Approaches to Integrated Safety Analyses

Gerd Rosenkranz
Novartis Pharma AG, Basel, Switzerland

GMDS Annual Conference, Leipzig 2006

Overview

- **Reporting and coding of adverse events**
- **Reduction of complexity**
- **Modeling of incidences of AEs (one study, one experimental treatment and a comparator)**
- **Example**
- **Generalization to several studies and treatments**
- **Discussion and outlook**

What are adverse events?

- **All signs and symptoms that occur or worsen after initiation of a treatment, regardless of their relationship to the drug**
- **The above definition implies that the number of reported adverse events can be large**
- **To relate adverse events to treatments they have to be investigated in a comparative setting**
- **Adverse events are reported verbatim; they are coded to enable a sensible analysis.**

Coding of adverse events using MedDRA 5.1

Level	Use	Number of categories
System organ class (SOC)	reporting	26
High level group term (HLGT)		332
High level term (HLT)		1683
Preferred term (PT)	reporting	16102
Low level term (LLT)	coding	56981

At the PT level, the data of one subject constitute a 16000 dimensional 0-1 vector (disregarding severity, time and repetition)

Analyzing adverse events

- Routinely, incidence of adverse events is summarized by SOC and preferred term.
- Sometimes, a p-value is calculated for the difference of incidences among an experimental drug and a comparator by preferred term
- Analyses of adverse events suffer inherently from multiplicity issues
- In contrast to efficacy, the increased degree of „conservatism“ caused by adjustment for multiplicity is not desirable
- Can safety information from several studies of a clinical development program be analyzed simultaneously (integrated safety analysis)?

Reduction of complexity

- Assume constant risk over time (Poisson model)
- Use summary statistics, e.g., number of AEs of a certain category (as defined by SOC or PTT)
- Consider total exposure time
- Notation
  - N_{ij} = number of AEs in category i under treatment j
  - E_j = total exposure under treatment j
  - λ_{ij} = risk rate in category i under treatment j
  - $p_{ij} = P[N_{ij}=n_{ij}] = Po(n_{ij}, \lambda_{ij}E_j)$
- Look at SOCs first (only 26 categories)

Models (one study, control and experimental)

- **Fixed effect model: identical treatment effects for all categories**

$$\log(\lambda_{i0}) = \mu_i$$

$$\log(\lambda_{i1}) = \mu_i + \theta$$

  - **Not realistic, except if $\theta = 0$.**

- **Fixed effects model: different treatment effects per category**

$$\log(\lambda_{i0}) = \mu_i$$

$$\log(\lambda_{i1}) = \mu_i + \theta_i$$

  - **large number of unrelated parameters**

- **Mixed effects model: random treatment effect**

$$\log(\lambda_{i0}) = \mu_i$$

$$\log(\lambda_{i1}) = \mu_i + U_i, \qquad U_i \sim N(\theta, \sigma^2)$$

  - **„Borrows strength“ from all categories**

Likelihoods

- **Fixed effects model**
  - $L_i(\mu,\theta) = Po(n_{i0},E_0\exp(\mu_i))\ Po(n_{i1},E_1\exp(\mu_i+\theta_i))$
  - θ_i determined by maximization of L_i alone
- **Mixed effects model**
  - $L_i(\mu,\theta,\sigma) = Po(n_{i0},E_0\exp(\mu_i)) \int Po(n_{i1},E_1\exp(\mu_i+u_i))\varphi((u_i-\theta)/\sigma)du_i$
  - θ contained in each factor of $L = \prod_i L_i$
  - Less parameters to estimate, but **estimation numerically more involved**
  - **Predicted effect** in category i is that u_i that maximizes

    $$Po(n_{i1},E_1\exp(\mu_i + u_i))\varphi((u_i-\theta)/\sigma)$$

    after fixed effects parameters are replaced by their mle's.

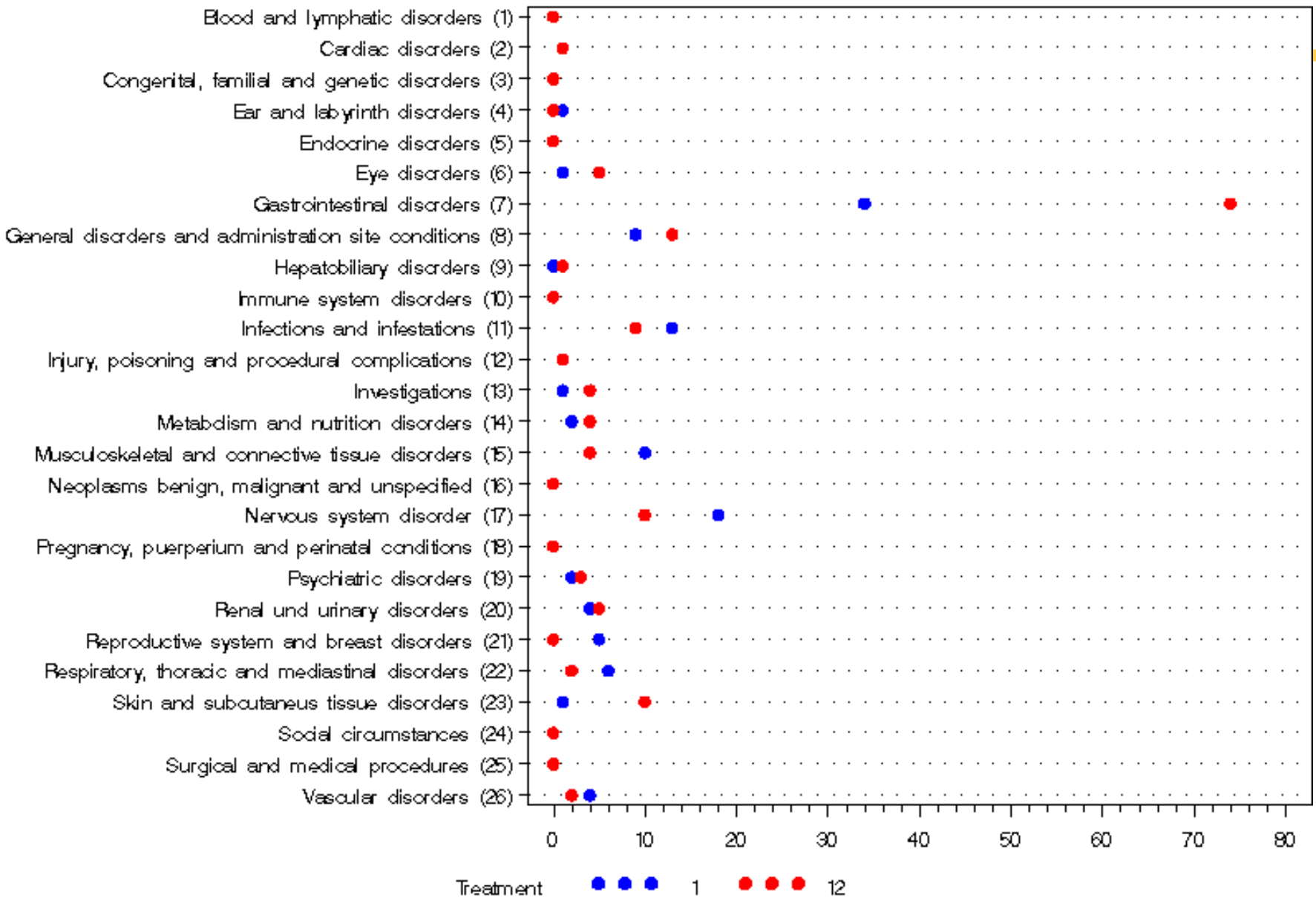
Example

- 28 days NSAID study in osteoarthritis patients
- 94 patients (2497 ptds) under Diclofenac
- 97 patients (2545 ptds) under placebo
- Analyses performed with PROC NLMIXED and PROC MULTTEST in SAS V8.2

Schnitzer et al., Arthritis and Rheumatism, 2004;51:549-557.

Note: the publication reports numbers of patients with at least one adverse event, we analyze numbers of adverse events

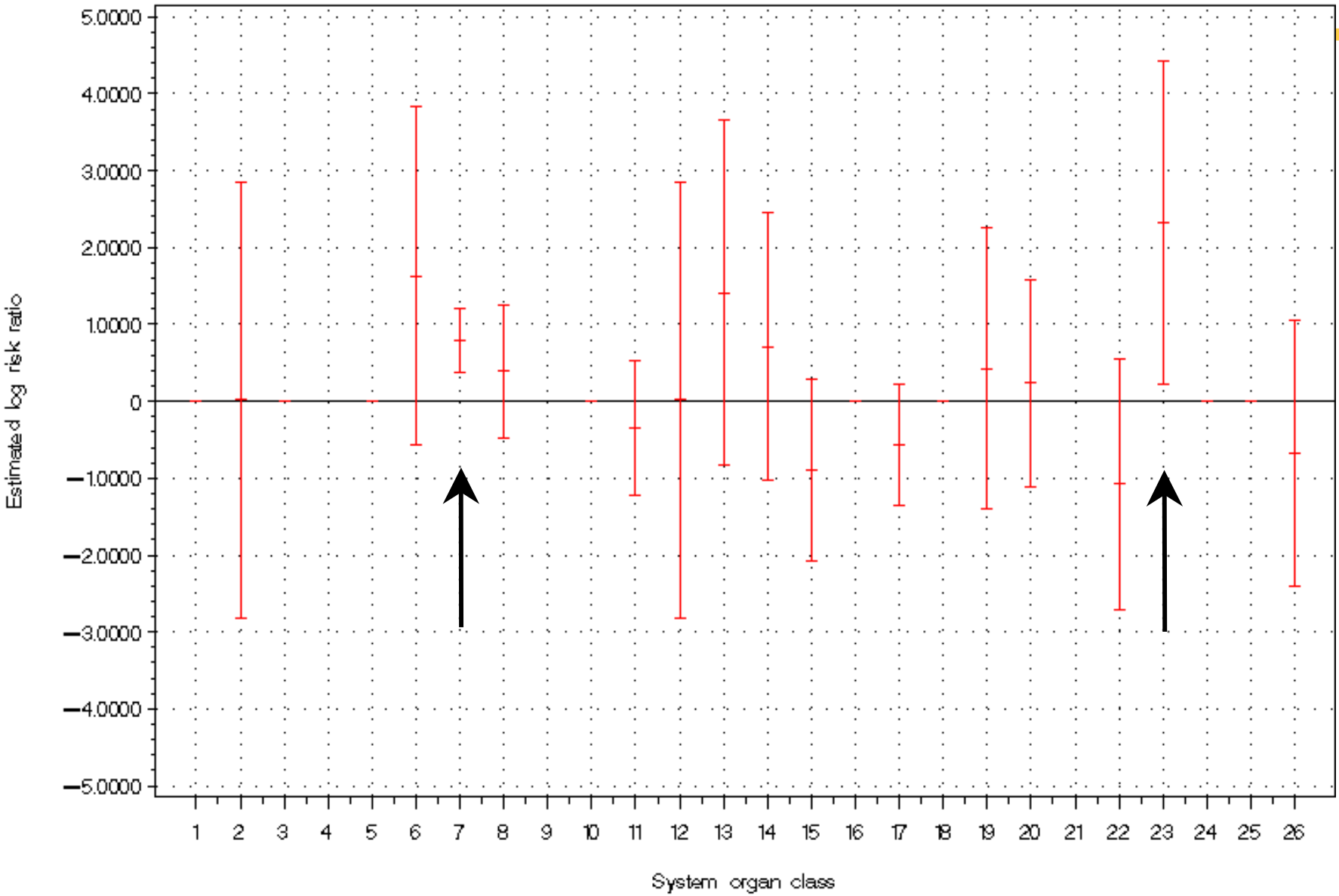
Number of AEs by system organ class



NOVARTIS

Fixed effects model

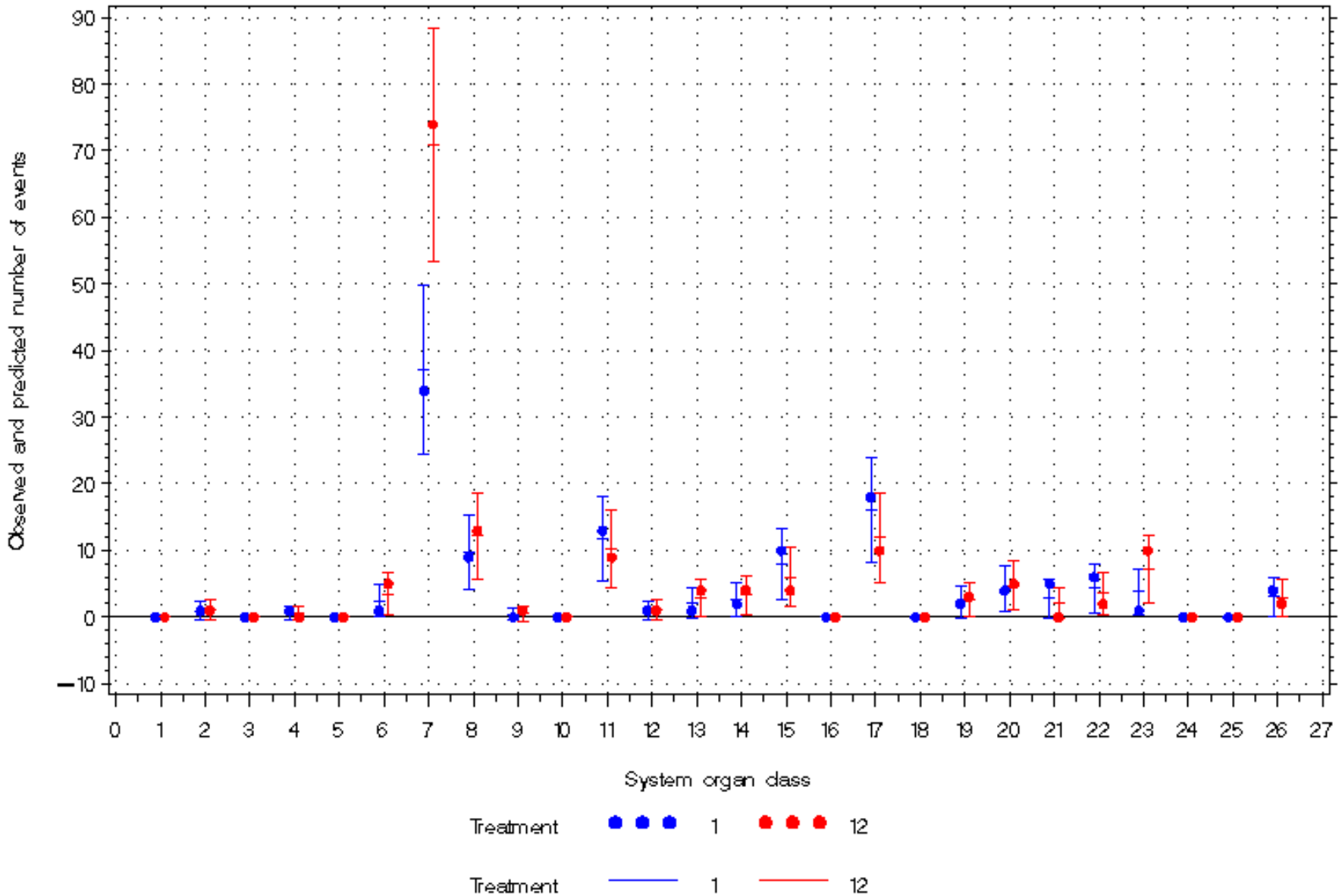
Estimated risk ratios and 95% confidence intervals



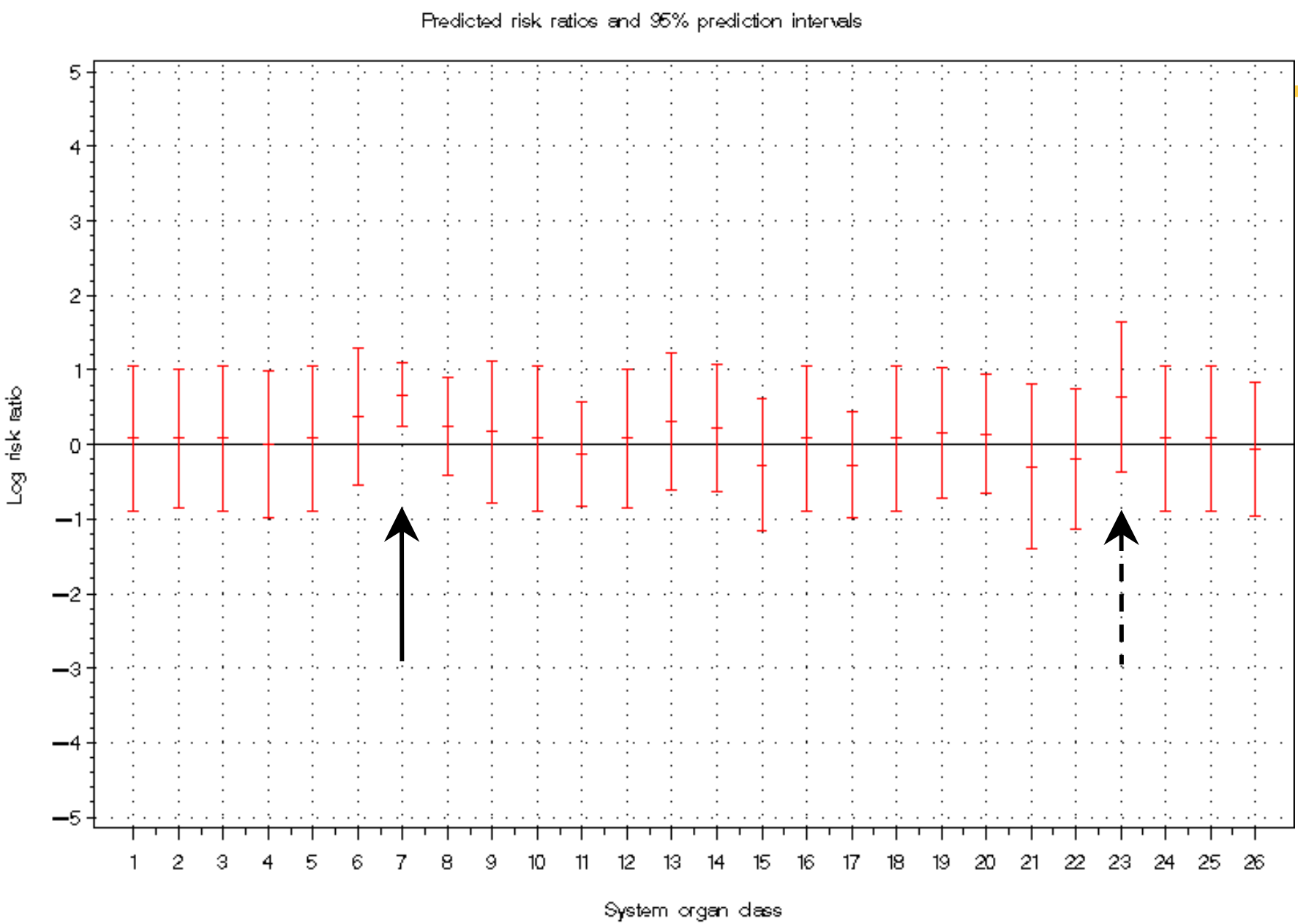
NOVARTIS

Mixed effects model

Observed and predicted number of adverse events and 95% prediction intervals



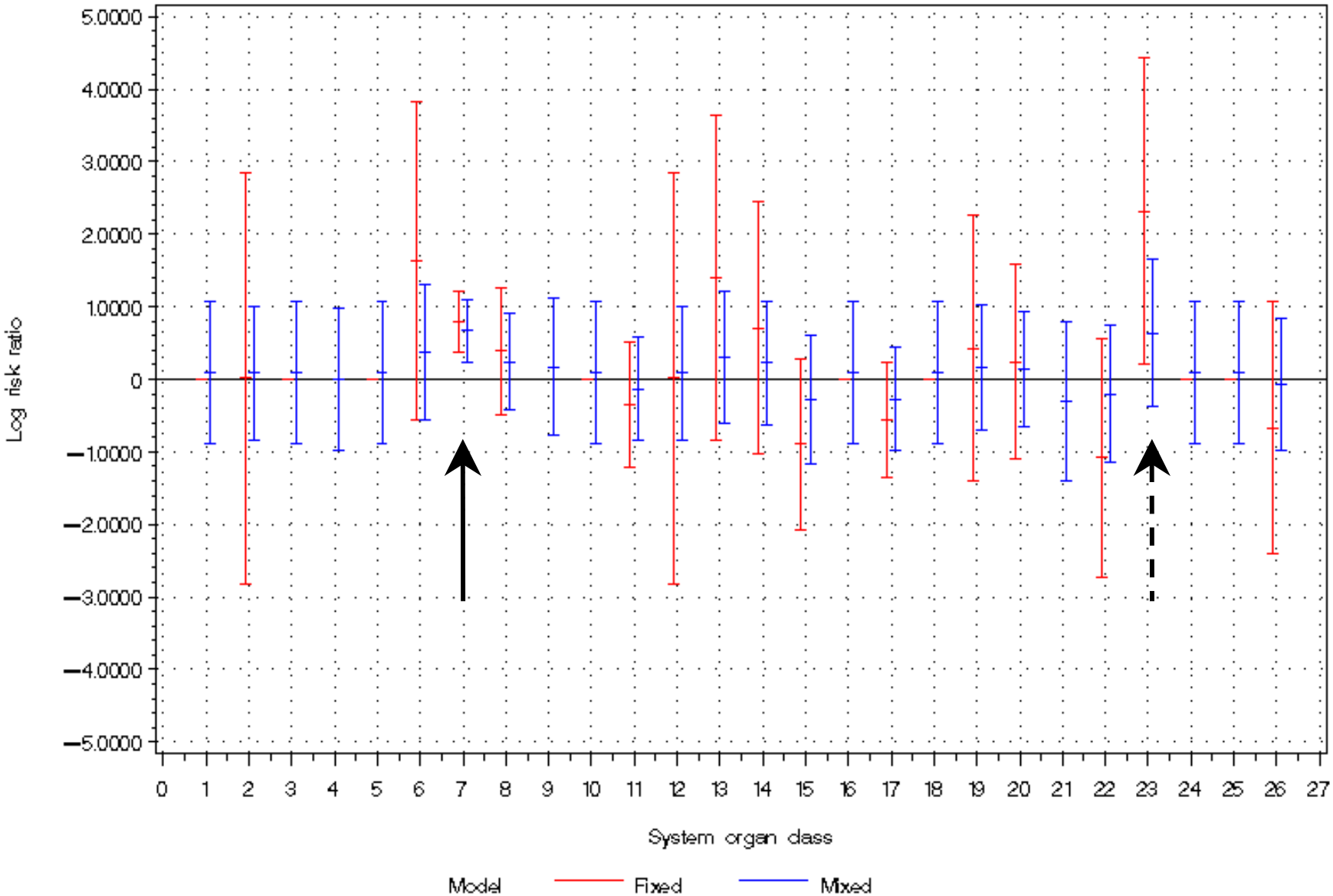
NOVARTIS


Mixed effects model
Predicted risk ratios and 95% prediction intervals
Log risk ratio
5
4
3
2
1
0
−1
−2
−3
−4
−5
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26
System organ class

NOVARTIS

Comparison of fixed and mixed effects analysis

Estimated/predicted risk ratios and 95% confidence/prediction intervals



NOVARTIS

Results

- The fixed effects model analysis suggests that differences occur for SOC 7 and SOC 23
- The corresponding raw p-values are p=0.0003 for SOC 7 and p=0.0313 for SOC 23
- With Bonferroni, Holm or FDR correction: only SOC 7 significant on 5% level
- The mixed effects analysis suggests a difference for SOC 7 only
- The predictions for SOCs with little information (small numbers of events) are reduced towards zero as compared to the fixed effect analysis

Generalization

- **The mixed model can be generalized to include more than one study („meta-analysis“):**

  $$L_i = \{\prod_k Po(n_{i0k}, E_{0k}\exp(\mu_{ik}))\} \int \{\prod_k Po(n_{i1k}, E_{1k}\exp(\mu_{ik}+u_i))\}\varphi((u_i-\theta)/\sigma)du_i$$

- **Additional treatments can be accounted for by additional factors in the likelihood**
- **Dataset with several drugs administered in several studies (not necessarily in all) has been analyzed**

Discussion and outlook

- Proposed methodology works reasonably well on the SOC level, mixture distributions for random effects are under consideration
- The random effects assumption more plausible within SOCs.
- Testing and modeling approach give consistent results – however, modeling provides additional insight (estimation)
- Implementation of dose-response models or covariates desirable
- Constant hazard assumption questionable for long-term studies
- Computational potential of algorithms as implemented in PROC NLMIXED of SAS probably not adequate for large problems
- Extension to lower level MedDRA terms might only be feasible in a stepwise manner, i.e., by dropping high level categories without signals and the corresponding lower level terms before proceeding

Acknowledgements/references

Special thanks to my colleagues Beat Neuenschwander for input and discussions during the course of this project and Bernd Mellein for providing the data and background information.

References:

Berry SM, Berry DA. Accounting for multiplicities in assessing drug safety: a three-level hierarchical mixture model. Biometrics 2004;60:418-426.

DuMouchel W. Bayesian data mining in large frequency tables, with an application to the FDA spontaneous reporting system. The American Statistician, 1999;53:177-199.