Bayesian Approaches for Benefit-Risk Assessment with Examples

Ram Tiwari and Chul Ahn
Division of Biostatistics
Center for Device and Radiological Health
Food and Drug Administration

2018 Joint Statistical Meetings
Vancouver Convention Center, CC East 17
Monday 7/30/2018, 10:30 AM- 12:20 PM

Acknowledgements

This presentation is based on collaboration with:

- Shiqi Cui, Graduate Student, University of Missouri
- Yueqin Zhao, PhD: Mathematical Statistician, DB-7, OB, CDER
- Jyoti Zalkikar, PhD: Team Leader, DB-5, OB, CDER
- Lisa LaVange, PhD: Director, OB, CDER

Papers:

- Zhao, Y., Zalkikar, J., Tiwari, R. C., and LaVange, L. M. (2014), A Bayesian approach for benefit-risk assessment, Statistics in Biopharmaceutical Research, online.
- Zhao, Y., Cui, s., and Tiwari, R.C. (2016), Bayesian approach to personalized benefit-risk assessment, Statistics in Biopharmaceutical Research, online

Outline of the talk

- Introduction
  - Some commonly-used Benefit-risk (BR) measures
- Data Format: Benefit-Risk Categories
  - Hydromorphone clinical trial data
- BR measures based on Global benefit-risk scores
- Bayesian approach based on multinomial-Dirichlet conjugate prior
  - Sequential updating
  - Power Prior
  - LPML model selection
- Bayesian approach based on Dirichlet process
- Conclusion

INTRODUCTION

- The benefit-risk assessment is the basis of regulatory decisions in the pre-market and post-market review processes.

- The evaluation of benefit and risk faces several challenges.
  - Benefit and risk are usually not measured on the same scales;
  - Patients withdrawal are not rare and usually due to many reasons;
  - Benefit and risk are not independent;
  - The benefit/risk trade-off may change over the course of a clinical trial.

COMMONLY USED BR MEASURES

- Benefit-risk ratio
- Time without symptoms of disease and toxic effects (TWiST)
- Ratio of number needed to treat for benefits (NNT(B)) to number needed to treat for risks (NNT(R))
- Multiple criteria decision analysis (MCDA)
- Integrated benefit-risk measure (MDIC)
- Global benefit-risk scores -- Chuang-Stein et al. (1991)

DATA FORMAT: BR CATEGORIES

- Chuang-Stein et al. (1991): the outcomes of a clinical trial can be classified into five mutually exclusive categories.

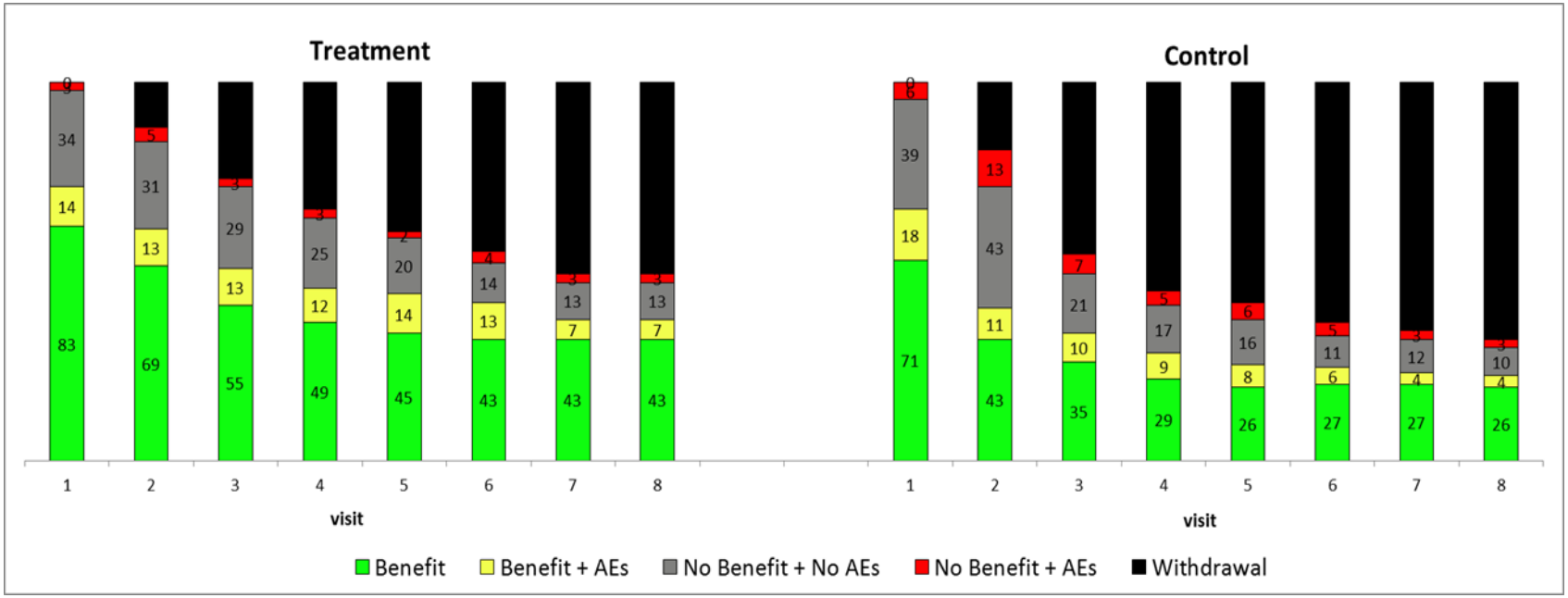
Table: Outcomes of a clinical trial with binary response data.

	Benefit	No benefit
No adverse event	Category 1	Category 3
Adverse event	Category 2	Category 4
Withdrawal	Category 5	

SUBJECT-LEVEL OUTCOMES OF HYDROMORPHONE TRIAL DATA

- 2 arms (treatment and control), 5 categories
- 134 individuals at 8 visits on each arm

EXAMPLE 1: HYDROMORPHONE DATA*



*Data was provided by Jonathan Norton.

SUBJECT-LEVEL RESPONSE ON TREATMENT ARM

Subject	Visit							
	1	2	3	4	5	6	7	8
1	1	1	1	1	1	1	1	1
2	1	1	1	1	1	2	5	5
…	…							
134	4	5	5	5	5	5	5	5

AGGREGATED DATA BY VISITS ON TREATMENT ARM

Category	Visit							
	1	2	3	4	5	6	7	8
1	82	65	58	50	48	43	43	44
2	15	12	10	11	15	15	8	7
3	25	30	23	20	20	15	15	17
4	12	5	3	4	2	4	3	3
5	0	22	39	49	49	57	65	63

CHUANG-STEIN ET AL. (1991): GLOBAL BR SCORES

$$BRScore\_Linear = \sum_{i=1}^{2} w_i p_i - \sum_{i=3}^{5} w_i p_i$$

$$BRScore\_Ratio = \frac{(\sum_{i=1}^{2} w_i p_i)^e}{\sum_{i=3}^{5} w_i p_i}$$

$$BRScore\_Cmp\_Ratio = \frac{w_1 p_1}{w_5 p_5} \left(\frac{w_2 p_2}{w_3 p_3 + w_4 p_4}\right)^f$$

$$where \quad w_1 = 2, \quad w_2 = 1, \quad w_3 = 0, w_4 = 1, \quad w_5 = 2$$

BR MEASURES

- BR measures based on the global scores proposed by Chuang-Stein et al. (1991)

$$BR\_Linear = (\sum_{i=1}^{2} w_i p_{i,T} - \sum_{i=3}^{4} w_i p_{i,T}) - (\sum_{i=1}^{2} w_i p_{i,C} - \sum_{i=3}^{4} w_i p_{i,C})$$

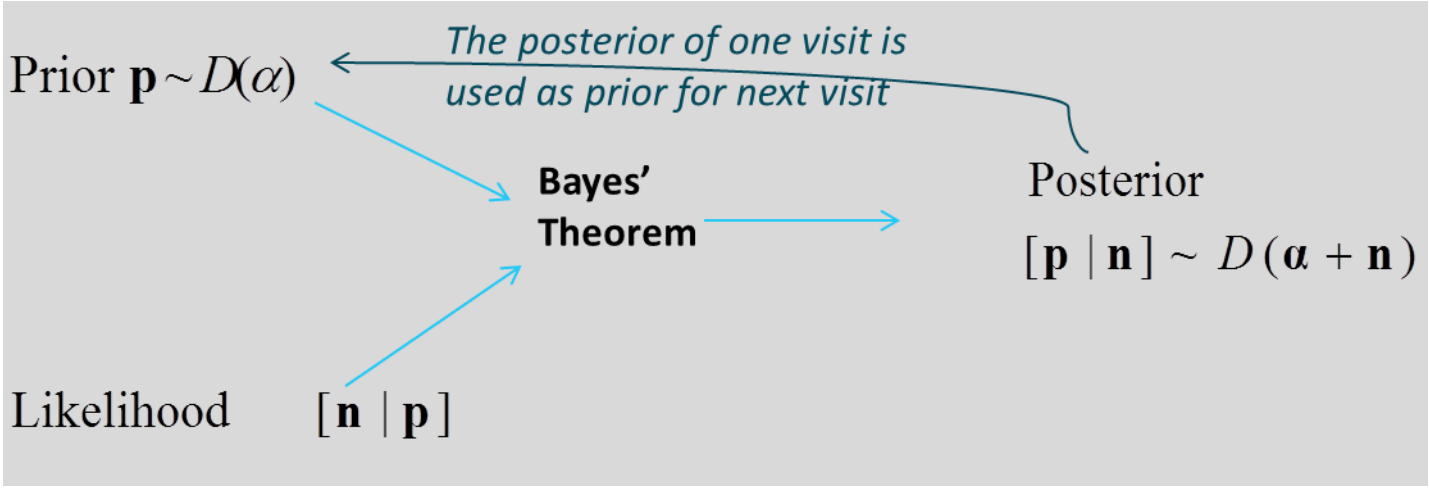
$$BR\_Ratio1 = \log(\frac{(\sum_{i=1}^{2} w_i p_{i,T})^e}{\sum_{i=3}^{4} w_i p_{i,T}}) - \log(\frac{(\sum_{i=1}^{2} w_i p_{i,C})^e}{\sum_{i=3}^{4} w_i p_{i,C}})$$

$$BR\_Ratio2 = \log(\frac{(\sum_{i=1}^{2} w_i p_{i,T})^e}{w_2 p_{2,T} + w_4 p_{4,T}}) - \log(\frac{(\sum_{i=1}^{2} w_i p_{i,C})^e}{w_2 p_{2,C} + w_4 p_{4,C}})$$

- BR measures based on the global scores are for each arm (treatment and comparator) separately.

BAYESIAN APPROACH: USE OF DIRICHLET DISTRIBUTION AS A CONJUGATE PRIOR

- Dirichlet distribution is used as the conjugate prior for multinomial distribution, and the posterior distribution of the five-category random variable is derived at each visit using sequentially updated posterior as a prior.



SEQUENTIAL UPDATING

- Sequential updating of the posteriors are given by:

Visit 1: $[\mathbf{p} \mid \mathbf{n}_1] \propto [\mathbf{n}_1 \mid \mathbf{p}]^*[\mathbf{p}]$

Visit 2: $[\mathbf{p} \mid \mathbf{n}_1, \mathbf{n}_2] \propto [\mathbf{n}_2 \mid \mathbf{p}]^*[\mathbf{p} \mid \mathbf{n}_1]$

.....

Visit *M*: $[\mathbf{p} \mid \mathbf{n}_1, \mathbf{n}_2, \ldots, \mathbf{n}_M] \propto [\mathbf{n}_M \mid \mathbf{p}]^*[\mathbf{p} \mid \mathbf{n}_1, \mathbf{n}_2, \ldots, \mathbf{n}_{M-1}]$

- The posterior mean (i.e., Bayes estimate) and 95% credible interval for each of the four measures are obtained using a Markov chain Monte Carlo (MCMC) technique.

DECISION RULES

- For a BR measure,
  - If the credible interval include the value zero, the benefit does not outweigh the risk;
  - If the lower bound of the credible interval is greater than zero, the benefit outweighs the risk;
  - If the upper bound of the credible interval is less than zero, the risk outweighs the benefit.

USE OF POWER PRIOR

- Power prior (Ibrahim and Chen, 2000) is used through the likelihood function to discount the information from previous visits, and the posterior distribution of the five-category random variable is obtained using the Dirichlet prior for **p** and a Beta (1, 1) as a power prior for .

Visit 1: $[\mathbf{p} \mid \mathbf{n}_1] \propto [\mathbf{n}_1 \mid \mathbf{p}]^*[\mathbf{p}]$

Visit 2: $[\mathbf{p}, a_0 \mid \mathbf{n}_1, \mathbf{n}_2] \propto [\mathbf{n}_2 \mid \mathbf{p}]^*[\mathbf{n}_1 \mid \mathbf{p}]^{a_0} *[\mathbf{p}]^*[a_0]$

.....

Visit *M*: $[\mathbf{p}, a_0 \mid \mathbf{n}_1, \mathbf{n}_2, \ldots, \mathbf{n}_M] \propto [\mathbf{n}_M \mid \mathbf{p}]^*[\mathbf{n}_1, \mathbf{n}_2, \ldots, \mathbf{n}_{M-1} \mid \mathbf{p}]^{a_0} *[\mathbf{p}]^*[a_0]$

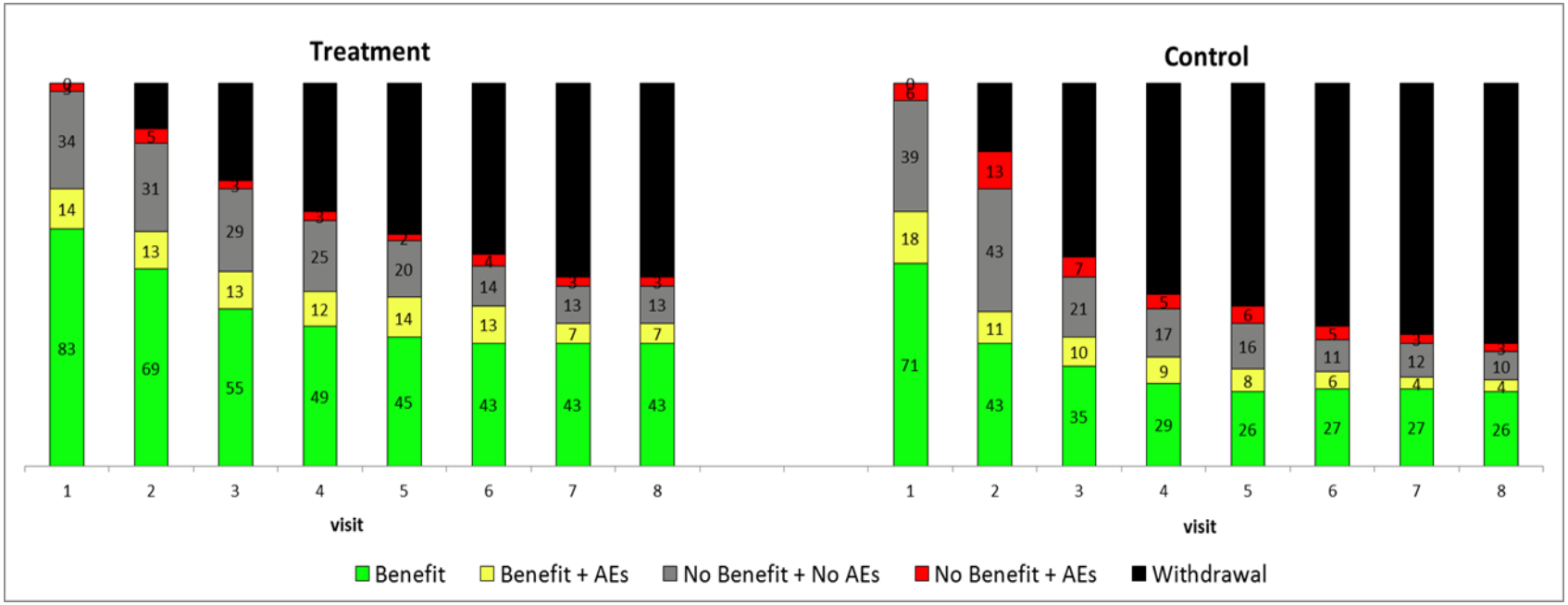
MODEL FIT

- The model fit of the two models (with and without power prior) is assessed through the conditional predictive ordinate (CPO) and the logarithm of the pseudo-marginal likelihood (LPML). The larger the value of LPML, the better fit the model is. Here, $\mathbf{n}_{(i)}$ is the data with $\mathbf{n}_i$ removed.

$$CPO_i = f(\mathbf{n}_i \mid \mathbf{n}_{(i)}) = \int [\mathbf{n}_i \mid \mathbf{p}, \mathbf{n}_{(i)}][p \mid \mathbf{n}_{(i)}] d\mathbf{p}$$

$$LPML = \sum_{i=1}^{5} \log f(\mathbf{n}_i \mid \mathbf{n}_{(i)}) = \sum_{i=1}^{5} \log CPO_i$$

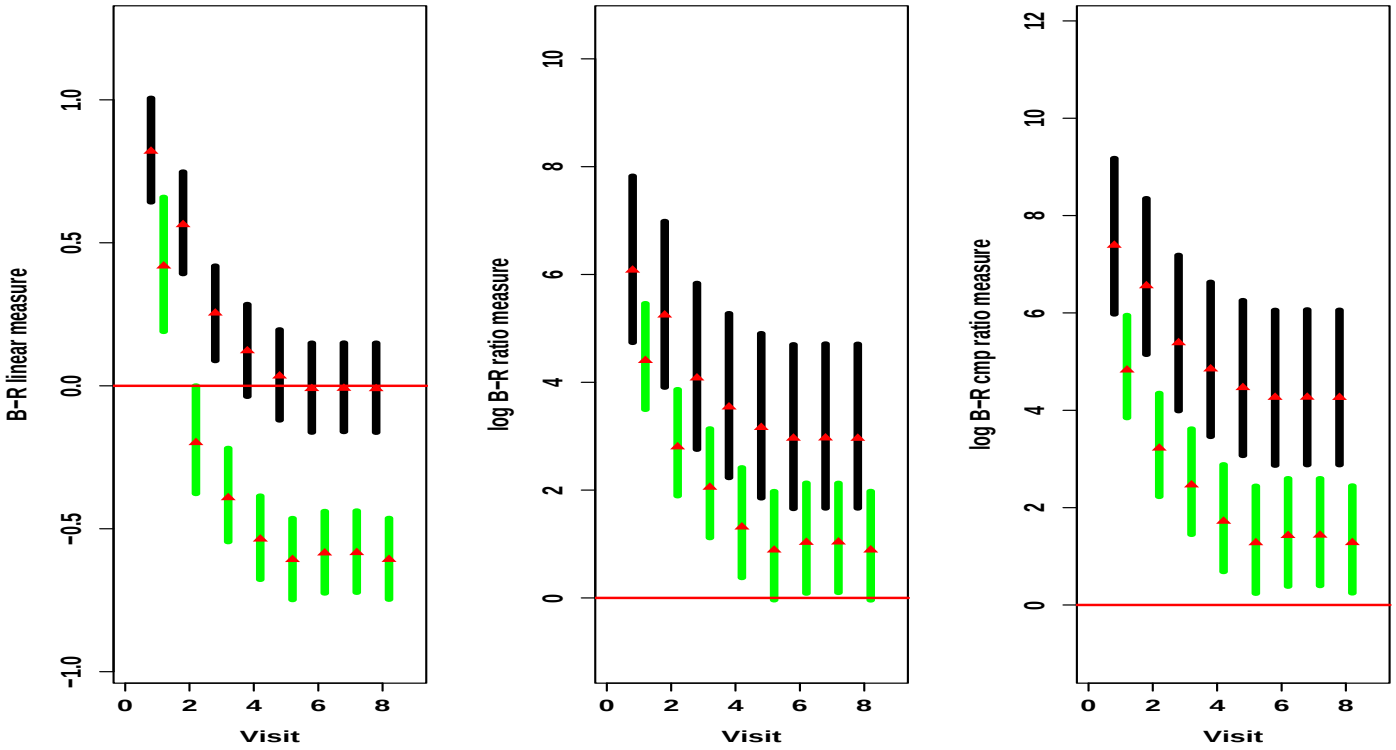
BACK TO OUR EXAMPLE 1: HYDROMORPHONE




FDA
U.S. FOOD & DRUG
ADMINISTRATION

AGGREGATED BR MEASURES

- Posterior means and posterior 95% credible intervals for BR-L, log(BR-R) and log(BR-CR) averaged over all subjects.



Red triangles indicates posterior means, black and green bars are credible intervals from treatment and control arms respectively.

TESTING EQUALITY OF BENEFIT-RISK MEASUREMENTS BETWEEN TREATMENT AND CONTROL ARMS

Table: 95% credible intervals of ABRD-L, ABRD-R and ABRD-CR at each visit.

	Visits							
	1	2	3	4	5	6	7	8
ABRD-L	(0.103, 0.695)	(0.499, 1.019)	(0.409, 0.874)	(0.438, 0.874)	(0.431, 0.857)	(0.363, 0.788)	(0.365, 0.787)	(0.386, 0.809)
ABRD-R	(-0.049, 3.612)	(0.726, 4.392)	(0.307, 3.981)	(0.506, 4.168)	(0.554, 4.229)	(0.207, 3.879)	(0.218, 3.883)	(0.362, 4.008)
ABRD-CR	(0.753, 4.594)	(1.505, 5.380)	(1.085, 4.975)	(1.298, 5.188)	(1.359, 5.254)	(1.007, 4.892)	(1.000, 4.889)	(1.153, 5.056)

There is significant evidence of increasing benefit-risk measurements on the treatment arm than the control arm.

PERSONALIZED BR ASSESSMENT: MOTIVATION

- Zhao et al. (2014) method
  - Works on the aggregated level
  - Ignore the subject-level response differences
  - Assume subjects are exchangeable at each visit
- The proposed model:
  - Incorporates subject-level effects
  - Uses Dirichlet process model to accommodate the similarity of subject responses
  - Provides subject-specific response probabilities and benefit-risk analysis

PERSONALIZED BR BAYESIAN MODEL

- A generalized linear model (McCullagh and Nelder, 1989) is used to define the log-odds for all categories with respect to the reference (the first category).
- Define the log-odds of the i-th subject's response falling into category j ($y_{ik} = j$) with respect to category 1, at visit k as

$$\eta_{ijk} = \log \frac{\Pr(y_{ik} = j)}{\Pr(y_{ik} = 1)} = \beta_{ij} + \gamma_k$$

$\boldsymbol{\beta_i} = (\beta_{i2}, , \ldots , \beta_{iJ})$: subject-level category effects ($i = 1, \ldots, N$)
γ_k ($k = 1, \ldots, K$): random longitudinal visit effects

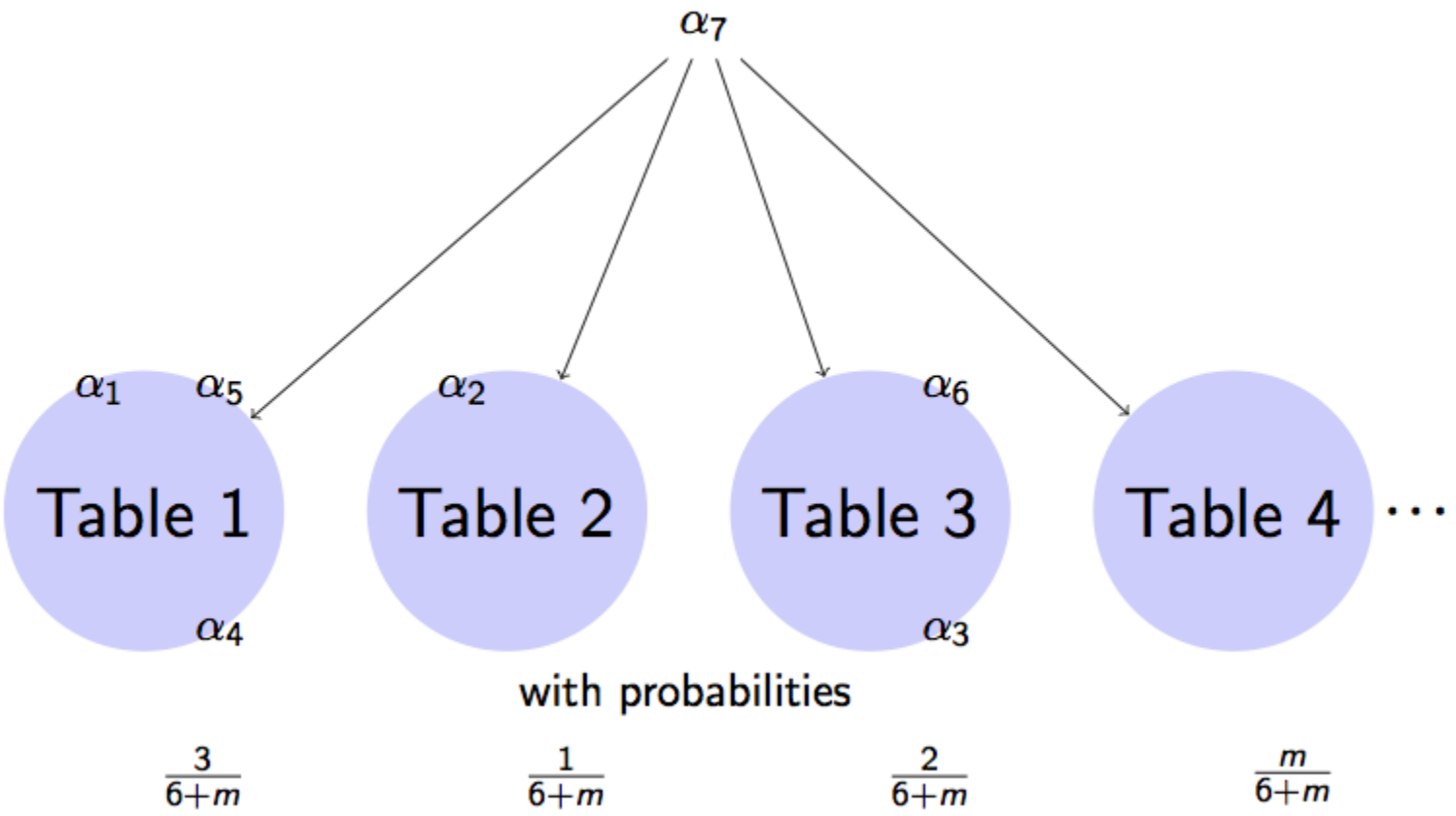
WHY USE DIRICHLET PROCESS?

- The proposed model considers subject-level differences
- There also exists similarity among the subject responses (clusters for subjects)
  - Within cluster, subject-level effects are same
  - Across clusters, subject-level effects are distinct
- Mixture density for subject-level differences
  - Unknown number of Clusters

PRIOR SPECIFICATION

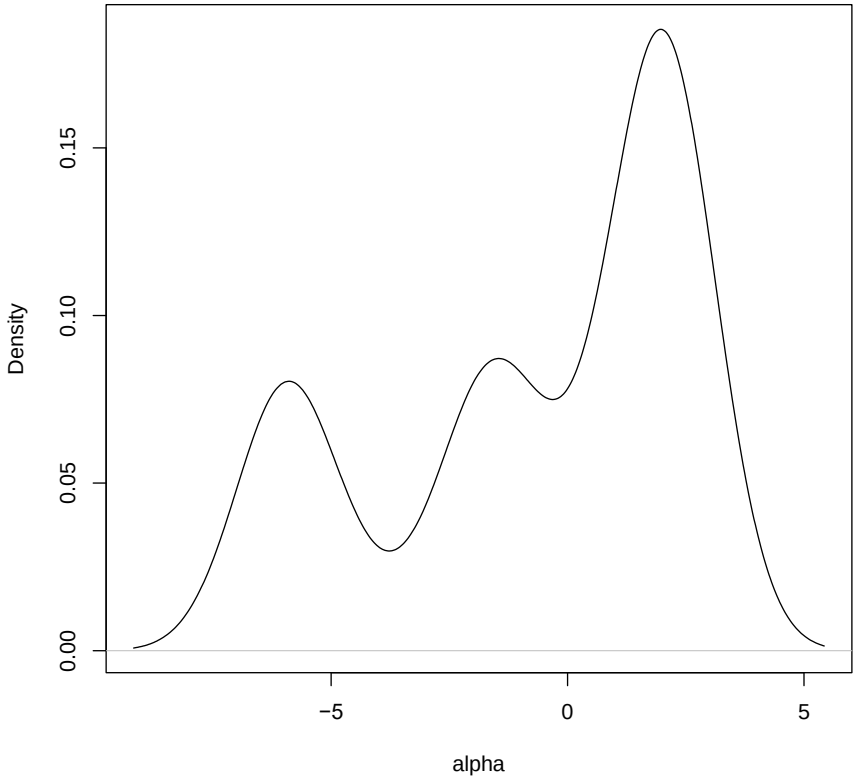
- Dirichlet process prior for the subject-level effects β_{ij}.
  - $\beta_{ij} \sim F, F \sim DP(m, F_0)$
  - $m \geq 0$: the concentration parameter
  - F_0: baseline distribution $N(0, \Sigma_\beta)$
- Visit effects $\gamma = (\gamma_1, ..., \gamma_K)' \sim N(0, \Sigma_\gamma)$

DP CLUSTERS: CHINESE RESTAURANT PROCESS

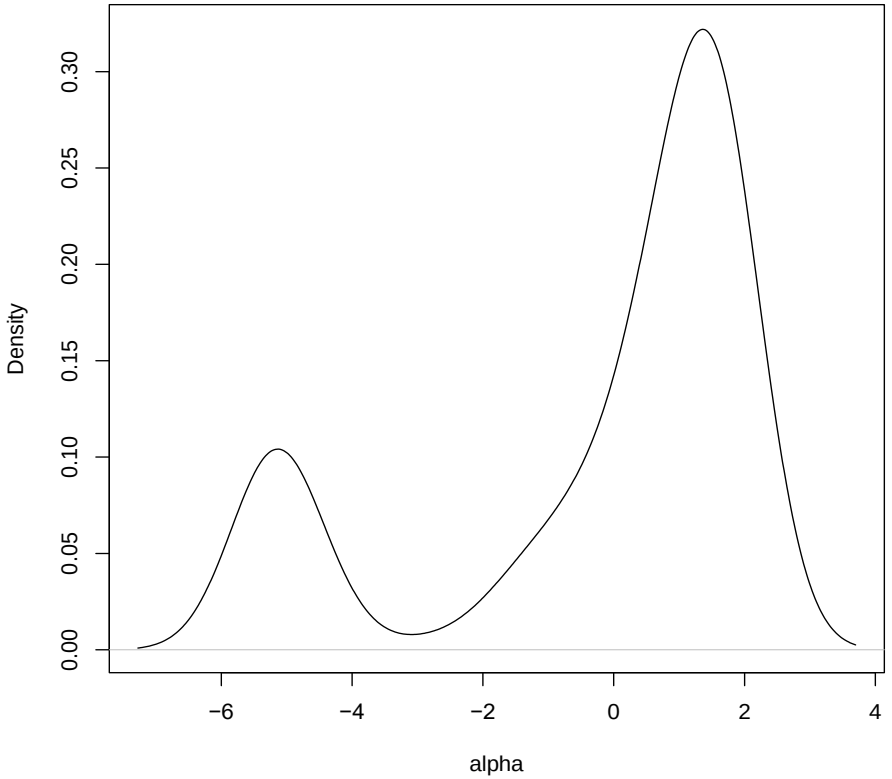


CLUSTERED SUBJECT-LEVEL EFFECTS

Treatment arm

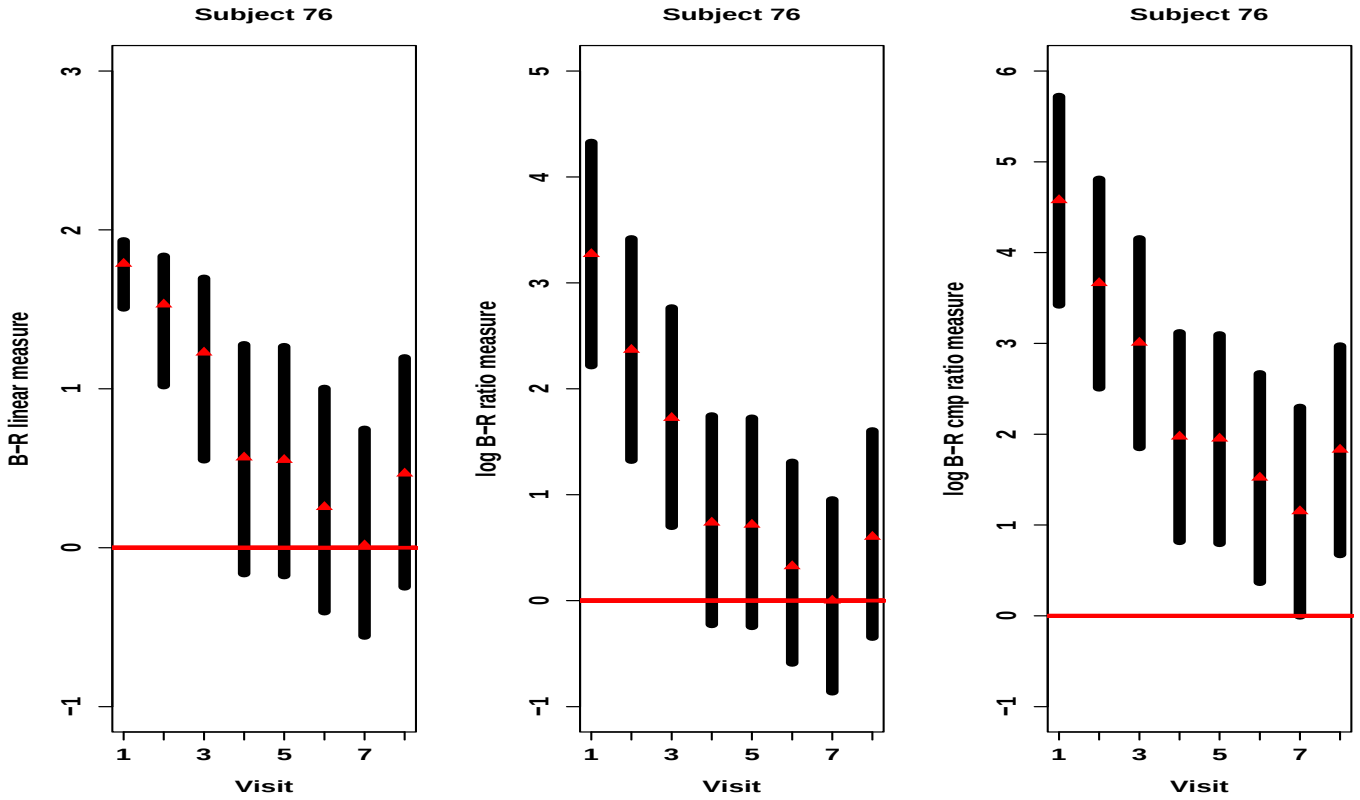


Control arm



PERSONALIZED BENEFIT-RISK MEASURES

Posterior profiles of three benefit-risk measures for subject, i = 76 from treatment arm. Vertical bars indicate 95% credible intervals of measurements, red triangles are corresponding posterior means.



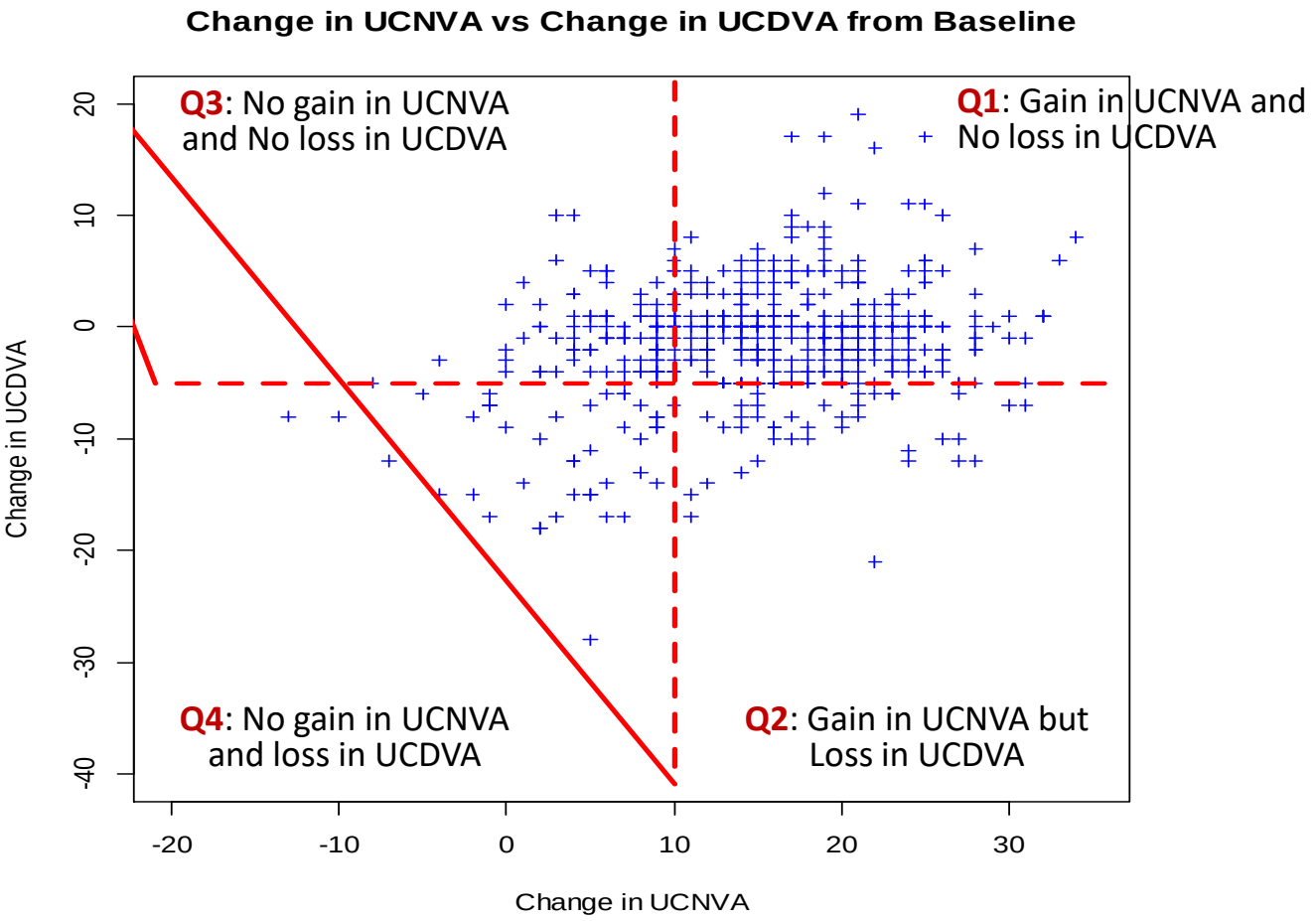
EXAMPLE 2: OPHTHALMIC DEVICE

- Benefit is measured by improvement in uncorrected (without glasses or contact lenses) **near** or intermediate visual acuity (UCNVA).

- Risk is measured by uncorrected **distance** visual acuity (UCDVA), since subjects may be giving up distance vision for some gain in near or intermediate vision.

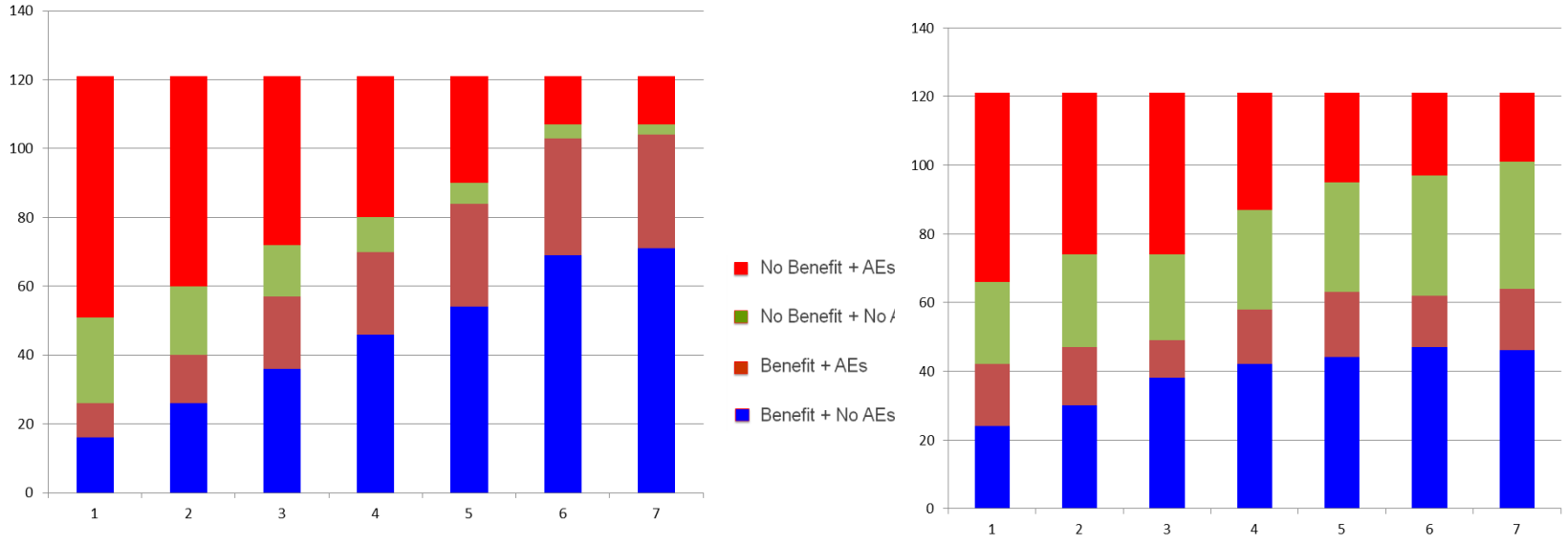
- The data on the joint evaluation of benefit and risk , (UCNVA, UCDVA), of the device is given on the next slide

EXAMPLE 2: OPHTHALMIC DEVICE



EXAMPLE 2: DATA FROM OPHTHALMIC STUDY: HYPOTHETICAL

- Two-arm study, 242 patients were randomly assigned to treatment and control arms. Patients were followed up for seven visits, and their outcomes were assigned to one of the four benefit-risk categories.



BENEFIT-RISK (BR) MEASURES

- Interested in three BR measures:

$$(Q_1 + Q_2) - (Q_3 + Q_4) \rightarrow \text{Difference between Benefit and No-Benefit}$$
$$(Q_1 + Q_2)/(Q_3 + Q_4) \rightarrow \text{Ratio of Benefit to No-Benefit}$$
$$Q_1 / Q_4 \rightarrow \text{Ratio of Best Quadrant to Worst Quadrant}$$

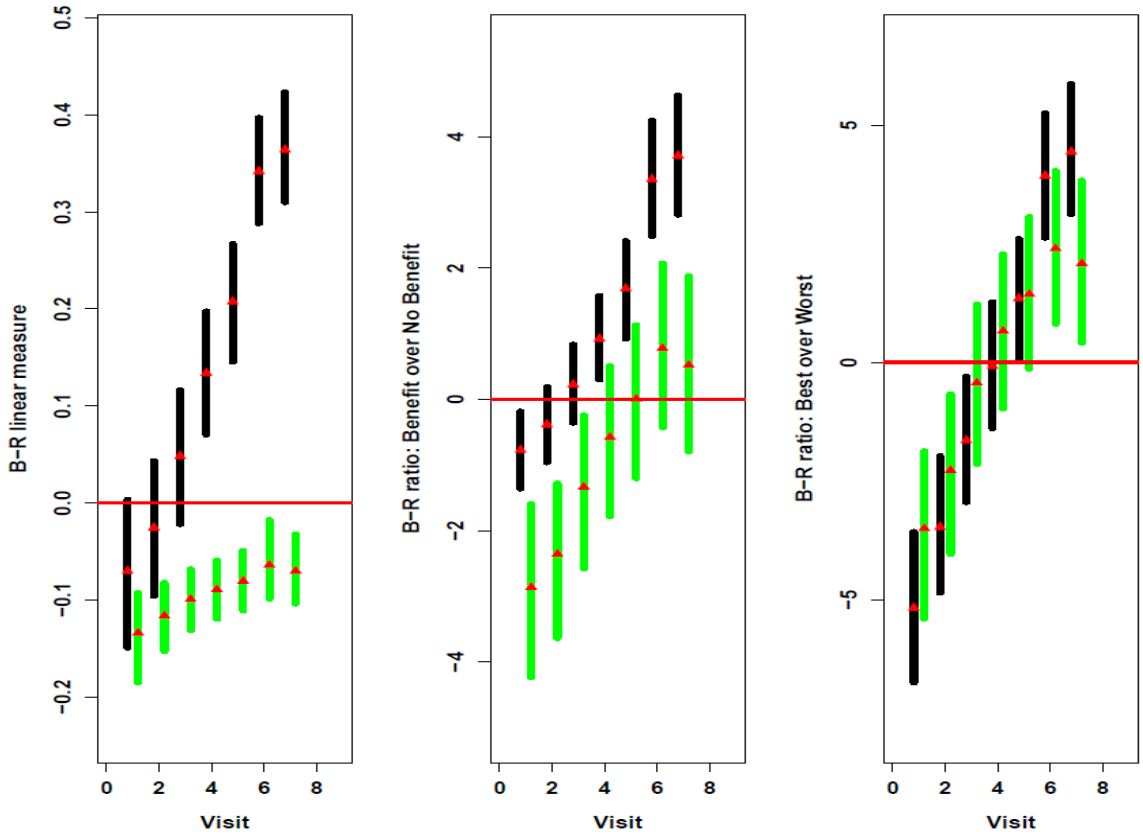
$$\text{BR-L} = \sum_{i=1}^{2} w_i p_{i,T} - \sum_{i=3}^{4} w_i p_{i,T} : (Q_1 + Q_2) - (Q_3 + Q_4)$$

$$\text{BR-R1} = \frac{\sum_{i=1}^{2} w_i p_{i,T}}{\sum_{i=3}^{4} w_i p_{i,T}} : (Q_1 + Q_2)/(Q_3 + Q_4)$$

$$\text{BR-R2} = \frac{w_1 p_{1,T}}{w_4 p_{4,T}} : Q_1 / Q_4$$

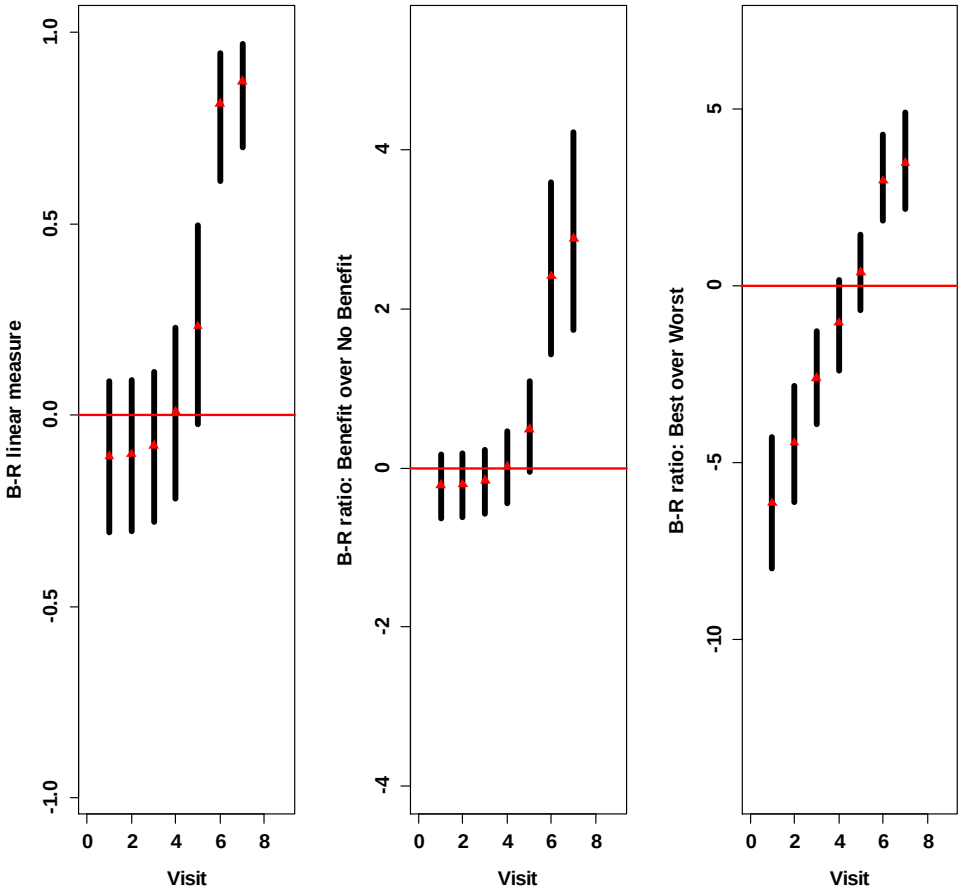
POPULATION-LEVEL BR ASSESSMENT

Posterior 95% CIs of BR-L, BR-R1, and BR-R2 at visits k = 1, . . . , 7 for treatment arm (black) and control arm (green)



PERSONALIZED BENEFIT-RISK ASSESSMENT

Posterior 95% CIs of BR-L, BR-R1, and BR-R2 at visits k = 1, . . . , 7 for one single subject i = 66, from the treatment arm



CONCLUSION

- Quantitative measures of benefit and risk are important aspect in the drug and device evaluation process
- The Bayesian method is a natural method for longitudinal data by sequentially updating the prior; Power prior can be used to discount information from previous visits.
- Use of Dirichlet process for modeling subject-level longitudinal BR categorical data seems to be a reasonable approach
- Model selection approaches can be incorporated to compare model fits
- More quantitative work on BR assessment is warranted

Other References

Gelber, R. D., Gelman, R. S., and Goldhirsch, A. (1989), A quality-of-life-oriented endpoint for comparing therapies, Biometrics, 781-795.

Holden, W. L., Juhaeri, J., and Dai, W. (2003a), Benefit-risk analysis: a proposal using quantitative methods, Pharmacoepidemiology and drug safety, 12, 611-616.

Spiegelhalter, D. J., Best, N. G., Carlin, B. P., and Van Der Linde, A. (2002), Bayesian measures of model complexity and fit, Journal of the Royal Statistical Society: Series B (Statistical Methodology), 64, 583-639.

Chuang-Stein, C., et al. (1991), Three measures for simultaneously evaluating benefits and risks using categorical data from clinical trials, *Statistics in medicine*, 10, 1349-1359

Ibrahim, J. G. and Chen, M.-H. (2000), Power prior distributions for regression models, *Statistical Science*, 46-60.

Sethuraman, J. and Tiwari, R. C. (1982) Convergence of Dirichlet measures and the interpretation of their parameter, *Statistical Decision Theory and Related Topics III* 2 305-315.

J. Sethuraman. A constructive definition of Dirichlet priors. Statistica Sinica, 4:639-650, 1994.

THANK YOU!