LMU LUDWIG-MAXIMILIANS-UNIVERSITÄT MÜNCHEN
IBE MEDIZINISCHE FAKULTÄT
INSTITUT FÜR MEDIZINISCHE INFORMATIONSVERARBEITUNG
BIOMETRIE UND EPIDEMIOLOGIE

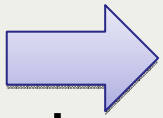
Signal Generation in Veterinary Pharmacovigilance Databases

Dipl.-Stat. Marietta Rottenkolber

Institute for Medical Informatics, Biometry and Epidemiology, Ludwig-Maximilians-Universität München, Munich, Germany

3RD INTERNATIONAL SYMPOSIUM ON VETERINARY PHARMACOVIGILANCE
Berlin
13 and 14 December 2010

Safety information from clinical trials

- Drug safety information from clinical trials is incomplete at the point of market authorization
  - Rare Adverse Drug Reactions
  - Drug-drug Interactions
  - Long-term use
  - No ‚real world' data
-  After market authorization of new pharmaceuticals is granted, the risk-benefit profile must be regularly re-evaluated.

Spontaneous Reporting System

- Spontaneous reporting data is analyzed in order to detect relevant unknown adverse drug reactions.
- Limitations of spontaneous reporting databases:
  - Number of exposed patients is unknown → no possibility to estimate incidence
  - Causality assessment not standardized
  - Numerous missing, inaccurate, or duplicate records
  - Underreporting

Data Mining

- Systematic application of statistical methods for pattern recognition

- Pharmacovigilance:
  - Systematic analysis of spontaneous reports with statistical methods
  - Calculation of a correlation score of a reaction and a drug
  - Application of this score for the detection of a safety problems

Disproportional analysis

- Evaluates the frequency of a certain drug-reaction combination based on the frequency of all other drugs and reactions in the database
- Example: In 2% of all reports erythromycin is the suspected drug and in 3% of all reports cardiac arrhythmia is the reported reaction; this means that the combination is reported in 0.06% (i.e., 0.02*0.03) of all reports
- Methods use different measures for the disproportionality
- All measured values can‘t prove the causal correlation between drug and reaction

Disproportional analysis

- Based on the following 2x2 table

Number of reports	Target reaction	Other reaction	Total
Target drug	a	b	a+b
Other drug	c	d	c+d
Total	a+c	b+d	a+b+c+d

Relative Reporting Ratio (RRR)

- Observed number of reports with the target combination **a**
- Calculation of the expected number of reports under the assumption that the target drug and target event are independent:
  - Proportion of reports, that containing reaction j: **(a+c)/(a+b+c+d)**
  - Reports with drug i: **(a+b)**
  - Expected number: **e=(a+c)(a+b)/(a+b+c+d)**

Number of reports	**Target reaction**	**Other reaction**	**Total**
Target drug	a	b	a+b
Other drug	c	d	c+d
Total	a+c	b+d	a+b+c+d

Relative Reporting Ratio (RRR)

- Relative Reporting Ratio

  a/e=a/((a+b)(a+c)/(a+b+c+d))
- Statistically instabil for small a and very small e
- Combination with RRR=100
  - a = 1000, e = 10
  - a = 100, e = 1
  - a = 10, e = 0.1
  - a = 1, e = 0.01

Proportional Reporting Ratio (PRR)

- Measure most commonly used in pharmacovigilance

	Target drug	Other drug
Target reaction	a	b
Other reaction	c	d

$$\mathrm{PRR} = \frac{a/(a+c)}{b/(b+d)}$$

Database

- German pharmacovigilance database

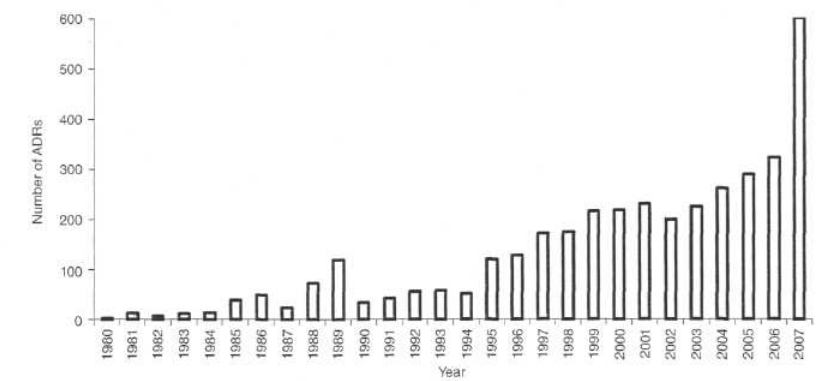


Fig. 5.1 Number of national ADRs per year reported to the BVL (without counting duplicates and follow-ups).

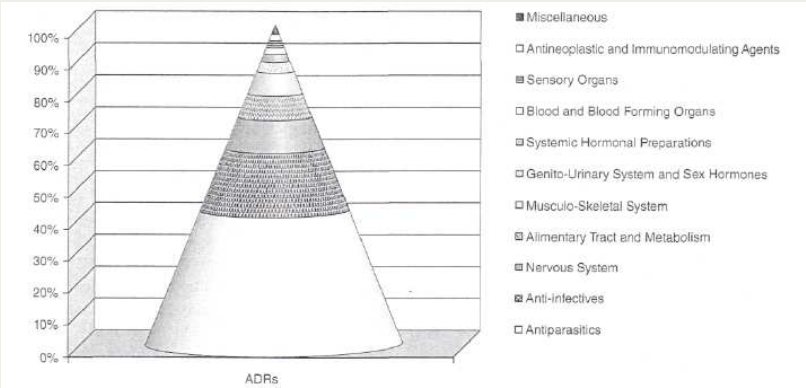


Fig. 5.4 ADRs that occurred in Germany in 2007, stratified by main ATCvet code of the products involved.

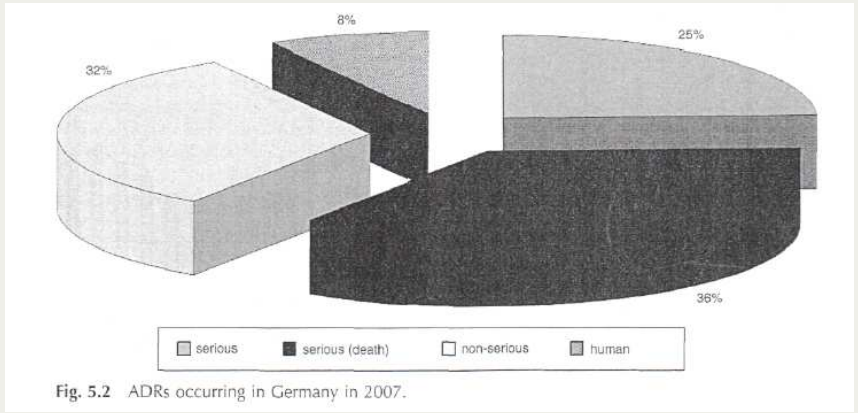


Fig. 5.2 ADRs occurring in Germany in 2007.

Source: Kevin Woodward (Editor):Veterinary Pharmacovigilance: Adverse Reactions to Veterinary Medicinal Products; Wiley-Blackwell; Chapter 5

Example PRR

	Meloxicam	Other drug
Elevated renal enzymes	**179**	**8**
Other reaction	**47**	**7,064**
Total	**226**	**7,072**

PRR= (179/226)/(8/7,072)=700.2

Reporting Odds Ratio (ROR)

	Target drug	Other drug
Target reaction	a	b
Other reaction	c	d

$$\mathrm{ROR} = \frac{a}{bc/d}$$

Example ROR

	Meloxicam	Other drug
Elevated renal enzymes	**179**	**8**
Other reaction	**47**	**7,064**
Total	**226**	**7,072**

ROR= 179/((8*47)/7,064)=3362.9

Threshold

- A value of 1 for PRR or ROR implies no correlation
- A signal occurs, if the value of PRR or ROR is greater than a predefined threshold.
  - Too low threshold: many false positive signals
  - Too high threshold: correct positive signals are missed
- The commonly used threshold for PRR and ROR is 2

Signal criteria

- According to *Evans et al.* a signal occurs if the following conditions are met:
  - Number of reports ≥ 3 and
  - PRR ≥ 2 and
  - $X^2 \geq 4$

Source: Evans et al. Pharmacoepidemiology Drug Safety, 2001:10:483-6

Pros and cons

Pros	Cons
Easy computation	PRR can only be calculated if the cells a and c ≠ 0; ROR only if the cells a, b, c, and d ≠ 0
Easy interpretation	Large databases -> Multiple comparison problem
	Difficulties to compare the results of particular combinations (is a drug with a PRR=10 and $\mathbf{X}^2$=5 more unsafe than a drug with PRR=5 and $\mathbf{X}^2$=20?)

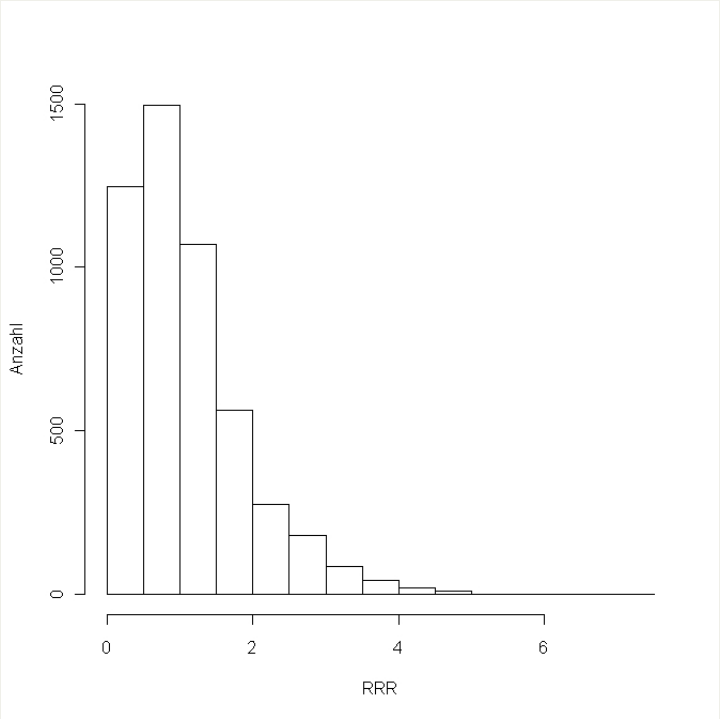
BAYESIAN PROCEDURES

Multi Gamma Poisson Shrinker (MGPS)

- Measure is the Empirical Bayes Geometric Mean (EBGM)
- Based on the Relative Reporting Ratio
- Idea: Identify combinations with $a > \frac{(a+b)(a+c)}{(a+b+c+d)}$
- First step: Computation of the RRR for each particular drug-reaction combination
  Second step: MGPS process is operated
- The drug-reaction combinations are ordered according to the difference of the estimated and expected values under independence

MGPS Process

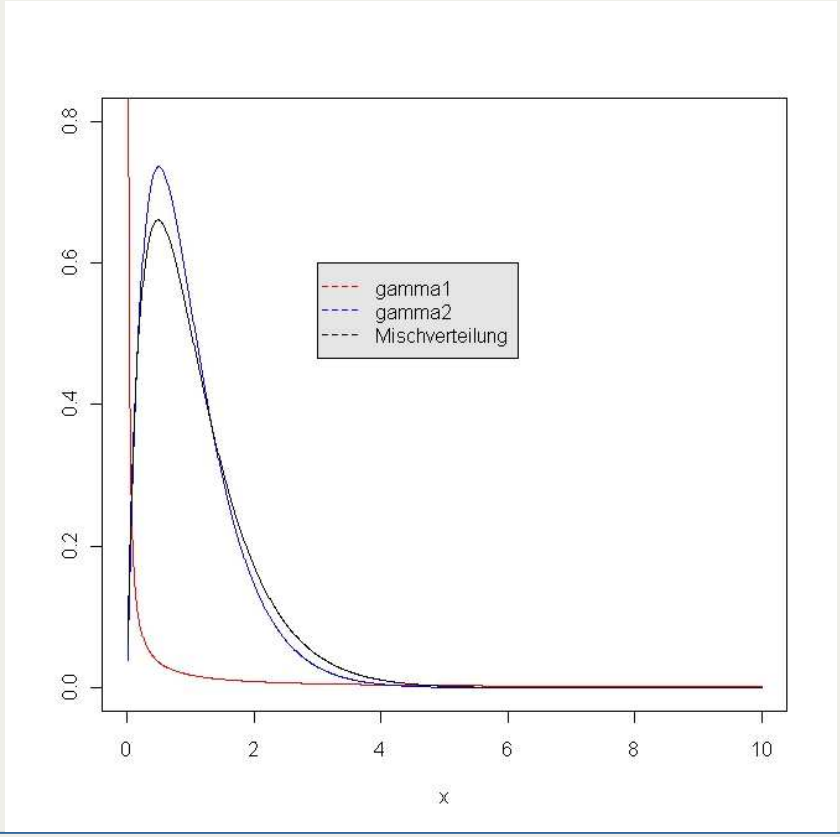
- The RRR is calculated for each particular drug-reaction combination



- N is distributed according to a Poisson density with mean $\mu=\lambda*e$

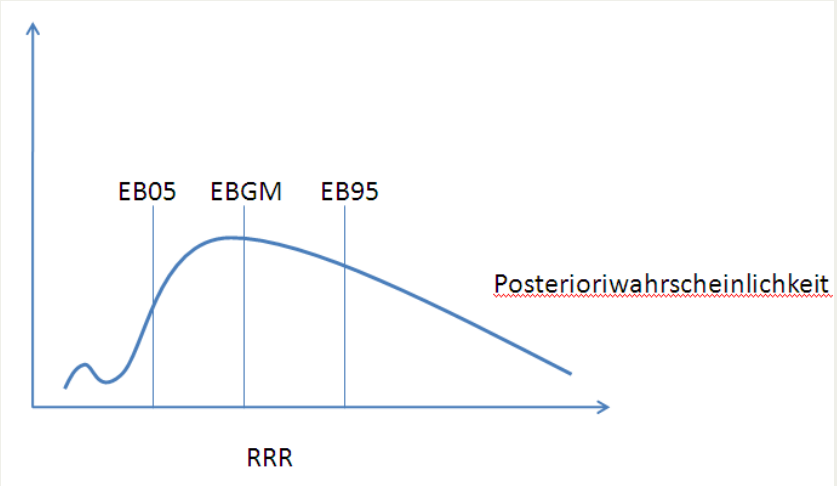
MGPS Process

- The a priori distribution for λ is a mixture of two gamma distributions



MGPS Process

- The unknown values are estimated with empirical Bayes methods
- The a posteriori density for RRR can be estimated by usage of the Bayes' therorem



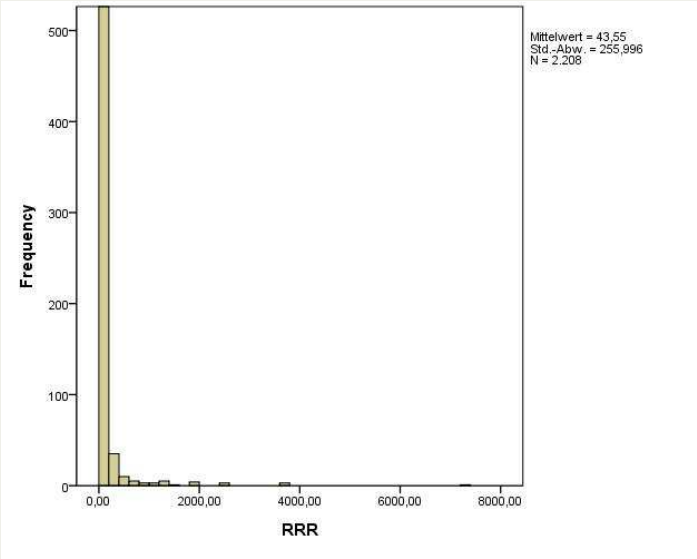
- The geometric mean of the a posteriori density is the estimator for RRR

Multi Gamma Poisson Shrinker

- EB05 and EB95 are the lower and upper limit of the 90 % confidence interval of the EBGM,
- The confidence interval can be estimated with the a posteriori density

- Interpretation
  - EBGM=6 (i.e., the drug-reaction combination occurs six times more often as expected)
  - EB05=3 (i.e., the drug-reaction combination occurs at least three times more often as expected)

Example MGPS



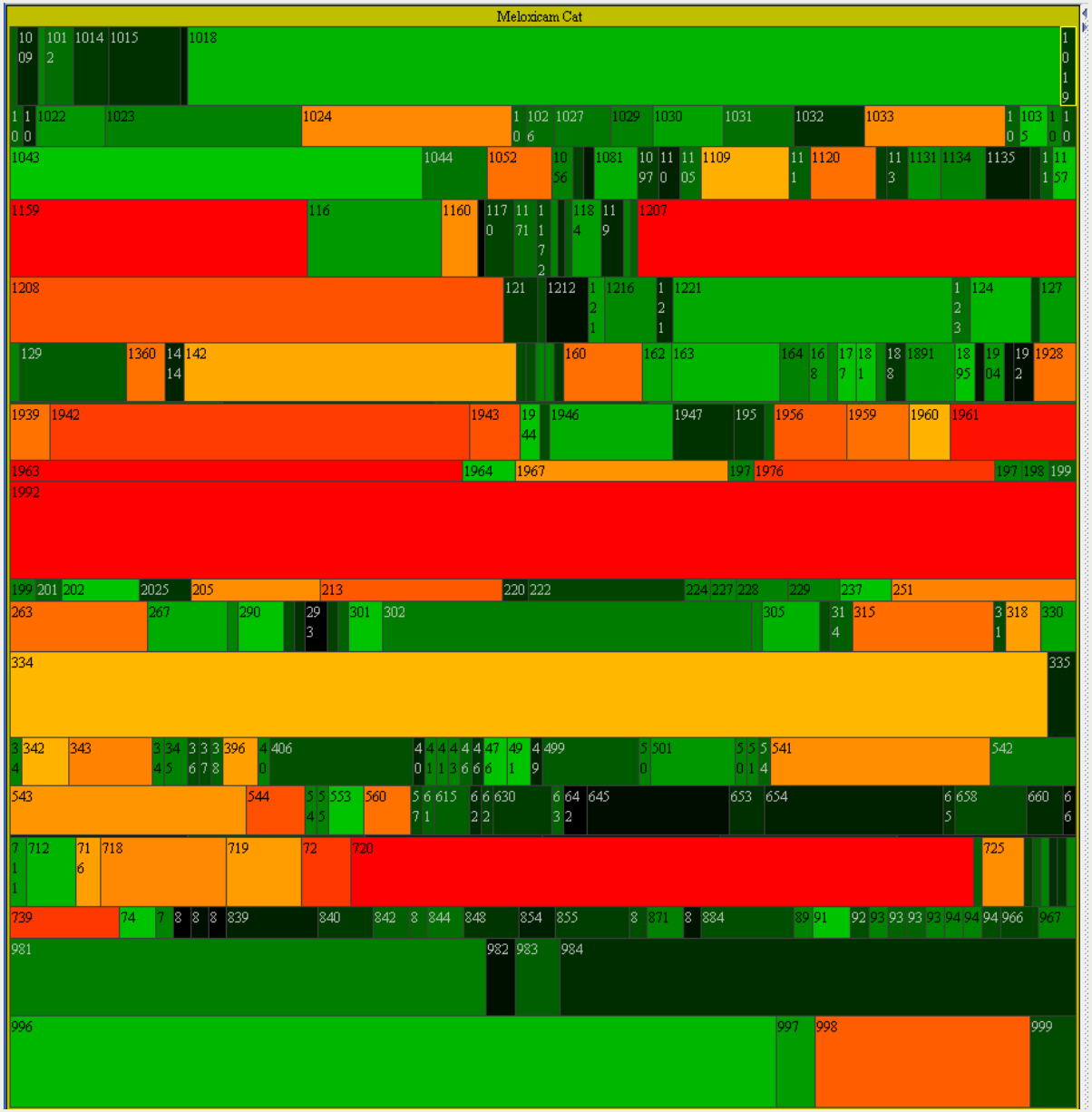
Estimated Hyperparameters	
α_1	0.0719
β_1	0.01809
α_2	0.5465
β_2	0.616
p	0.07284

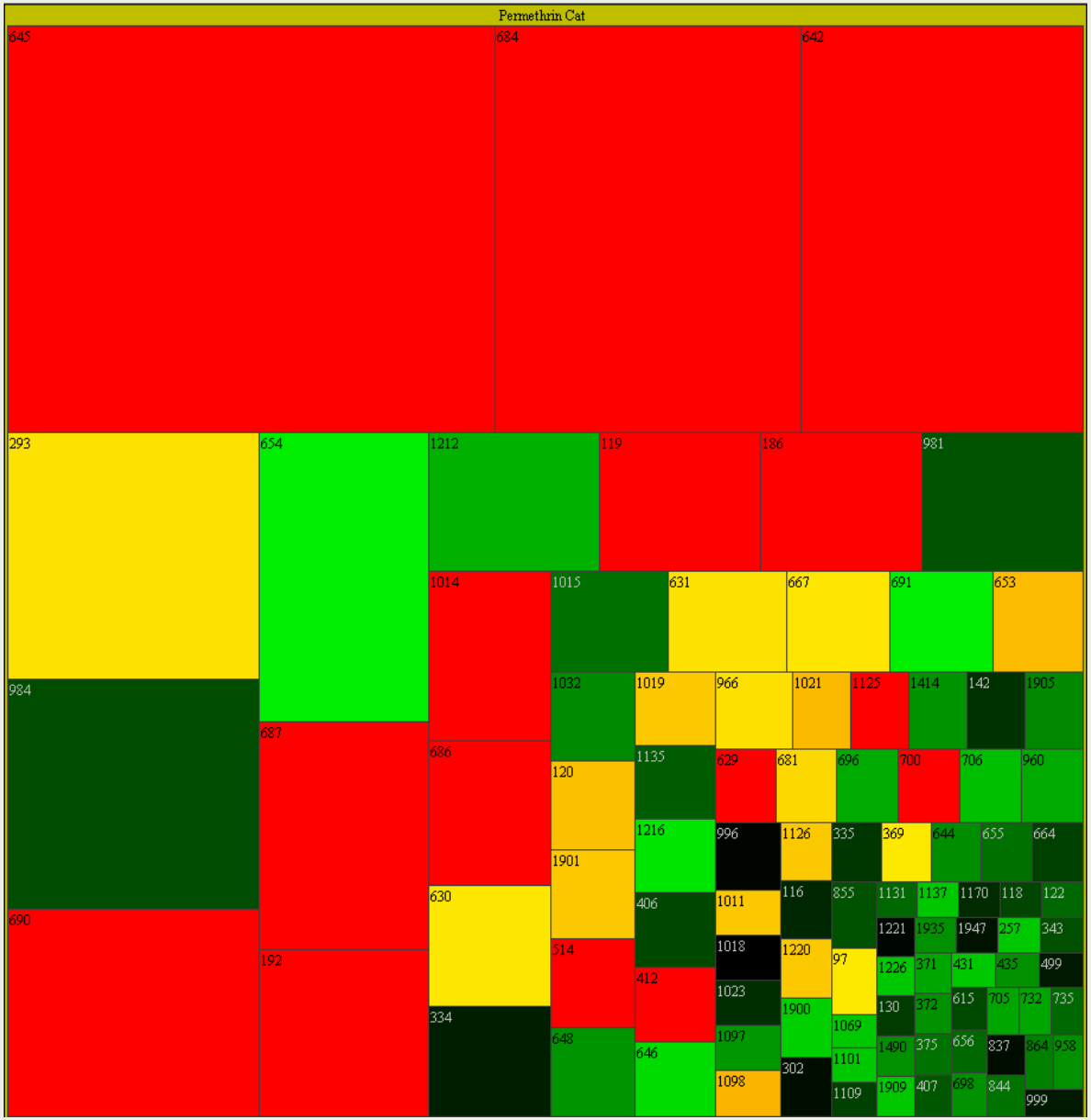
Drug	Reaction	Number	E	RR	EBGM
Meloxicam	Elevated renal enzymes	179	51,01	3,51	3.47

Graphical Visualization

- Treemap
  - Hierarchical data visualization technique of Shneiderman
- Each ADR is represented as a distinct rectangular
- The size of the rectangular is proportional to the frequency of the ADR
- The colour represents the size of the EBGM-measure (the colour black denotes EBGM = 0, green rectangulars show an EBGM from >0 to 2 the colour red means EBGM > 2)

Source: Shneiderman B. ACM Transactions on Graphics (TOG). 1992; Volume 11;Issue1:92 - 99


Meloxicam Cat


Permethrin Cat
645
684
642
293
654
1212
119
186
981
1014
1015
631
667
691
653
984
1032
1019
966
1021
1125
1414
142
1905
687
686
120
1135
629
681
696
700
706
960
1216
996
1126
335
369
644
655
664
1901
630
406
1011
116
855
1131
1137
1170
118
122
690
514
1018
1220
97
1221
1935
1947
257
343
192
412
1023
1226
371
431
435
499
334
1900
130
372
615
705
732
735
648
1097
1069
1490
375
656
837
864
958
646
1098
302
1101
1109
1909
407
698
844
999

Disproportional analysis

Pros	Limitations
No additional data are needed	Single drug-reaction combinations can be determined only
A priori hypotheses are not necessary	Sensitive to Reporting Bias
Adequate for explorative analyses	Correction for confounding by concomitant medication is impossible
	Computations are dependent from the background Rate – Masking problem
	Multiple comparison problem -> numerous false positive signals

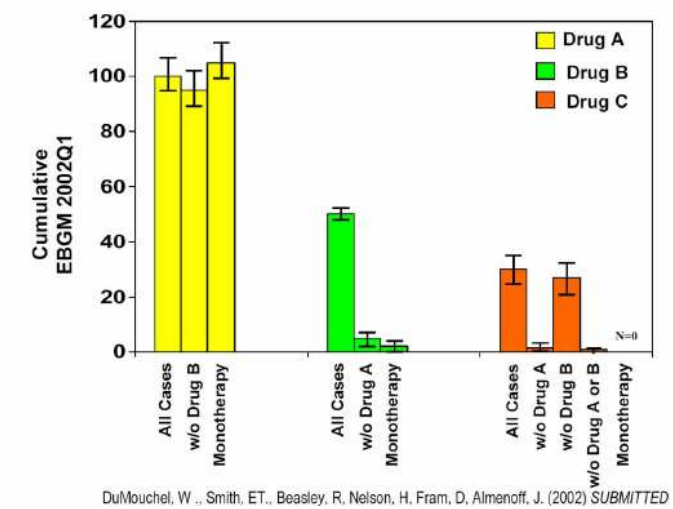
Multiple Comparison Problem and False Discovery Rate

- Concerning the multiple comparison problem the α-error can‘t be controlled -> many false positive results
- Application of the false discovery rate i.e., proportion of the false positive results on all positive results)
- Can be used with all disproportional methods.

Confounding by concomitant medication

- First approach:
  - Comparison of the signals of drug A, drug B, as well as drug A and B combined
  - Very complex if many drugs are suspected at the same time

- Second approach:
  - Application of regression models



Regression models

- Dependent variable = reaction
- As this is a binary variable, a logistic regression model is used
- For each reaction an own model is built

$$\log\left(\frac{\mathrm{P}(y_i = 1)}{1 - \mathrm{P}(y_i = 1)}\right) = \beta_0 + \beta_1 x_{i1} + \ldots + \beta_k x_{ik}$$

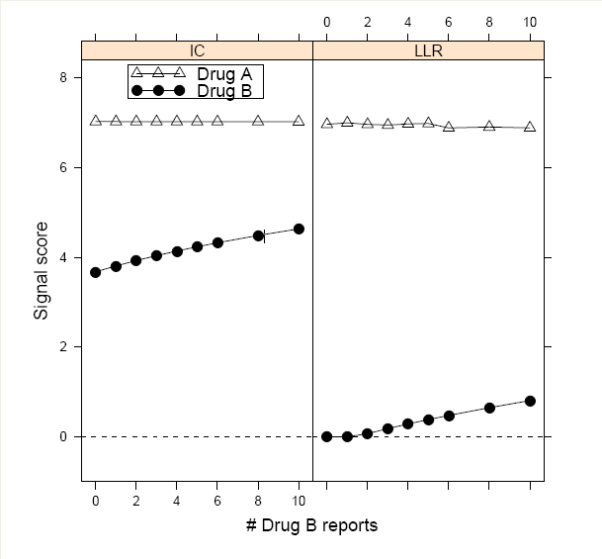
- P denotes the probability, that a report with that drug also contains the predefined reaction

Example regression

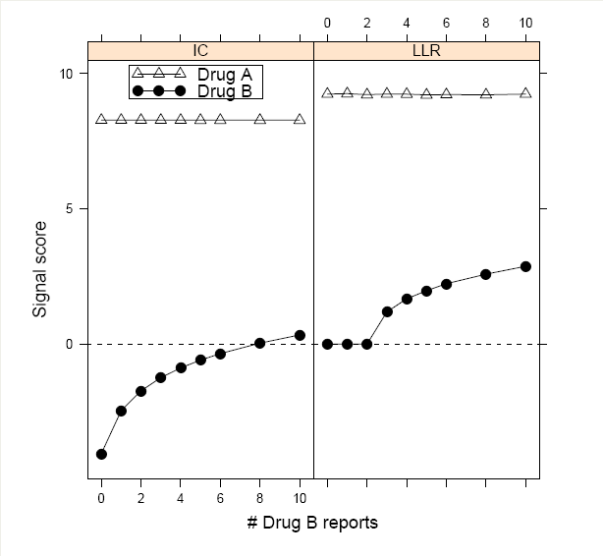
Scenario	A/X	$A, B/X$	B/X	$A/\neg X$	$B/\neg X$	$A, B/\neg X$	$\neg A, \neg B/X$ (spread)
1	100	10	0-10	1000	1000	100	1000 (100)
2	5000	0	0-10	5000	5000	0	1000 (100)

A and B are made up drugs and X is a made up ADR. The numbers indicate how many reports that were added on certain combinations in the different scenarios. The 'spread' is the number of drugs that the reports are spread out on.

Szenario 1



Szenario 2



Simulationsstudie aus O. Caster: Mining the WHO Drug Safety Database Using Lasso Logistic Regression

Conclusions

- Data mining is a frequently used approach in order to evaluate these data, to detect statistical correlations based on algorithms, and to develop new hypotheses for future research.
- Treemaps are an important improvement for a two dimensional view of an overall safety profile of a drug
- Different methods differ in terms of their results → no gold standard exists
- A causal correlation is just one explanation for the signal
- As detected signals are mere statistical correlations, it is important to validate these signals by clinical reviews and epidemiological evidence.

Literatur

- Evans et al., Use of Proportional Reporting Ratios (PRRs) for signal generation from spontaneous adverse drug reaction reports, Pharmacoepidemiol Drug Saf, 2001:10:483-6
- DuMouchel. Bayesian data mining in large frequency tables, with an application to the FDA Spontaneous Reporting System. The American Statistician. 1999;53(3):177-190
- Shneiderman B. Tree visualization with tree-maps: 2-d space-filling approach. ACM Transactions on Graphics (TOG). 1992;Volume 11;Issue1:92 - 99

LMU LUDWIG-MAXIMILIANS-UNIVERSITÄT MÜNCHEN

Thank you for your attention!