Data Mining in Pharmacovigilence

Aimin Feng, David Madigan, and Ivan Zorych

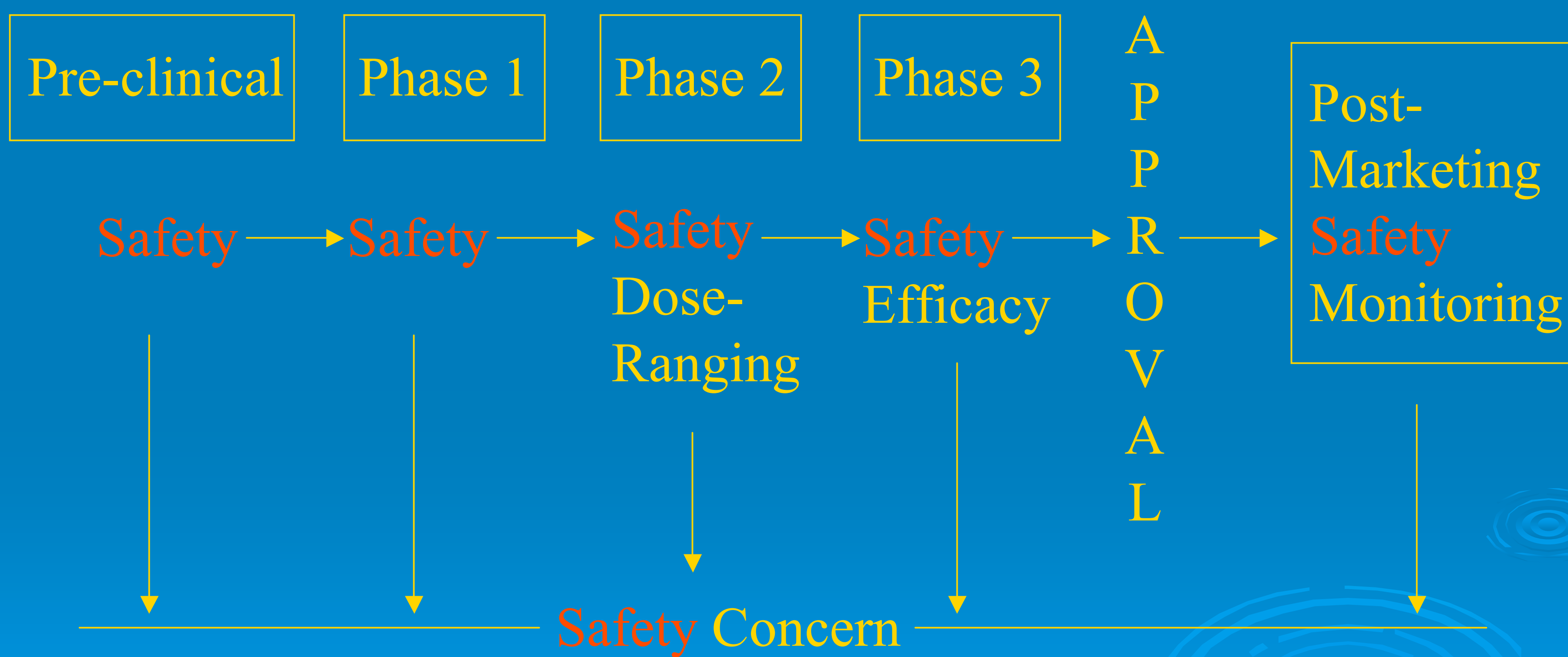
dmadigan@rutgers.edu

http://stat.rutgers.edu/~madigan

Overview

- Intro. to Post-marketing Surveillance
- SRS Databases
- Existing Analysis Methods
- Our Approaches
  - Bayesian Logistic Regression
  - Propensity Score
- Conclusions

Safety in Lifecycle of a Drug/Biologic product



Why Post-marketing Surveillance

- Limitations on pre-licensure trials
  - Size
  - Duration
  - Patient population: age, comorbidity, severity
- Fact
  - Several hundred drugs have been removed from market in the last 30 years due to safety problems which became known after approval

Databases of Spontaneous ADRs

- FDA Adverse Event Reporting System (AERS)
  - Online 1997 – replace the SRS
  - Over 250,000 ADRs reports annually
  - 15,000 drugs - 16,000 ADRs
- CDC/FDA Vaccine Adverse Events (VAERS)
  - Initiated in 1990
  - 12,000 reports per year
  - 50 vaccines and 700 adverse events
- Other SRS
  - WHO - international pharmacovigilance program

Weakness of SRS Data

- Passive surveillance
  - Underreporting
- Lack of accurate “denominator”, only “numerator”
  - “Numerator”: No. of reports of suspected reaction
  - “Denominator”: No. of doses of administered drug
- No certainty that a reported reaction was causal
- Missing, inaccurate or duplicated data

Existing Methods

- Multi-item Gamma Poisson Shrinker (MGPS)
  - US Food and Drug Administration (FDA)
- Bayesian Confidence Propagation Neural Network
  - WHO Uppsala Monitoring Centre (UMC)
- Proportional Reporting Ratio (PRR and aPRR)
  - UK Medicines Control Agency (MCA)
- Reporting Odds Ratios and Incidence Rate Ratios
  - Other national spontaneous reporting centers and drug safety research units

Existing Methods (Cont'd)

- Focus on 2X2 contingency table projections

	AE j = Yes	*AE j = No*	*Total*
Drug *i* = Yes	*a*=20	*b*=100	120
Drug *i* = No	*c*=100	*d*=980	1080
Total	*120*	*1080*	1200

- 15,000 drugs * 16,000 AEs = 240 million tables
- Most N_{ij} = 0, even though $N_{..}$ very large

The Different Measures

Measure of Association	*Formula*	*Probabilistic Interpretation*
RR Relative Risk*	$\frac{a * (a+b+c+d)}{(a+c) * (a+b)}$	$\frac{\Pr(ae \mid drug)}{\Pr(ae)}$
PRR Proportional Reporting Ratio	$\frac{a / (a+b)}{c / (c+d)}$	$\frac{\Pr(ae \mid drug)}{\Pr(ae \mid \neg drug)}$
ROR Reporting Odds Ratio	$\frac{a / c}{b / d}$	$\frac{\Pr(ae \mid drug) / \Pr(\neg ae \mid drug)}{\Pr(ae \mid \neg drug) / \Pr(\neg ae \mid drug)}$
Information Component	$\log_2 \frac{a * (a+b+c+d)}{(a+c) * (a+b)}$	$\log_2 \frac{\Pr(ae \mid drug)}{\Pr(ae)}$

These Measures not "Robust"

	AE = Yes	AE = No
D1 = Yes	a=1	b=100
D1 = No	c=5	d=1080

	AE = Yes	AE = No
D2 = Yes	a=2	b=100
D2 = No	c=5	d=1080

Measure	Drug D1	Drug D2
PRR	2.1	4.3
ROR	2.2	4.3
IC	1.0	1.7
RR	2.0	3.3

Reverend Bayes to the rescue!

Bayesian Statistics

The Bayesian approach has deep historical roots but required the algorithmic developments of the late 1980's before it was of any use

The old sterile Bayesian-Frequentist debates are a thing of the past

Most data analysts take a pragmatic point of view and use whatever is most useful

Think about this…

	Hospital											
	A	**B**	**C**	**D**	**E**	**F**	**G**	**H**	**I**	**J**	**K**	**L**
No. of ops. n	27	148	119	810	211	196	148	215	207	97	256	360
No. of deaths r	0	18	8	46	8	13	9	31	14	8	29	24

Denote by θ the probability that the next operation in Hospital A results in a death

Use the data to estimate (i.e., guess the value of) θ

Hospital Example (0/27)

$$f(\theta \mid data) = \frac{f(data \mid \theta) f(\theta)}{f(data)} \propto f(data \mid \theta) f(\theta)$$

posterior distribution

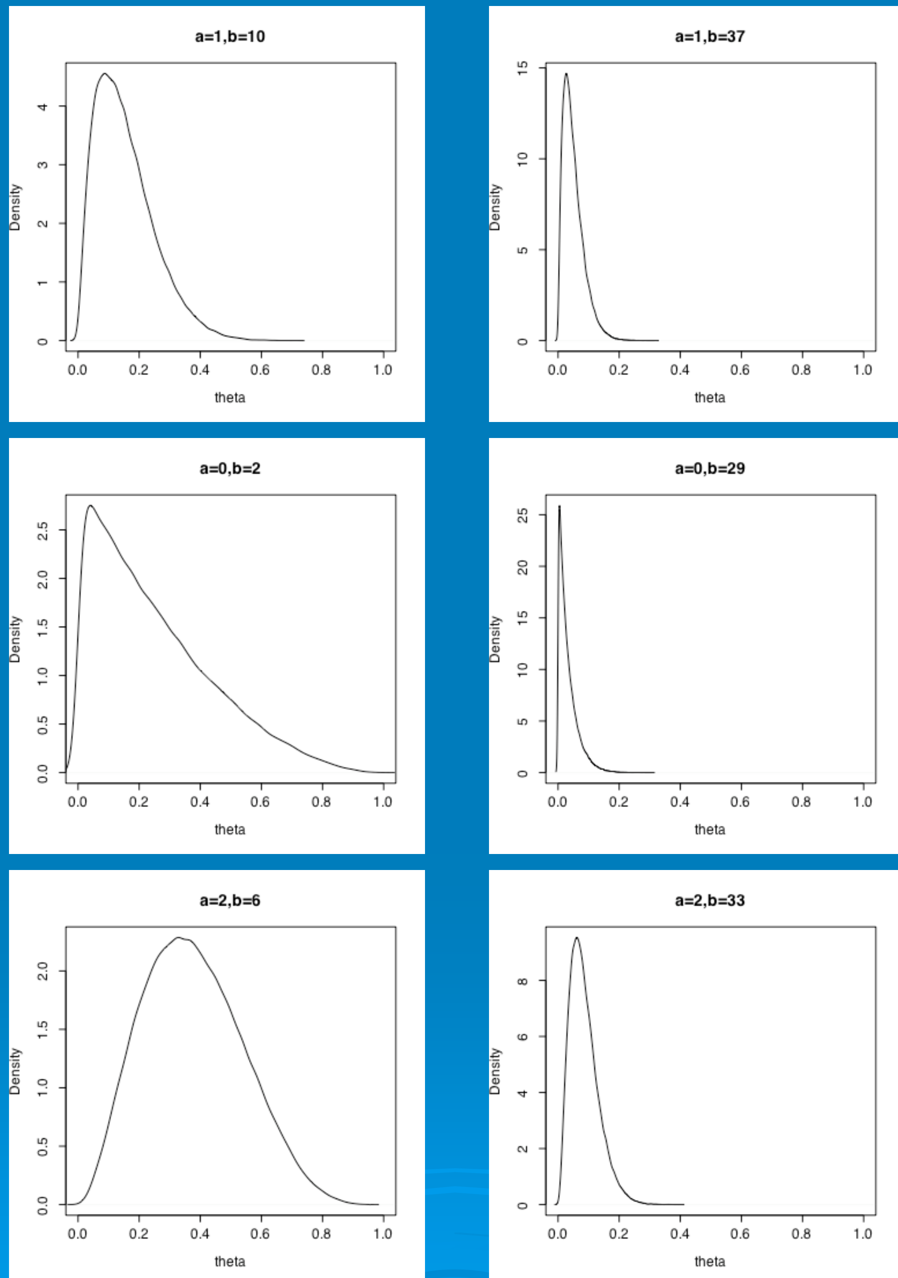
likelihood

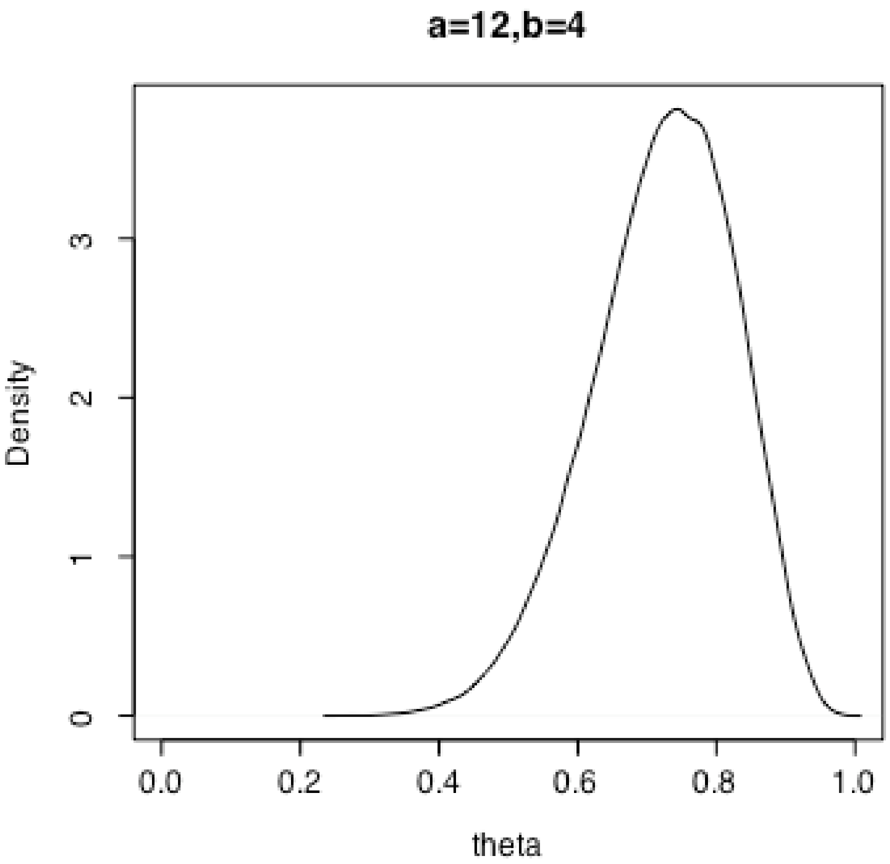
$$\binom{27}{0} \theta^{0} (1-\theta)^{27}$$

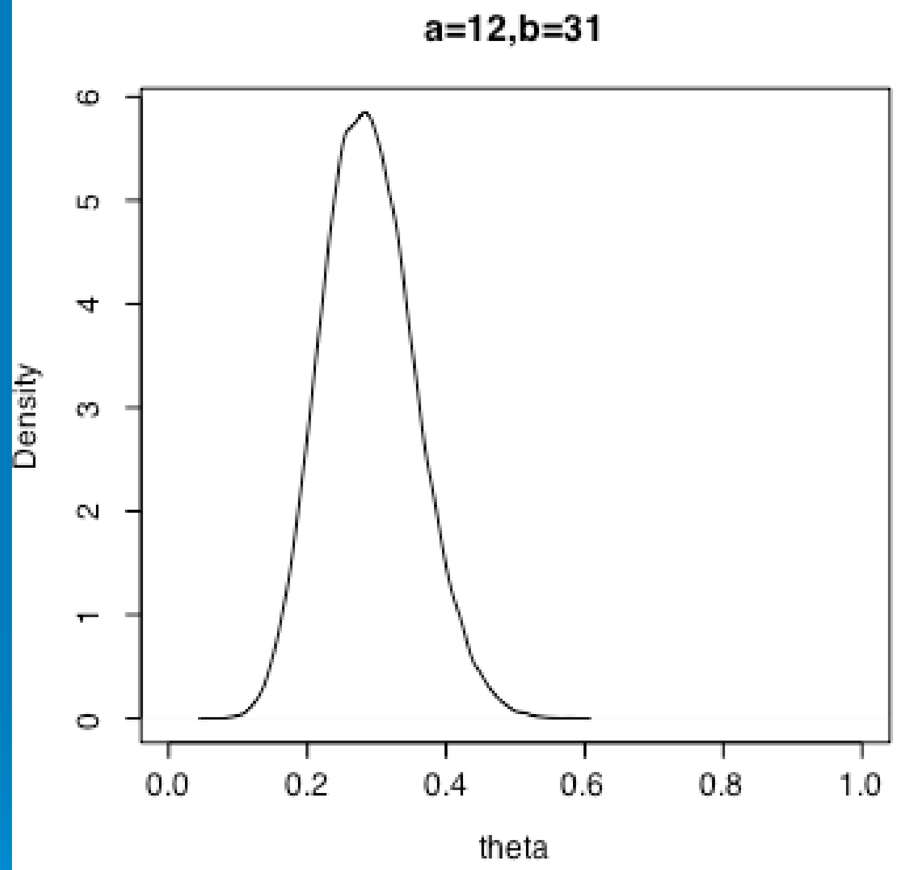
prior distribution

$$c\theta^{a} (1-\theta)^{b}$$

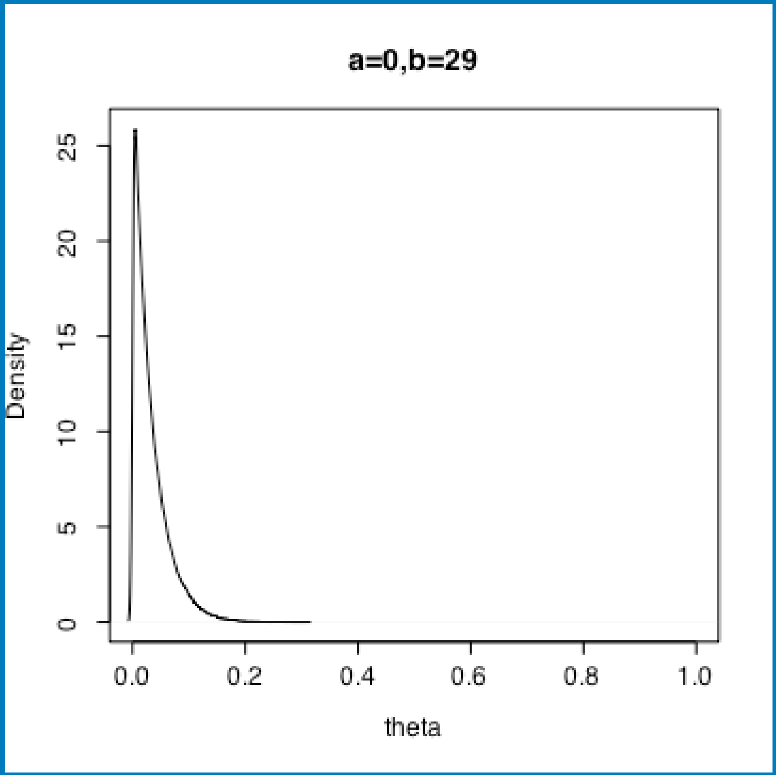
$$\propto \theta^{a+0} (1-\theta)^{b+27}$$


a=1,b=10
Density
0 1 2 3 4
0.0 0.2 0.4 0.6 0.8 1.0
theta
a=1,b=37
Density
0 5 10 15
0.0 0.2 0.4 0.6 0.8 1.0
theta
a=0,b=2
Density
0.0 0.5 1.0 1.5 2.0 2.5
0.0 0.2 0.4 0.6 0.8 1.0
theta
a=0,b=29
Density
0 5 10 15 20 25
0.0 0.2 0.4 0.6 0.8 1.0
theta
a=2,b=6
Density
0.0 0.5 1.0 1.5 2.0
0.0 0.2 0.4 0.6 0.8 1.0
theta
a=2,b=33
Density
0 2 4 6 8
0.0 0.2 0.4 0.6 0.8 1.0
theta





Unreasonable prior distribution implies unreasonable posterior distribution



What to report? Mode? Mean? Median?

Mode → 0.032 ; Median → 0.023

Posterior probability that theta exceeds 0.2? → 0.013

theta* such that Pr(theta > theta*) = 0.05 → 0.095

theta* such that Pr(theta > theta*) = 0.95 → 0.002

Posterior probability that theta is in (0.002,0.095) is 90%

More formal treatment…

	Hospital											
	A	**B**	**C**	**D**	**E**	**F**	**G**	**H**	**I**	**J**	**K**	**L**
No. of ops. n	27	148	119	810	211	196	148	215	207	97	256	360
No. of deaths r	0	18	8	46	8	13	9	31	14	8	29	24

Denote by θ_i the probability that the next operation in Hospital i results in a death

Assume $\theta_i \sim \text{beta}(a,b)$

Compute joint posterior distribution for all the θ_i simultaneously

	Hospital											
	A	B	C	D	E	F	G	H	I	J	K	L
No. of Ops (n)	27	148	119	810	211	196	148	215	207	97	256	360
Raw Rate (x/n)	0.00	12.16	6.72	5.68	2.37	6.63	6.08	14.42	6.76	8.25	11.33	6.67
Post. Mean	5.77	10.50	7.01	5.88	4.15	6.86	6.58	12.58	6.94	7.85	10.34	6.81
Post. S.D.	2.3	2.3	1.8	0.8	1.3	1.5	1.6	2.2	1.5	2.1	1.8	1.2
Raw Rank	1	11	7	3	2	5	4	12	8	9	10	6
Post. Rank	2	11	8	3	1	6	4	12	7	9	10	5

"Borrowing strength"

Shrinks estimate towards common mean (7.4%)

Technical detail: can use the data to estimate a and b

This is known as "empirical bayes"

Relative Reporting Ratio

N_{ij} →

	AE_j	Not AE_j
$Drug_i$	$a=20$	$b=100$
Not $Drug_i$	$c=100$	$d=980$

- If the Drug and the AE were independent, what would you expect *a* to be?
  - Overall $(a+c)/(a+b+c+d)$=120/1200=10% have the AE
  - So, 10% of the “Drug” reports should have the AE
  - That is $(a+b)*((a+c)/(a+b+c+d))$=120*10%=12=E_{ij}
  - Note $N_{ij}/E_{ij}=a/\ (a+b)*((a+c)/(a+b+c+d))$=RR
  - RR = 20/12 = 1.67 = N/E = Pr(AE|Drug)/Pr(AE)

Relative Reporting Ratio

$(RR_{ij}=N_{ij}/E_{ij})$

$$E_{ij}=N_{ij}*N_{..}/N_{i.}N_{.j}$$

	AE_j	Not AE_j	
$Drug_i$	N_{ij}		$N_{i.}$
Not $Drug_i$			
	$N_{.j}$		$N_{..}$

- Advantages
  - Simple
  - Easy to interpret
- Disadvantages
  - Extreme sampling variability when baseline and observed frequencies are small

    (N=1, E=0.01 v.s. N=100, E=1)
  - GPS provides a shrinkage estimate of RR that addresses this concern.

Same Relative Reporting Ratio!

	AE_j	Not AE_j
$Drug_i$	*a=1*	*b=5*
Not $Drug_i$	*c=5*	*d=49*

Chi-square = 0.33

	AE_j	Not AE_j
$Drug_i$	*a=20*	*b=100*
Not $Drug_i$	*c=100*	*d=980*

Chi-square = 6.58

	AE_j	Not AE_j
$Drug_i$	*a=200*	*b=1000*
Not $Drug_i$	*c=1000*	*d=9800*

Chi-square = 65.8

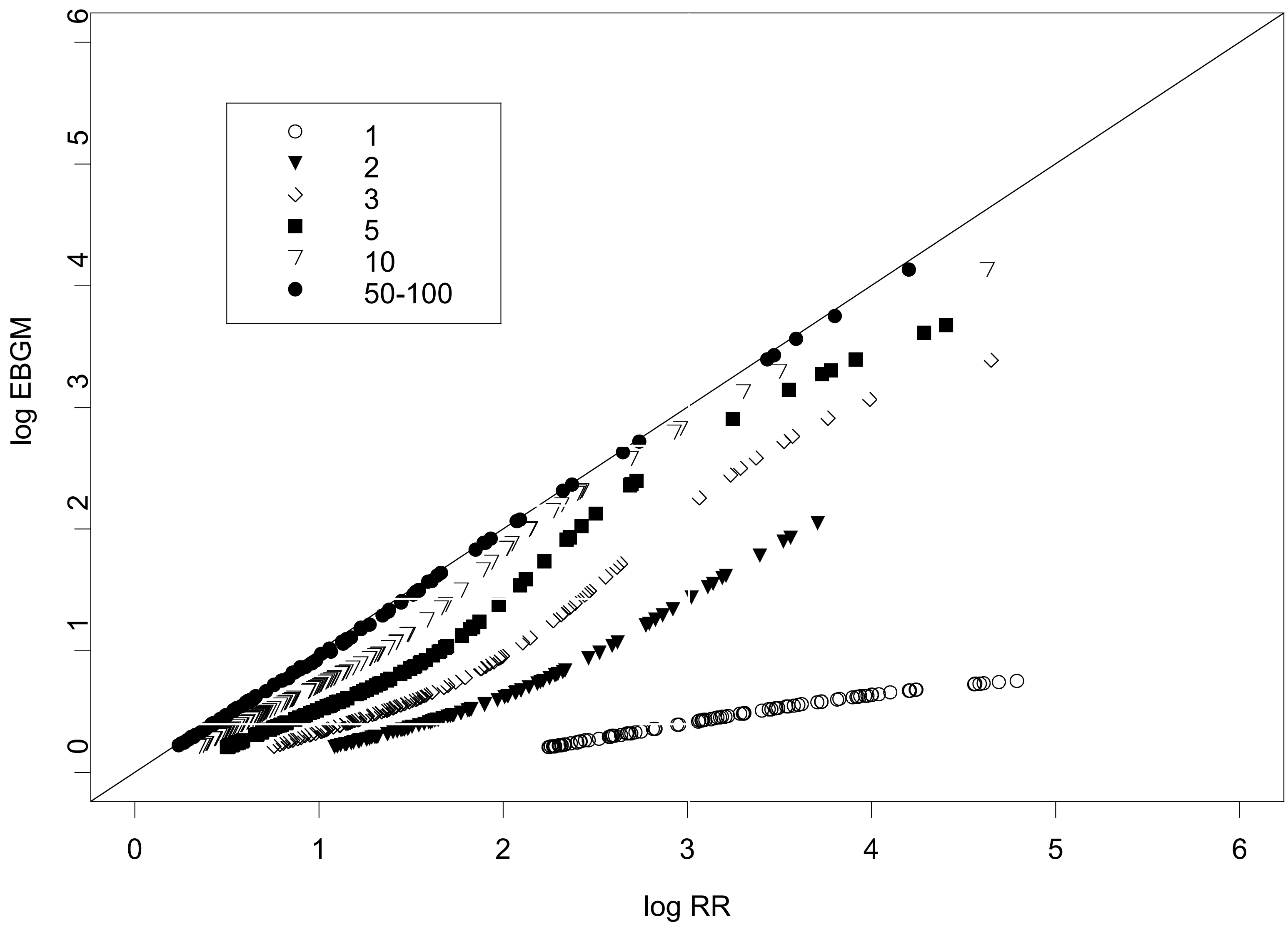
GPS/MGPS

- GPS/MGPS follows the same recipe as for the hospitals
- Denote by ρ_{ij} the true RR for Drug *i* and AE *j*
- Assumes the ρ_{ij}'s arise from a particular 5-parameter distribution
- Use empirical Bayes to use the data to estimate these five parameters.

GPS-EBGM

- Define $\lambda_{ij} = \mu_{ij} / E_{ij}$, where
  - $N_{ij} \sim \text{Poisson}(\mu_{ij})$
  - $\lambda_{ij} \mid \lambda \sim p * g(\lambda; \alpha_1, \beta_1) + (1-p) * g(\lambda; \alpha_2, \beta_2)$

    a mixture of two Gamma Distributions
- EBGM = Geometric mean of Post-Dist. of λ_{ij}
  - Estimates of μ_{ij} / E_{ij}
  - "Shrinks" $N_{ij} / E_{ij} \rightarrow 1$
  - Smaller variances than N_{ij} / E_{ij}

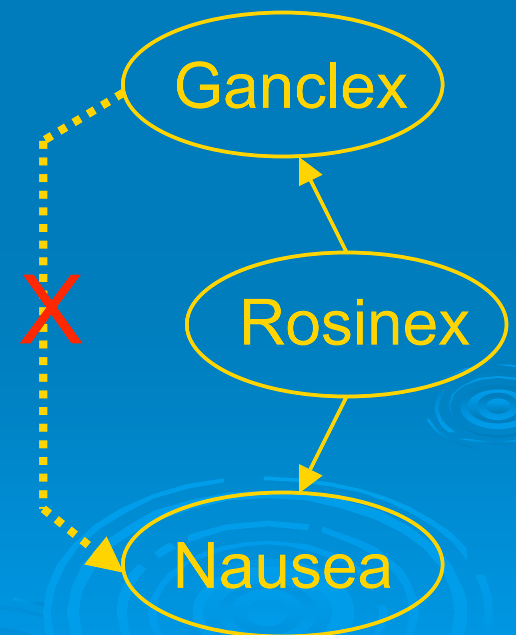


log EBGM
log RR
0
1
2
3
4
5
6
1
2
3
5
10
50-100

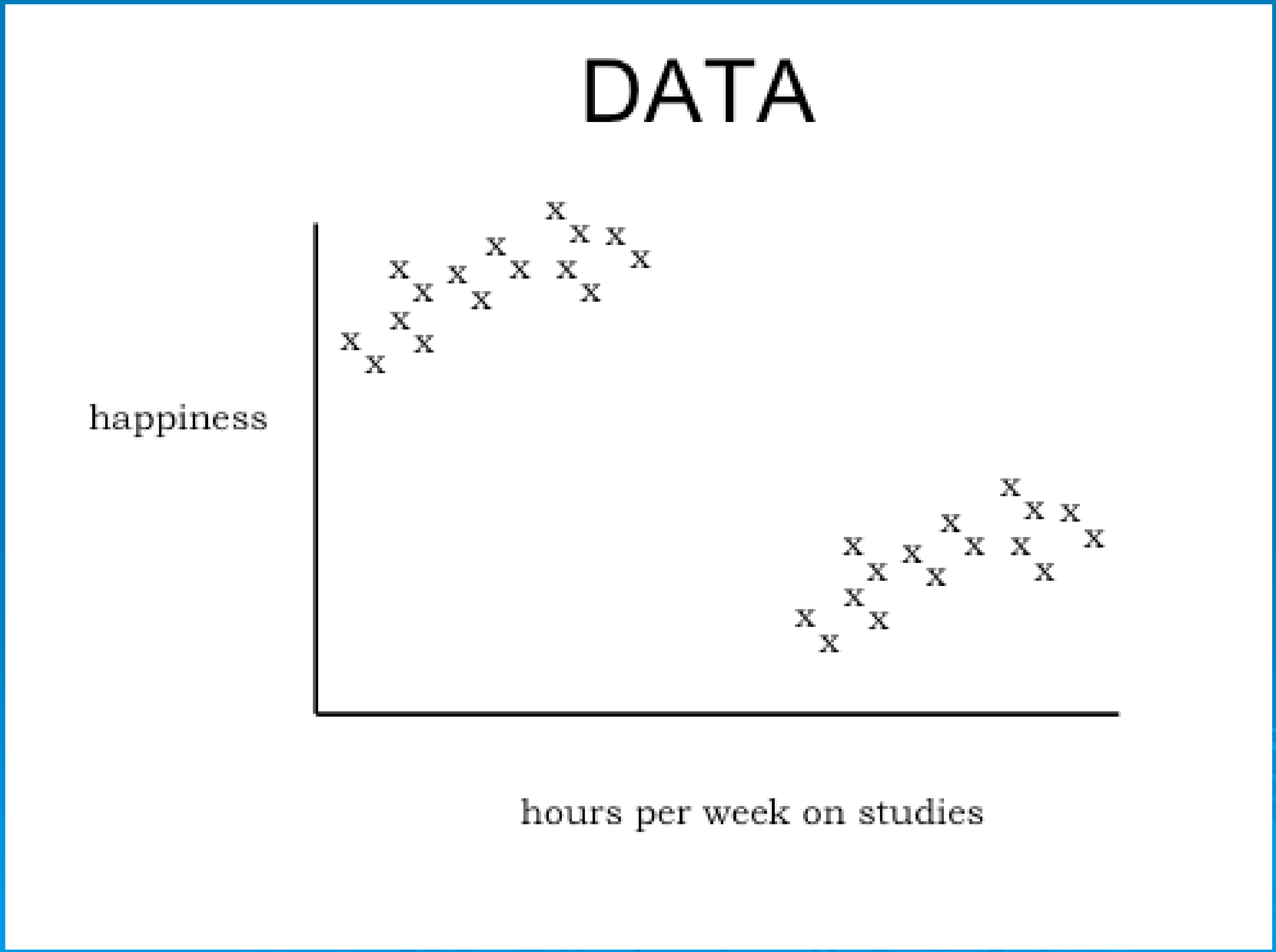
Simpson's Paradox

- Contingency table analysis ignores effects of drug-drug association on drug-AE association
- Simpson's Paradox

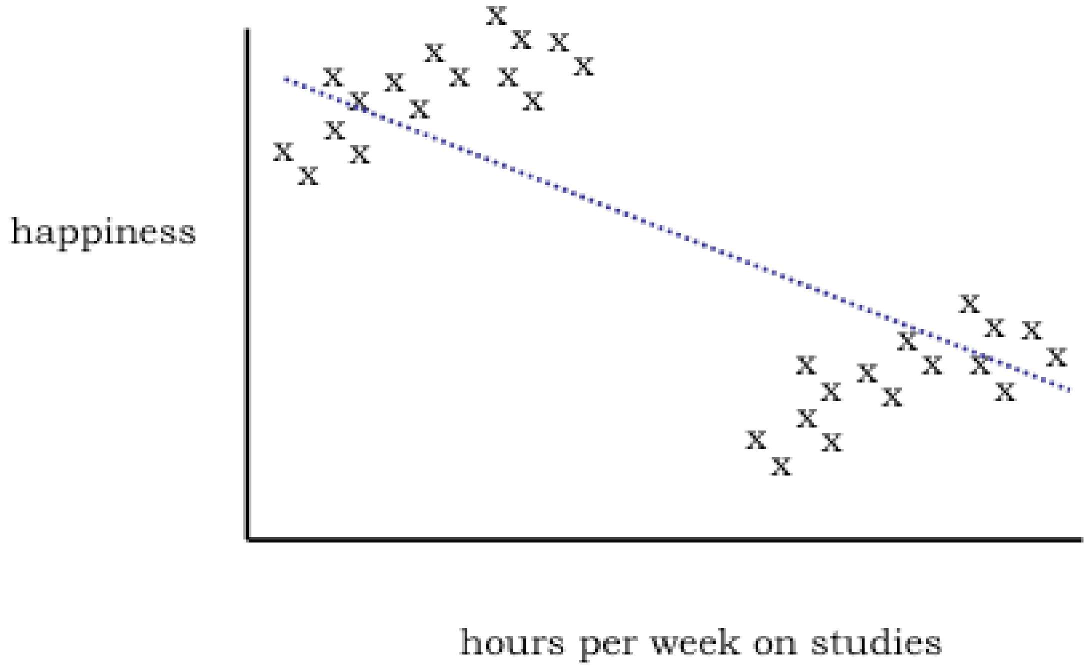
	Rosinex		No Rosinex		Total	
	Nausea	No Nausea	Nausea	No Nausea	Nausea	No Nausea
Ganclex	81	9	1	9	82	18
No Ganclex	9	1	90	810	99	811
RR	1		1		4.58	



Bad Things Can Happen…

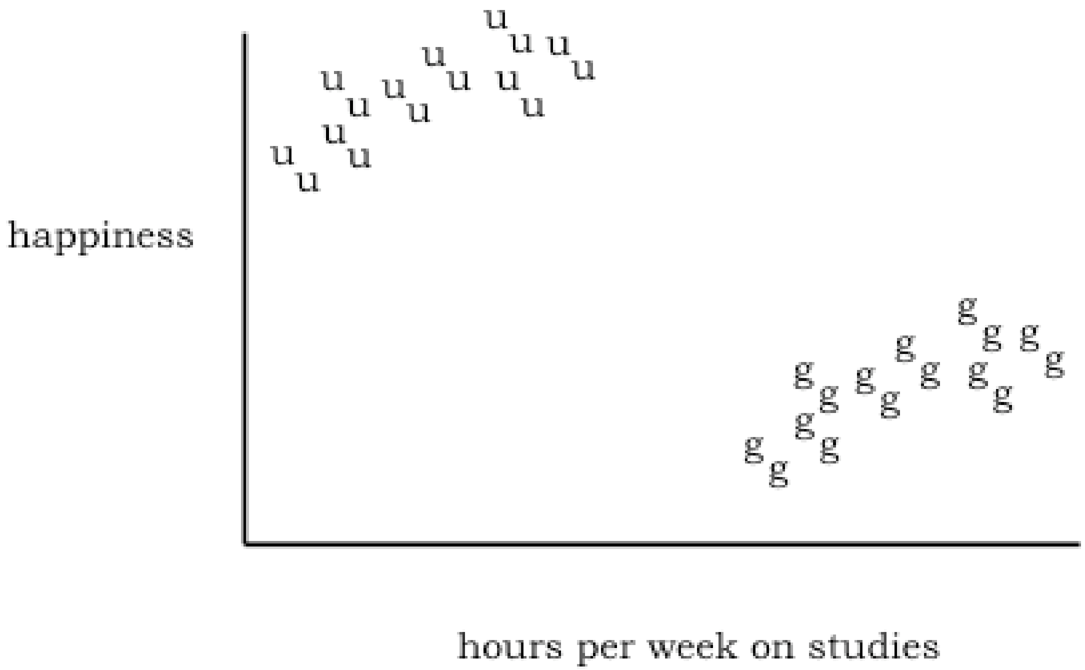


simple regression line

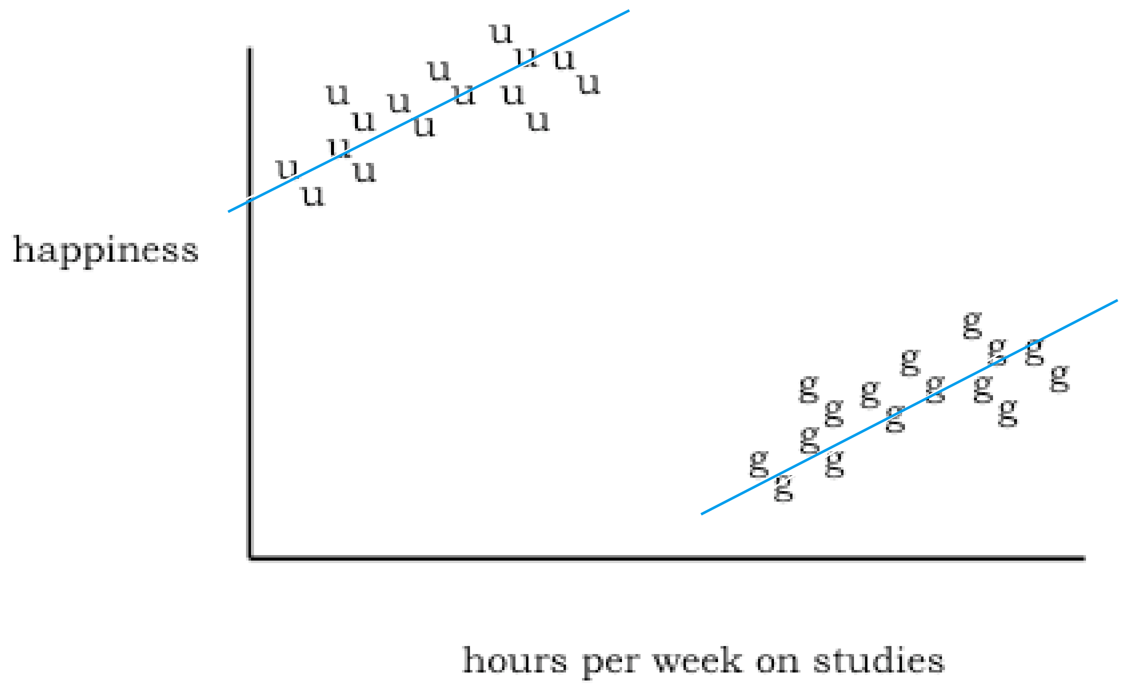


HAP = $\beta_0 + \beta_1$ x HOURS, β_1 will be estimated to be negative

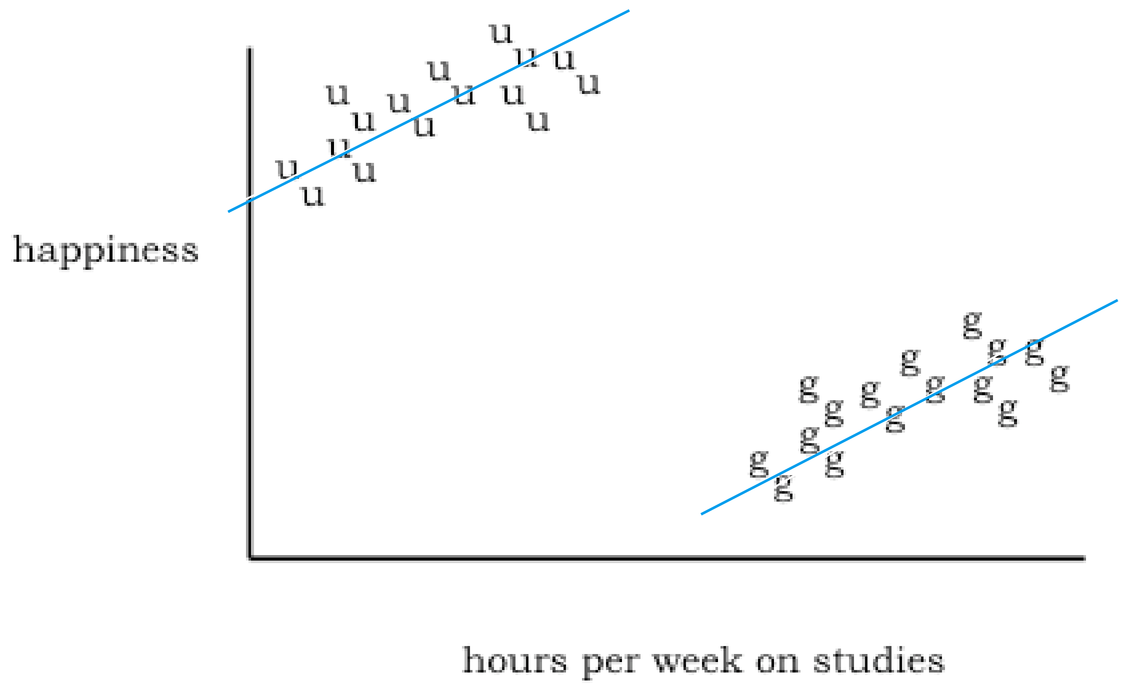
A 2nd Look at the DATA



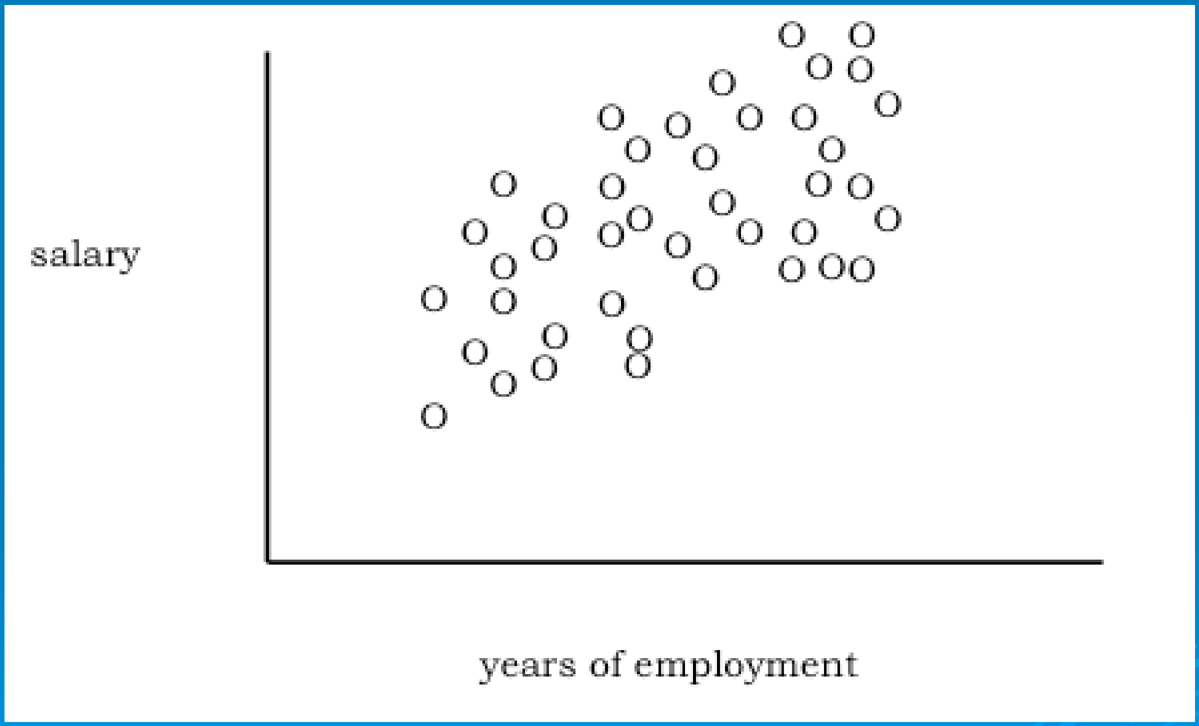
A 2nd Look at the DATA



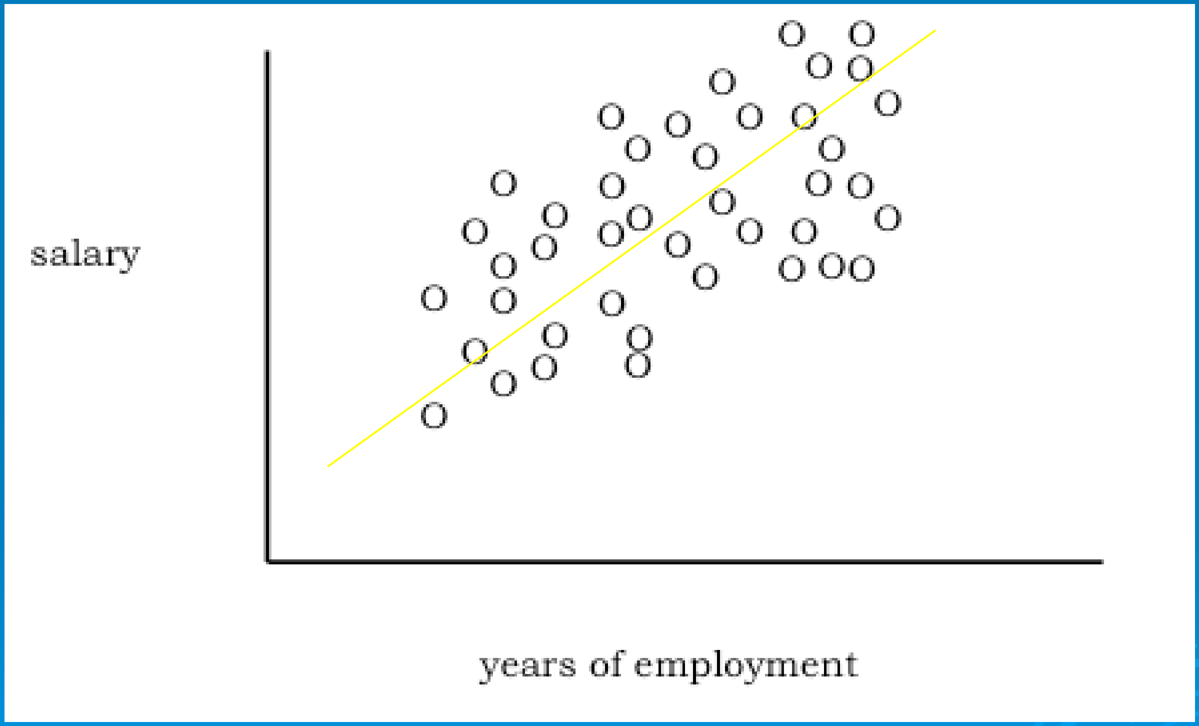
A 2nd Look at the DATA

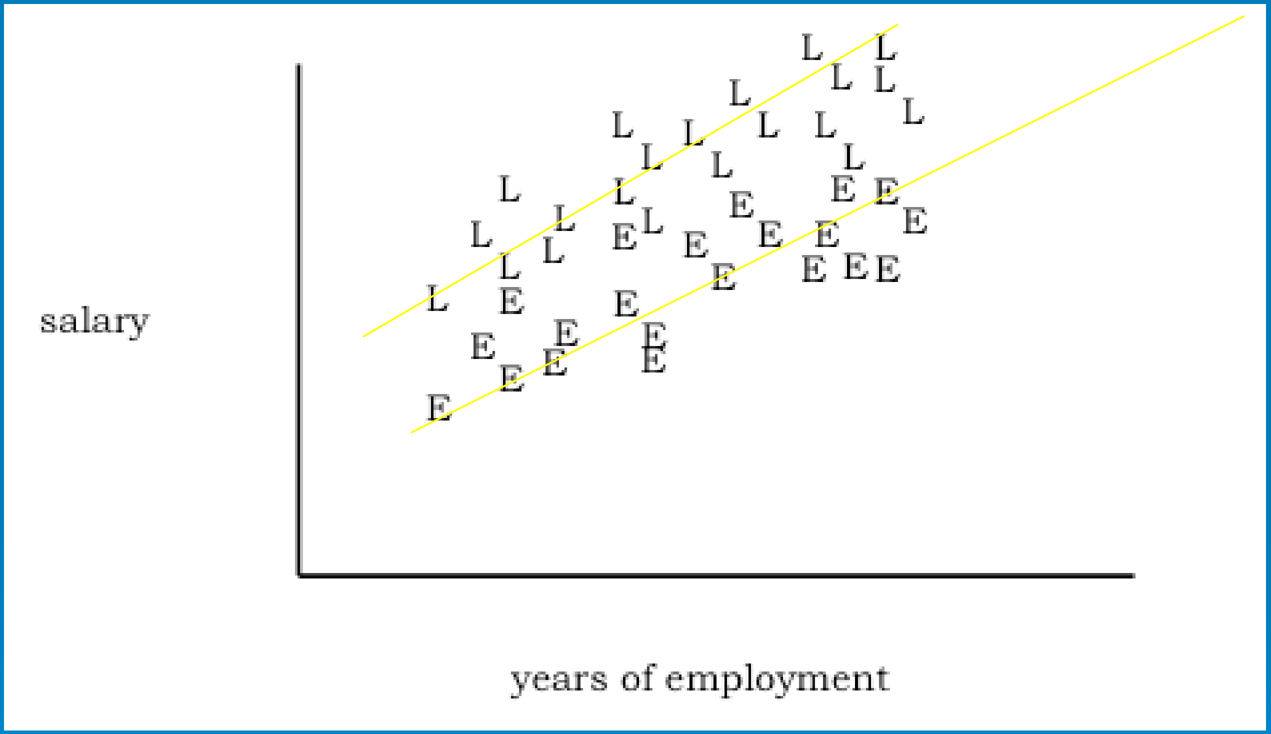


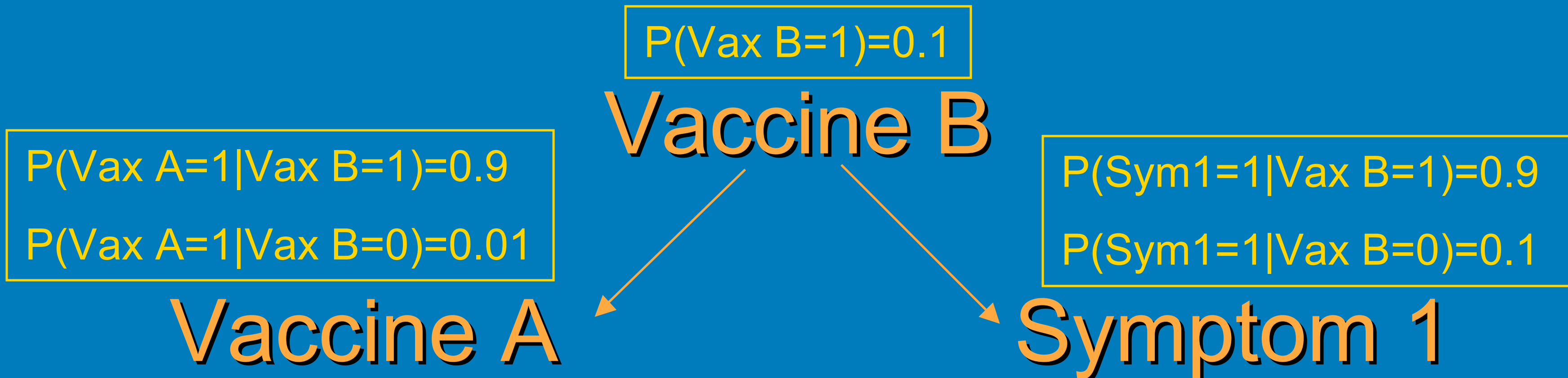
Other Odd Things Can Happen…



Other Odd Things Can Happen…




salary
years of employment



		Sym1 vs Vax A		Sym1 vs Vax B	
		Value	Rank	Value	Rank
N		1673	2	1826	1
Bayesian Logistic Method	Normal	-3.05E-02	4194	4.69	5
	Normal-CV	0.885	151	3.44	6
	Laplace	-3.00E-02	9136	4.69	13
	Laplace-CV	0.00	9127	3.99	7
GPS EBGM		2.84	73	3.02	68
Observed RR		2.84	744	3.03	681

Logistic Regression

- log [*P*/(1-*P*)] = intercept + ∑ (each drug effect)
  - *P* = Pr (report with these drugs will have the AE)
- Classic logistic regression hard to scale up
  - Huge number of predictors (drugs)
- Bayesian Logistic Regression (Shrinkage Method)
  - Put a prior on coefficients ($\beta_1, \ldots, \beta_p$), and shrink their estimates towards zero
    - Stabilize the estimation when there are many predictors
    - Bayesian solution to the multiple comparison problem

Bayesian Logistic Regression

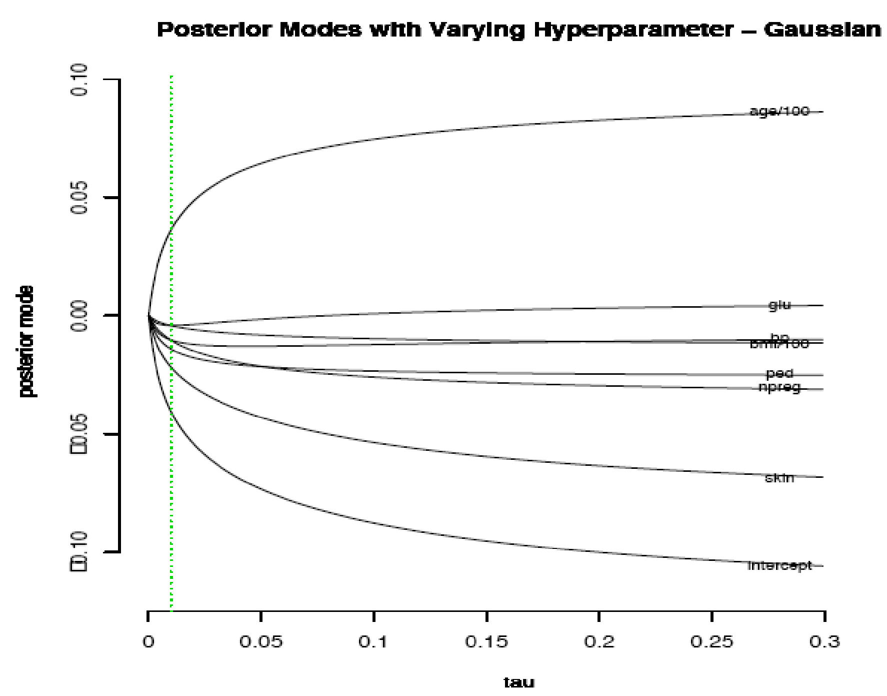
- Two shrinkage methods
  - Ridge regression - Gaussian prior

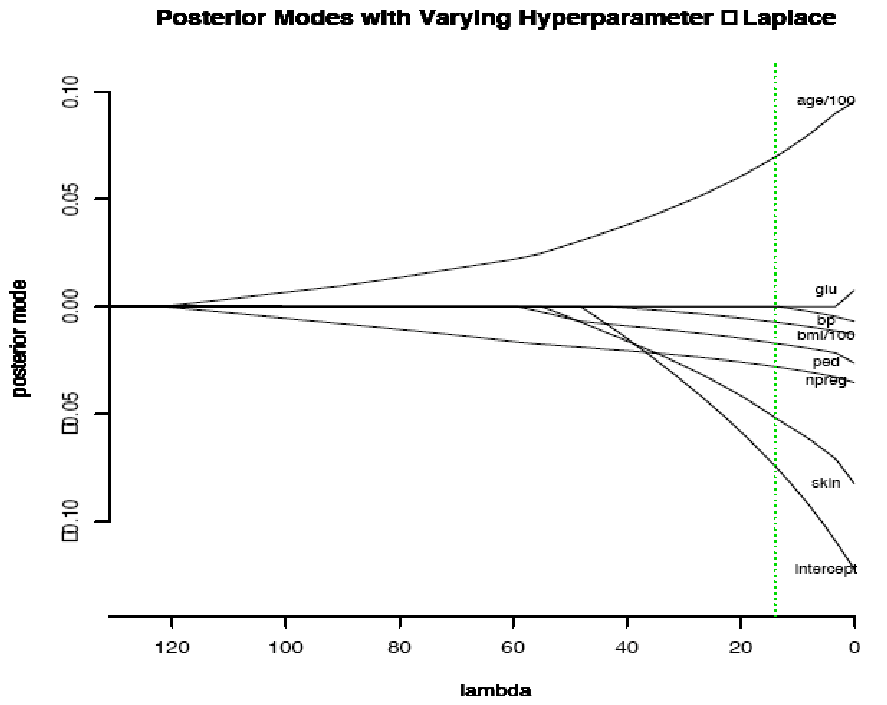
    $$\beta_j \sim N(0,\lambda)$$

  - Lasso regression - Laplace prior

    $$f(\beta_j) \propto \exp\{-\lambda\, |\beta_j|^{\lambda}\}$$

- Choosing hyperparameter λ
  - Decide how much to shrink
  - Cross-validation: choose prior to fit left-out data
  - Aggregation method by Bunea and Nobel (2005)


Posterior Modes with Varying Hyperparameter – Gaussian
posterior mode
tau
age/100
glu
bp
bmi/100
ped
npreg
skin
intercept


Posterior Modes with Varying Hyperparameter Laplace
posterior mode
lambda
age/100
glu
bp
bmi/100
ped
npreg
skin
intercept

Bayesian Logistic Regression

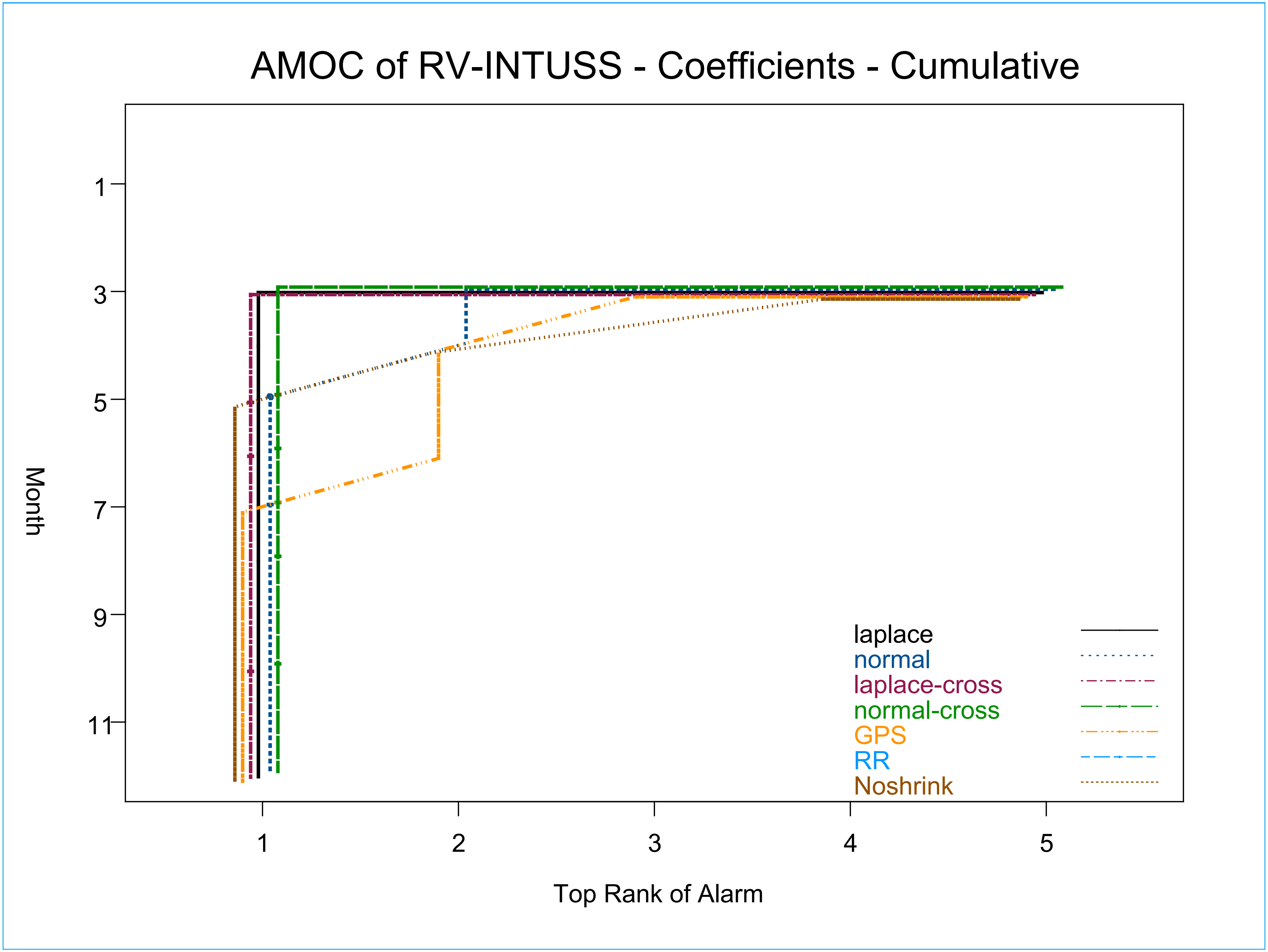
- Software: Bayesian Binary Regression (BBR)
  - http://stat.rutgers.edu/~madigan/BBR
  - Two priors: Gaussian and Laplace
  - Hyperparameter: fixed, default and CV
  - Handles millions of predictors efficiently
- Safety Signal: an apparent excess of an adverse effect associated with use of a drug
  - Coefficients β's – logs of odds ratios
  - $Pr(AE_j \mid drug_i) - Pr(AE_j \mid \text{not } drug_i)$

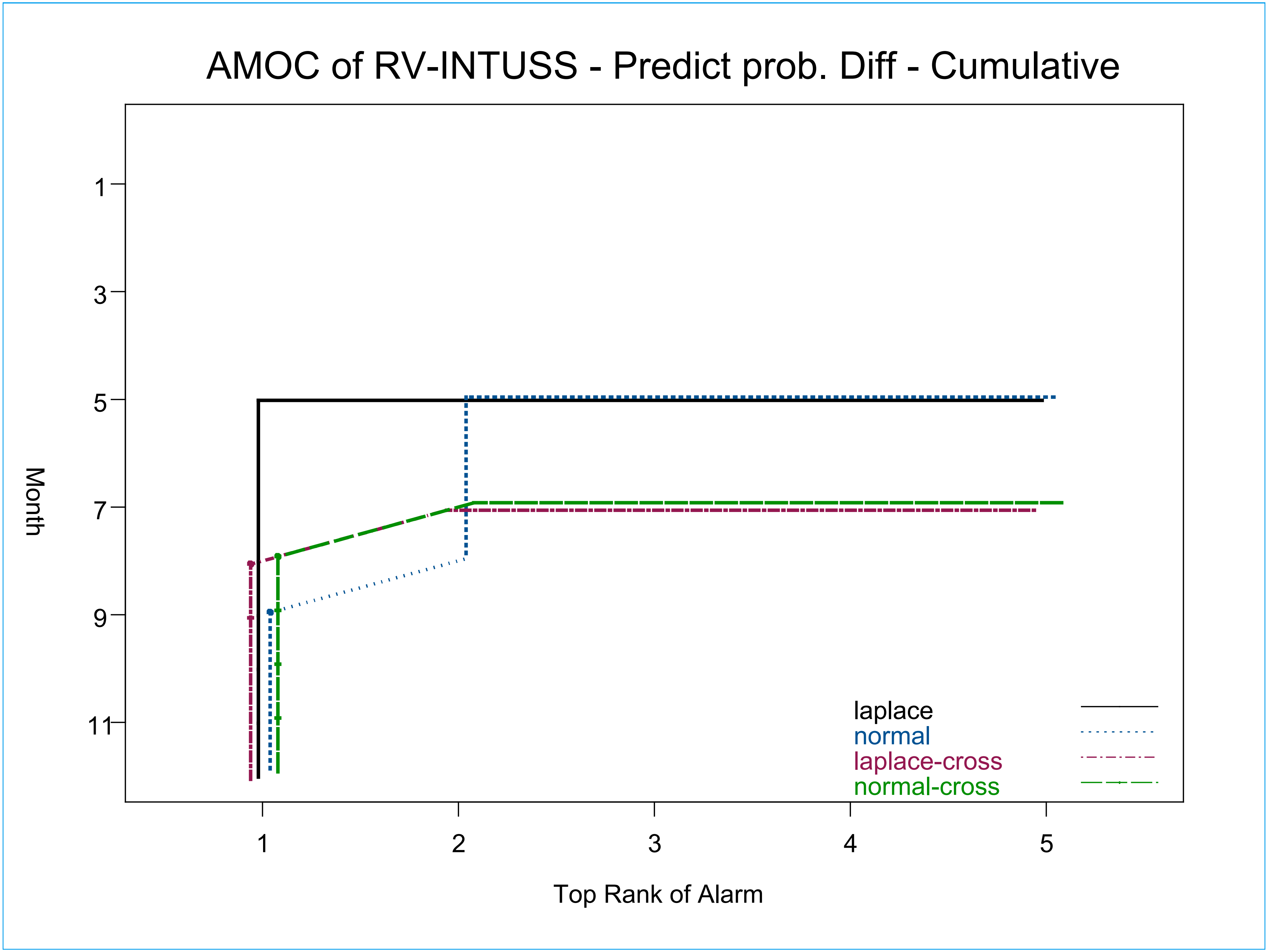
Evaluation Strategies

- Top-Rank Plot for Safety Signal
  - To compare the timeliness of outbreak detection
  - Similar to AMOC (Activity Monitor Operating Characteristic) curve in fraud detection
  - Y – window (month in 1999)
  - X – Top rank of association from window 1 to corresponding window

RV v.s. INTUSS

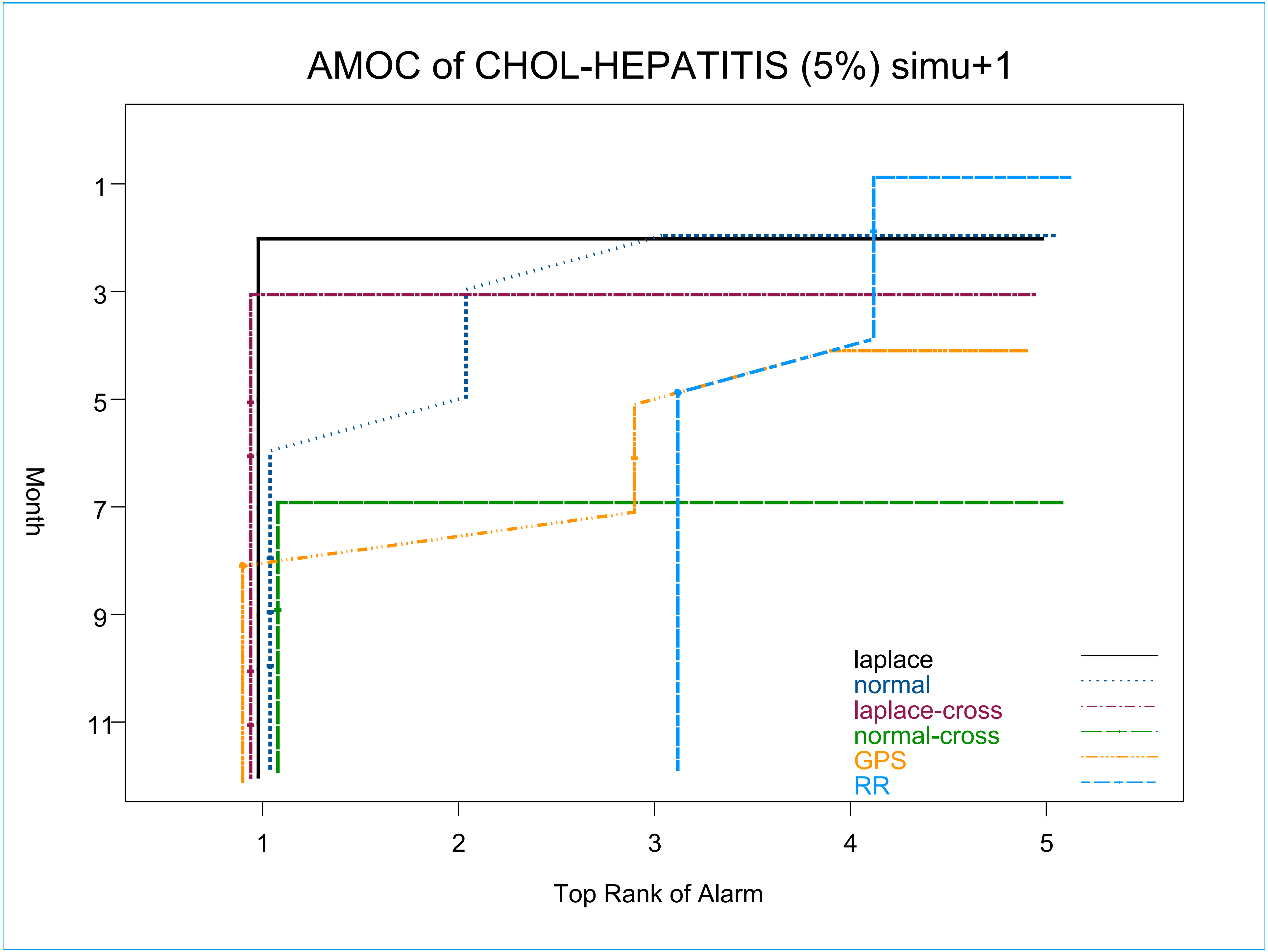
- Rotavirus
  - Severe diarrhea (with fever and vomiting)
  - Hospitalize 55,000 children each year in US
- Intussusception (INTUSS)
  - Uncommon type of bowel obstruction
- RotaShield (RV)
  - Licensed on 8/31/1998 in US
  - Recommended for routine use in infants
  - Increased the risk for intussusception
    - 1 or 2 cases among each 10,000 infants
  - On 10/14/1999, the manufacturer withdrew RV


AMOC of RV-INTUSS - Coefficients - Cumulative
Month
1
3
5
7
9
11
Top Rank of Alarm
1
2
3
4
5
laplace
normal
laplace-cross
normal-cross
GPS
RR
Noshrink

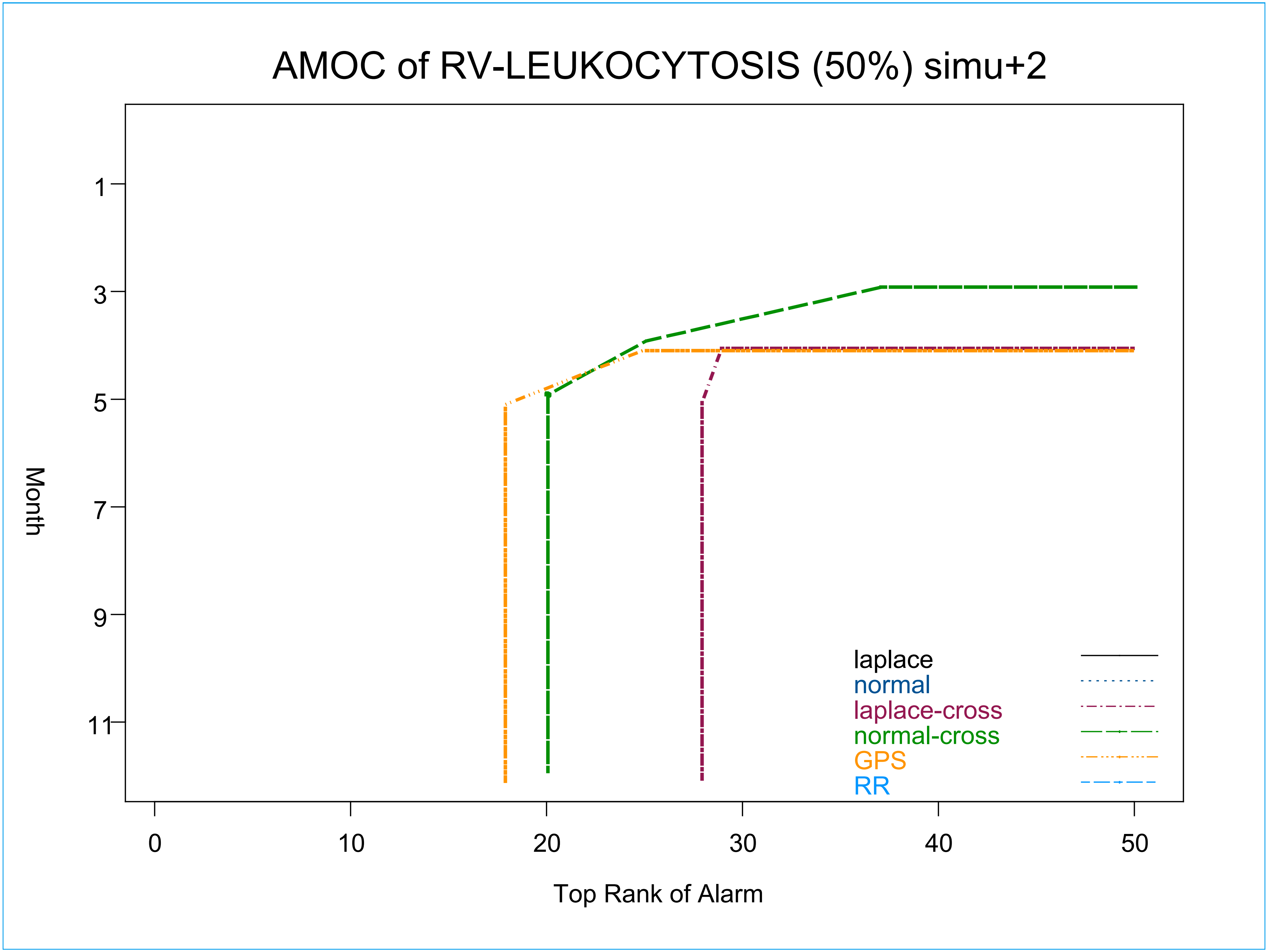

AMOC of RV-INTUSS - Predict prob. Diff - Cumulative
Month
1
3
5
7
9
11
Top Rank of Alarm
1
2
3
4
5
laplace
normal
laplace-cross
normal-cross

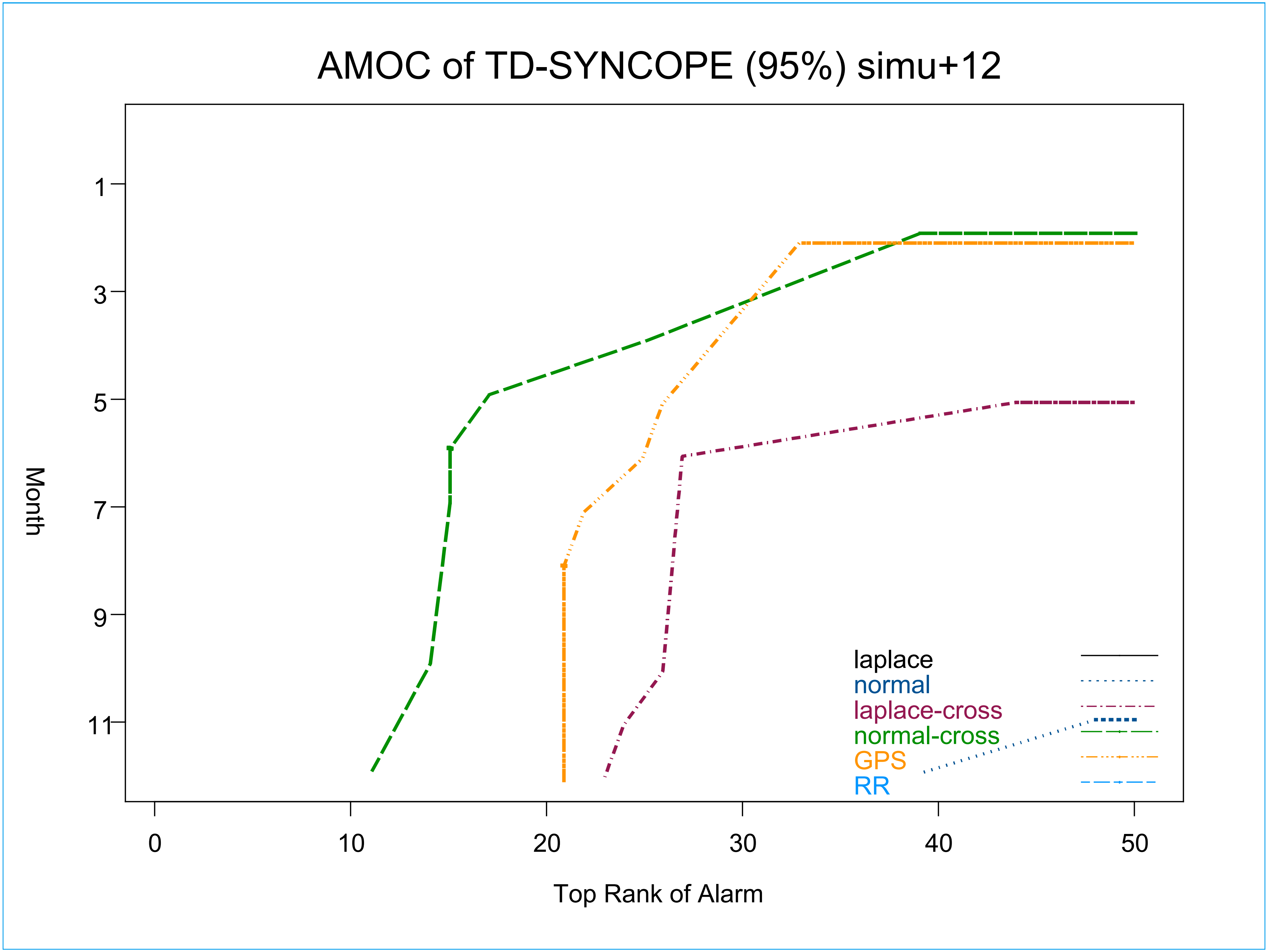
Simulation

- Step-by-step procedure
  - Choose either a rare (5%, 1), intermediate (50%, 3), or common (95%, 100) vaccine - adverse event (V-A) combination
  - Use year 1998 data as baseline
  - Add extra report(s) per month of 1999 containing the chosen V-A combination
  - Generate the AMOC curve

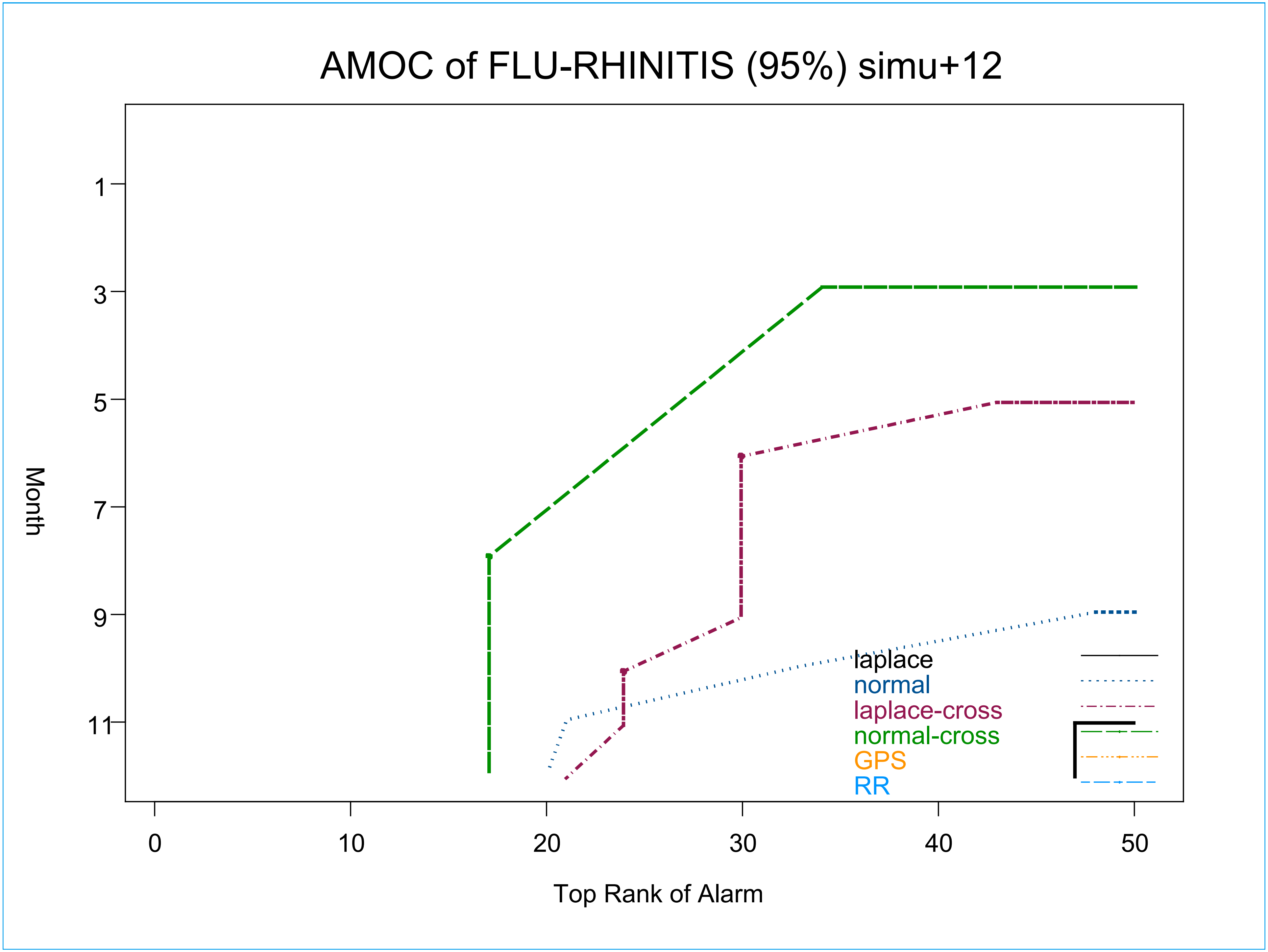


AMOC of CHOL-HEPATITIS (5%) simu+1
Month
1
3
5
7
9
11
Top Rank of Alarm
1
2
3
4
5
laplace
normal
laplace-cross
normal-cross
GPS
RR


AMOC of RV-LEUKOCYTOSIS (50%) simu+2
Month
1
3
5
7
9
11
0
10
20
30
40
50
Top Rank of Alarm
laplace
normal
laplace-cross
normal-cross
GPS
RR



AMOC of TD-SYNCOPE (95%) simu+12
Month
1
3
5
7
9
11
0
10
20
30
40
50
Top Rank of Alarm
laplace
normal
laplace-cross
normal-cross
GPS
RR



AMOC of FLU-RHINITIS (95%) simu+12
Month
1
3
5
7
9
11
0
10
20
30
40
50
Top Rank of Alarm
laplace
normal
laplace-cross
normal-cross
GPS
RR

Conclusions of Simulation

- The Bayesian Logistic Regressions (Normal-CV and Laplace-CV) signal consistently, and are at least as good as GPS method
- Simple RR cannot signal for intermediate and common cases
- GPS is relatively good on rare and intermediate cases, but not stable on common cases

Discussion of Logistic Method

- Advantages over low-dimensional tables
  - Correct confounding and mask effect
  - Analyze multiple drugs/vaccines simultaneously
- Limitations
  - Build separate model for each AE
    - Ignore dependencies between AEs
  - Fail to adjust for unmeasured/unrecorded factors
    - health status, unreported drugs, etc.
  - Model-based approach
    - Require model assumptions

Causal Inference View

- Rubin's causal model
  - Potential outcomes

    Factual outcome

    I took an aspirin and my headache went away

    Counterfactual outcome

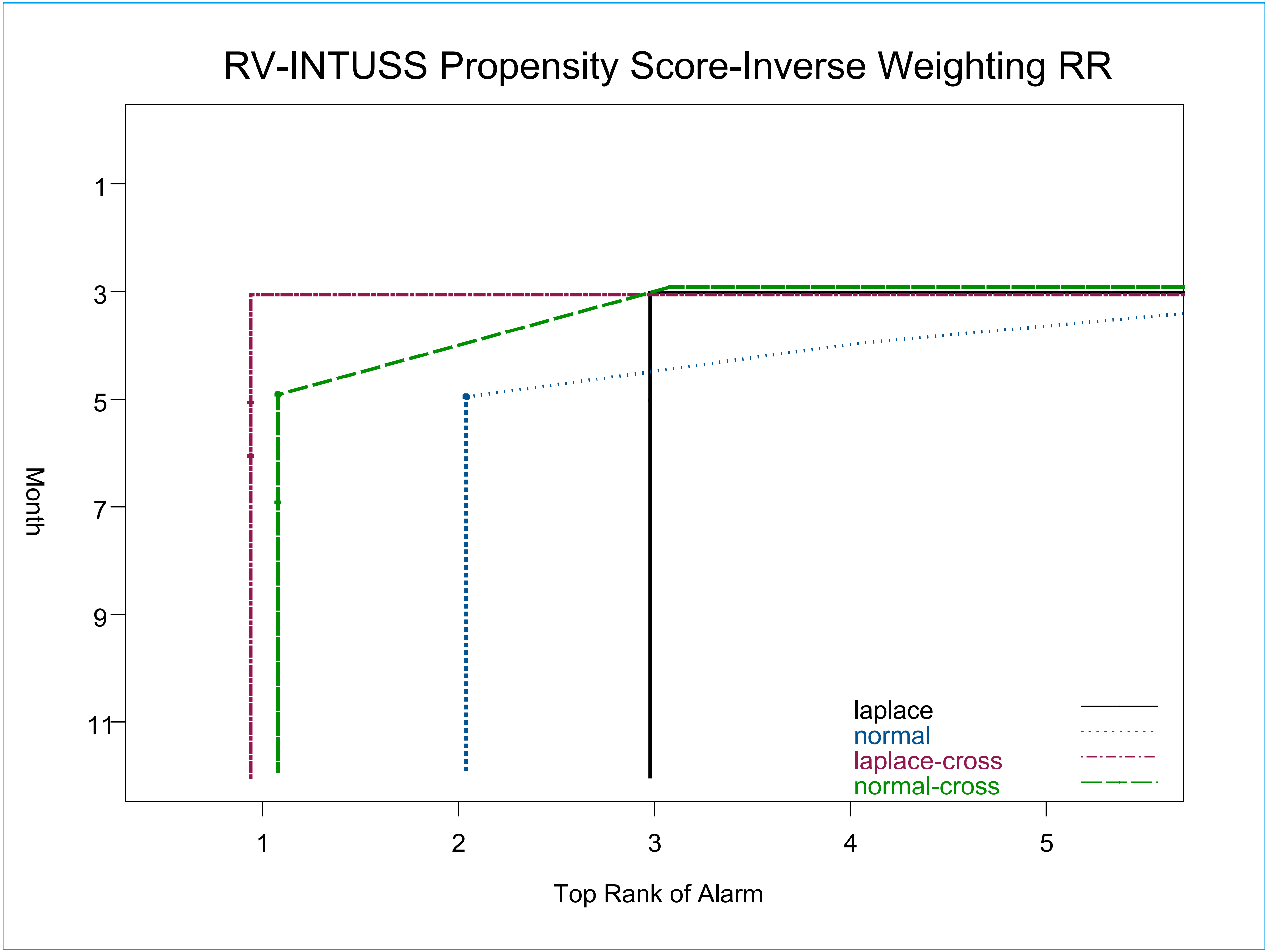
    If I hadn't taken an aspirin, I'd still have a headache
- Define:
  - Z_i : treatment applied to unit i (0=control, 1=treat)
  - $Y_i(0)$: response for unit *i* if $Z_i = 0$
  - $Y_i(1)$: response for unit *i* if $Z_i = 1$
  - Unit level causal effect: $Y_i(1) - Y_i(0)$
  - Fundamental problem: only see one of these!

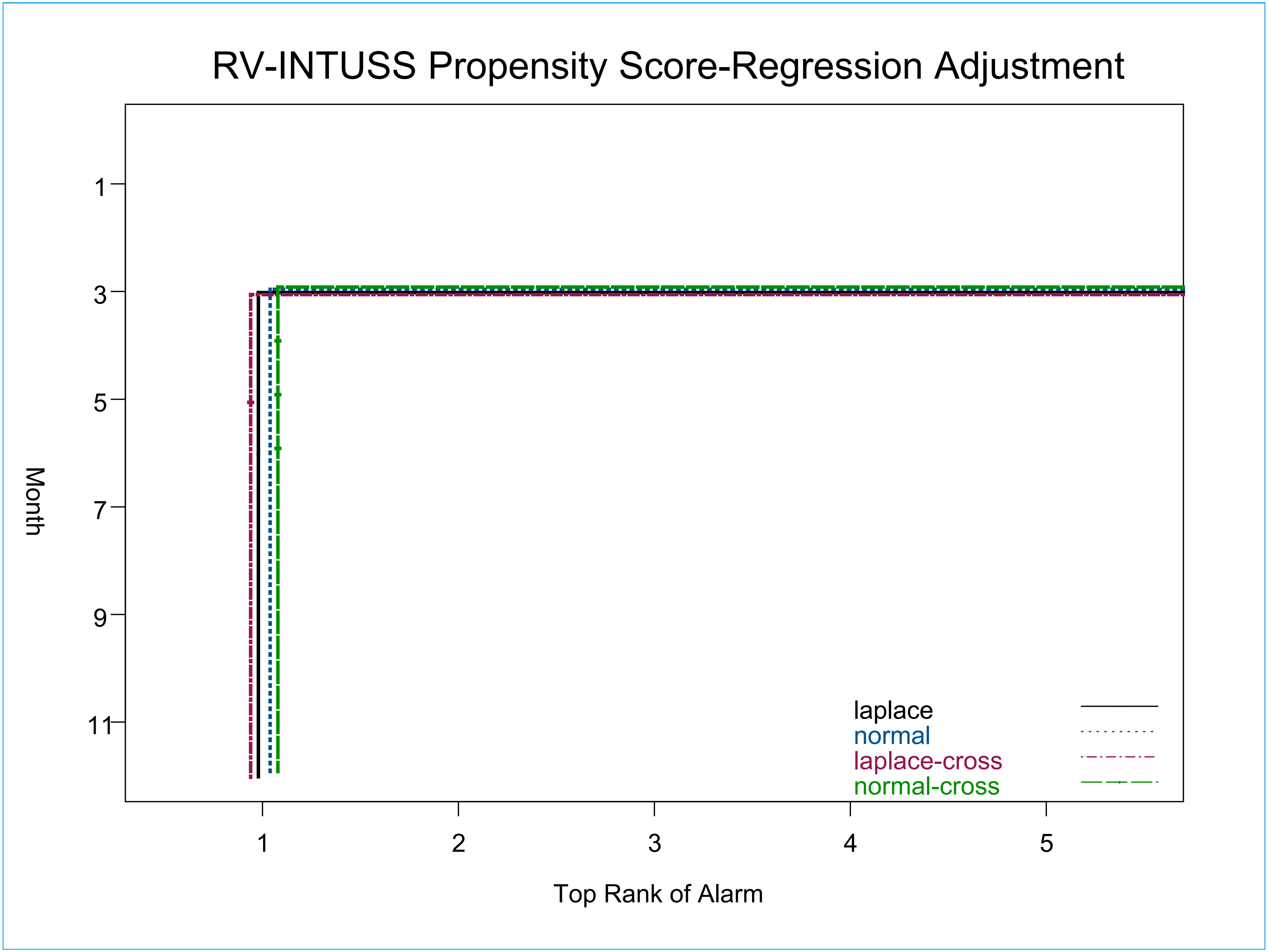
Bias Due To Confounding

- Individuals are observed already under their respective conditions
- The two groups may differ in ways other than just the observed condition
- Average effects may be biased due to confounding between covariates and group condition
- We can simulate randomization or counterfactual world using information from observational study…sort of

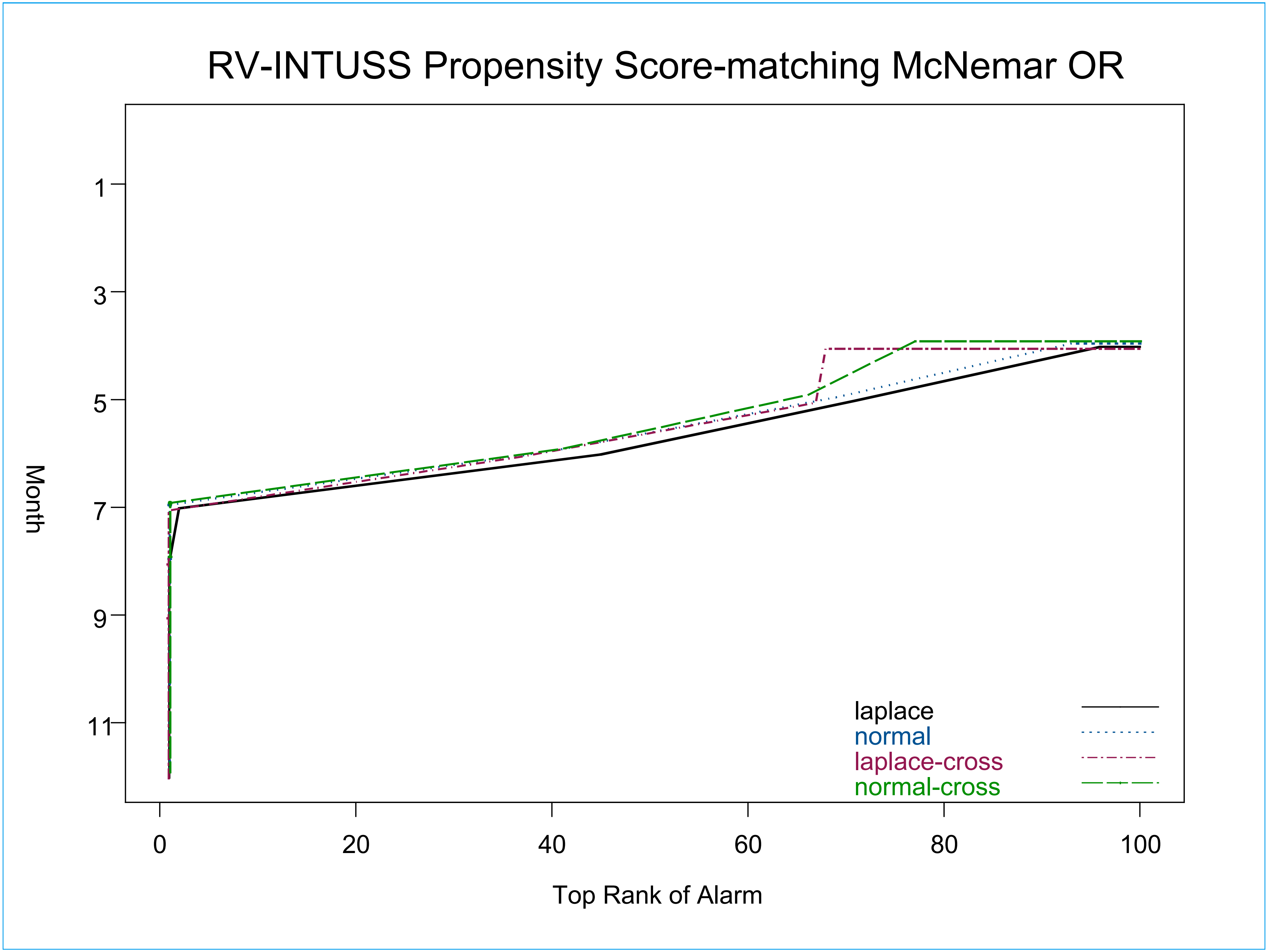
Propensity Score Method

- Definition
  - $e(x_i) = P(Z_i=1 \mid X_i=x_i)$
    Conditional probability of assignment to test treatment $Z_i=1$ given observed covariates
  - Assuming no unmeasured confounders, stratifying on $e(x_i)$ leads to causal inferences just as valid as in randomized trials
- Methods with propensity scores:
  - Inverse weighting
  - Regression adjustment
  - Matching


RV-INTUSS Propensity Score-Inverse Weighting RR
Month
1
3
5
7
9
11
Top Rank of Alarm
1
2
3
4
5
laplace
normal
laplace-cross
normal-cross



RV-INTUSS Propensity Score-Regression Adjustment
Month
1
3
5
7
9
11
1
2
3
4
5
Top Rank of Alarm
laplace
normal
laplace-cross
normal-cross


RV-INTUSS Propensity Score-matching McNemar OR
Month
1
3
5
7
9
11
0
20
40
60
80
100
Top Rank of Alarm
laplace
normal
laplace-cross
normal-cross

Conclusion

- "First generation" Method
  - Contingency table methods
  - Deal with each drug and each adverse event in isolation
- "Second generation" Method
  - Bayesian logistic regression
  - Propensity score
  - Deal with large numbers of drugs jointly and with multi-drug interactions
- Ultimate Method
  - Not only interactions and relationships among drugs , but also adverse events
  - Question: which sets of drugs cause which sets of adverse events?