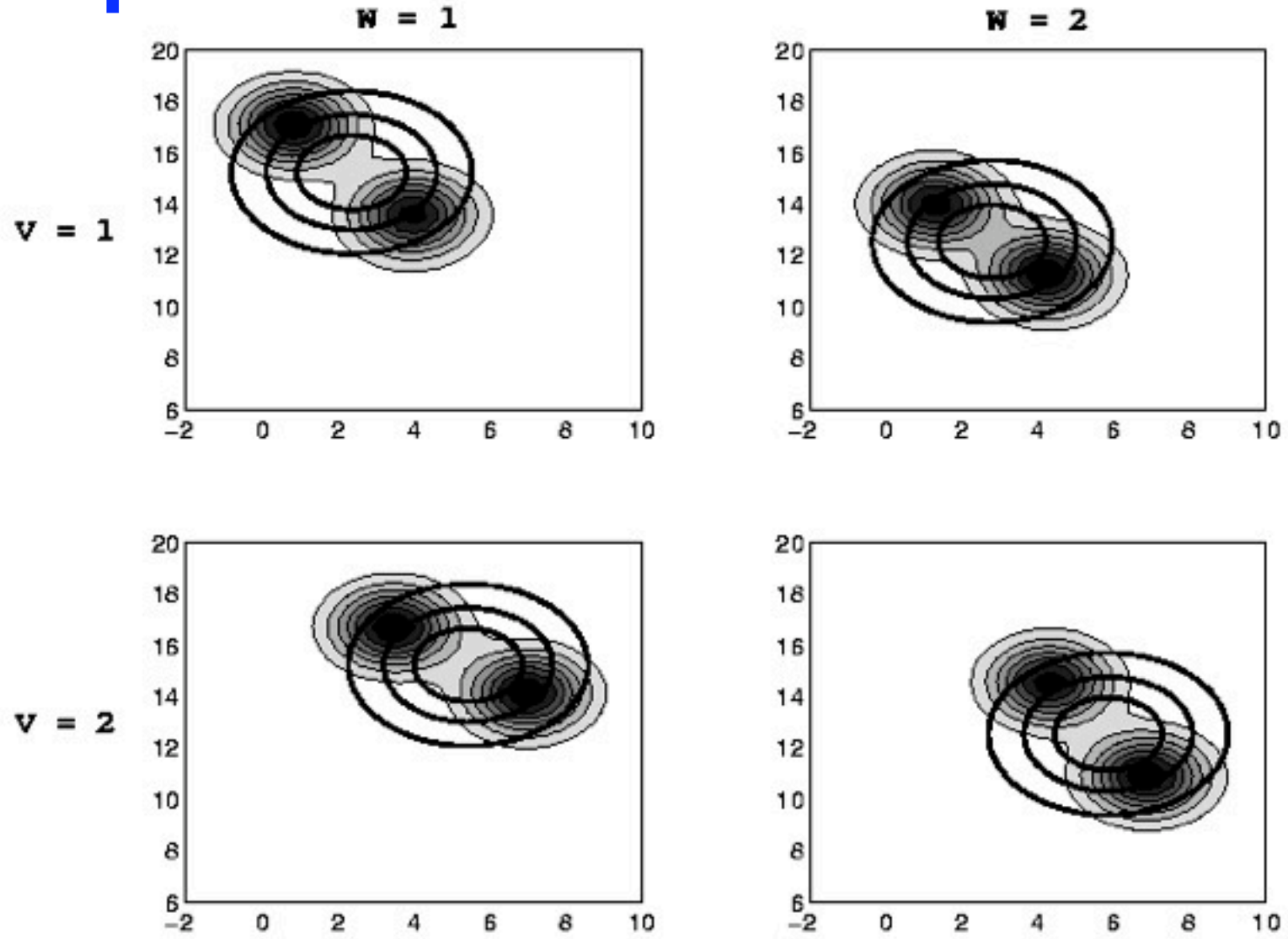
- **Locations are functions of x**

$$m_{xh} = \mu_h + A_{v,h} + B_{w,h}$$

2 Factors, 2 Levels Each:



Compare Mixture to ML Estimation



Covariates

- **Dependent Dirichlet Process (DDP)**
  - **Decompose locations as sums of random measures**
- **Categorical covariates**
  - **ANOVA DDP**
    - **Two factors v and w:** $x = (v, w)$

$$m_{xh} = \mu_h + A_{v,h} + B_{w,h}$$

- **Categorical & continuous covariates**
  - **Linear DDP**
    - **One categorical & 1 continuous** $x = (v, z)$

$$m_{xh} = \mu_h + A_{vh} + \beta_h z$$

Back to Transplant Study

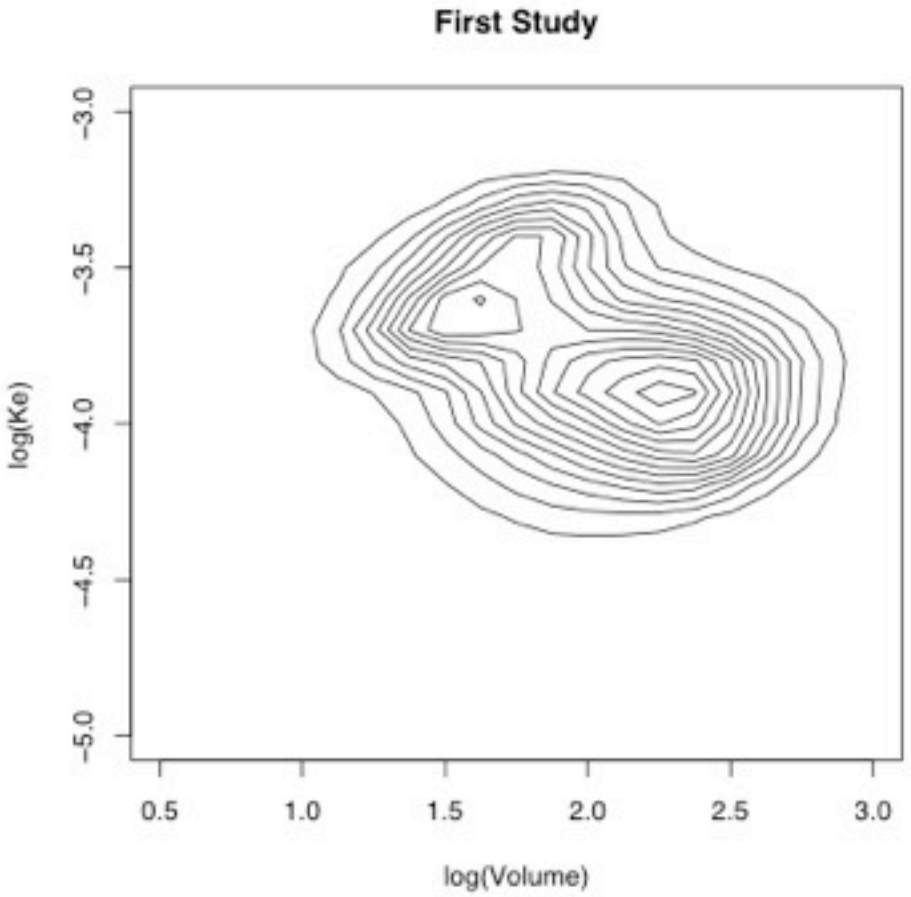
- **Multiple studies**
  - **Hierarchical extension**
    - G **is 2x3 matrix-variate dist'n**

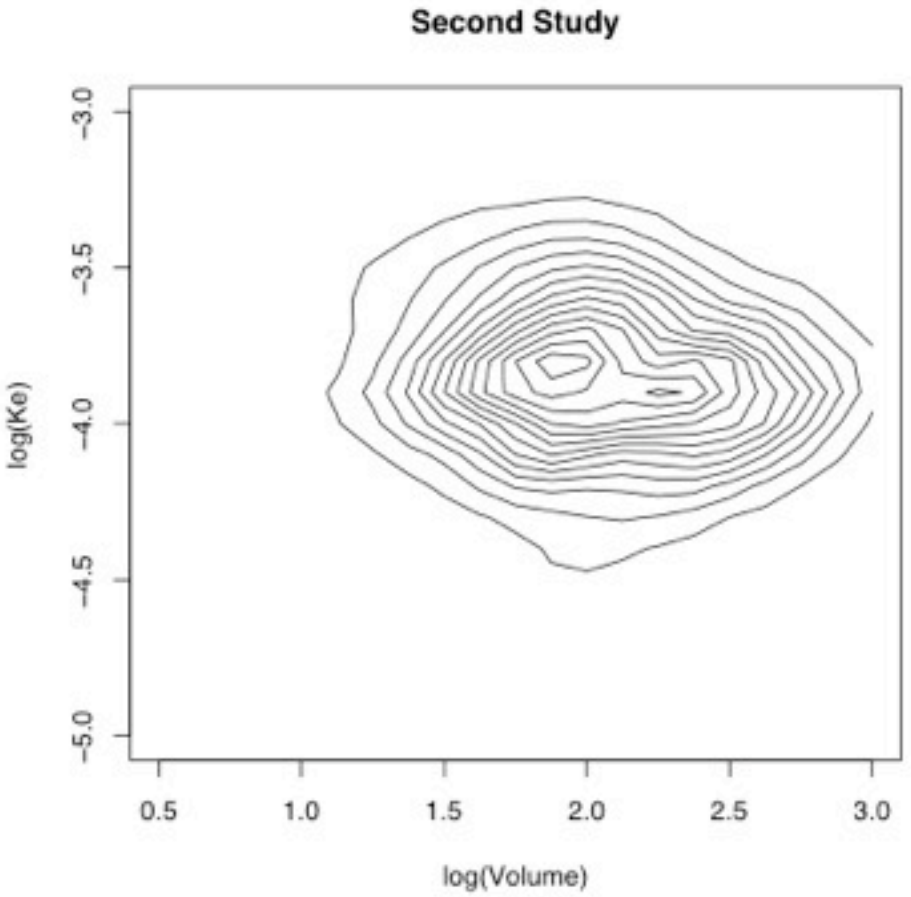
$$\theta_{ki} \sim \int N(\mu_{ki}, S) dG_{ki}(\mu_{ki}), k = 1,2,3$$

$$A \sim G$$

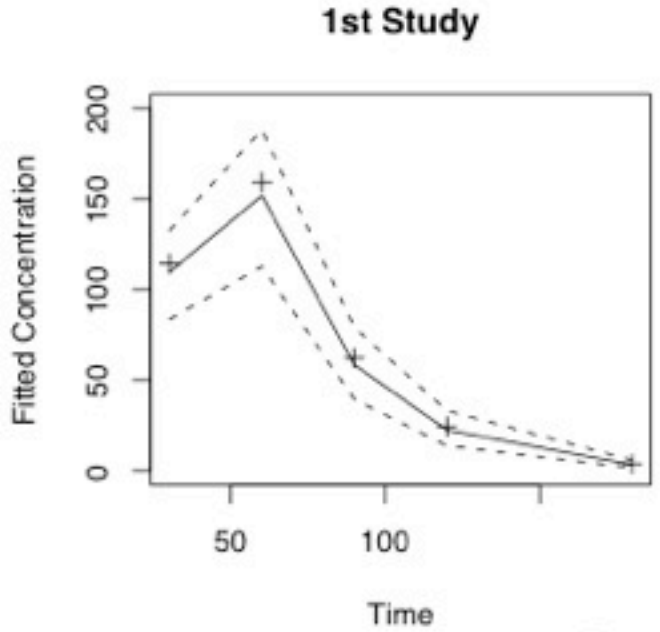
$$A = [m, ST_2, ST_3] = \begin{array}{c} \begin{array}{ccc} \text{Main} & \text{Study2} & \text{Study3} \end{array} \\ \begin{pmatrix} \log V_0 & \log V_2 & \log V_3 \\ \log k_0 & \log k_2 & \log k_3 \end{pmatrix} \end{array}$$

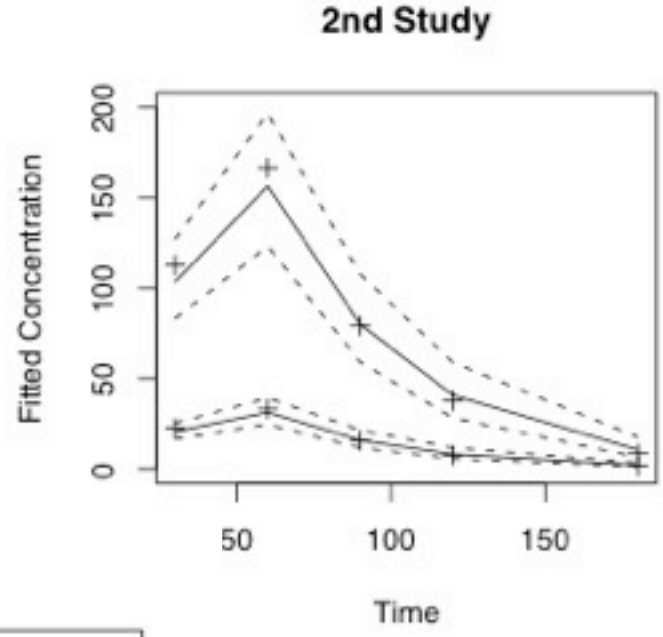
Have 2 Studies

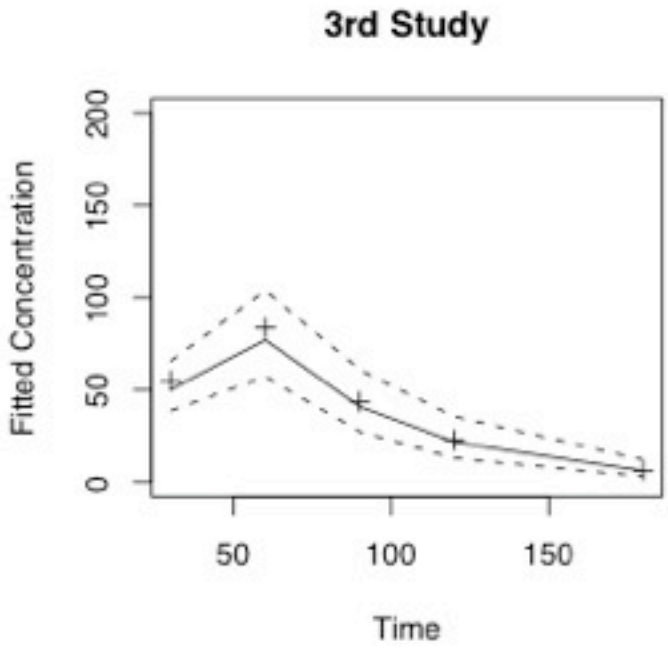




Fitted Profiles







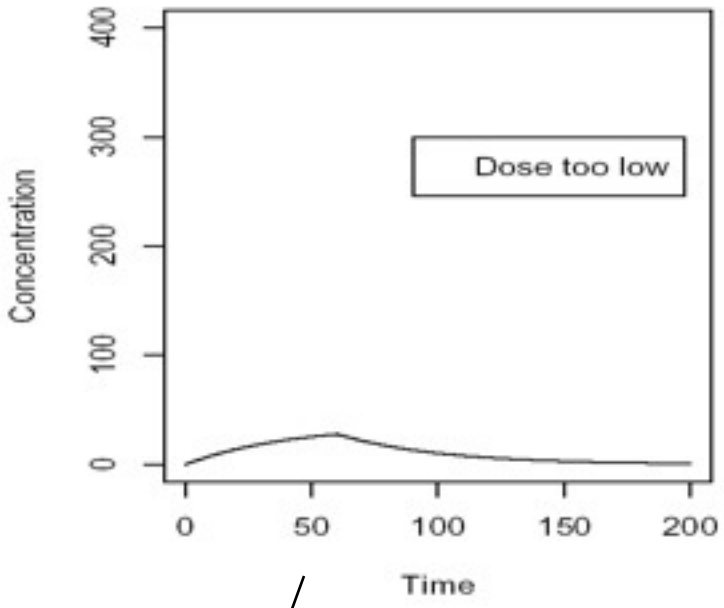
Incorporate Test Dose

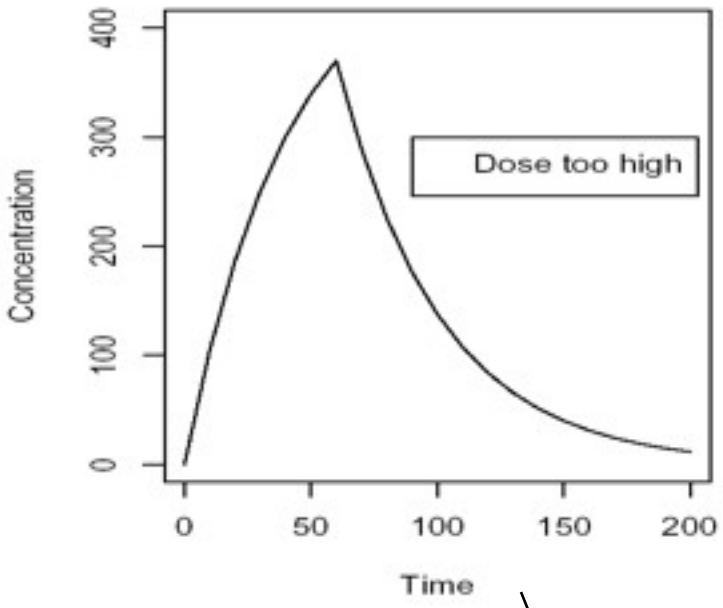
Back to Transplant Example

- Have
  - Model for drug's PK
  - Probability model for data

- Need utility function to optimize

Dose Optimization





AUC too low!

AUC too high!

Too Low | **TARGET** | Too High

Bayesian Optimal Design

- **Let** $u(y,d,\theta) = L(AUC),\ AUC = f(d,V,k)$

- **Pick dose d* that minimizes**

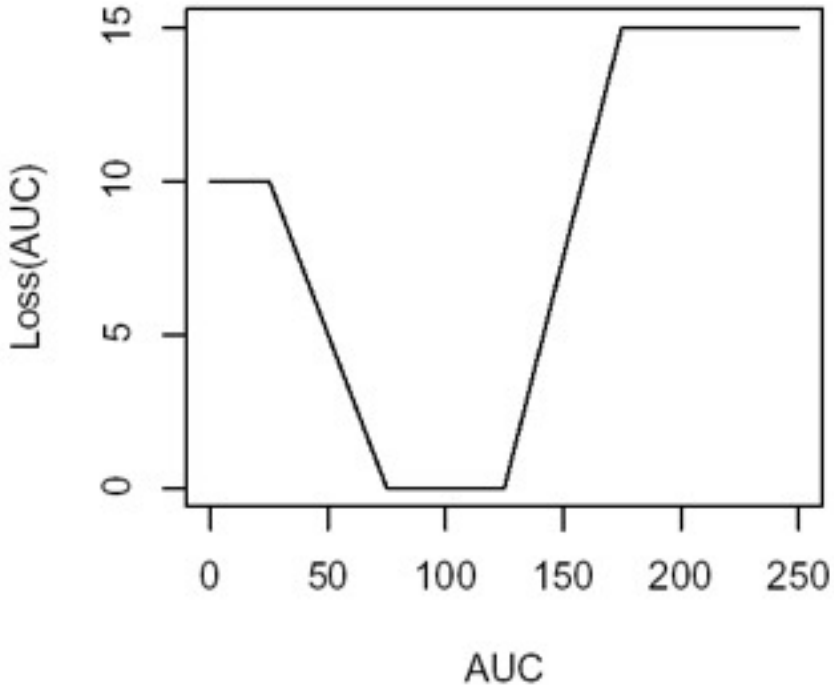
$$E\left[u(y,d,\theta)\right] = \iint u(y,d,\theta)\, p_d(y,\theta)\, dy\, d\theta$$

- **With data:**
  - **Studies 1 & 2 plus new patient's low dose data**

$$E\left[u(y,d,\theta)\right] = \int L\left[AUC(\theta,d)\right] p(\theta|D)\, d\theta$$

Asymmetric Loss Function

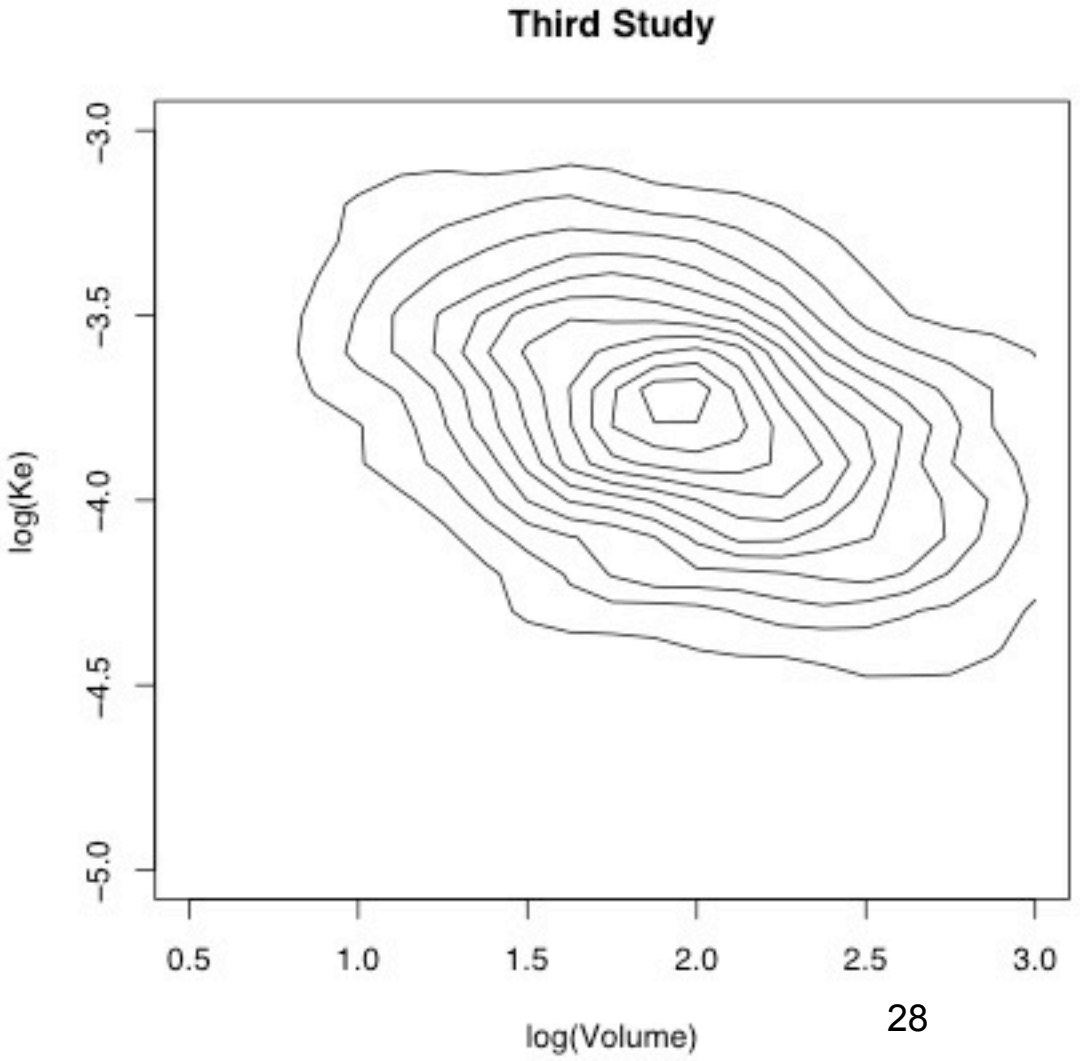
- **Want AUC in "optimal" range**



- **Loss function**

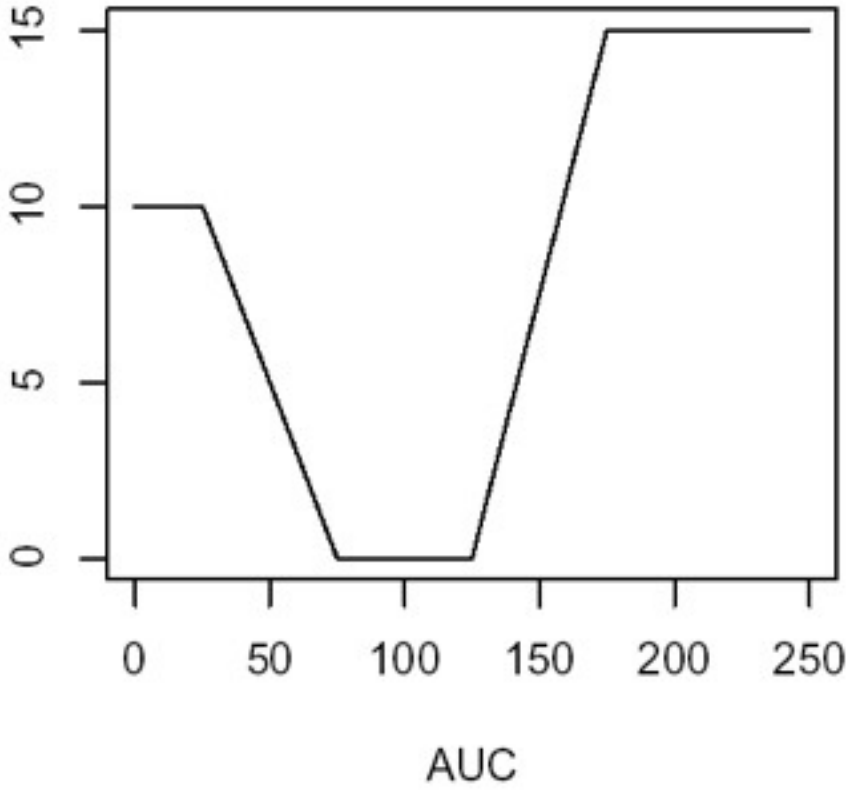
$$L(auc) = \begin{cases} L^{-}(auc, AUC_{ll}) & \text{if } auc < AUC_{ll} \\ 0 & \text{if } AUC_{ll} < auc < AUC_{ul} \\ L^{+}(auc, AUC_{ul}) & \text{if } auc > AUC_{ul} \end{cases}$$

For Study 3: Posterior for this patient's PK parameters

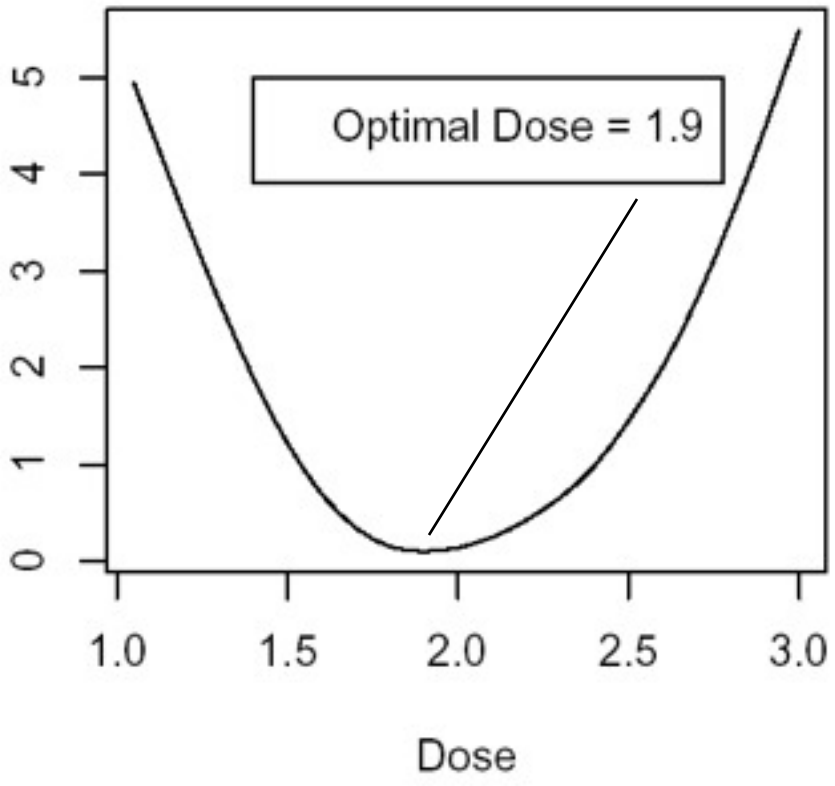


Optimal Dose w.r.t. Posterior

Loss Function



Expected Loss



Summary

- Mixture models
  - flexible inference
- DDP structure
  - Categorical and continuous covariates
- Optimal design for PK

- Currently evaluating our strategy against covariate-based dosing

Спасибо