Data Mining for Drug Safety

Statistical Analyses of Spontaneous Reports and Clinical Safety Data

William DuMouchel, PhD
Chief Statistical Scientist

DIMACS Symposium
Rutgers University
February 12, 2007

Lincoln TECHNOLOGIES, INC.
A PHASE•FORWARD Company

Data Mining for Drug Safety

- Spontaneous reports
  - Disproportionality and Bayesian smoothing
  - Computer environment—Drilldown, signal management
  - Drug interaction signals
  - Logistic regression on spontaneous report databases
- Clinical trial safety data analyses
  - CDISC standard—pooling events from many trials
  - Multiple comparisons—Many events, many subgroups
  - Bayesian smoothing of rare-event probabilities
  - Searching for unexpected syndromes due to treatment
  - Subgroup analyses using Bayesian logistic regression

Databases of Spontaneous Adverse Drug Reaction Reports

- US FDA Spontaneous Report System (SRS/AERS)
  - Post-Marketing Surveillance of all Drugs since 1969
  - Renamed AERS (Adverse Event Reporting System) in 1997
    - New ADR Coding System (COSTART vs. MedDRA)
  - Version without Identifiers Available Publicly
- US FDA/CDC Vaccine Adverse Events (VAERS)
  - Stricter US Laws for Vaccine Adverse Event Reporting
- Other Databases for Medical Devices, etc.
- World Health Organization VIGIBASE
  - Includes Data from many Countries
  - ADR Coding System WHOART

Objectives and Limitations of Analyses of Spontaneous Reports

- Explore for Drug-Event Associations
  - Estimate a Measure of Association for every Combination
  - How Can a Rate Be Defined without a Denominator?
    - Matching External Sales or Prescription Counts Not Feasible
    - Construct Internal Denominators from Independence Model
  - Screening Objective – All Findings Require Follow-up
- Severe Limitations of Data Reliability
  - No Research Protocol
  - Adverse Event Report Rates Vary from Year to Year
  - Report Rates Vary by Drug and by ADR Type
  - No Certainty that a Reported Reaction Was Causal

Constructing a Denominator for N

- For every D_iE_j pair = (Drug of Interest, Event of Interest)
  - Use the database to tabulate a 2 x 2 table of report counts
  - Compute an *expected* or *baseline* count *e* from (*a, b, c, d*)
    - Based on assumption of no association between Drug and Event
    - $e = b(a + c)/(b + d)$ [Proportional Reporting Ratio method]
    - $e = bc/d$ [Reporting Odds Ratio method]
    - $e = (a + b)(a + c)/(a+b+c+d)$ [Relative Report Rate: MGPS method]
      - This method works best when adjusting for trend or demographic covariates in computation of *e*
  - n/e = Measure of Disproportionality for this Drug and Event

	Reports With Drug i	Reports W/O Drug i	Total
Reports With Event j	$n_{ij} = a$	b	$a + b$
Reports W/O Event j	c	d	$c + d$
Total	$a + c$	$b + d$	$a+b+c+d$

Disproportionality Analyses

- Although the idea of computing *n/e* ratios for all or some drug-event combinations is simple, its widespread use is very recent
  - Computer and database advances enabled ease of use and evaluation
- Biostatisticians were uncomfortable with performing formal analyses on tabulations of spontaneous reports
  - Unknown reporting mechanism can lead to reporting biases
  - Frequent noncausal associations with indications and comorbidities
  - All large values of *n/e* require follow-up for medical validity
- Small values of *n* and/or *e* require statistical sophistication
  - PRR requires threshold values of *n* and 2 x 2 table chi-squared value
  - Bayesian statistical methods produce “shrinkage” values of *n/e*
    - Help avoid the “multiple comparisons” fallacy
  - US FDA, UK MHRA and WHO UMC have each adopted Bayesian disproportionality methods

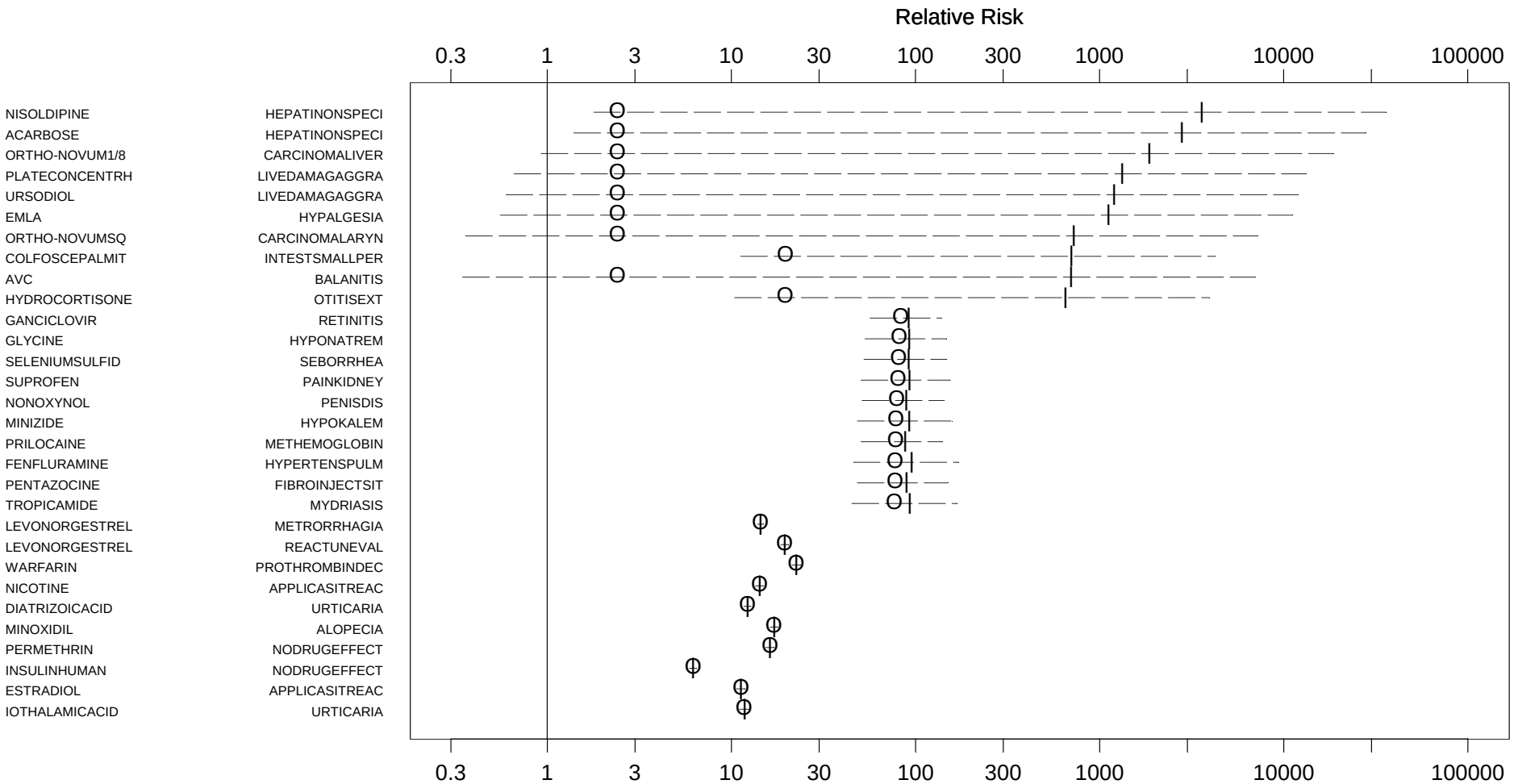
Combined Analysis of Drug-Event Counts in a Database

- Large Two-Way Table with Possibly Millions of Cells
  - One Column for each Drug, One Row for each Event
  - Rows and Columns May Have Thousands of Categories
  - Most Cells Are Empty, even though $N_{..}$ Is very Large
- "Bayesian Data Mining in Large Frequency Tables"
  - *The American Statistician* (1999) (with Discussion)
  - SRS Database with 1398 Drugs and 952 AE Codes
  - N_{ij} = Count of Reports Containing Drug *i* and Event *j*
  - Only 386 000 out of 1 331 000 Cells Have $N_{ij} > 0$
  - 174 Drug-Event Combinations Have $N_{ij} > 1000$
  - Develops and Illustrates Bayesian Estimation Method "GPS"

Empirical Bayes Gamma-Poisson Shrinker (GPS Method)

- Estimate $\lambda_{ij} = \mu_{ij} / E_{ij}$, where $N_{ij} \sim \text{Poisson}(\mu_{ij})$
- Assume Superpopulation Model for λ
  - Prior Distribution Is Mixture of 2 Gamma Distributions
  - Estimate the 5-Parameter Prior from All the (N_{ij} , E_{ij}) Pairs
- Posterior Distributions of each λ_{ij} Are Used to Create "Shrinkage" Estimates
  - EBGM = Empirical Bayes Geometric Mean of Posterior Dist.
    - Estimate of μ_{ij} / E_{ij} Has Smaller Variance than N_{ij} / E_{ij}
  - Rank Cells by $EB05_{ij}$ = Lower 5% Point of Posterior Dist.
  - More "Interesting" than Ranking Cells Based on "P-Values"
    - Compare (N = 10, E = 0.1) to (N = 2000, E = 1000)

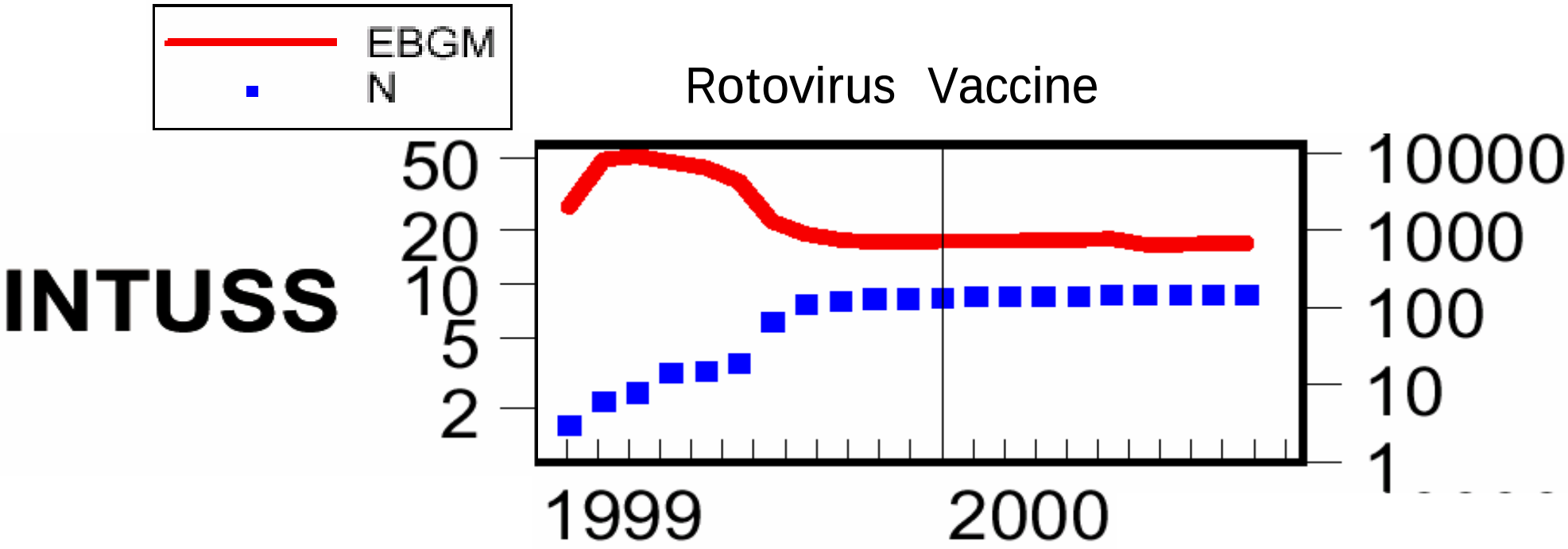
Plot of Classical Estimate with Conf. Int. and Bayesian “Shrinkage” Estimates [O]



O Empirical Bayes Geometric Mean of RR ---|--- Observed RR with 99.9% Classical Conf. Int.

Lincoln TECHNOLOGIES, INC.
A PHASE•FORWARD. Company

Example of Large Signal with Small N

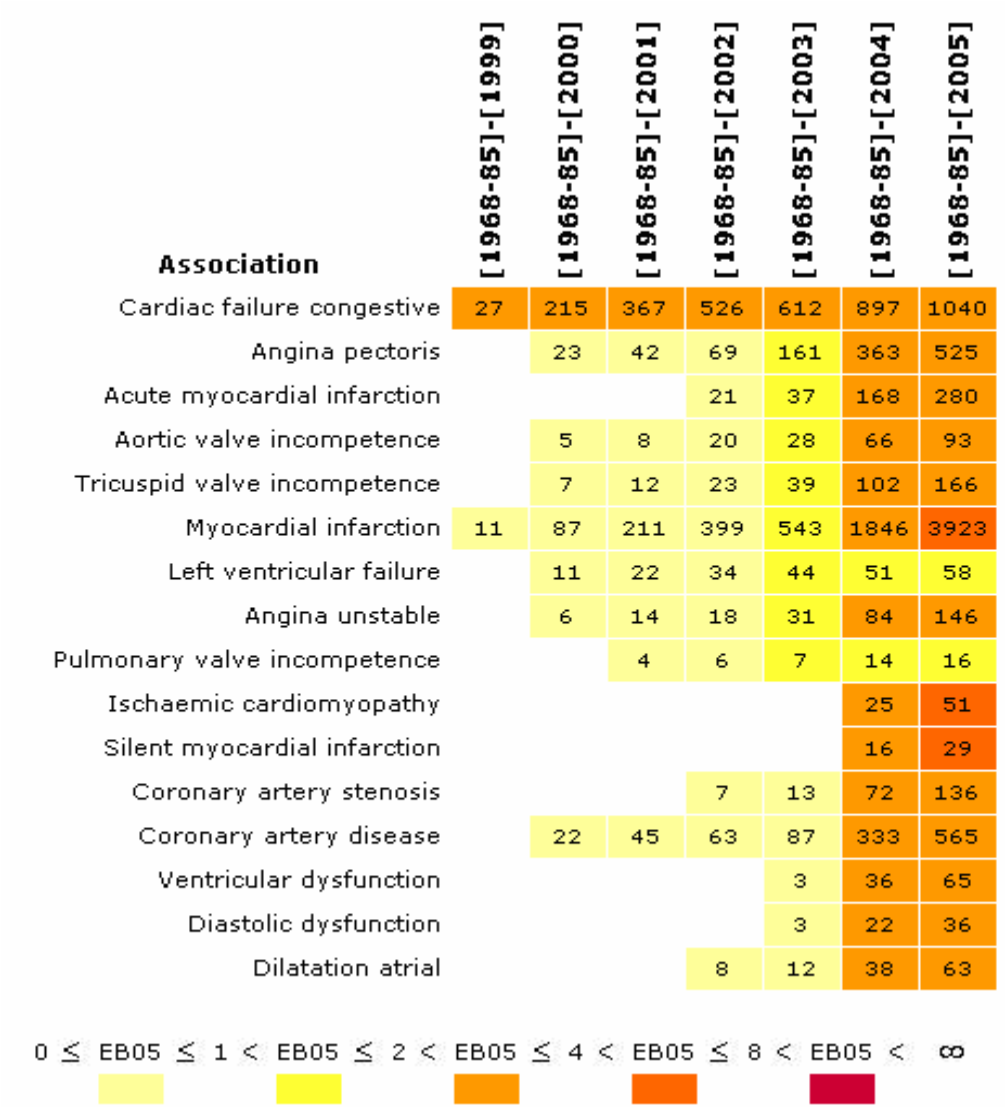


- The RV Vaccine Was Used in U.S. in 1998-99 and Was Withdrawn from the Market when the Association with Intussusception, a Severe GI Condition, Was Confirmed.

Computerized Safety Signal Analysis

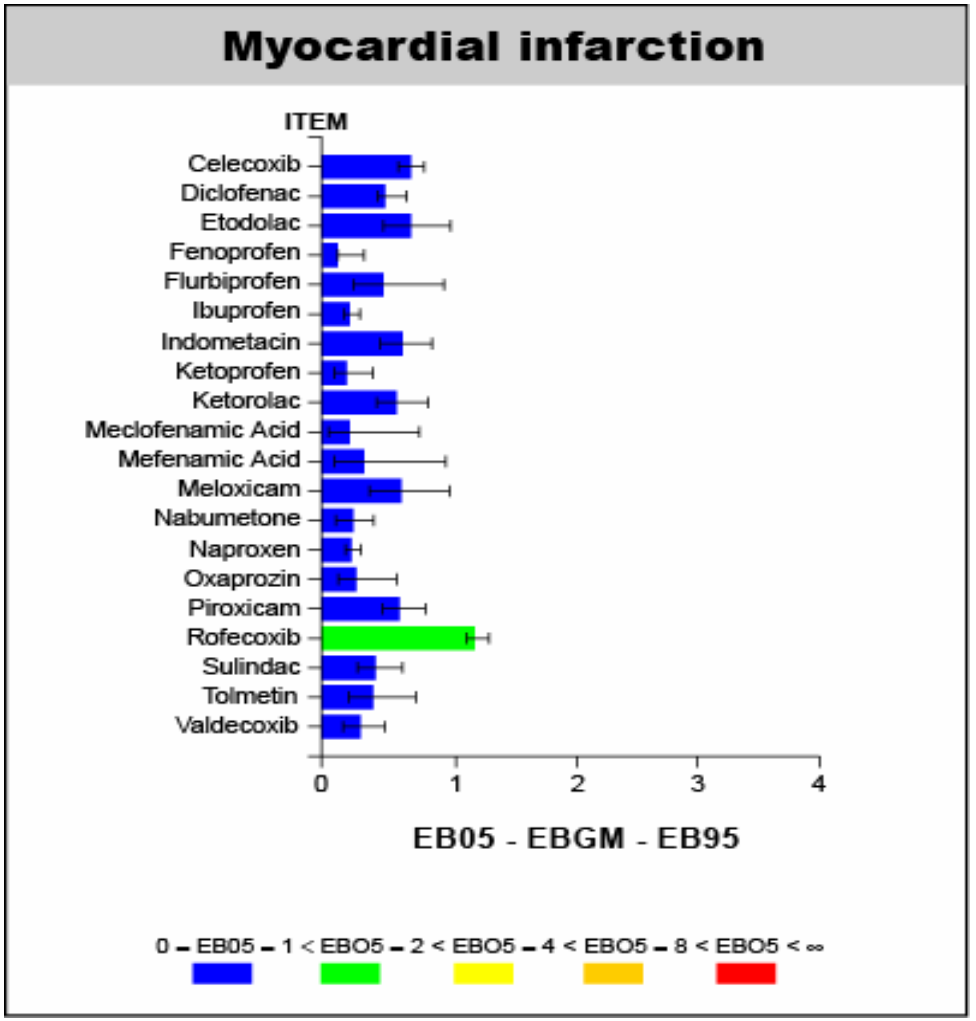
- Data mining and statistical signal detection
  - Empirical Bayes estimates of relative reporting rate (EBGM)
  - Proportional Reporting Ratios (PRR)
  - Graphical and tabular presentation
- Access to safety report data
  - Drilldown to case data from data mining associations
  - A case series feature to organize and record evaluations
  - Analyses of both public and private safety databases
- Modern web-based software
  - All processing on server
  - Simple web browser client – operates through firewalls
- In-House Data Can Be Combined with Public Databases

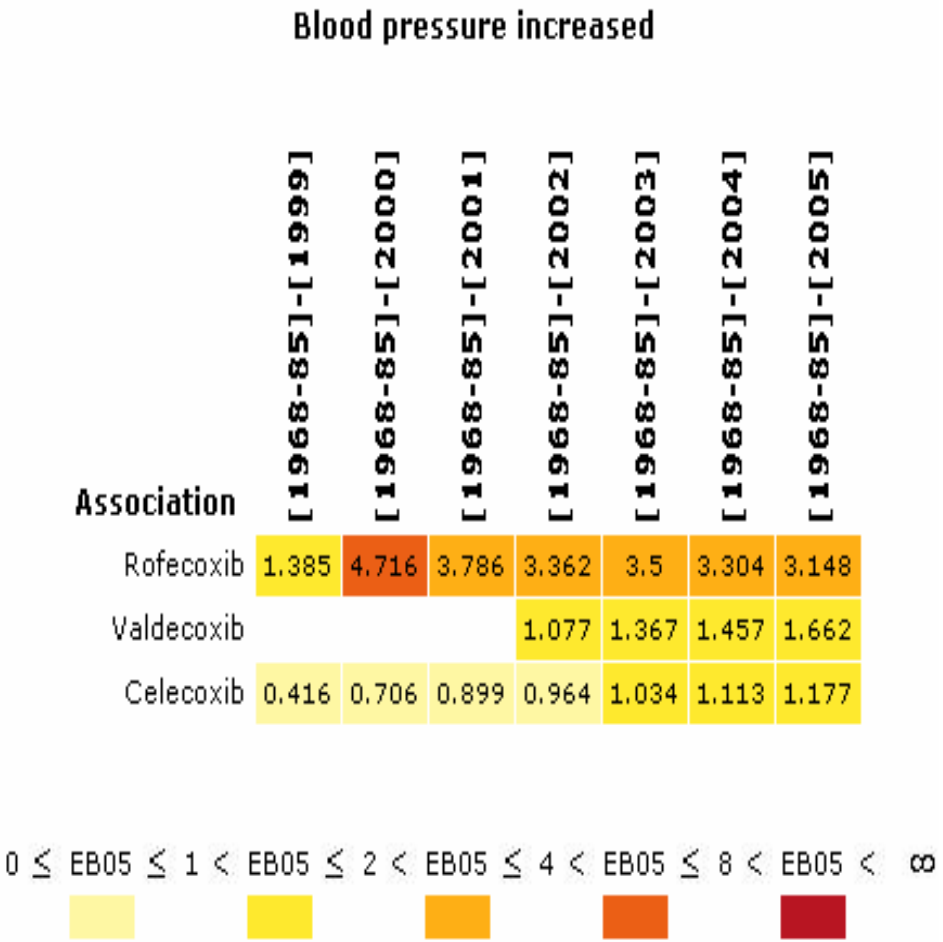
Cardiac Signals for Rofecoxib



Association	[1968-85]-[1999]	[1968-85]-[2000]	[1968-85]-[2001]	[1968-85]-[2002]	[1968-85]-[2003]	[1968-85]-[2004]	[1968-85]-[2005]
Cardiac failure congestive	27	215	367	526	612	897	1040
Angina pectoris		23	42	69	161	363	525
Acute myocardial infarction				21	37	168	280
Aortic valve incompetence		5	8	20	28	66	93
Tricuspid valve incompetence		7	12	23	39	102	166
Myocardial infarction	11	87	211	399	543	1846	3923
Left ventricular failure		11	22	34	44	51	58
Angina unstable		6	14	18	31	84	146
Pulmonary valve incompetence			4	6	7	14	16
Ischaemic cardiomyopathy						25	51
Silent myocardial infarction						16	29
Coronary artery stenosis				7	13	72	136
Coronary artery disease		22	45	63	87	333	565
Ventricular dysfunction					3	36	65
Diastolic dysfunction					3	22	36
Dilatation atrial				8	12	38	63

Comparisons of NSAIDS in AERS





AERS to 3Q03 (Suspect drugs)

Lincoln TECHNOLOGIES, INC.
A PHASE•FORWARD. Company

Importance of Studying Drug Interactions

Drugs withdrawn due to severe drug interactions

Brand Name	*Generic Name*
Baycol	Cerivastatin
Propulsid	Cisapride
Seldane	Terfenadine
Hismanal	Astemizole
Posicor	Mibefradil

Drug interactions cause up to 2.8% of hospital admissions

(Grymonpre et al. J Am Geriatr Soc 1988)

50% of elderly take 5 drugs/week; 12% take 10 drugs/week

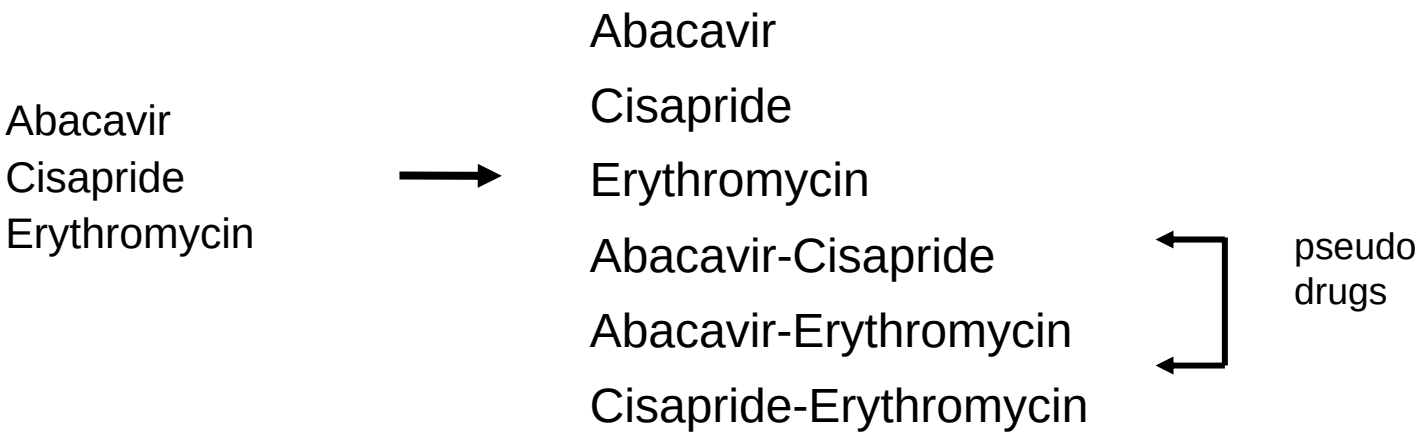
(Harvard Health Letter: March 2004)

The more medications you take,
the greater your chance of a drug interaction

Interaction Analysis Approach

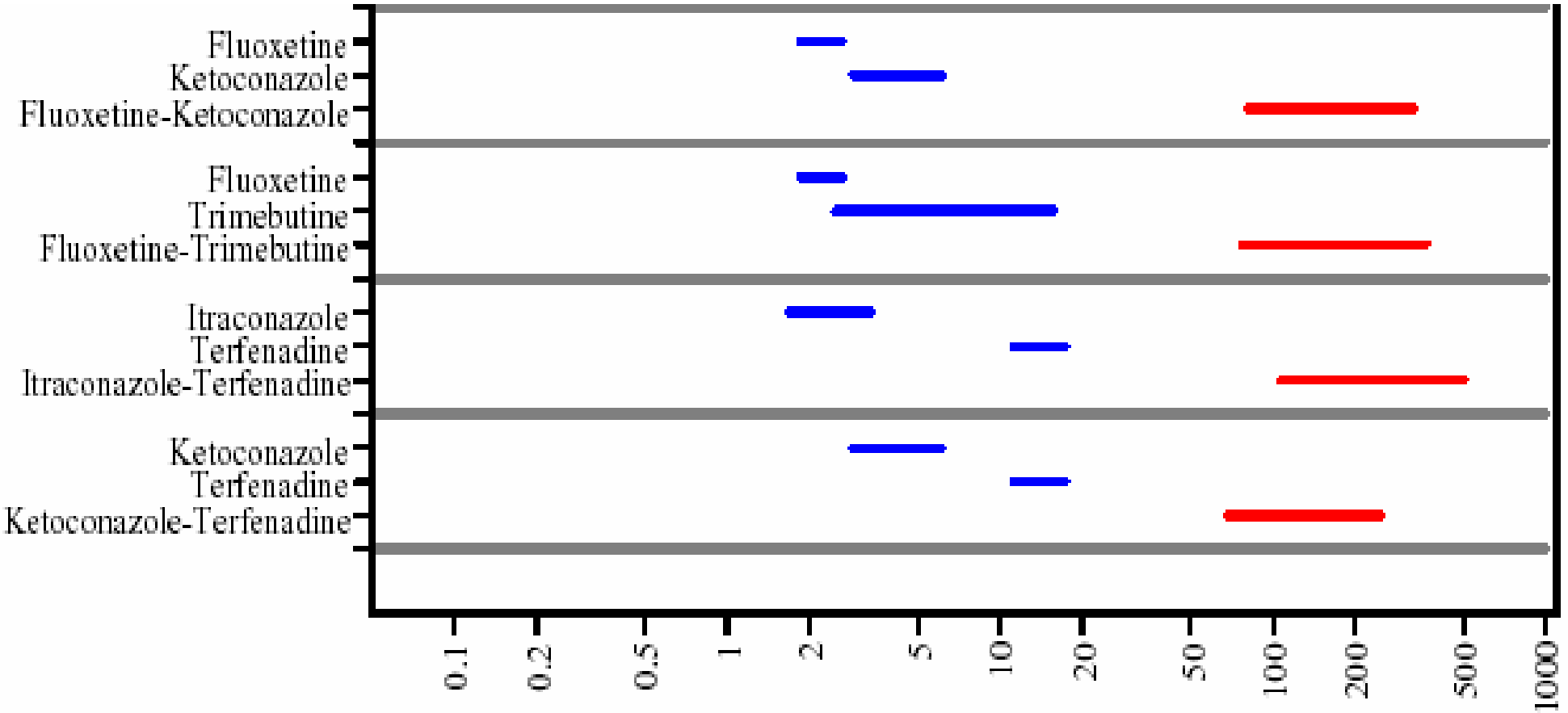
Overview of Method

- Yang, XM, *Pharmacoepidemiology and Drug Safety* 2004, 13, suppl 1: S247
- Introduce observed drug pairs as additional “pseudo-drugs”
- Reports with 3 drugs can be treated as reports with 6 “drugs”



Drug	Event	N	EBGM	PRR
Cisapride	**Torsade de pointes**	**92**	**19.525**	**69.919**
Erythromycin	**Torsade de pointes**	**58**	**20.425**	**13.227**
Cisapride-Erythromycin	**Torsade de pointes**	**18**	**228.733**	**755.355**

More Torsade de Pointes Examples



Torsade de pointes signal scores where the difference between the signal score for each drug alone and for the combination is large

Masking – Effect of Background Rate

- A relative reporting rate needs a denominator
  - This background or “noise” rate should ideally exclude effects of predictors having very large signals
- MGPS, PRR and similar methods naively assume that all reports excluding the one drug being focused on are background noise
  - The “control group” may include other drugs with very high signals for the event of interest
  - Analysis should estimate the effects of more than one drug at a time
    - 2 x 2 Table analysis is too simple

Confounding

- Unless multiple predictors are themselves uncorrelated, one-predictor-at-a-time analyses can be biased
  - GPS, PRR and similar methods don't account for effect of Drug-Drug associations on Drug-Event associations
  - Drugs that are often prescribed together can be confounded
    - Co-prescribed drugs partially inherit each other's associations
  - Synonymous terms
    - Signal leakage
    - Innocent bystander effect
- Need a multivariate methodology

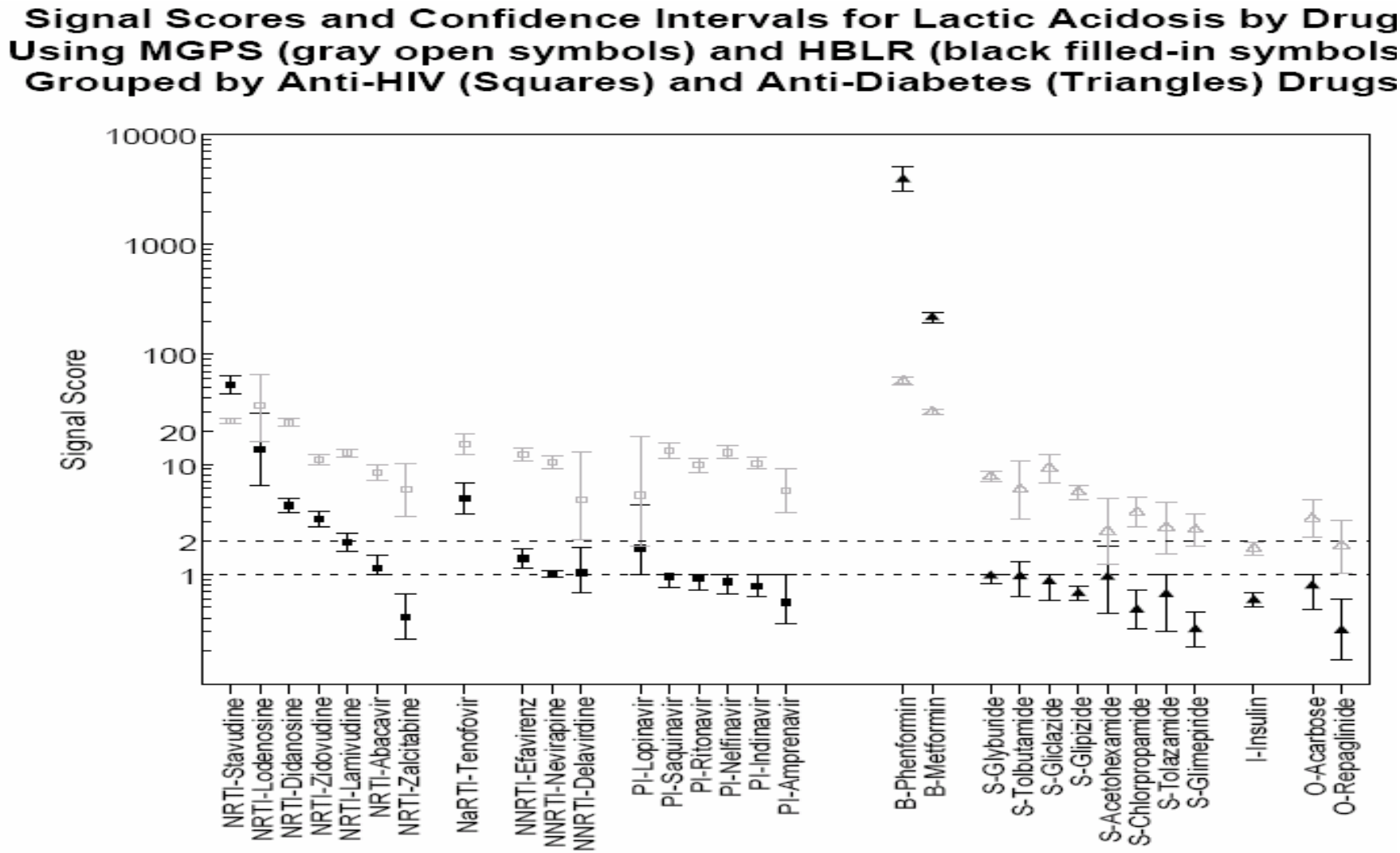
Multiple Regression Analysis

- Standard method for deconfounding predictors
  - Adverse event as response (dependent variable)
  - Stratification variables and drug presence/absence as predictors
  - Background noise rate automatically estimated
  - Can be extended to estimate drug interactions
- Several variants of multiple regression are available
  - Logistic regression most often used when response is event presence/absence
    - Assumes log[P/(1-P)] is a sum of predictor effects
    - Coefficients are interpreted as logs of odds ratios
  - Must verify assumptions of particular regression model
- Different regression fit for each adverse event
  - Each regression has coefficients for many different drugs

Shrinkage Methods for Regression

- **Put a Prior Distribution on the Coefficients (B_1, B_2, …)**
  - The Prior Distribution Moves the Estimated Bs Toward 0
    - Stabilizes the estimation when there are very many predictors
    - Bayesian solution to the multiple comparisons problem
  - Combining regressions on different responses (events)
    - Medically related events (e.g. same SOC) may have similar coefficients
    - Bayesian hierarchical model can allow “borrowing strength” across events
- **Common Ways to Choose the Prior Distribution**
  - Must Decide How Much to Shrink
  - Cross-Validation: Choose Prior to Fit Left-Out Data
  - Bayes: Fit Prior Parameters at Same Time as the Bs
  - Methods Are Well-Developed in Data Mining Literature

Szarfman (2005 FDA Science Forum)



Clinical Trial Safety Data

- Similarities with spontaneous report data
  - Could use 2 x 2 table analyses or logistic regression
  - Must cope with multiple comparisons
- Differences
  - Smaller sample sizes
  - Usually just two treatments vs. thousands in database
  - Prospective study
  - Randomized allocation of treatment
  - More valid comparison group
- Active surveillance studies
  - Large longitudinal database of medical records
  - Attempt to match users of two drugs with propensity scores
  - Maybe closer to clinical trial data than to spontaneous reports

Pooling Data Across Trials

- Combined analysis of multiple trials comparing the same treatment
  - May be the best option for studying rare adverse events
  - Analyses can adjust for varying background rate per trial
  - Similar to pooled-data meta-analyses
- CDISC data format standard for clinical trials
  - Consortium of FDA, drug companies, and software firms
  - FDA has announced eventual requirement for all NDAs
  - Several such NDAs have been submitted already
  - Some FDA reviewers are now doing partial conversions to CDISC SDTM format to allow combined safety reviews
  - Software to take advantage of the data standard available or under development by various vendors

Bayesian Shrinkage Models

- Statistical validity of searching for extreme differences
  - Most significant adverse event or patient subgroup
- Classical approach to post-hoc interval estimates
  - Maintain centers of CI at observed differences
  - Expand widths of every CI
  - Expansion is greater the more differences you look at
  - If you look at too many, the CI's are too wide to be useful
- Bayesian approach
  - Requires a prior distribution for differences
    - Can estimate it from the multiple observed differences available
  - Centers of CI's are "shrunk" toward average or null difference
    - High-variance differences shrink the most
  - Widths of CI's usually shrink a little too
  - The more you look at, the better you can model the prior dist.

Searching for Event Clusters

- An *event cluster* (associated with treatment) is a set of at least three AEs for which *all pairs* of said AEs tend to show up in Treatment patients more often than Comparator patients and also more than expected if the AEs are independent within each arm of the study
  - Defining potential syndromes by event frequency distributions rather than by theoretical medical mechanisms
- We declare a potential syndrome if *all pairs within a cluster* meet some distributional threshold
  - Syndromic Odds Ratio for 2 events (Treatment vs. Comparator)
    - SOR(E1,E2) = OR(E1*E2)/max[OR(E1), OR(E2), 1]
  - Bayesian statistical methods estimate smoothed probabilities for AEs and pairs of AEs for each arm of the studies
    - EB versions of Beta-binomial model seem to work well
  - Clustering algorithms find groups of events having high SORs

Empirical Bayes Beta-Binomial Model

- Assume K different binomial distributions
  - $N_k \sim \text{Binomial}(n_k, P_k)$ $\quad k = 1, \ldots, K$
  - $P_k \sim \text{Beta}(\beta X_{k1}, \beta X_{k2})$ $\quad$ N, n, X known; P_k unknown
- Suppose you want to shrink N_k/n_k toward $X_{k1}/(X_{k1} + X_{k2})$
  - Estimate β by maximum likelihood for beta-binomial distribution
  - Only one parameter to estimate
  - Posterior mean of $P_k = (N_k + \beta X_{k1})/(n_k + \beta X_{k1} + \beta X_{k2})$
- Various choices of X for different applications
  - $X_{k1} = p_0, X_{k2} = 1 - p_0$ $\quad$ [Shrink every N_k/n_k toward p_0]
- The shrinkage estimators are useful when many of the counts are 0 and you want to estimate odds ratios
  - Multiple comparisons protection when searching for extreme deviations

Bayes Model for Event Probabilities

- Events 1 to K with Treatment and Comparator Groups
  - n_t patients in treatment group, $(n - n_t)$ in comparator group
  - N_{kt} treatment patients with event k, $(N_k - N_{kt})$ in comparator group
  - P_k = probability of event k in treatment group [= N_{kt} / n_t ??]
  - Q_k = probability of event k in comparator group [= $(N_k - N_{kt})/(n - n_t)$??]
  - Bayes model shrinks both P_k and Q_k toward N_k / n
  - Equivalently, shrink every N_{kt}/N_k toward n_t/n
  - "Beta-binomial" Bayesian model for proportions
  - $P_k = (N_{kt} + \beta n_t/n) / (N_k + \beta n_t/N_k)$ [estimate β by EB method]
  - $Q_k = (N_k - N_{kt} + \beta (n-n_t)/n) / (n-N_k + \beta(n-n_t)/N_k)$ [same β for all k]

- Odds Ratios $OR.EB_k = P_k(1 - Q_k) / Q_k(1 - P_k)$
  - 90% confidence intervals ($OR.05_k$, $OR.95_k$)

Bayes Model for Event Pairs

- N_{jk} = Number of patients with both AE j and AE k
  - N_{jkt} and $(N_{jk} – N_{jkt})$ in treatment and comparator groups
  - P_{jk} = probability of both event j and k in treatment patient
  - Q_{jk} = probability of both event j and k in comparator patient
  - If AEs are independent, $P_{jk} = P_jP_k$ and $Q_{jk} = Q_jQ_k$
  - Another beta-binomial model to shrink N_{jkt}/n_t toward P_jP_k and yet another to shrink $(N_{jk} – N_{jkt})/(n – n_t)$ toward Q_jQ_k
- Odds Ratios for AE pairs
  - $OR_{jk} = P_{jk}(1 – Q_{jk}) / Q_{jk}(1 – P_{jk})$
- Syndromic Odds Ratio
  - $SOR_{jk} = OR_{jk} / \max(1, OR_j, OR_k)$
  - AE pairs occur together preferentially in treatment group more strongly than can be explained by single-AE associations

Finding Potential Syndromes

- Cluster AEs using distance measure $d_{jk} = 1/SOR_{jk}$
- Standard Hierarchical Clustering Methods
  - Use either complete or average linkage methods
  - Example uses average $d_{jk} < 2/3$ [SOR > 1.5]
  - Discard clusters with fewer than three AEs
- Example from a Trial with n = 902, n_t = 676
  - Most frequent 50 AEs are analyzed

EVENT	*N (Overall)*	*N (Treatment)*	*P (Overall)*	*P.EB.Treatment*	*P.EB.Comparator*	*OR.EB*	*OR.05*	*OR.95*
THIRSTPT	152	147	0.1685	0.2104	0.0431	5.92	3.38	10.37
CCFAGGPT	119	91	0.1319	0.1341	0.1253	1.08	0.74	1.58
DIZZYPT	100	69	0.1109	0.1039	0.1318	0.76	0.52	1.12
URINFREQPT	94	89	0.1042	0.1258	0.0396	3.49	1.93	6.31
DRYMOUTHPT	91	82	0.1009	0.1168	0.0532	2.36	1.39	3.99
NAUSEAPT	71	51	0.0787	0.0763	0.0859	0.88	0.56	1.39
CHESTPAINPT	66	50	0.0732	0.0737	0.0715	1.03	0.63	1.69
FATIGUEPT	65	51	0.0721	0.0745	0.0648	1.16	0.70	1.93
DYSPNOPT	62	46	0.0687	0.0683	0.0702	0.97	0.59	1.59
HEADACHEPT	60	46	0.0665	0.0676	0.0633	1.07	0.64	1.80

Potential Syndromes

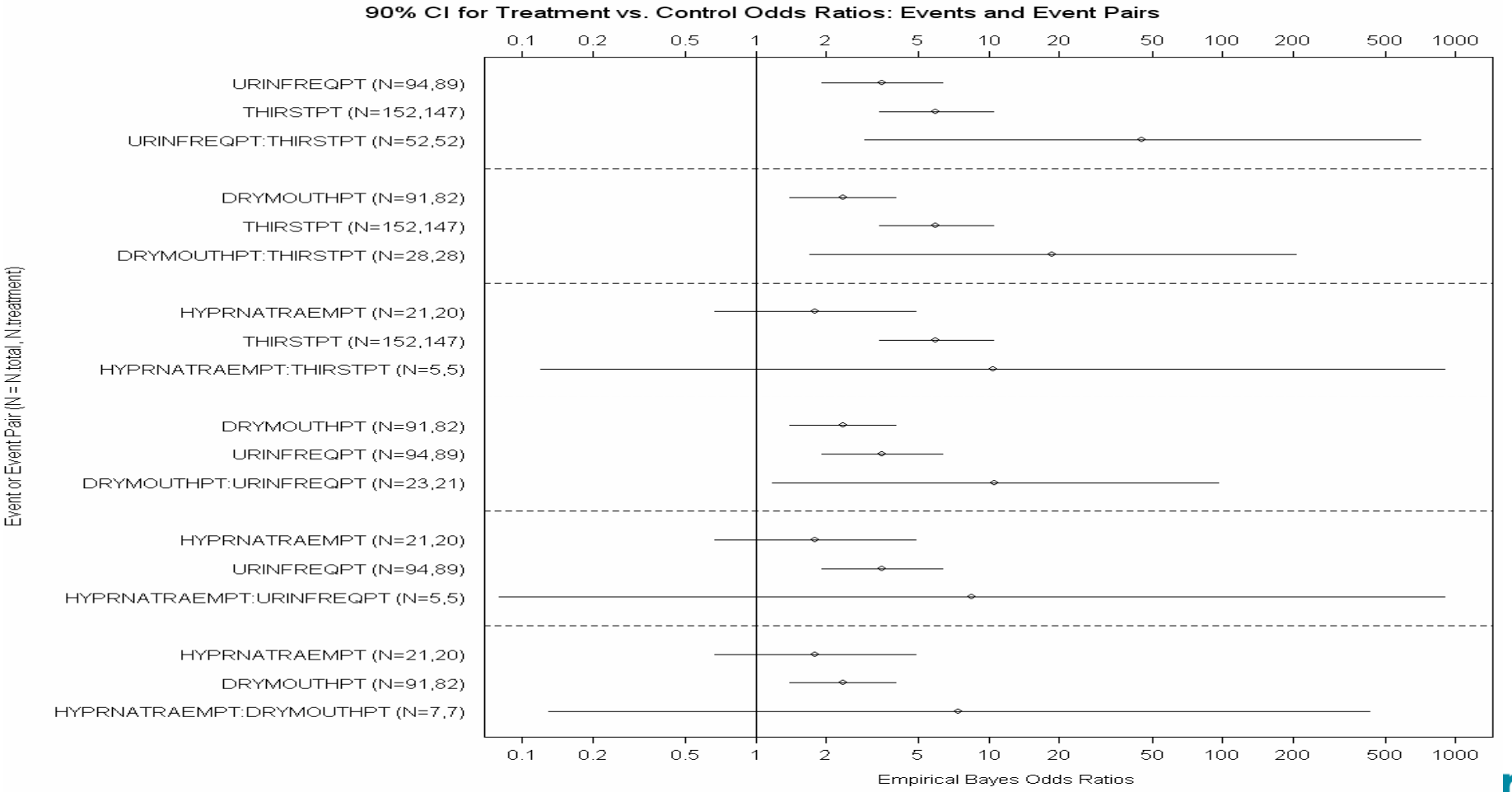
- Example from a Trial with n = 902, n_t = 676
  - Most frequent 50 AEs are analyzed
  - Example cluster cut height < 2/3 [SOR > 1.5]
- Six Clusters returned (Each cell corresponds to a *PAIR* of events)

Row	Syndrome	Counts of Event Pairs (Treatment/Comparator) for Events Belonging to Potential Syndromes.			
1	1	HPRKALPT			
2	1	6/0	RENALIMPPT		
3	1	2/0	3/0	HYPOKALAEMPT	
4	1	3/0	3/0	2/0	ANEMIAPT
5	2	THIRSTPT			
6	2	52/0	URINFREQPT		
7	2	28/0	21/2	DRYMOUTHPT	
8	2	5/0	5/0	7/0	HYPRNATRAEMPT
9	3	AGGRFTGPT			
10	3	6/1	WGTINCRPT		
11	3	4/1	3/0	ORTHOPNPT	
12	4	HEADACHEPT			
13	4	7/0	UTINFPT		
14	4	5/0	3/0	ARTHRALGIAPT	
15	5	NAUSEAPT			
16	5	8/0	DIARRHPT		
17	5	15/2	7/1	VOMITPT	
18	6	CCFAGGPT			
19	6	12/4	CONSTIPT		
20	6	16/2	8/0	INSOMPT	

4 ≤ SOR < 8
2 ≤ SOR < 4
1 < SOR < 2

Example for Cluster of 4 Events

- Note that Odds Ratio for Event Pair > Odds Ratio for each Event [902 patients, 676 Treated]
  - Syndromic OR measures how much OR for pair exceeds maximum of OR for component events
  - When SOR threshold is exceeded for all pairs within a group of 3+ events, call it a potential syndrome



Logistic Regression for Subgroup Analyses of Multiple Events

- Start from a set of *Medically Related* events to study
  - Set of events from potential signal
  - Set of events from SOR clusters (potential syndromes)
  - Set of ad-hoc events, or all events within a MedDRA SOC
- Fit Logistic Regressions to each AE as a response
  - Use exactly the same predictor model for each AE
    - Age, gender, concomitant medication, medical history, etc.
  - Include treatment and interactions with treatment as predictors
  - Generate parameter estimates for predictors and interactions
- Empirical Bayes shrinkage of estimated coefficients
  - Coefficients of each predictor borrow strength across AEs
  - Overall treatment and interaction effects shrink toward 0

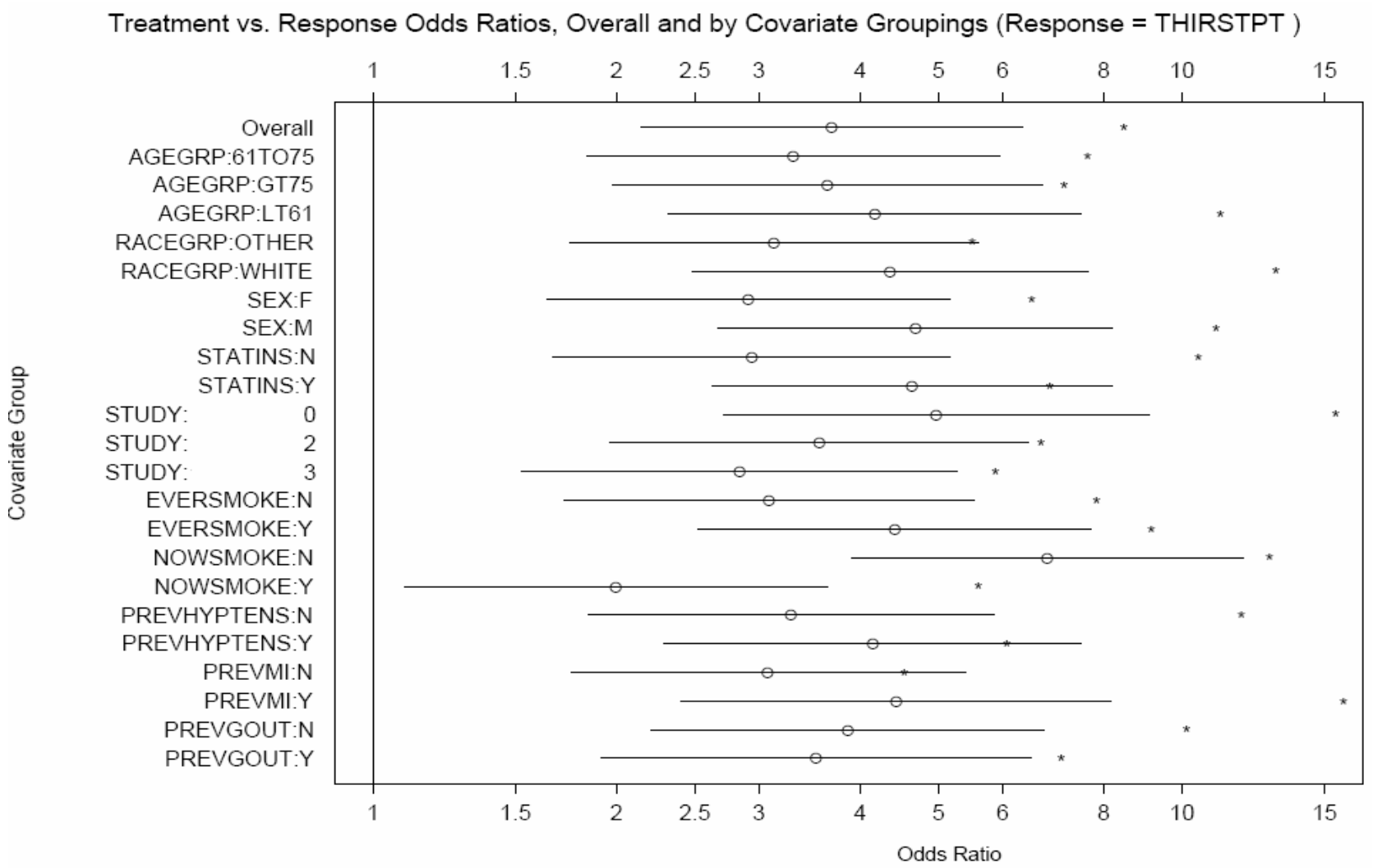
Rationale for EB Model Across Events

- Coping with fine grain of adverse event data
  - Compare T vs. C on 20 varieties of hepatic issues
  - Approach 1—separate analyses of all 20 events
    - Small counts lead to non significant comparisons
    - Adjustment for multiple comparisons further reduces sensitivity
  - Approach 2—define a single event as union of the 20 events
    - Significant differences may be washed out by the pooling
    - Even if significant, little information about original 20 differences
- Compromise approach—EB hierarchical model
  - 20 individual estimates that “borrow strength” from each other
  - Let B_{jk} = coefficient of jth treatment effect/interaction on kth AE
    - $B_{jk} \sim N(\mu_j, \sigma_j^2)$ [prior distribution shrinks AEs toward each other]
    - $\mu_j \sim N(0, \tau^2)$ [prior for overall treatment effects shrinks toward 0]
  - Estimated prior variances σ_j^2 and τ^2 control amount of shrinkage
    - Appropriate amount of shrinkage avoids multiple comparisons fallacy

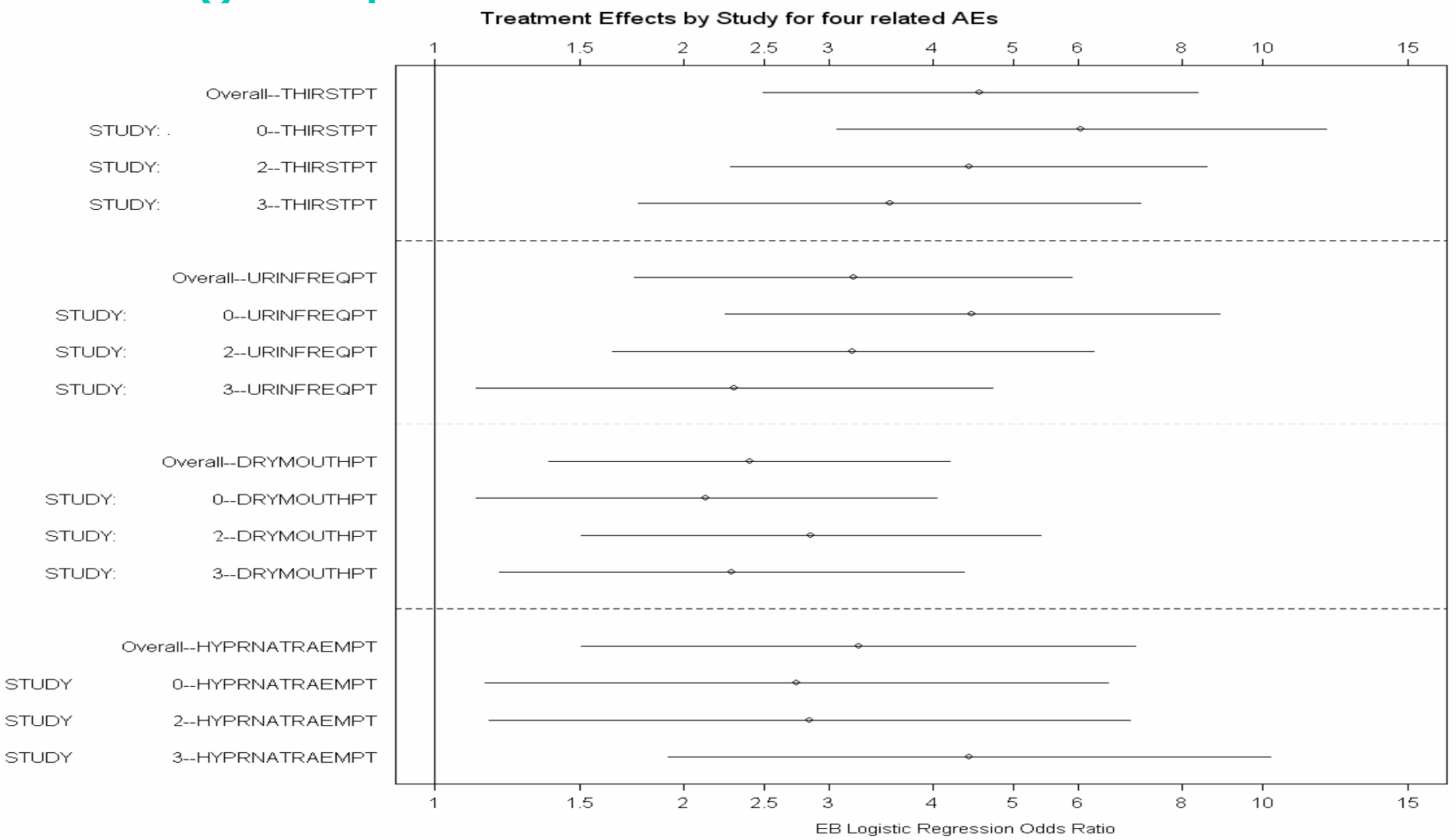
Display of Subgroup Effects

- Logistic Regression Coefficients Are Interpreted as Logs of Odds Ratios
  - Graphs of confidence intervals for each subgroup
  - Confidence intervals that do not overlap are interpreted as significant differences in subgroups
- Separate graph for each covariate and AE
  - Different layouts possible
  - Compare original and shrinkage estimates
  - Compare overall treatment effects across AEs
  - Compare subgroup effects across medically related AEs

Display of Subgroup Effects



Comparing Different AEs and Subgroups



Safety Analyses of Clinical Data

- Data mining of clinical trial safety data has many of the same challenges as analysis of spontaneous reports
  - Although the data will be cleaner, there will be less of it and the multiple comparisons issues are just as significant
- Combined analyses of multiple trials is important
  - CDISC data standards make pooling data easier
  - This is a form of pooled-data meta-analysis
- Bayesian models can be useful here too
  - Multivariate estimation of many possibly related AEs
  - Searching for potential syndromes (different AEs in the same patients) that are associated with treatment
  - Searching for subgroup effects
  - Borrowing strength across medically related AEs

Safety Data Mining References (1)

- Almenoff, J. S., W. DuMouchel, A. Kindman, X. Yang and D. M. Fram (2003). "Disproportionality analysis using empirical Bayes data mining: a tool for the evaluation of drug interactions in the post-marketing setting." Pharmacoepidemiology and Drug Safety 12(6): 517-521.
- Almenoff JA, LaCroix KK, Yuen NA, Fram D, DuMouchel W (2006) Comparative Performance of Two Quantitative Safety Signalling Methods - Implications for Use in a Pharmacovigilance Department. *Drug Safety*, 29 (10): 875-887.
- Bate A, Lindquist M, Edwards IR, Orre R (2002) A Data Mining Approach for Signal Detection and Analysis. *Drug Safety*, 25(6): p. 393-397.
- DuMouchel W (1999) Bayesian data mining in large frequency tables, with an application to the FDA Spontaneous Reporting System. *The American Statistician*, 53(3):177-190.
- DuMouchel W, Pregibon D (2001) Empirical Bayes screening for multi-item associations. in *7th ACM SigKDD Intl Conference on Knowledge Discovery and Data Mining*. San Francisco: ACM Press.
- DuMouchel W, Smith ET, Beasley R, Nelson H, Yang X, Fram D, Almenoff JA (2004) Association of asthma therapy and Churg-Strauss syndrome: an analysis of postmarketing surveillance data, *Clinical Therapeutics*, vol. 26, no. 7.
- Evans S (2002) Statistical Methods of Signal Detection in Mann R, Andrews E, *Pharmacovigilance,* Wiley, ISBN 0-470-49441-0
- Fram DM, Almenoff JS, DuMouchel W (2003) Empirical Bayesian Data Mining for Discovering Patterns in Post-Marketing Drug Safety Data Proc. ACMSIGKDD 2003 Intl. Conf. on Knowledge Discovery from Data.

Safety Data Mining References (2)

- Niu MT, Erwin DE, Braun MM (2001) Data mining in the US Vaccine Adverse Event Reporting System (VAERS): early detection of intussusception and other events after rotavirus vaccination. *Vaccine*, 19: 4627-37.
- O'Neill RT, Szarfman A (1999) Discussion: Bayesian data mining in large frequency tables, with an application to the FDA Spontaneous Reporting System by William DuMouchel. *The American Statistician* 53(3):190-6.
- O'Neill RT, Szarfman A (2001) Some FDA perspectives on data mining for pediatric safety assessment. *Curr Ther Res Clin Exp* 62:650-663.
- Solomon R, DuMouchel W (2006) Contrast Media and Nephropathy: Findings from Systematic Analysis and Food and Drug Administration Reports of Adverse Effects. *Investigative Radiology*, 41(8): 651-660.
- Szarfman A, Machado SG, O'Neill RT (2002) Use of Screening Algorithms and Computer Systems to Efficiently Signal Higher-Than-Expected Combinations of Drugs and Events in the US FDA's Spontaneous Reports Database. *Drug Safety* 25(6): p. 381-392.
- van Puijenbroek EP, Diemont WL, van Grootheest K (2003) Application of Quantitative Signal Detection in the Dutch Spontaneous Reporting System for Adverse Drug Reactions *Drug Safety* 26 (5): 293-301
- Yang XM (2004) Using disproportional analysis as a tool to explore severe drug-drug interactions in AERS database *Pharmacoepidemiology and Drug Safety* 13, supplement 1: S247