

# Bayesian Monitoring of Drug Safety Data

---

Jouni Kerman, Beat Neuenschwander, Michael Branson  
*Novartis Pharma AG, Statistical Methodology Group*  
ROeS meeting, Bern, September 10, 2007

# Outline

---

- Objectives & Challenges
- Data Structure & Modeling Strategy
- Illustration
- Summary

# Objectives

---

- Statistical assessment of drug safety data
  - Safety data = adverse event counts
  - Exploratory; for **internal monitoring**

# Goal

---

- **Monitoring** of safety information; accumulation of evidence
- **Early detection** of potential safety signals
- **Prediction**: what are we likely to see in the next study?

# Challenges

---

Monitoring and interpreting safety data is not easy

- **Rare events**
  - Example: 0 vs. 1 events in 35000 person-days
- **Multiplicity of information**
  - Studies
  - Treatment groups
  - Patient subgroups
  - Occasional false signals in safety profiles are unavoidable

# Tackling the multiplicity problem

- **Testing approach: p-value adjustments**
  - leaves original estimates unchanged, with more uncertainty
- **Hierarchical approach (evidence synthesis, meta-analysis)**
  - **shrinkage**: extreme signals are pulled towards an ‘overall average’
  - **more precise estimates** compared to stratified (by-study) analyses
  - offers a **compromise between pooled and stratified analyses**
  - needs a ‘reasonable’ **modeling framework** (no ‘unique’ solution!)
  - is essentially **Bayesian**

# Bayesian hierarchical approaches

- Bayesian hierarchical approaches to safety data
  - [1] DuMouchel, W (1999) Bayesian Data Mining in Large Frequency Tables, With an Application to the FDA Spontaneous Reporting System. *The American Statistician* 53:3
  - [2] Berry, S and Berry, D (2004) Accounting for Multiplicities in Assessing Drug Safety: A Three-Level Hierarchical Mixture Model. *Biometrics* 60:418-426.
- Our approach is inspired by 2)

# Test Data Set

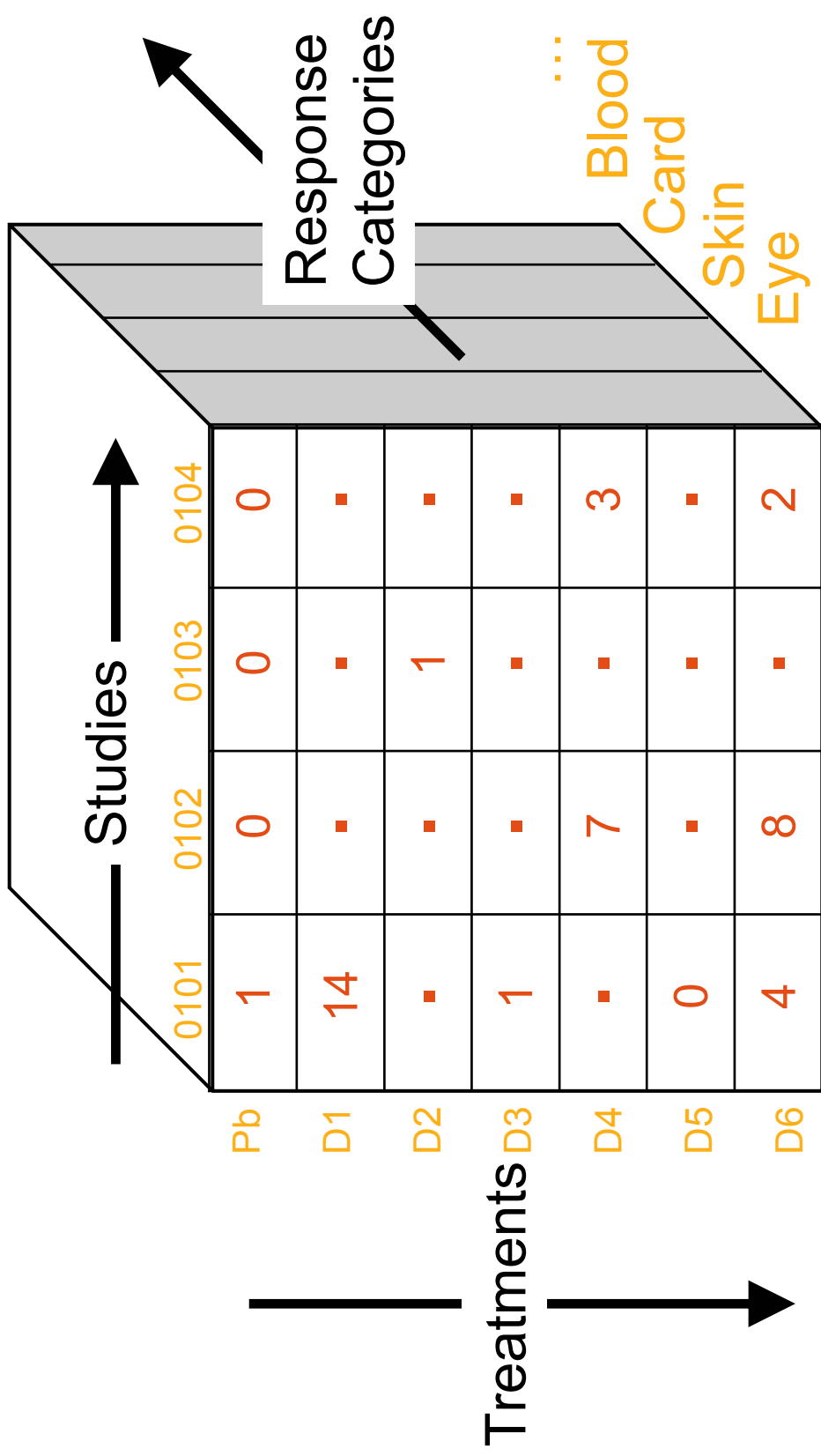
| study | trt   | events | exposure | category |     |
|-------|-------|--------|----------|----------|-----|
| 1     | 0101  | Pb     | 1        | 2545     | Eye |
| 2     | 0102  | Pb     | 0        | 16710    | Eye |
| 3     | 0103  | Pb     | 0        | 3287     | Eye |
| 4     | 0104  | Pb     | 0        | 5455     | Eye |
| 5     | 0101  | Dose 1 | 14       | 2655     | Eye |
| 6     | 0103  | Dose 2 | 1        | 3434     | Eye |
| 7     | 0101  | Dose 3 | 1        | 2685     | Eye |
| 8     | 0102  | Dose 4 | 7        | 36570    | Eye |
| 9     | 0104  | Dose 4 | 3        | 5793     | Eye |
| 10    | 0101  | Dose 5 | 0        | 2609     | Eye |
| 11    | 0101  | Dose 6 | 4        | 2778     | Eye |
| 12    | 0102  | Dose 6 | 8        | 36381    | Eye |
| 13    | 0104  | Dose 6 | 2        | 5390     | Eye |
| 14    | ..... |        |          |          |     |

Adverse event counts and exposures are pooled over individuals for each

- treatment  $t$ ,
- study  $s$ ,
- response category

Full data set has an identical format for each response category.

# Multilevel structure



# Modeling the rate of adverse events

## Basic model

- Adverse events  $y_{ts}$  follow a Poisson process with rate  $\lambda_{ts}$ 
  - Treatment  $t$ , study  $s$
  - Fit the model separately for each adverse event category
- $y_{ts} \sim \text{Poisson}(\lambda_{ts} E_{ts})$ 
  - Exposure (per 100 days):  $E_{ts}$
  - Mixed-effects (Poisson regression) model for  $\log(\lambda_{ts})$ 
    - accounts for between-study variation
    - accounts for overdispersion (due to nonconstant Poisson rates)
    - allows for prediction of rates in future studies

# Modeling strategy

*Fit several models; don't forget the simple ones*

- Fit a series of models of increasing complexity for  $\lambda_{ts}$ 
  - Compare estimates from complex models to those from the simplest models
  - View estimates graphically
  - View estimates from several models in a single graph

# Modeling the rate of adverse events

*Two reference models and a multilevel model*

- 1. No-pooling model**  
effects across studies are not exchangeable at all – look at them separately
- 2. Complete-pooling model**  
effects are fully exchangeable across studies– you can pool events together
- 3. Multilevel (hierarchical) model**  
effects may be exchangeable – but let us estimate from the data to what extent this is plausible

# Models

## *Decomposition of the log-rate*

### 1. No-pooling model

$$\log(\lambda_{ts}) = \mu_s + \eta_{ts} \text{ for all } t, s.$$

Fixed effects.  $\eta_{0s} = 0$  (placebo).

### 2. Complete-pooling model

$$\log(\lambda_{ts}) = \mu + \theta_t \text{ for all } s.$$

Fixed effects.  $\theta_0 = 0$  (placebo).

### 3. Multilevel model

$$\log(\lambda_{ts}) = \mu + \psi_s + \theta_t + \eta_{ts}$$

$\mu, \theta_t$  are fixed effects.  $\theta_0 = 0$ .

$\psi_s \sim N(0, \sigma^2)$ , random effect.  $\sigma \sim$  fixed prior.

$\eta_{ts} \sim N(0, \tau_t^2)$ , random effect.  $\tau_t \sim$  hierarchical prior.

# Multilevel model

## Fitting strategies

### 1. No-pooling model

$\lambda_{ts} \sim$  independent Gamma( $\alpha, \beta$ )  
compute contrast  $\log(\lambda_{ts}/\lambda_{1s})$  by direct simulation

### 2. Complete-pooling model

First pool data over studies

$\lambda_t \sim$  independent Gamma( $\alpha, \beta$ )  
compute contrast  $\log(\lambda_t/\lambda_1)$  by direct simulation

### 3. Multilevel model: Poisson regression

$\log(\lambda_{ts}) = \mu + \psi_s + \theta_t + \eta_{ts}$

- Use iterative simulation (MCMC)

# Safety signals and the case of zero observations

- “Signal if  $\Pr(\theta_t > 0 \mid y) > 0.95$ ”
  - Zero events ( $y=0$ ) contributes almost no information: **sensitivity of the safety signal is determined by the choice of prior.**
  - There is no agreement what *the* noninformative prior should be
  - You *shouldn't* just take a generic off-the-shelf prior
- We must calibrate the model so that our inferences are reasonable.
  - Introduce **weakly informative priors** such that the contrast estimate will have a reasonable range of values when  $y=0$ .

## Calibration of the prior distribution: example

- Given a desired “threshold event count”  $y$ , choose the Gamma shape parameter  $\alpha$  such that  **$y/E$  vs.  $0/E$**  gives a signal when the no-pooling model is used.
  - $\Pr(\text{treatment rate} > \text{placebo rate} \mid y_1=2, y_0=0) \approx 0.95$ 
    - For the Gamma model,  $\alpha=0.3$  would give a signal at 95% level for  $2/E$  vs  $0/E$
    - Translate this to log scale
      - E.g. prior 95% uncertainty interval of the rate : (1/400, 400).
- Research on this topic is still ongoing

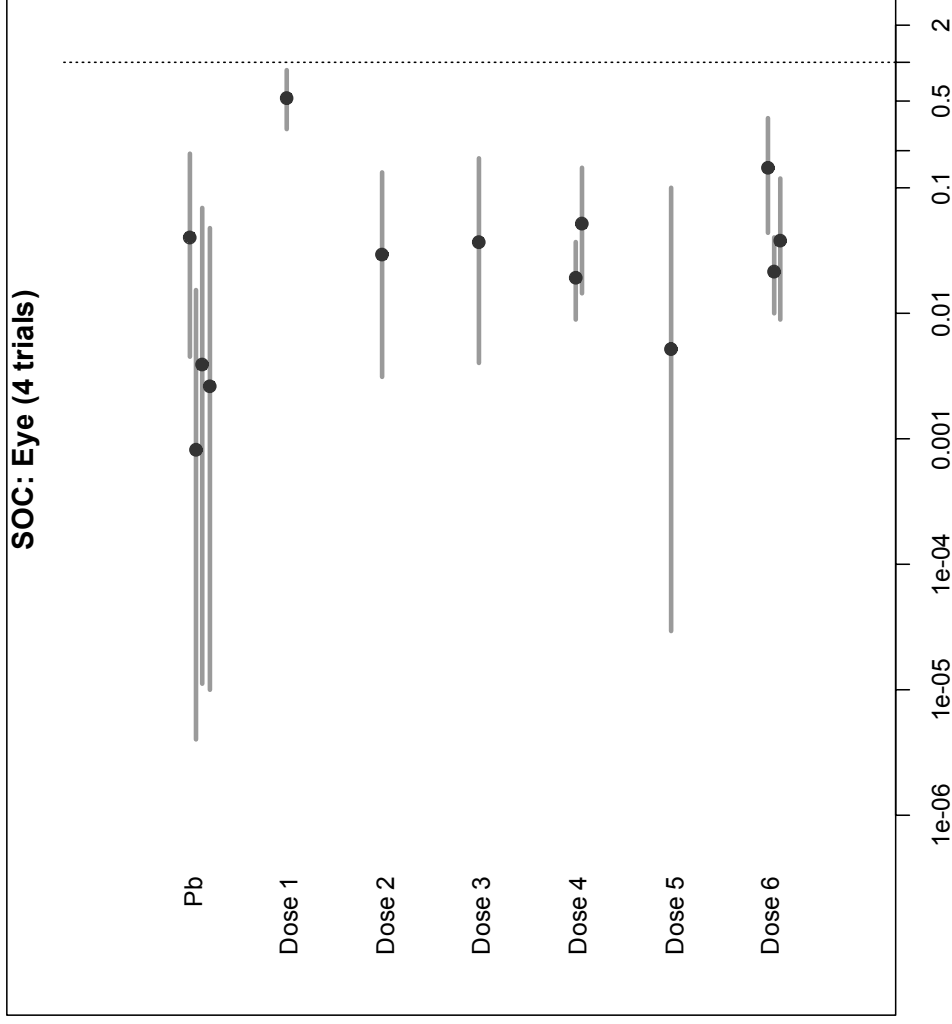
# Illustration: visualisation of the process

---

- Safety monitoring
- Signal detection
- Prediction

# Log-rates: $\log(\lambda_{ts})$ from No-Pooling Model (gray)

Treatment rate estimates



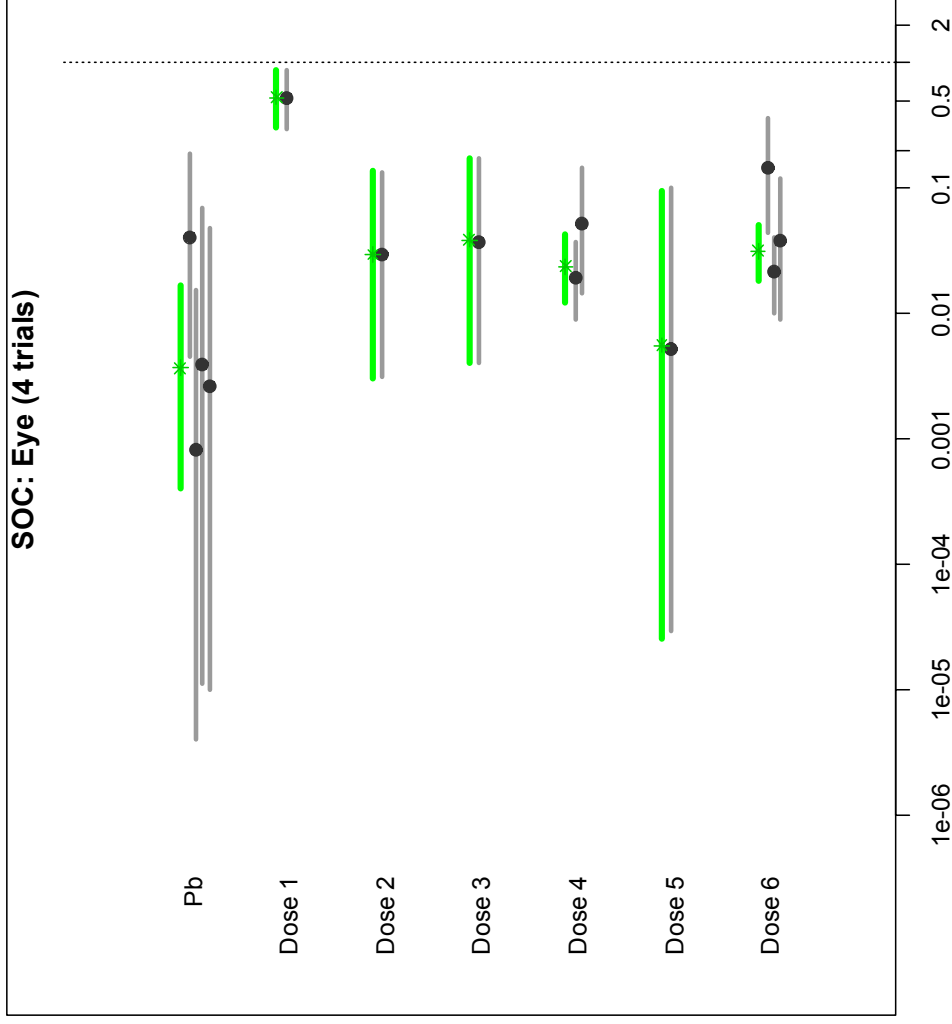
No-pooling model yields separate estimates for each study.

Shown: 95% posterior intervals for each of the clinical studies.

Dots = posterior means.

# $\log(\lambda_t)$ from Complete-Pooling Model (green)

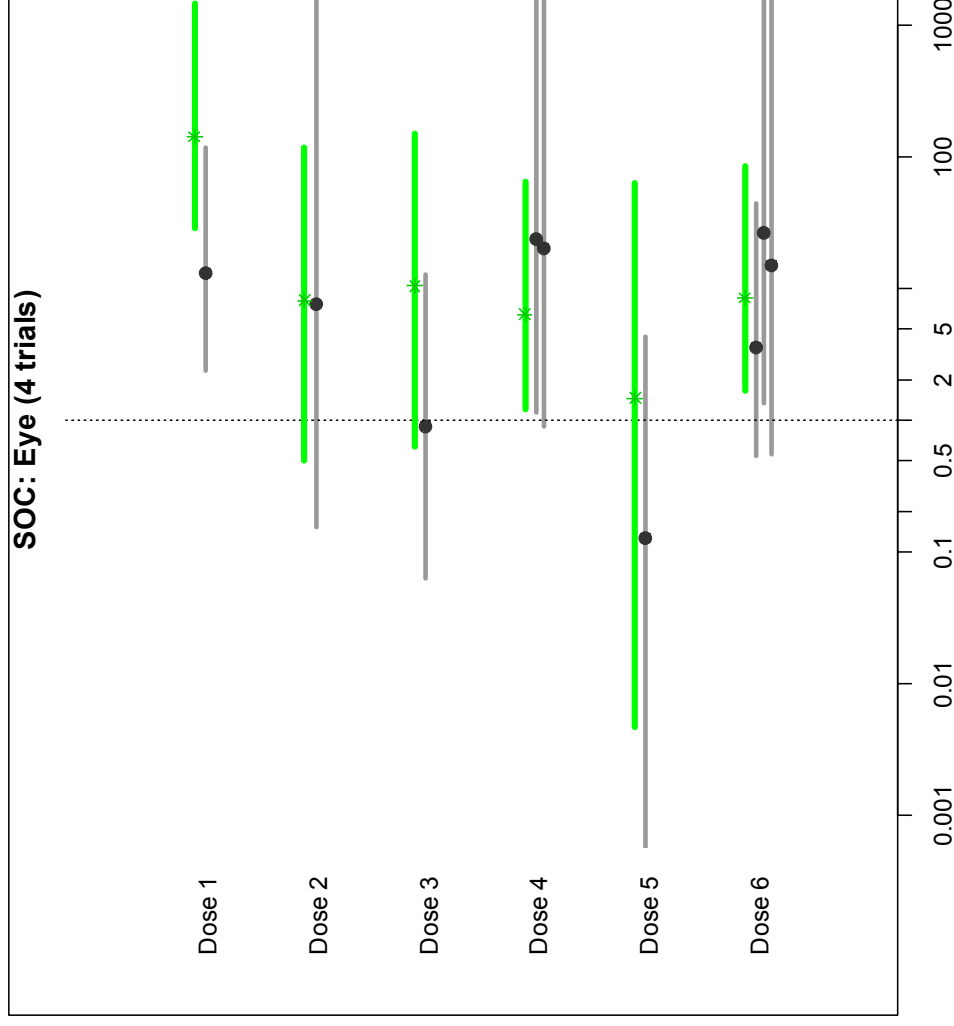
Treatment rate estimates



Complete-pooling model yields too narrow uncertainty intervals, assuming no between-study variation.

# Contrasts: treatment vs. placebo rate

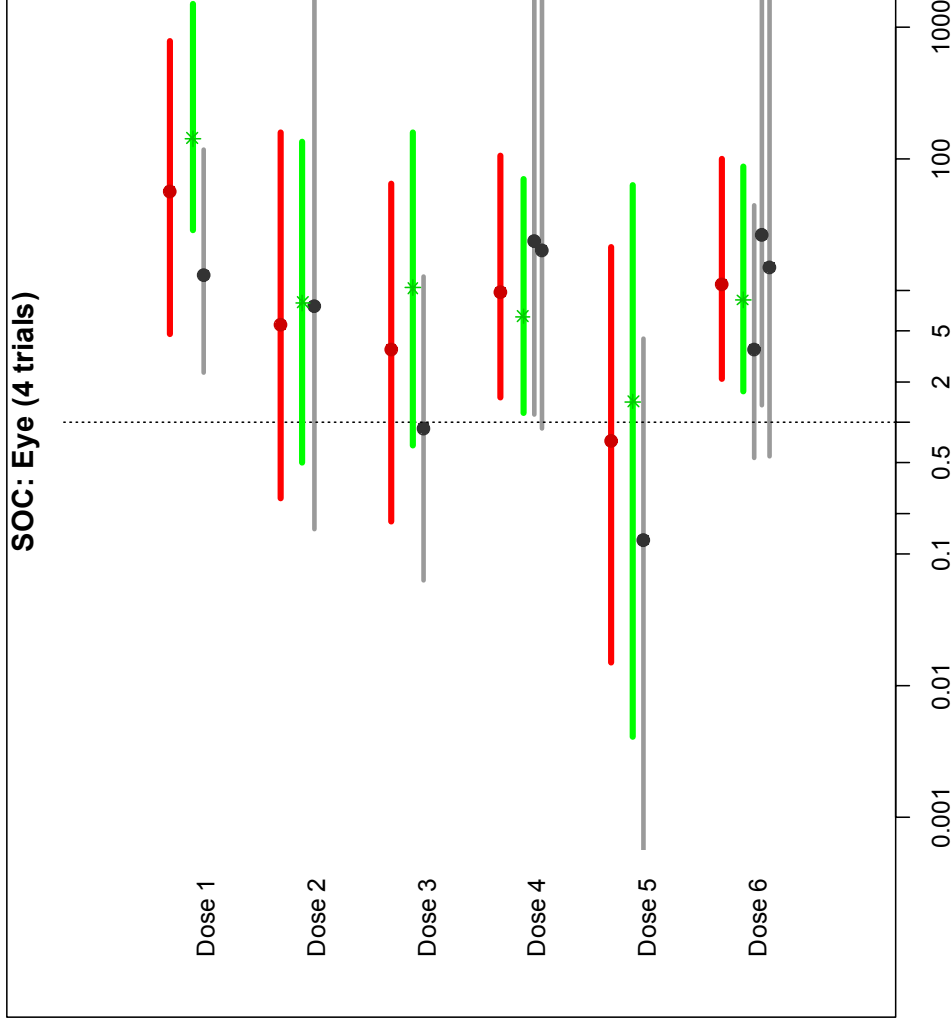
Treatment contrasts vs. Placebo



Complete-pooling model may give exaggerated contrast estimates.

# Multilevel estimates of contrasts (red)

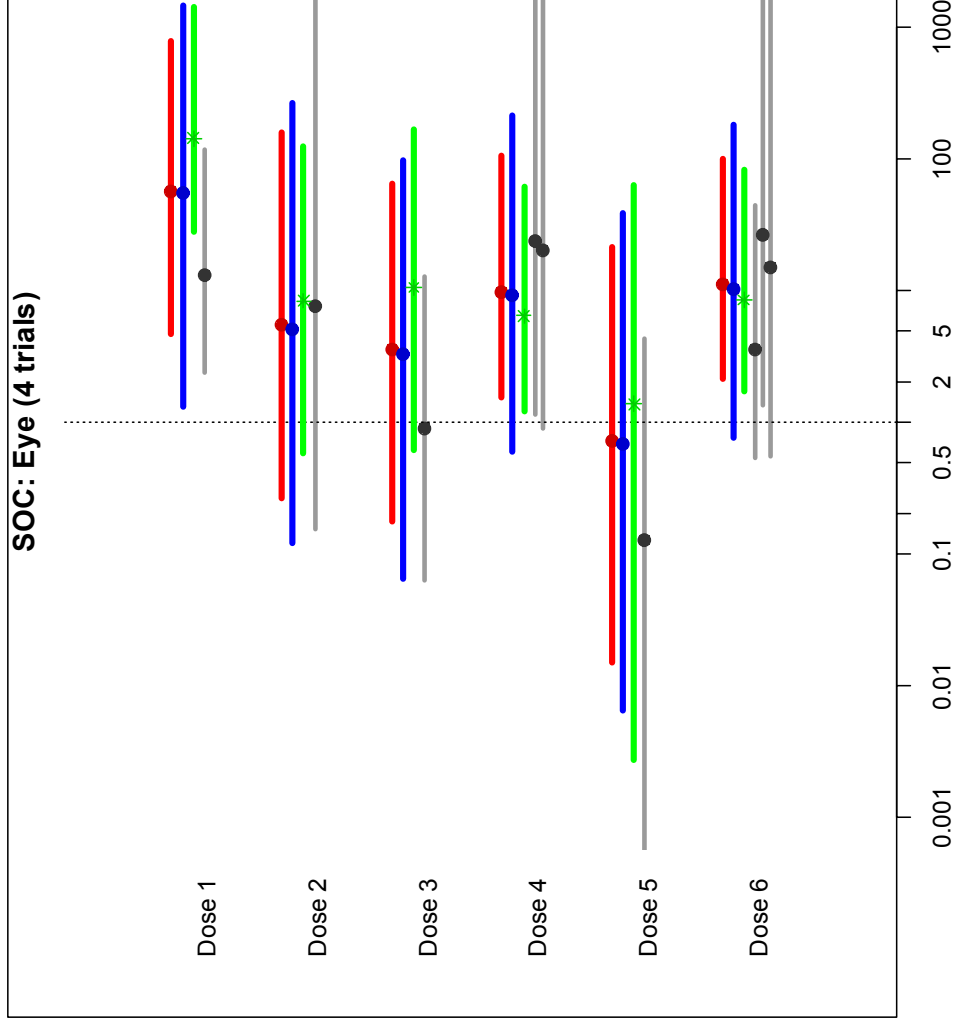
Treatment contrasts vs. Placebo



The multilevel model borrows information across studies and treatments, yielding more realistic estimates.

# Predicted rate ratios (blue)

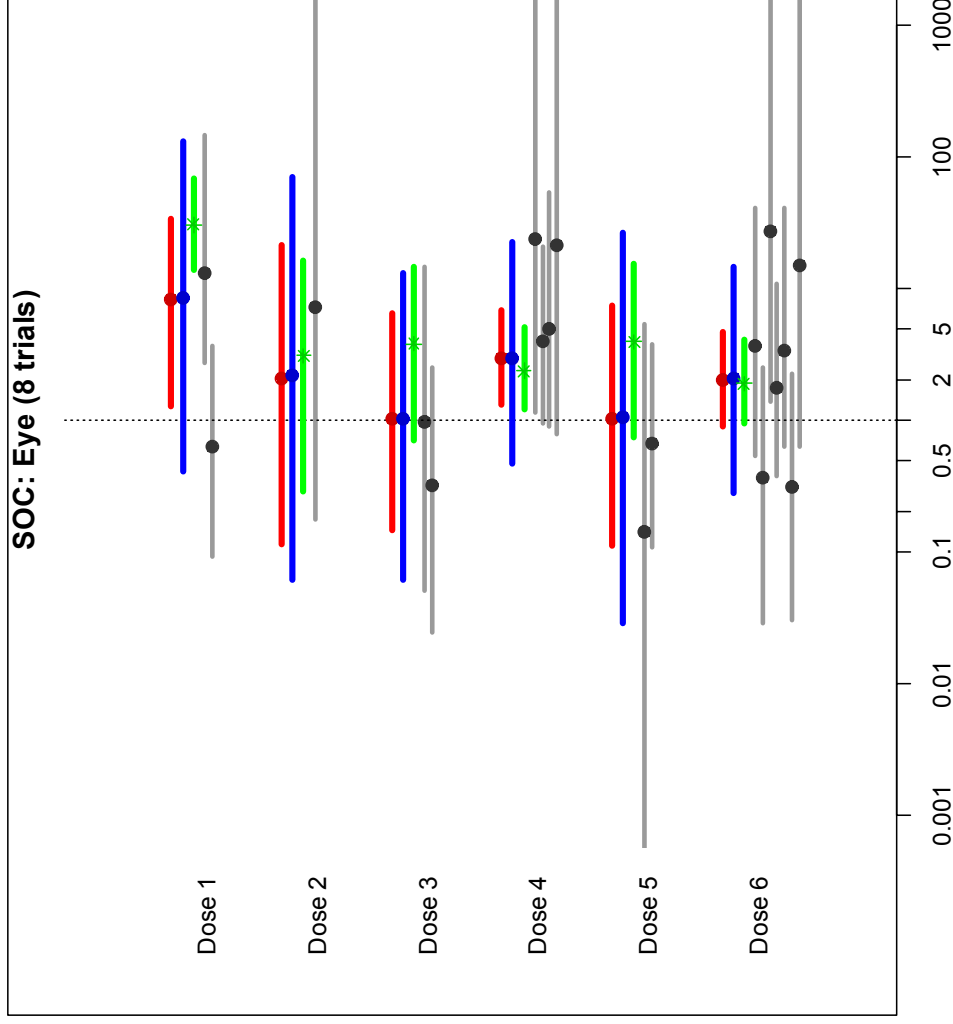
Treatment contrasts vs. Placebo



Looking forward:  
what can we expect  
in a future trial?

# Up to now, 4 trials ... now, adding 4 more trials ...

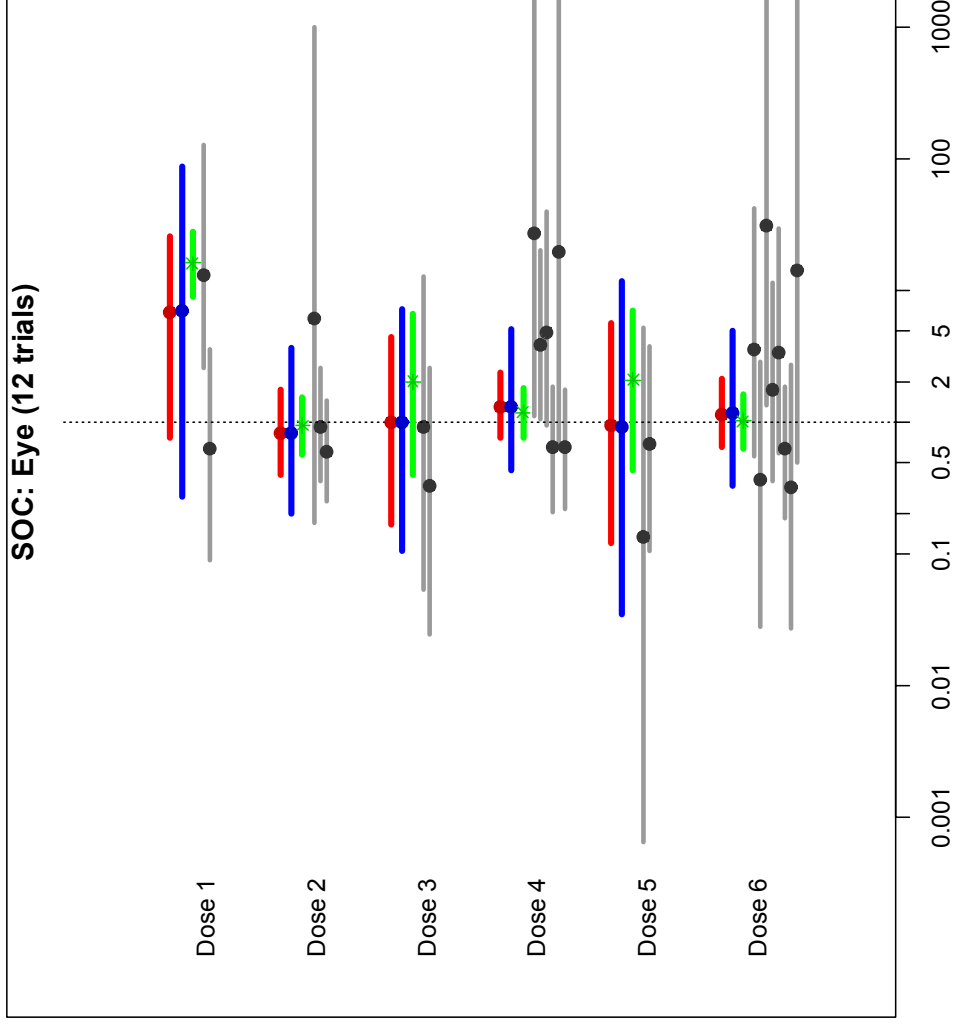
Treatment contrasts vs. Placebo



As more information is added, uncertainty intervals get narrower.

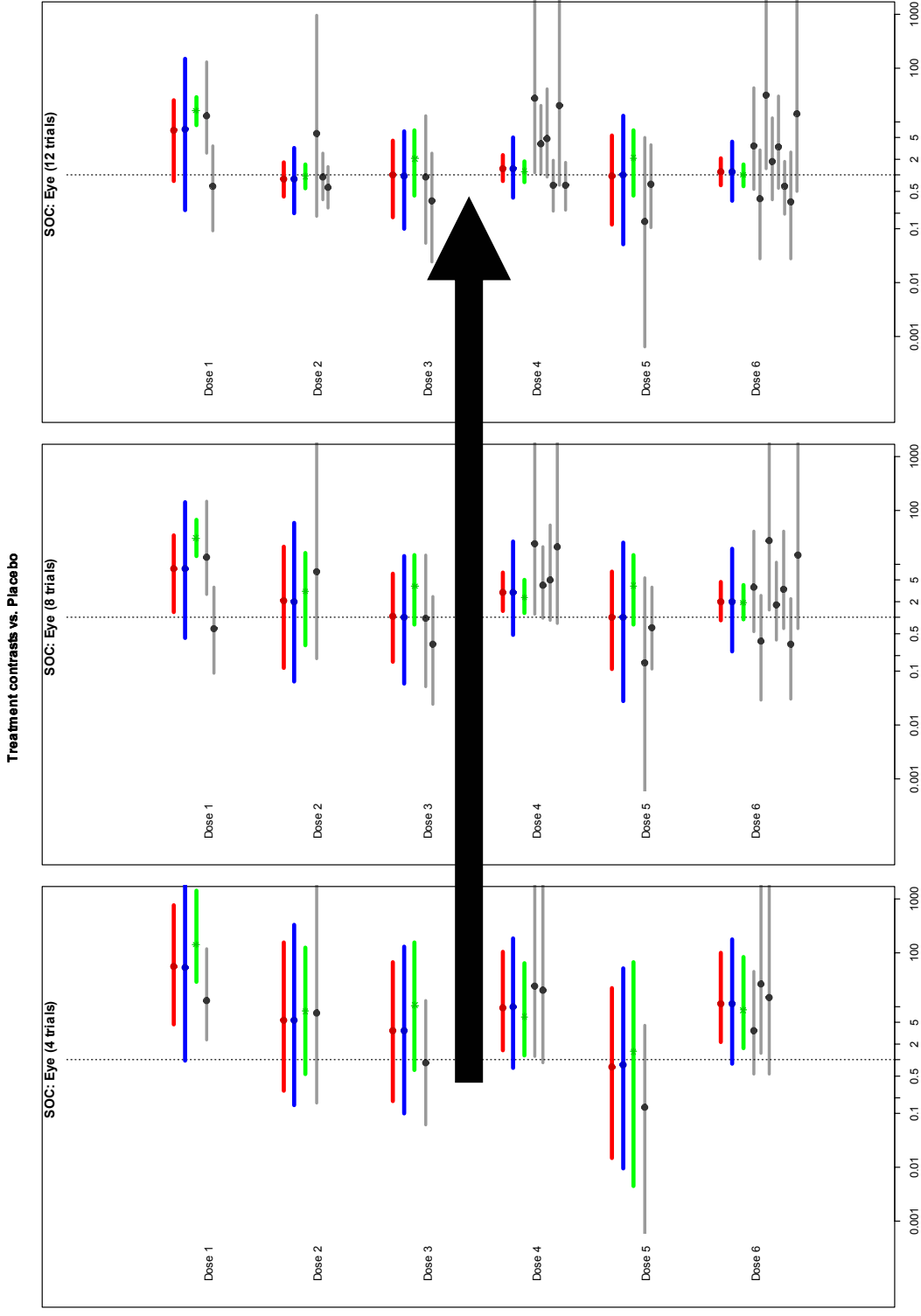
# Adding more trials: now, a total of 12

Treatment contrasts vs. Placebo



As more information is added, uncertainty intervals get narrower.

# Process of safety monitoring



# Summary

- Advantages of the Bayesian multilevel modeling approach
  - Easily expandable; complexity reflects reality
  - Readily interpretable estimates
  - Straightforward predictions
  - Generates safety signals (given a “level of proof”)
  - Makes also use of trial data with zero observed events
- Caveats
  - Choice of weakly informative priors is important
  - Possible computational challenges with complex models
- Don't forget simple models
  - Display them side by side to the estimates from complex models

# Acknowledgements

---

- Beat Neuenschwander
- Michael Branson
- David Spiegelhalter (MRC Cambridge)
- Gerd Rosenkranz
  
- Developers of WinBUGS and R
- Andrew Gelman (R2WinBUGS)