Likelihood Ratio-Based Tests for Longitudinal Safety Data

October 24, 2014

Ram Tiwari and Lan Huang

Office of Biostatistics, CDER, FDA

Disclaimer

The views expressed by the speakers of this talk are their own and do not necessarily represent those of FDA.

Useful references

1. *Lan Huang, Jyoti Zalkikar, and Ram Tiwari. A likelihood based method for signal detection with application to FDA's drug safety data. Journal of the American Statistical Association (JASA), 106 (496), 1230-1241, 2011*

2. *Lan Huang, Jyoti Zalkikar, Ram Tiwari. Likelihood ratio tests for longitudinal drug safety data. Statistics in Medicine, 33(14), 2408-2424, 2014*

3. *Lan Huang, Ted Gou, Jyoti Zalkikar, Ram Tiwari. A review of statistical methods for safety surveillance. Therapeutic Innovation & Regulatory Science, 48 (1), 98-108, 2014*

Outline

- Basic LRT method for large post-market safety database
- OB in-house tool development and illustration
- Longitudinal LRT method for data with exposure information
- Discussion

Background

- Large databases: AERS, Vigibase, MAUDE
- Data mining methods (frequentist, Bayesian, OMOP/IMEDS)
- Objective of the safety exploration
  - Signal detection in large safety database
  - Clinical trials database
  - Passive/active

IxJ Safety Data-Matrix for AERS Database

Drugs

AEs

	1	…	j	…	J	**Row total**
1	n_{11}	…	n_{1j}	…	n_{1J}	$n_{1.}$
2	n_{21}	…	…	…	n_{2J}	$n_{2.}$
…	…	…	…	…	…	…
i	…	…	n_{ij}	…	n_{iJ}	$n_{i.}$
…	…	…	…	…	…	…
I	n_{I1}	…	n_{Ij}	…	n_{IJ}	…
Col. total	$n_{.1}$	…	$n_{.J}$	…	$n_{.J}$	$n_{..}$

2x2 Table of nij

	$Drug_j$	Other drugs	
AE_i	n_{ij}	Subtracted	$n_{i.}$
Other AEs	Subtracted	Subtracted	Subtracted
	$n_{j.}$	Subtracted	$n_{..}$

- Fix a drug, say Drug j, and construct a 2X2 table for each AE: If there are, say, 16,000 AEs, then there are 16,000 such 2x2 tables
- Most of the frequentist's methods and some Bayesian methods work with 2X2 tables

Statistical model and hypothesis

$$n_{ij} \sim Pois(n_{i.} p_i)$$

$$n_{.j} - n_{ij} \sim Pois((n.. - n_{i.}) \times q_i)$$

$H_0 : p_i = q_i = p^*$ for all AEs, i

$H_a : p_i > q_i$ for all at least one AE, i

RR=p/q, AE i vs. other AEs

$$LR = \max L_a / \max L_0$$

Likelihood Ratio Test (LRT) statistic

$$LR_{ij} = \frac{L_a(\hat{p},\hat{q})}{L_0(\hat{p}^*)} = \left[\frac{\left(\frac{n_{ij}}{n_{i.}}\right)^{n_{ij}} \left(\frac{n_{.j}-n_{ij}}{n_{..}-n_{i.}}\right)^{n_{.j}-n_{ij}}}{\left(\frac{n_{.j}}{n_{..}}\right)^{n_{.j}}} \right],$$

- Test Statistic is MaxLR=max of LR_ij (i=1 to I) over i=1,…,I AEs.
- LogLR and MaxLogLR can be used for faster computation

Re-parametrization and adjustment

$$LR_{ij} = \left(\frac{n_{ij}}{n_{i.} \times \frac{n_{.j}}{n_{..}}} \right)^{n_{ij}} \left(\frac{n_{.j} - n_{ij}}{(n_{..} - n_{i.}) \times \frac{n_{.j}}{n_{..}}} \right)^{n_{.j} - n_{ij}} = \left(\frac{n_{ij}}{E_{ij}} \right)^{n_{ij}} \left(\frac{n_{.j} - n_{ij}}{n_{.j} - E_{ij}} \right)^{n_{.j} - n_{ij}}$$

$$E_{ij} = \frac{n_{i.} n_{.j}}{n_{..}} = n_{i.} \frac{n_{.j}}{n_{..}}$$

To adjust for a covariate (such as age or gender)(stratified analysis), we simply calculate the age-adjusted or gender adjusted expected cases. We first calculate the E_ijk, k=1,2 (by gender), then we combine them together.

$$E_{ij} = \sum_k E_{ij}^k = \sum_k [n_{i.}^k \times \frac{n_{.j}^k}{n_{..}^k}]$$

Theory behind the multinomial simulation for the null data

Assume that the marginal totals n1., …, nI., are fixed. Under H0, assume that n1j, …, nIj are ind distributed as

$$n_{1j} \sim Poisson(n_{1.}\, p).$$

....

$$n_{Ij} \sim Poisson(n_{1.}\, p),\ p > 0, unknown.$$

Then,

$$(n_{1j},...,n_{Ij}) \mid n_{.j} \sim mult(n_{.j}, (\frac{n_{1.}}{n..},...,\frac{n_{I.}}{n..})).$$

Hypothesis testing

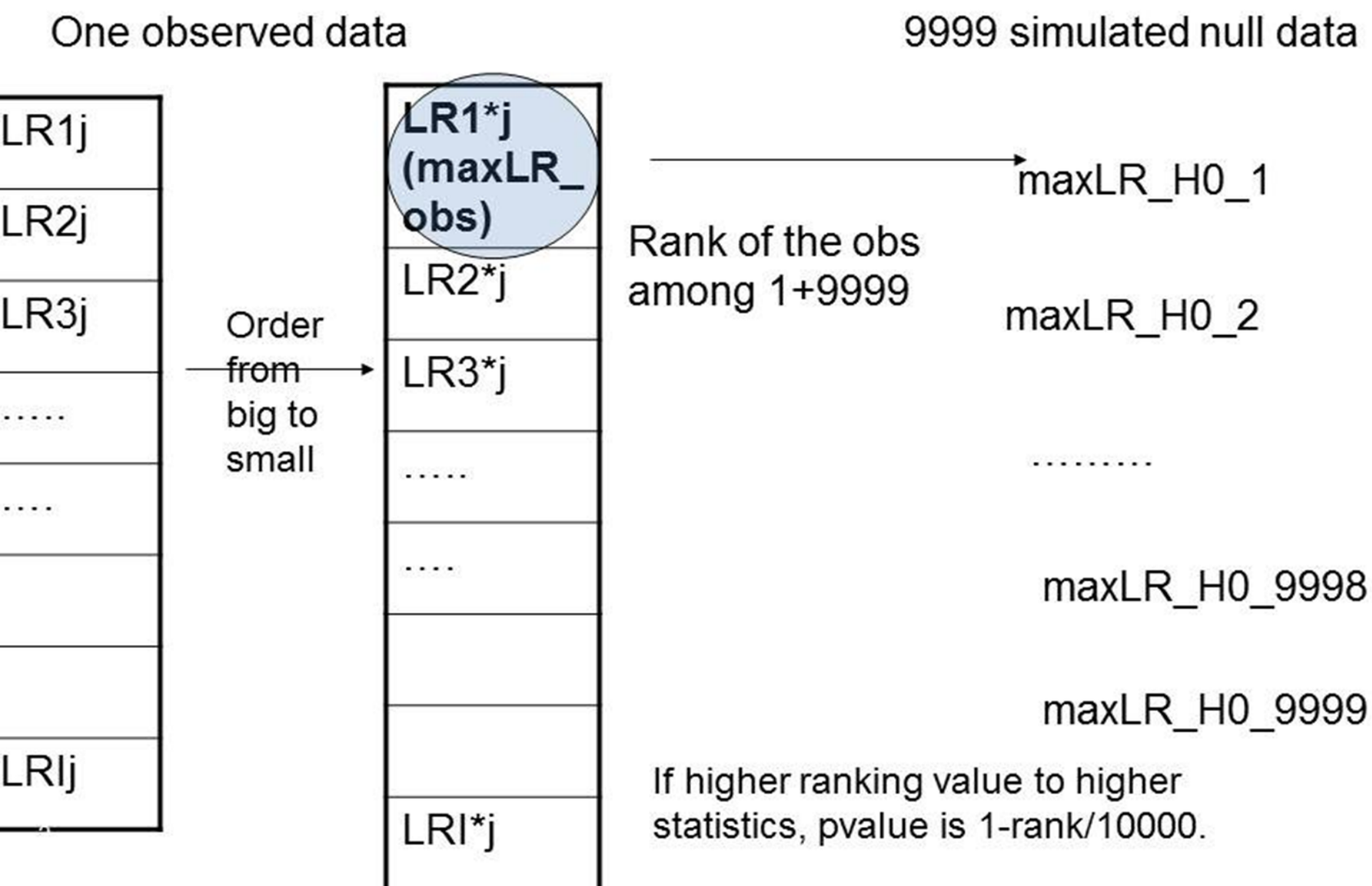
- The distribution of MaxLR under H0 is not analytically tractable, we use Monte Carlo method to obtain the empirical dist.

- Cases can be generated using multinomial distribution (n.j, (n1./n..), (n2./n..), …., (nI./n..)) assuming homogeneous reporting rate.

P-value calculation

- Calculate MaxLR from observed data (one)
- Calculate MaxLRs from the 9,999 simulated null data.
- Threshold at alpha=0.05 is 95 percentile of the 10,000 (=1+9999) MaxLRs.
- Reject H0 if obs MaxLR> threshold.
- Compare the observed MLR and the ones from simulation ---→ p-value = P(MLR> obs MaxLR)= Max # of times simulated MaxLR> obs MaxLR /10000.
- Gatekeeping step-down process (1^{st}, 2^{nd}, 3^{rd}, …)

LRT process for a drug (jth column)



OB in-house tool for LRT method

Signal Detection Using LRT Method

New Analysis | Create Report

Know what you are doing

New Analysis
You can start a new analysis using LRT. You are about to enter parameters for the analysis. Think about what time period you want to focus on, what data set you want to analyze, and what drugs or AEs you are interested in, and so on so forth.

Create Report
You might want to come back to see what the analysis results look after you submitted the last run of LRT. Normally, after you submit a LRT run, you take a short break or take care of other things while the time-consuming computation runs by itself, which involve a lot of simulations.

LRT: The Right Tool for the Righ Job

Quick Start Guide

Developed by Ted Guo, Lan Huang, Jyoti Zalkikar and Ram C. Tiwari of Office of Biostatistics/CDER/FDA, this tool (software solution) named LRT is aimed to help FDA's medical and statistical reviewers detect potential adverse events that are statistically associated with the drug of interest. In doing so, the false-positive rate is controlled at the 5% significance level.

This tool also help detects harmful drugs associated with the adverse event of interest. The false positive rate is controled in the same manner.

User: Please follow the instructions on the next screen. You will be able to define and fine tune your own analysis.

Example (Myocardial infarction)

PT	#Drug	n.j	PRR025 (>1)	sB05 (>2)	BCPNN025 (>0)	EB05 (>2)	LRT (p<0.05)
Myocardial infarction	1416	26,848	242	36	137	35	51

N	#Drug (Generic)	Nij	PRR025 (>1)	LRT (P<0.05)	sB05 (>2)	BCPNN025 (>0)	EB05 (>2)
1	Rosiglitazone	2231	✓	✓	✓	✓	✓
2	Metformin And Rosiglitazone	322	✓	✓	✓	✓	✓
3	Calcium Chloride And Glucose And Magnesi	637	✓	✓	✓	✓	✓
4	Clopidogrel	419	✓	✓	✓	✓	✓
5	Rosuvastatin	398	✓	✓	✓	✓	✓
6	Atorvastatin	506	✓	✓	✓	✓	✓
7	Calcium Chloride And Icodextrin And Magn	150	✓	✓	✓	✓	✓
8	Ticagrelor	109	✓	✓	✓	✓	✓
9	Glimepiride And Rosiglitazone	46	✓	✓	✓	✓	✓
10	Glyceryl Trinitrate	175	✓	✓	✓	✓	✓

LRT to longitudinal LRT (Motivation)

LRT	Longitudinal LRT
Count data	Count data with exposure information
Large post-market observational safety data	Observational or clinical trial data
Drug signals for one AE Or AE signals for one drug	Same
Multiple AEs and drugs	Same
Fixed time analysis	Same
Analysis over time using cumulative count data without planned alpha control	Use alpha-spending for analysis over time
covariate adjustment by stratification	same

Longitudinal LRT method (sequential LRT) for active surveillance

- General
  - Compare multiple AEs by drug
  - Compare two drugs for one AE of interest (1st occurrence or without recurrence)
  - Compare multiple drugs for one AE of interest (may have recurrence or combined AE terms)

- Control error rates and false discovery rate (FDR)

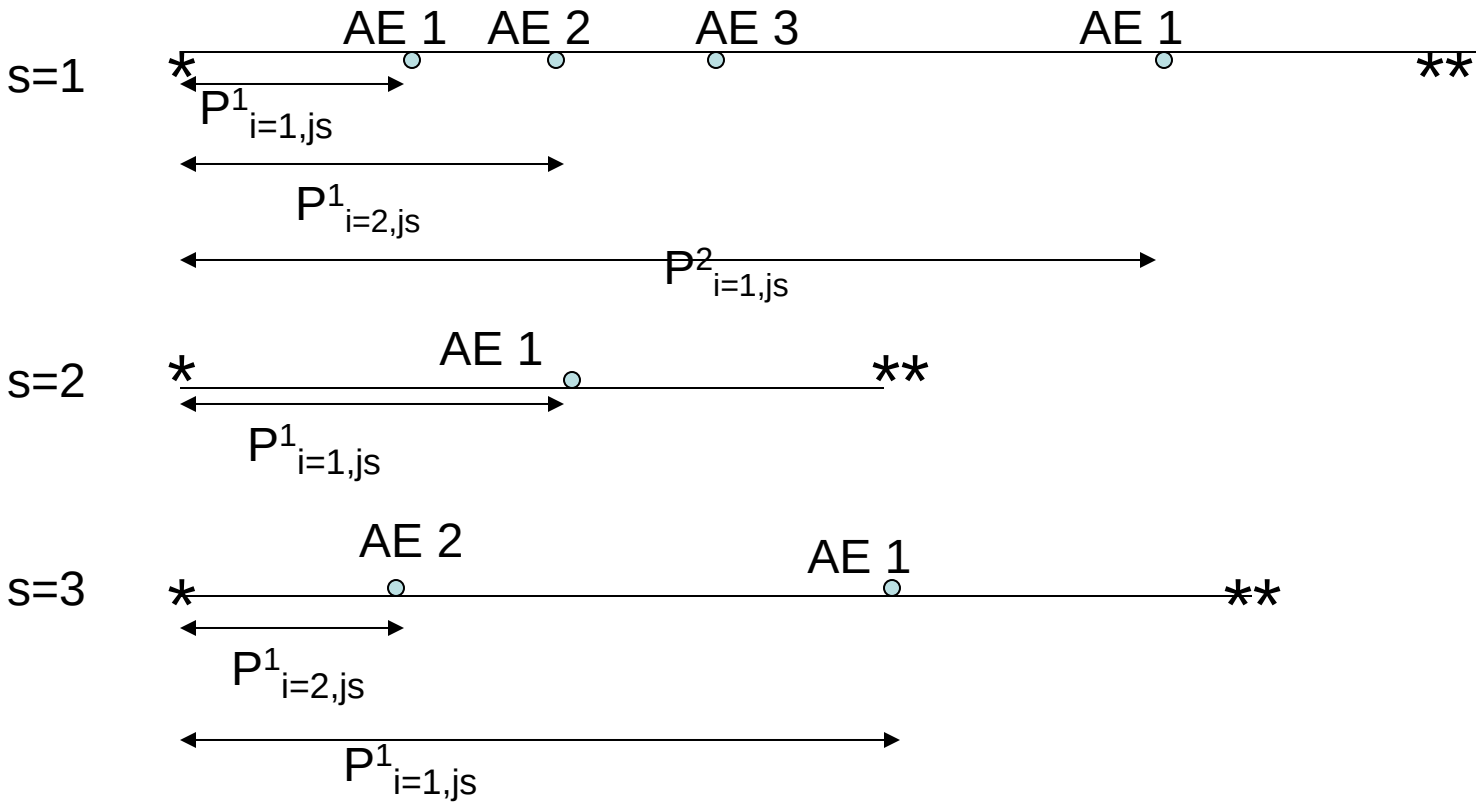
Define countable cases and drug exposure

- Countable cases: AEs that occur during the exposure period (other definitions: AEs occur several days after the drug exposure)

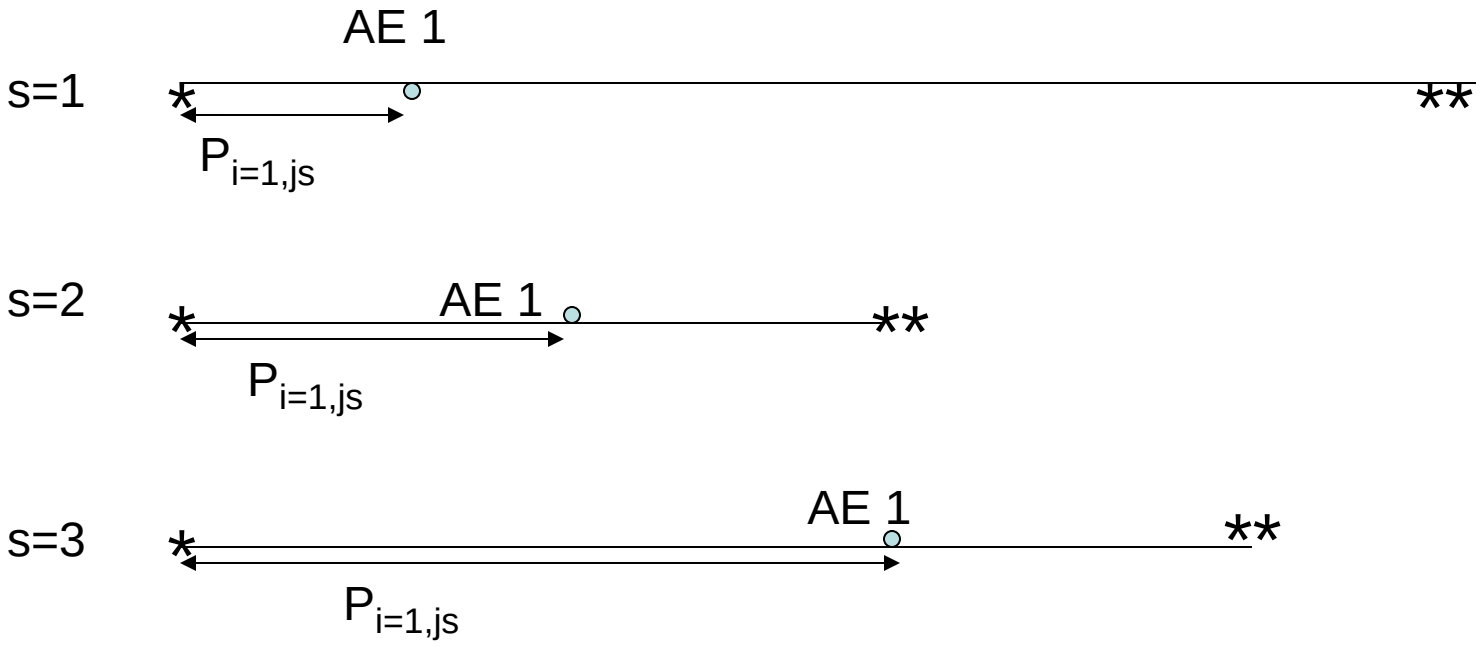
- Drug exposure
  - Event-time
  - Person-time
  - Exposure-time

- Time
  - calendar time
  - time after drug exposure

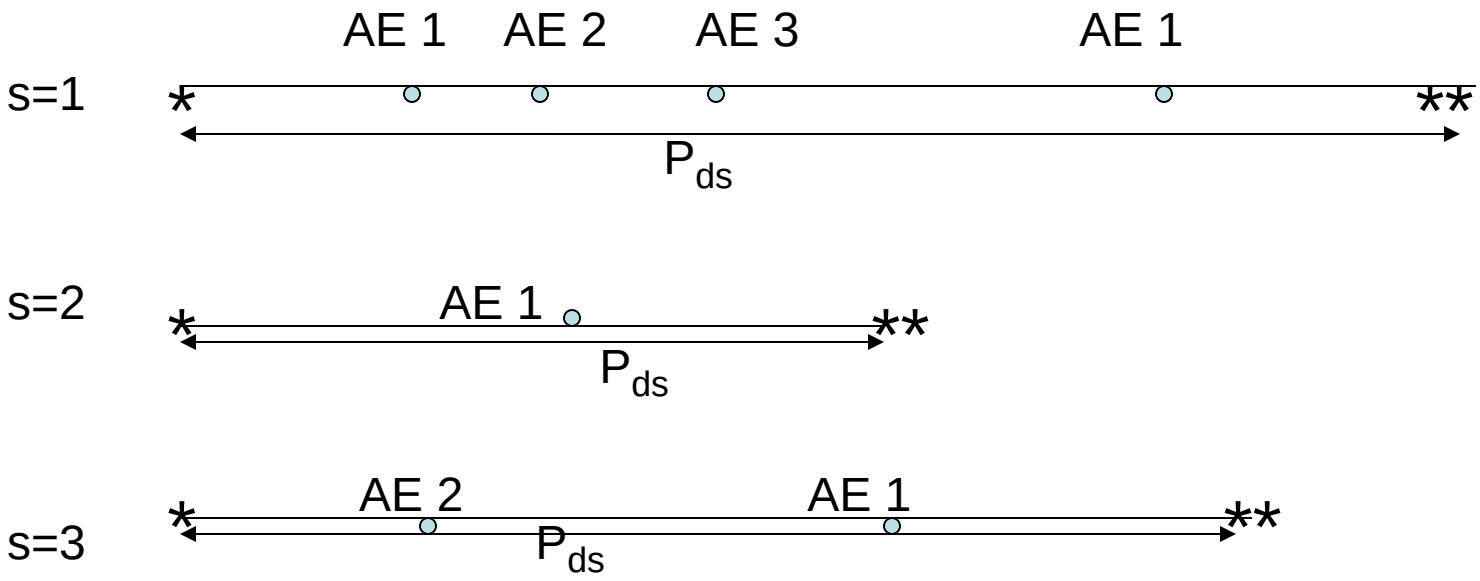
Definition of event-time



Definition of person-time



Definition of exposure-time



Working data structure (with event-time)

Drugs

AEs	1	2	3	…	14
1	n11	…	n1j	…	n1J
2	n21	…	…	…	n2J
…	…	…	…	…	…
i	…	…	nij	…	niJ
…	…	…	…	…	…
I	nI1	…	nIj	…	nIJ
Col total	n.1	…	n.j	…	n.J

Drugs

	1	2	3	…	14	Row total
1	P11	…	P1j	…	P1J	P1.
2	P21	…	…	…	P2J	P2.
…	…	…	…	…	…	…
i	…	…	Pij	…	PiJ	PiJ
…	…	…	…	…	…	…
I	PI1	…	PIj	…	PIJ	PIJ
Col total	P.1	…	P.j	…	P.J	P.J

J=14 in the above table. At look k (k=1,…, K=5), there are two tables constructed from the individual level data. Pij is the event-time (unit here is day) for the AE i and drug j. We suppress k in the notation.

Exposure-based longitudinal LRT Methods using event-time

$$n_{ijk} \sim Poisson(p_{ijk} \times P_{i.k}),$$

$$(n_{.jk} - n_{ijk}) \sim Poisson(q_{ijk} \times (P_{..k} - P_{i.k}))$$

$$i = 1,..., I,\ j = 1,..., J, K = 1,..., K.$$

H0: $p_i=q_i$ over i=1,…I, if J (drug) is fixed.
$RR_i=p_i/q_i$,=1 under H0, i=1,…, I, which is relative event-rate of ith AE vs. other AEs for fixed drug j.

The likelihood ratio is then

$$LR_{ijk} = \frac{(\hat{p}_{ijk,H_a})^{n_{ijk}} (\hat{q}_{ijk,H_a})^{n_{.jk}-n_{ijk}}}{(\hat{p}_{ijk,H_0})^{n_{.jk}}}$$

$$LR_{ijk} = \frac{(\frac{n_{ijk}}{P_{i.k}})^{n_{ijk}} (\frac{n_{.jk}-n_{ijk}}{P_{..k}-P_{i.k}})^{n_{.jk}-n_{ijk}}}{(\frac{n_{.jk}}{P_{..k}})^{n_{.jk}}}$$

$$= (\frac{n_{ijk}}{E_{ijk}})^{n_{ijk}} (\frac{n_{.jk}-n_{ijk}}{n_{.jk}-E_{ijk}})^{n_{.jk}-n_{ijk}} ; E_{ijk} = \frac{n_{.jk} \times P_{i.k}}{P_{..k}}$$

Test statistic is $\max LR_{jk} = \max_i LR_{ijk}, i = 1,..., I$

Working data structure (with person-time)

AE of interest

J=1

drug

1	n11k	
I=2	n21k	
Col total	n.1k= n11k+n21k	

1	P11k
2	P21k
Col total	P..= P11k+P21k

Sequential LRT (with person-time)

$$n_{11k} \sim Poisson(p_{11k} \times P_{11k}),$$
$$(n_{21k}) \sim Poisson(q_{21k} \times P_{21k})$$

$$n_{11k} \mid n_{.1k} \sim Binomial(n_{.1k}, \frac{RR_{11k} P_{11k}}{RR_{11} P_{11k} + (P_{21k})})$$

H0: $p_1=q_2$, Ha: p1>q2

$RR_1=p_1/q_2$, is relative risk of ith AE vs. the other AE for fixed drug j; or relative risk of ith drug vs. the other drug for fixed AE j.

The likelihood ratio is then

$$LR_{11k} = \frac{(\hat{p}_{11k,H_a})^{n_{11k}} (\hat{q}_{11k,H_a})^{n_{.1k}-n_{11k}}}{(\hat{p}_{11k,H_0})^{n_{.1k}}}$$

$$LR_{11k} = \frac{(\frac{n_{11k}}{P_{11k}})^{n_{11k}} (\frac{n_{21k}}{P_{21k}})^{n_{21k}}}{(\frac{n_{11k}+n_{21k}}{P_{11k}+P_{21k}})^{n_{11k}+n_{21k}}} = (\frac{n_{11k}}{E_{11k}})^{n_{11k}} (\frac{n_{21k}}{n_{.1k}-E_{11k}})^{n_{21k}}; E_{11k} = \frac{n_{.1k} \times P_{11k}}{(P_{11k}+P_{21k})}$$

Test statistic is $\max LR_{jk} = \max_i LR_{ijk}, i = 1,2.$

Relationship seqLRT with CSSP and maxSPRT

- CSSP statistic is the number of adverse events. seqLRT statistic is maxLR; same null data simulation process and assumption

- For maxSPRT: Let $n_{.k}=n_{.1k}$ be the total # cases up to time-interval k. $n^{drug}_k=n_{11k}$ is the # cases for drug i=1.

$$E_{11k} = \frac{n_{.1k} \times P_{11k}}{P_{11k} + P_{21k}} = \frac{n_{.k}}{M+1}$$

$$LRT_k = \frac{\left(\frac{n_k^{drug}}{n_{.k}}\right)^{n_k^{drug}} \left(\frac{n_{.k} - n_k^{drug}}{n_{.k}}\right)^{n_{.k} - n_k^{drug}}}{\left(\frac{1}{M+1)}\right)^{n_k^{drug}} \left(\frac{M}{M+1}\right)^{n_{.k} - n_k^{drug}}}$$

Working data structure (with exposure-time)

drugs

	Jth AE
1	n_{1j}
2	…
…	…
i	n_{dj}
…	…
I	n_{Dj}
Col total	$n_{.j}$

1	P_1
2	P_2
…	…
i	…
…	…
I	P_D
Col total	P.

Longitudinal LRT (with exposure-time)

For a fixed j*th AE, Assume $p_{dj^*s} = p_{dj^*}$,

The dist of the events: $n_{dj^*s} \sim ind \;\; Poisson(p_{dj^*} P_{ds})$

$$n_{dj^*} = \sum_s n_{dj^*s} \sim ind \;\; Poisson(p_{dj^*} P_d), P_d = \sum_s P_{ds}.$$

$$(n_{.j^*} - n_{dj^*}) \sim ind \;\; Poisson(q_{dj^*}(P_. - P_d))$$

$$RR_{dj^*} = \frac{p_{dj^*}}{q_{dj^*}} \text{ is relative risk of j*th AE for drug d vs. other drugs}$$

Test is $p_d = q_d$ over d=1,..,D if J (AE) is fixed

$RR_d = p_d/q_d$, i=1,…,D is relative risk of dth drug vs. other drugs for fixed AE j*.

The likelihood ratio is then

$$LR_{dj^*k} = \frac{(\hat{p}_{dj^*k,H_a})^{n_{djk}} (\hat{q}_{dj^*k,H_a})^{n_{.jk}-n_{djk}}}{(\hat{p}_{dj^*k,H_0})^{n_{.jk}}}$$

$$LR_{dj^*k} = \frac{(\frac{n_{dj^*k}}{P_{dk}})^{n_{dj^*k}} (\frac{n_{.j^*k}-n_{dj^*k}}{P_{.k}-P_d k})^{n_{.j^*k}-n_{dj^*k}}}{(\frac{n_{.j^*k}}{P_{.k}})^{n_{.jk}}}$$

$$= (\frac{n_{dj^*k}}{E_{dj^*k}})^{n_{dj^*k}} (\frac{(n_{.j^*k}-n_{dj^*k})^{n_{.j^*k}-n_{dj^*k}}}{n_{.j^*k}-E_{dj^*k}}, E_{dj^*k} = \frac{P_d \times n_{.j^*k}}{P_.}.$$

Test statistic is $\max LR_{jk} = \max_d LR_{djk}, d = 1,...,D$

Distribution of LRT under H0

- Test Statistic is $MaxLR_{jk}$ (discussed in earlier slides)
- The distribution of $MaxLR_{jk}$ under H0 is not analytically tractable
- We use Monte Carlo method to obtain the empirical dist for each k
- Cases can be generated “cumulatively’ using multinomial distribution.
- Cases can also be generated using multinomial for each individual time-period and then summing-up over time

Null data generation (cumulatively)

For event-time and person-time cases

$$(n_{1jk},...,n_{Ijk}) \mid n_{.jk} \sim Multinomial(n_{.jk},(\frac{P_{1.k}}{P_{..,}},...,\frac{P_{I.k}}{P_{..k}})), i = 1,...,I.$$

I is the total # of AEs or drugs under comparison.
For exposure-time cases,

$$(n_{1j*k},...,n_{Dj*k}) \mid n_{.j*k} \sim Multinomial(n_{.j*k},(\frac{P_{1k}}{P_{.k}},...,\frac{P_{Dk}}{P_{.k}})), d = 1,...,D.$$

D is total # of drugs under comparison.
The parameters in the multinomial distribution are from the observed data.

P-value calculation (for each period k)

- Calculate MaxLR from observed data (one) over time (k=1, 2, 3,…, K)

- Calculate MaxLRs from the 9,999 simulated null data for each time period

- At each time period, compare the observed MLR and the ones from simulation ---→ p-value = 1-rank of the observed maxLR among the 10000 maxLRs.

- If p-value< alpha(k) -→ reject H0 and identify signals

Alpha-Spending Functions and Decision Rules

- Specify the error rate to be spend at look k=1,…K. This can be monotonic power functions such as alpha spending functions:

$$\alpha(k) = \frac{1}{K}\alpha \qquad \alpha(k) = \alpha\frac{1}{2^k}$$

$$cum\,\alpha(k) = \frac{k}{K}\alpha \le \alpha \qquad cum\,\alpha(k) = \alpha\sum_{r=1}^{k}\frac{1}{2^r} \le \alpha$$

The second formulation does not depend on K.

Alpha-Spending Functions and Decision Rules

- Other choices of alpha-spending boundary functions are: O'Brien-Fleming, Pocock, Lan-DeMets, etc.
- At look k, the AE associated with the maxLR in the obs data is a signal for the particular drug if the p-value is < alpha(k).
- There could be secondary signals with next lower ordered values of LR, after maxLR, in the observed data, that have p-value <alpha(k): LR2, LR3….. (step-down procedure).

Application of LRT methods to Pooled clinical trial data for PPIs

- PPIs are a class of drugs that decrease gastric acid secretion through inhibition of the proton pump. It helps in the secretion of acid from the stomach glands.

- In a recent study, it has been found that proton pump inhibitors (PPIs) are associated with increased risk of hip fractures (side effect) (Yang et al. 2006). The increased risk of hip fractures is attributed to osteoporosis caused by proton pump inhibitors.

- Pooled clinical trial data from FDA/OTS/OCS legacy database

- PPIs were concomitantly used with test drugs for treating osteoporosis among targeted patients.

Application of LRT methods to Pooled clinical trial data for PPIs

- Pooled data with 10 trials (sample sizes from hundreds to thousands). # Subjects using concomitant PPIs is about 10% of the total sample size.
- A total of 14 drugs (7 test drugs, and 7 test drug+PPIs) are included in the exploration.
- Use calendar time (1996-97, -99, -2001, -03, -07). K=5.
- Alpha=0.05
- With alpha(k)=alpha/2*k, alpha(1)=0.025, alpha(2)=0.0125, alpha(3)=0.00625, alpha(4)=0.003125, alpha(5)=0.001563.

AEs signals with p<alpha(k) by drug (event-time)

		k=1	2	3	4	5
Placebo	ndotj	1251	4703	8282	29731	50364
	AE signals	3	6	34	43	74
	muscle cramp (rr)			4.1	2.3	4.4
	Bone pain (rr)					2.2
Placebo +PPIs	ndotj	95	273	1094	4833	9043
	AE signals	0	0	23	26	30
	muscle cramp (rr)					6.8
	muscle spasms (rr)					3.9

Comparison of PL vs. PL+PPIs (I=2) using Sequential LRT method (person-time)

- AE of interest is a composite AE (AEOST) including many AE terms associated with osteoporosis (J=1), 1st occurrence of AEOST.
- PL+PPIs vs. PL (I=2)
- Sample sizes Ndotj for j=1 are 57, 163, 232, 439, and 500 for analysis periods 1, 2, 3, 4, and 5, respectively.
- Rr values are 4.7, 2.4, 2.5, 1.9, 1.7 for k=1 to 5.
- When k=1, the p-value from seqLRT is 0.001. PL+PPIs had higher relative risk vs. PL for AEOST (<alpha(1)=0.025), stop the search by sequential method.

Safety signals for multiple occurrences of AEOST (PL+PPIs vs. PL, I=2) by longLRT (exposure-time)

	k=1	2	3	4	5
ndotj	65	195	286	647	787
rr	5.7	3.2	2.9	2.5	2.4
pvalue	**0**	**0**	**0**	**0**	**0**

PL+PPIs is a signal for AEOST for k=1 to 5 periods.

Do not stop monitoring the signals over time

Safety signals for multiple occurrences of AEOST (I=14 drugs) by longLRT (exposure-time)

		k=1	2	3	4	5
	ndotj	174	549	815	3902	6041
PL+PPIs	rr	5.9	3.4	3	1.8	1.6
	pvalue	**0**	**0**	**0**	**0**	**0**
Lasoxifene+ PPIs	rr			1.1	1.9	2
	pvalue			0.99	**0**	**0**
PTH+PPIs	rr			5.9	2	2.3
	pvalue			0	**0**	**0**
Bazedoxifene+PPIs	rr				2.2	0.9
	pvalue				**0**	**0.99**

Discussion

- One AE or multiple AEs, one drug and multiple drugs
- Drug class and AE group
- Count data or data with exposure
- Fixed time analysis or analysis over time
- Rate and risk
- Different definitions of exposure
- Method for incorporating the study effect

References

- Cook, A. J., Wellman, R.D., Tiwari, R.C., Li, L., Heckbert, S., March, T., Heagerty, P., and Nelson, J.C. (2012), ``Statistical approaches to group sequential monitoring of postmarket safety surveillance data: Current state of the art for use in the mini-sentinel pilot," *Pharmacoepidemiol Drug Saf*, 21 suppl, 72-81.
- Li, L. (2009), ``A conditional sequential sampling procedure for drug safety surveillance," *Statistics in Medicine*, 28, 3124-3138.
- Lieu, T.A., Kulldorff, M., Davis, R.L., Levis, E.M., Weintraub, E., Yih, K., Yin, R., Brong, J.S., Platt, R. (2007), ``Real-time vaccine safety data surveillance," *Medical Care*, 45 (10 Supl 2): S89-95.
- Noren et al. 2010 (See Lecture # 1)
- O'Brien, P.C. and Fleming, T.R. (1979), ``A multiple testing procedure for clinical trials," *Biometrics*, 35, 549-556.
- Platt R., Carnahan R.M., Brown J.S. et al. 2012, “The U.S. Food and Drug Administration’s Mini-Sentinel program: status and direction”, *Pharmacoepidemiology and Drug Safety,* *http://onlinelibrary.wiley.com/doi/10.1002/pds.3230/pdf*
- Pocock, S.J. (1977), ``Group sequential methods in the design and analysis of clinical trials," *Biometrika*, 64, 191-200.
- Schuemie, M.J. (2011), ``Methods for drug safety signal detection in longitudinal observational databases: LGPS and LEOPARD," *Pharmacoepidemiology and Drug Safety*, 20 (3), 292-299
- Yang, Y.X., Lewis, J.D., Epstein, S., Metz, D.C. (2006), ``Long term proton pump inhibitor therapy and risk of hip fracture,“ *JAMA*, 296 (24): 2947-53.

Appendix

Exposure-based longitudinal Methods

- Binomial LRT

$$LRT_k = \frac{\left(\frac{n_{ijk}}{P_{i.k}}\right)^{n_{ijk}} \left(1-\frac{n_{ijk}}{P_{i.k}}\right)^{n_{i.k}-n_{ijk}} \left(\frac{n_{.jk}-n_{ijk}}{P_{..k}-P_{i.k}}\right)^{n_{.jk}-n_{ijk}} \left(1-\frac{n_{.jk}-n_{ijk}}{P_{..k}-P_{i.k}}\right)^{n_{..k}-n_{.jk}-n_{i.k}+n_{ijk}}}{\left(\frac{n_{.jk}}{P_{..k}}\right)^{n_{.jk}} \left(1-\frac{n_{.jk}}{P_{..k}}\right)^{n_{..k}-n_{.jk}}}$$

$$\rightarrow Poisson, LRT$$

Notations for the following plots on drug exposure

- * Indicates start date of drug j (or d) and ** indicates stop date of drug j (or d).
- Assume that each subject takes a single drug (or drug combination); different subject may take different drugs or drug combinations.
- Circled dots indicate occurrences of AEs (AE i, i=1, 2, 3,…..). Only AEs between * and ** are *countable* cases and are shown in the plots over time.
- P^1_{ijs} is the **event-time** for sth subject taking jth drug and having 1st occurrence of ith AE. P^2_{ijs} is the event-time for sth subject taking jth drug and having 2nd occurrence of ith AE.
- $P_{ijs}=P^1_{ijs}$ is the event-time for sth subject taking jth drug and having 1^{st} occurrence of ith AE, which is also **person-time** when we only consider AE without recurrence or the 1^{st} occurrence of one AE with repeated occurrences
- P_{ds} is the **exposure-time** for sth subject taking dth drug and having multiple AEs during the exposure duration

Exposure for ith AE and jth (dth) drug (aggregation of subject-level information)

For event time, $P_{ij} = \sum_{s} \sum_{l(i,s)} P_{ijs}^{l(i,s)}, s = 1,..., S; l(i,s) = 1,..., L(i,s).$

S is the total # of subjects, and L(i,s) is the total # of occurrences of ith AE for sth subject.

$$P_{i.} = \sum_{j} P_{ij}, P_{.j} = \sum_{i} P_{ij}, P_{..} = \sum_{i} \sum_{j} P_{ij}.$$

For person-time, $P_{ij} = \sum_{s} P_{ijs}.$

For exposure-time, $P_d = \sum_{s} P_{ds}, P_{.} = \sum_{d} P_d = \sum_{d} \sum_{s} P_{ds}.$

Simulation using the information from the Pooled clinical trial data

- seqLRT for 1st occurrence of AEOST (J=1), PL+PPIs vs. PL only (I=2)

$$n_{ijk} \mid n_{.jk} \sim Binomial(n_{.jk}, \frac{RR_{ijk} P_{i.k}}{RR_{ij} P_{i.k} + (P_{..k} - P_{i.k})})$$

- longLRT for any occurrences of AEOST (J=1), multiple drugs

$$(n_{1j*k},...,n_{Dj*k}) \mid n_{.j*k} \sim Multinomial(n_{.j*k}, (RR_{1j} rr_0 \frac{P_{1k}}{P_{.k,}},..., RR_{Dj} rr_0 \frac{P_{Dk}}{P_{.k}})), d = 1,..., D.$$

$$P_{.k} = \sum_d P_{dk},\ RR_{dj} \geq 1,\ \sum_{d=1}^{D} \mathrm{RR}_{dj} rr_0 \frac{P_{dk}}{P_{.k}} = 1.$$

Simulation setup

- For H0 data, $RR_{i1}=1$
- For Ha data, $RR_{i1}=c$ (PL+PPIs vs. PL).
- c=1.2, 1.5, 2, 4, 6, 10

- Sample size as $z*n_{.1k}$
- c=1, 2, 4, 10

- Simulation 1000 data for each case

Performance evaluation

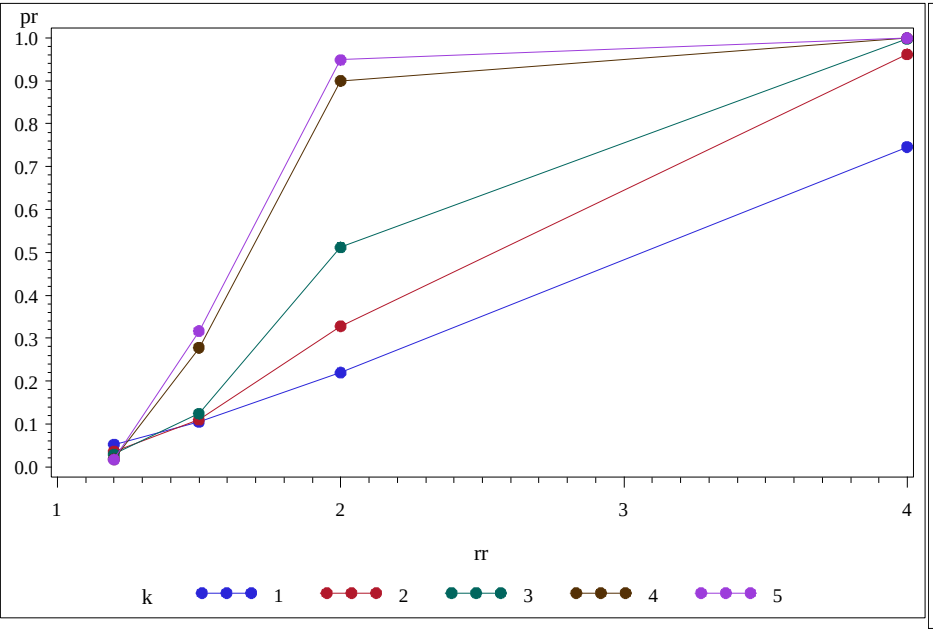
- Conditional power:
  - Pr(k)=#rejecting H0 at kth period/1000, k=1, …, 5.
- Unconditional power for seqLRT
  - Power(k)=pr(1)+…+(1-pr(1))×…×(1-pr(k-1)) ×pr(k)
- When data is generated under H0, pr(k) is conditional error rate and power-→ type-I error rate for seqLRT
- For longLRT without stopping the procedure, we use cumulative error rate cumer(k)=pr(1)+pr(2)…+pr(k)

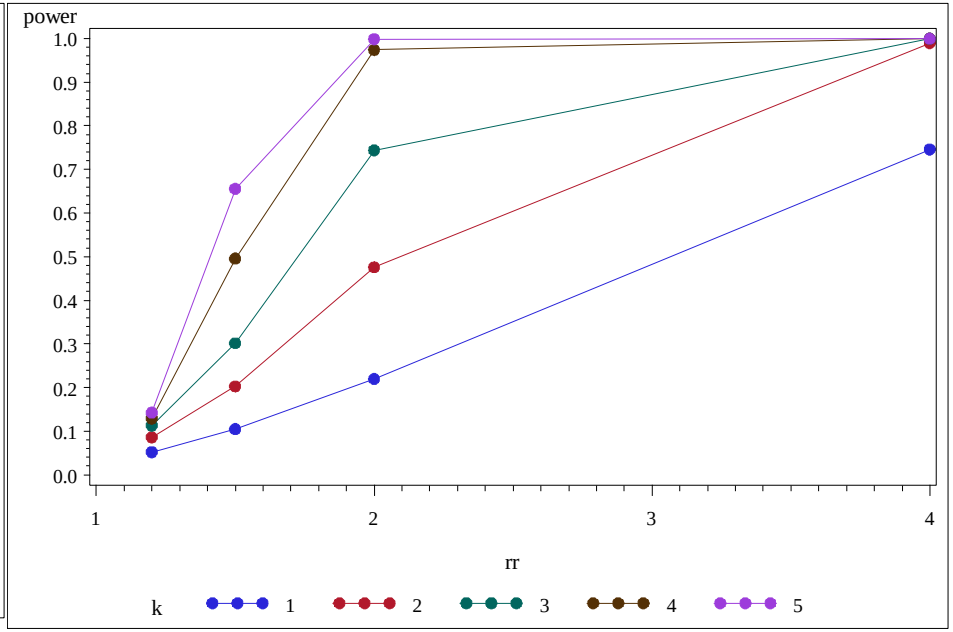
Patterns on error rate (rr=1 H0 data, z=1, similar for z=2, 4))

		k=1	2	3	4	5
seqLRT	ndotj	57	163	232	439	500
	pr(k)	0.027	0.012	0.007	0.003	0.001
	type-I error(k)	0.027	0.039	0.045	0.048	0.049
longLRT	ndotj	65	195	286	647	787
	pr(k)	0.018	0.005	0.007	0.004	0.001
	cumer(k)	0.018	0.023	0.03	0.034	0.035
	alpha(k)	0.025	0.0125	0.0063	0.00313	0.00156
	cum alpha	0.025	0.0375	0.0438	0.0469	0.0484

Patterns on power (seqLRT), z=1, rr=1.2 to 4

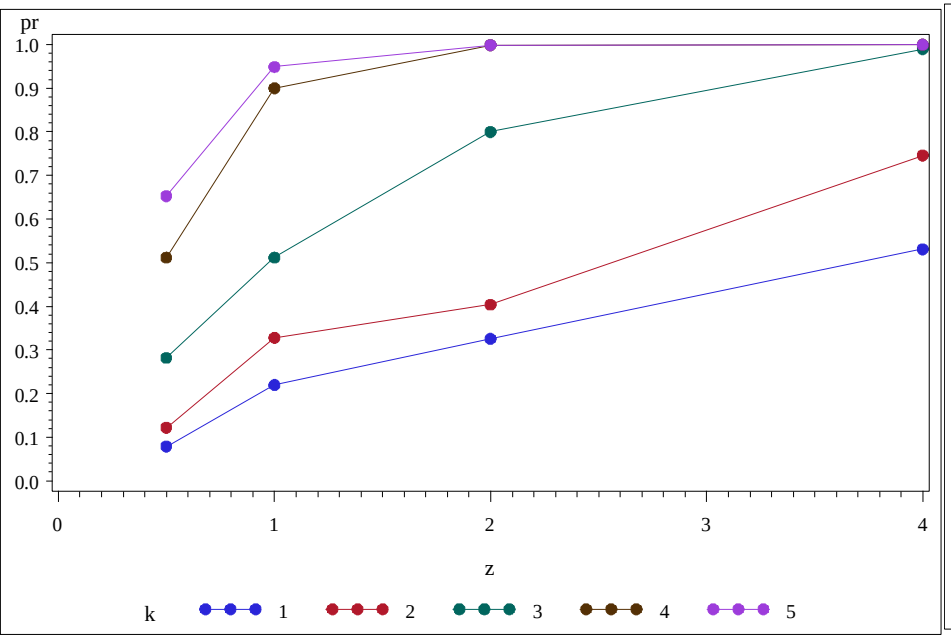
Conditional power	Unconditional power

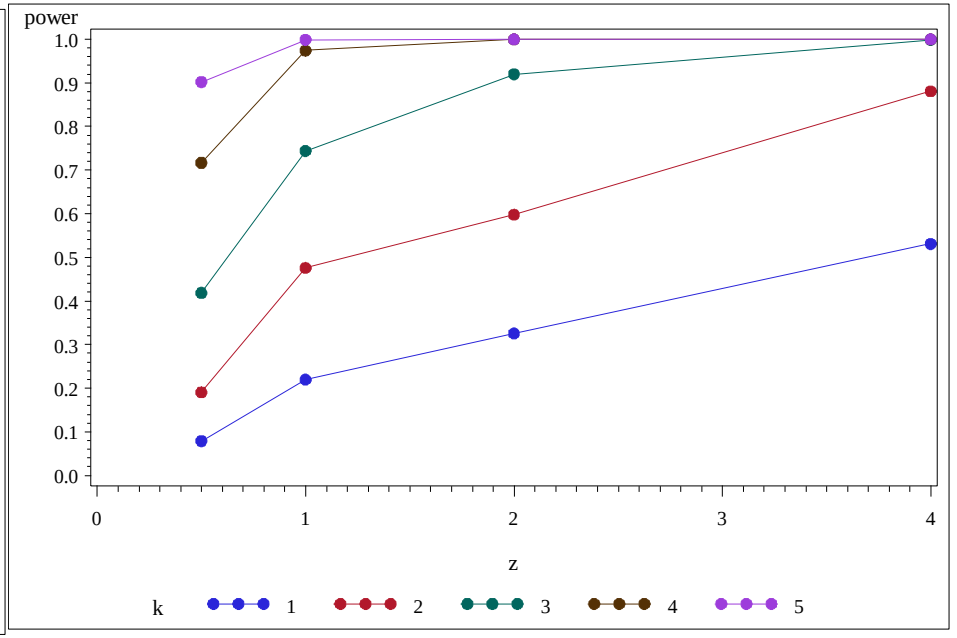




Patterns on power (seqLRT), rr=2, z=0.5 to 4

Conditional power	Unconditional power





Patterns on (conditional)power (longLRT)

Z=1, relative risk from 1.2 to 4	RR=2, z from 0.5 to 4