

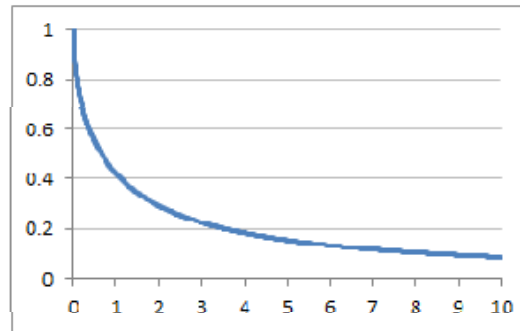
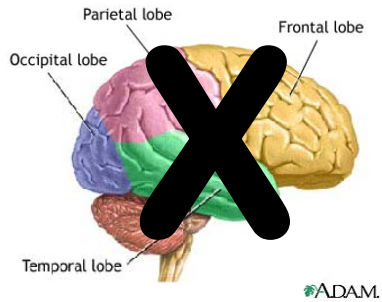


Bayes to the Rescue

David Madigan
Columbia University



Bayesian Learning Paradigm



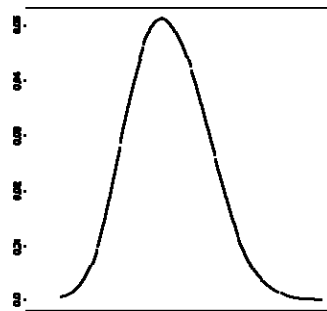
prior



Name	Thread pitch (mm)	Minor diameter tolerance	Nominal diameter (mm)	Head shape	Price for 50 screws	Available at factory outlet?	Number in stock	Flat or Phillips head?
M4	0.7	4c	4	Pan	\$10.08	Yes	276	Flat
M5	0.8	4c	5	Round	\$13.69	Yes	100	Both
M6	1	5c	6	Button	\$10.42	Yes	1043	Flat
M8	1.25	5c	8	Pan	\$11.68	No	296	Phillips
M10	1.5	6c	10	Round	\$16.74	Yes	486	Phillips
M12	1.75	7c	12	Pan	\$18.26	No	996	Flat
M14	2	7c	14	Round	\$21.19	No	231	Phillips
M16	2	8c	16	Button	\$23.57	Yes	292	Both
M18	2.1	8c	18	Button	\$26.67	No	664	Both
M20	2.4	9c	20	Pan	\$29.00	Yes	496	Both
M24	2.55	9g	24	Round	\$33.01	Yes	982	Phillips
M28	2.7	10g	28	Button	\$35.66	No	1067	Phillips
M36	3.2	12g	36	Pan	\$41.32	No	434	Both
M50	4.5	15g	50	Pan	\$44.72	No	740	Flat

data

Bayes' Rule

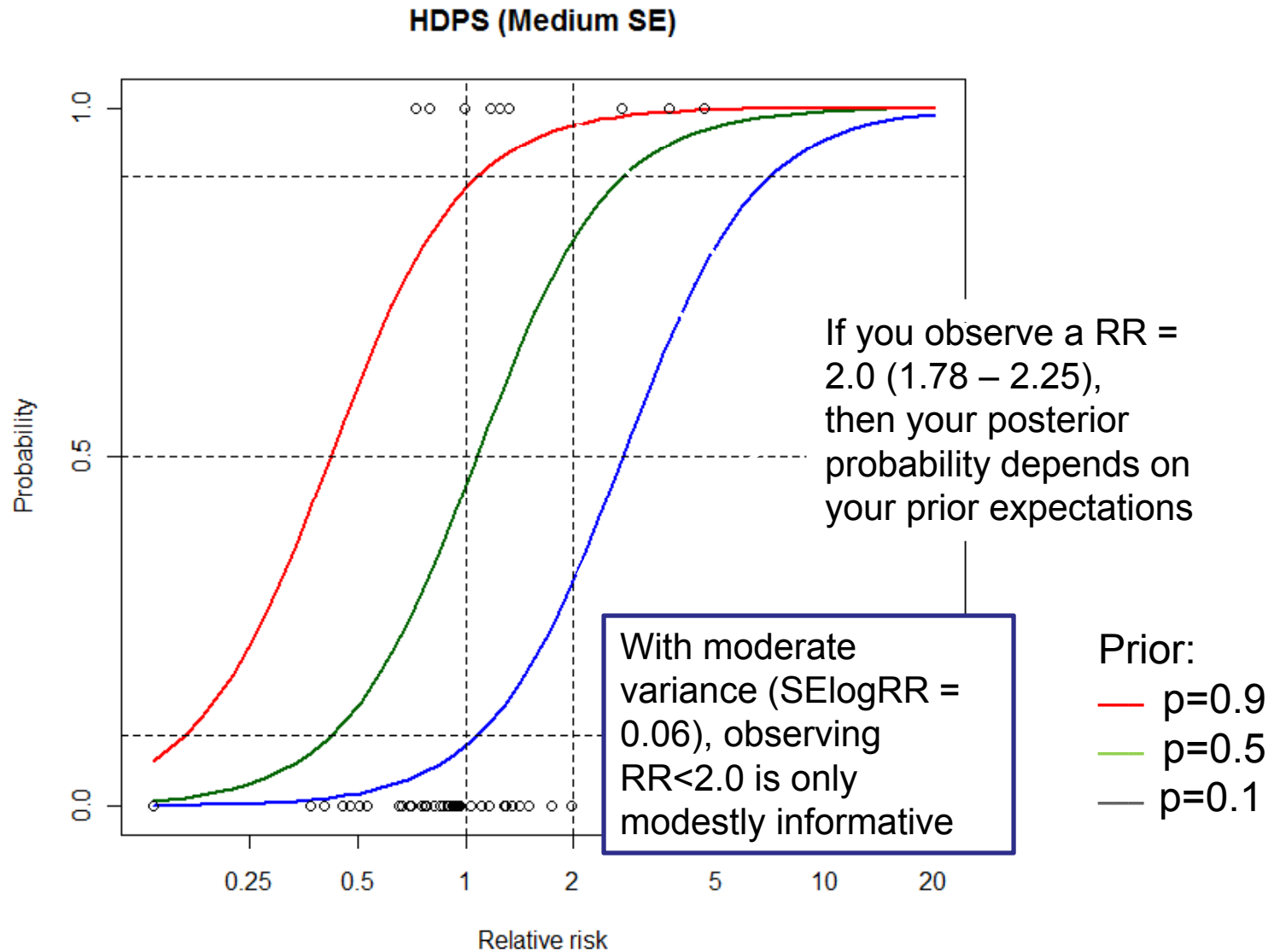


posterior

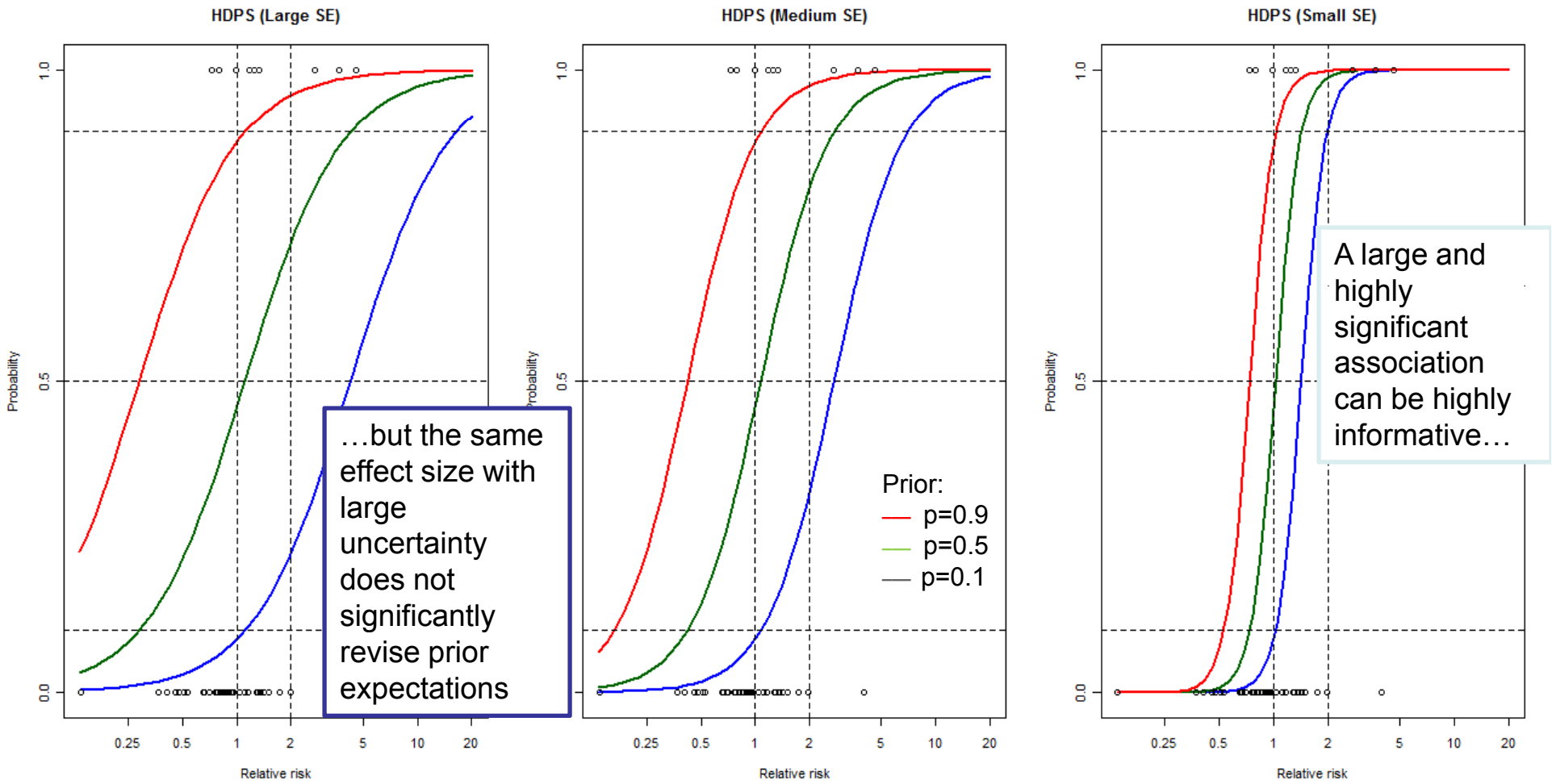
“Data”

Effect estimates of HDPS against CCAE (RR, SE)	Angioedema #1	Aplastic Anemia #1	Acute Liver Failure #1	Bleeding #1	Acute myocardial Infarction #1	Hip Fracture #1	Mortality after Myocardial Infarction #1	Acute Renal Failure #1	Upper GI Ulcer Hospitalization #1
OMOP ACE Inhibitor	1.80 (0.15)	0.40 (0.05)				0.91 (0.12)			0.87 (0.03)
OMOP Amphotericin B		3.30 (0.99)	1.05 (0.24)					4.01 (0.99)	
OMOP Antibiotics		1.22 (0.08)	1.00 (0.01)	1.14 (0.01)	1.06 (0.03)	1.05 (0.09)		1.44 (0.06)	
OMOP Antiepileptics	1.74 (0.38)	4.60 (0.80)						1.63 (0.21)	0.54 (0.05)
OMOP Benzodiazepines	0.13 (0.01)	1.10 (0.06)	0.98 (0.01)	1.11 (0.01)	1.18 (0.03)	1.41 (0.12)		1.06 (0.05)	
OMOP Beta blockers	0.81 (0.07)	0.63 (0.06)	0.95 (0.02)			1.69 (0.19)		0.78 (0.04)	0.88 (0.03)
OMOP Bisphosphonates		0.27 (0.05)	0.85 (0.03)		0.82 (0.07)			0.40 (0.04)	0.90 (0.06)
OMOP Tricyclic antidepressants		0.63 (0.07)	1.02 (0.02)	0.96 (0.01)	0.80 (0.04)			0.82 (0.06)	
OMOP Typical antipsychotics					0.96 (0.08)			1.97 (0.16)	3.46 (0.21)
OMOP Warfarin	0.53 (0.11)	0.47 (0.04)		2.13 (0.04)		1.2 (0.09)	0.49 (0.07)	0.76 (0.05)	

Revising prior expectations in light of new evidence from an active surveillance system



Impact of precision of observed estimates



...but the same effect size with large uncertainty does not significantly revise prior expectations

A large and highly significant association can be highly informative...

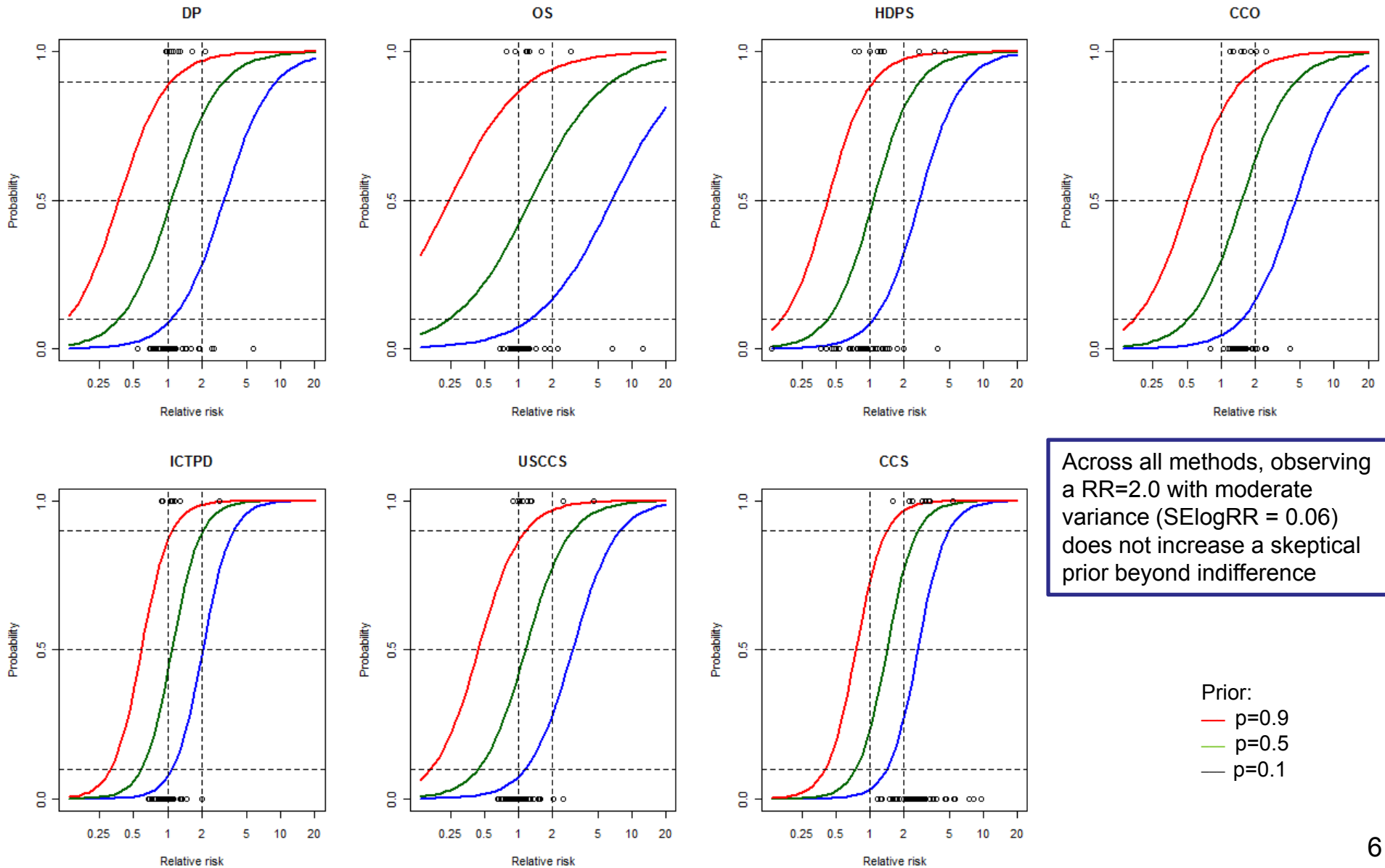
Scenarios: You observe RR=2.0 with confidence intervals based on standard error (SE):

Large SE: (1.01 – 3.97)

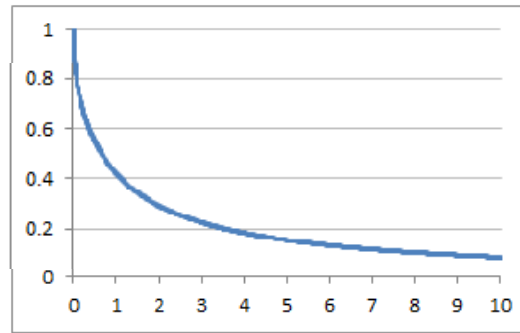
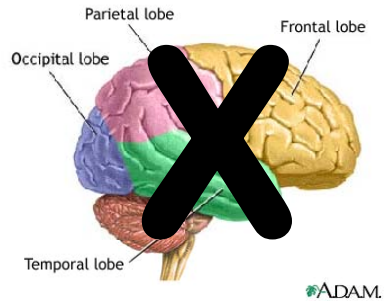
Medium SE: (1.78 – 2.25)

Small SE: (1.96 – 2.04)

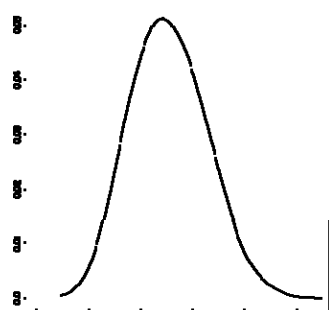
Impact of using estimates from different methods



Bayesian Learning Paradigm



prior

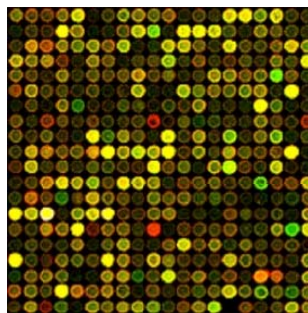


prior

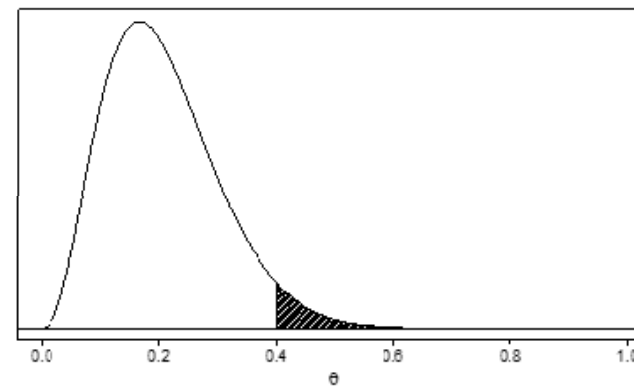
Name	Thread pitch (mm)	Minor diameter tolerance	Nominal diameter (mm)	Head shape	Price for 50 screws	Available at factory outlet?	Number in stock	Flat or Phillips head?
M4	0.7	4g	4	Pan	\$10.08	Yes	276	Flat
M5	0.8	4g	5	Round	\$13.89	Yes	183	Both
M6	1	5g	6	Button	\$10.42	Yes	1043	Flat
M8	1.25	5g	8	Pan	\$11.98	No	298	Phillips
M10	1.5	6g	10	Round	\$16.74	Yes	488	Phillips
M12	1.75	7g	12	Pan	\$18.26	No	998	Flat
M14	2	7g	14	Round	\$21.19	No	235	Phillips
M16	2	8c	16	Button	\$23.57	Yes	292	Both
M18	2.1	8c	18	Button	\$25.67	No	664	Both
M20	2.4	8c	20	Pan	\$29.09	Yes	486	Both
M24	2.55	9c	24	Round	\$33.01	Yes	662	Phillips
M28	2.7	10c	28	Button	\$35.66	No	1067	Phillips
M36	3.2	12c	36	Pan	\$41.22	No	434	Both
M50	4.5	15c	50	Pan	\$44.72	No	740	Flat

data

Bayes' Rule



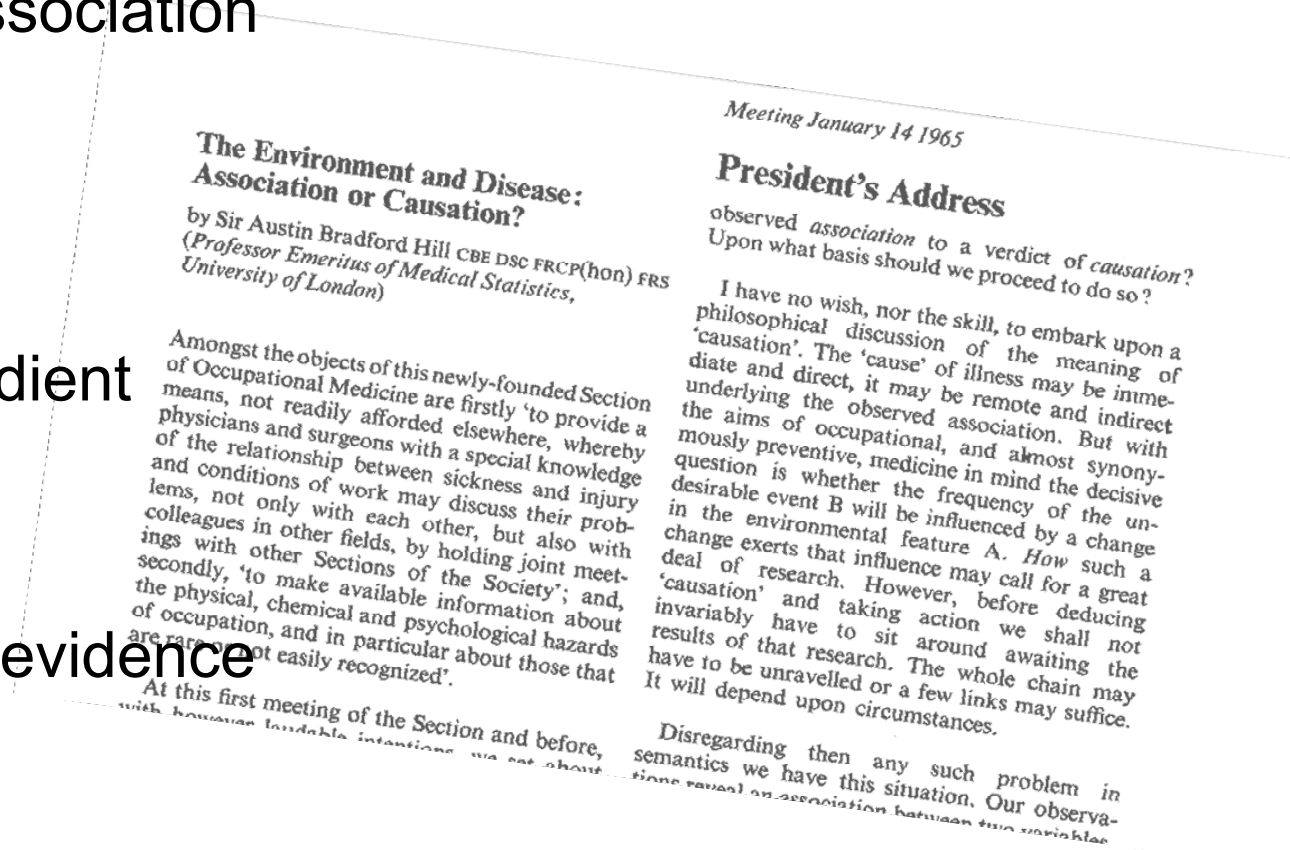
data



θ

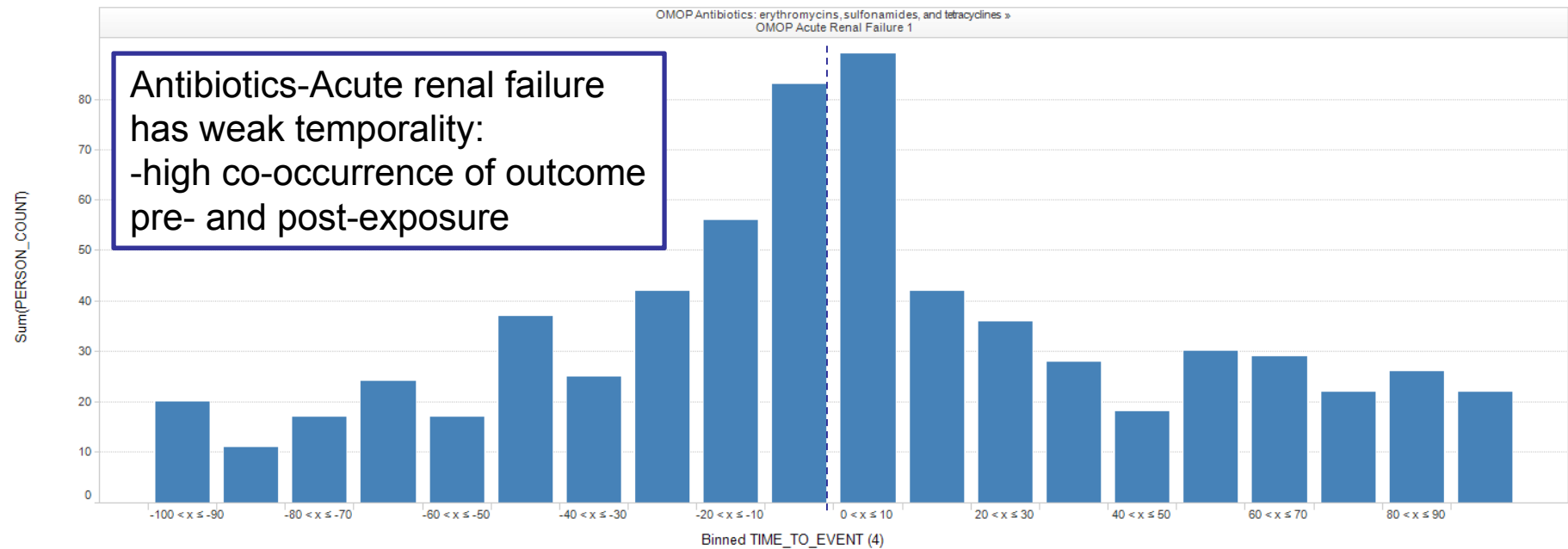
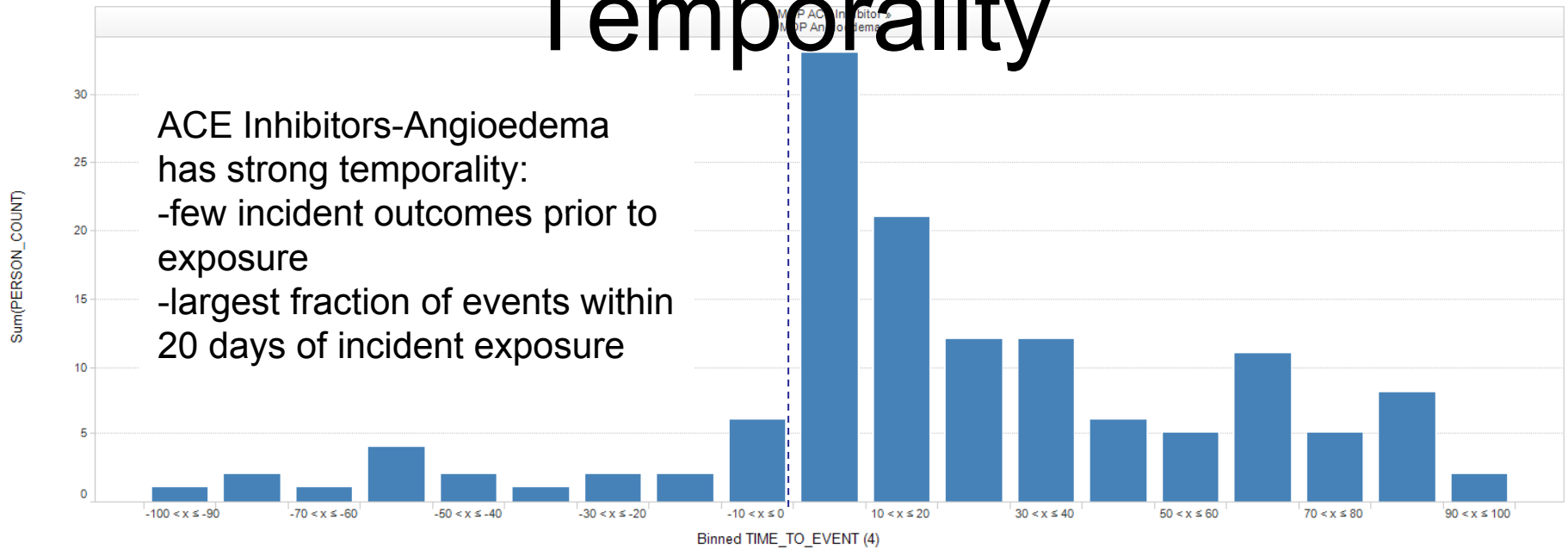
Hill's causality viewpoints

- Strength of association
- Consistency
- Specificity
- Temporality
- Biological gradient
- Plausibility
- Coherence
- Experimental evidence
- Analogy



Austin Bradford Hill, "The Environment and Disease: Association or Causation?," *Proceedings of the Royal Society of Medicine*, 58 (1965), 295-300.

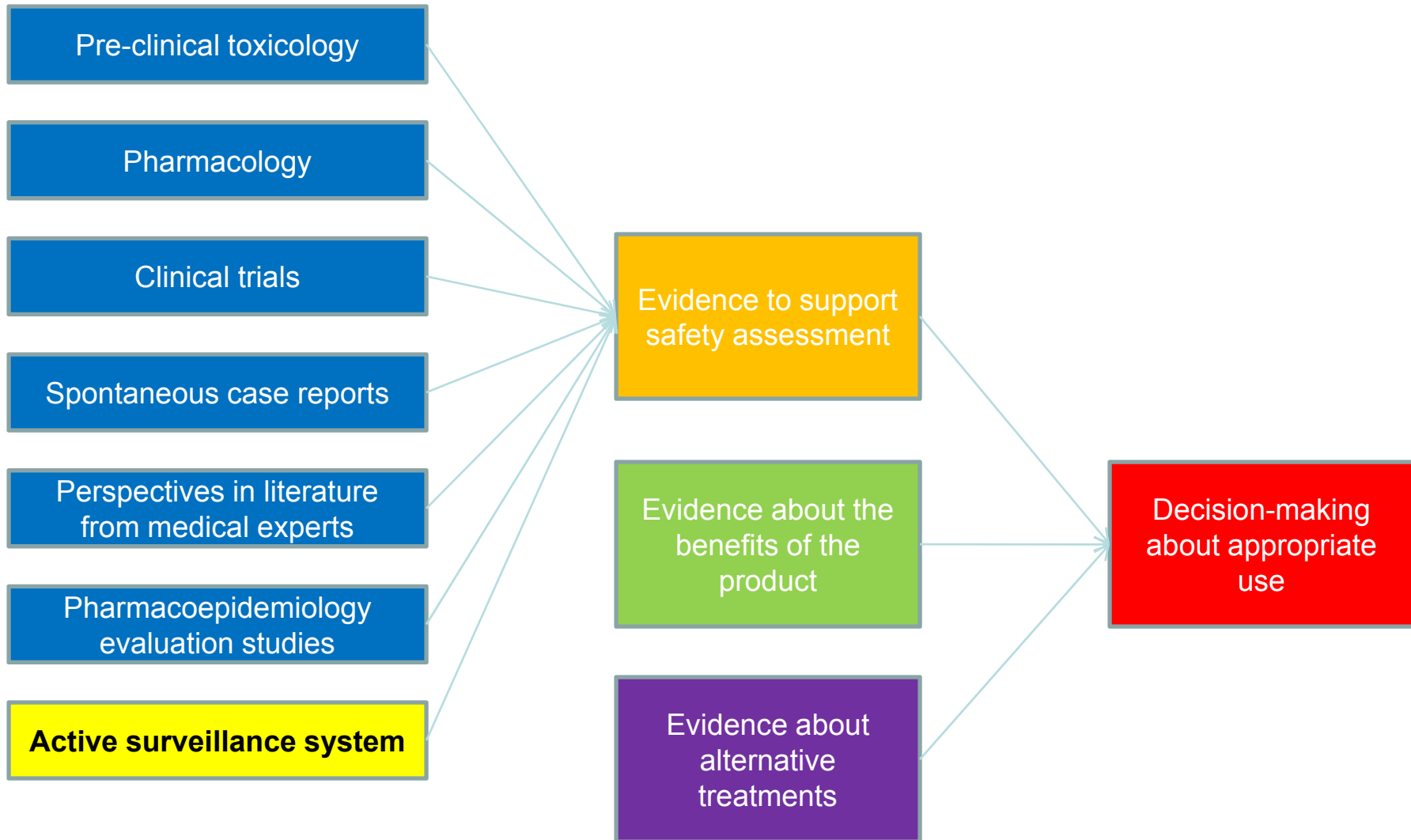
Temporality



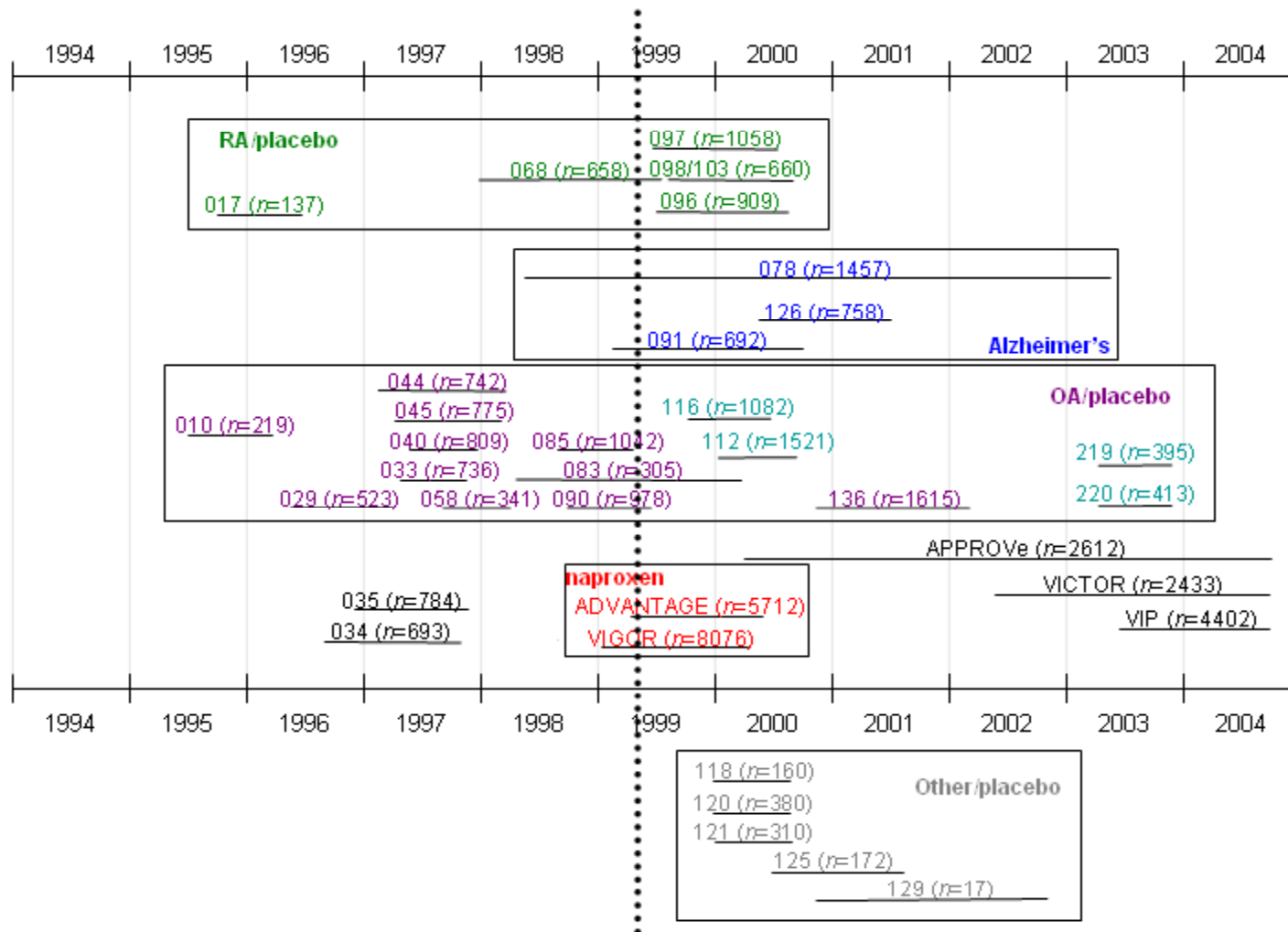
Harnessing Hill

- **Previously $p(\text{true} \mid \text{RR}, \text{SE})$**
 - Logistic regression with 2 predictors
- **Using Hill: $p(\text{true} \mid \text{RR}, \text{SE}, \text{temporality}, \text{coherence}, \text{consistency}, \text{etc.})$**
 - Logistic regression with many predictors
- **Thus we have a framework to formally integrate diverse evidence into the causal judgment**

Active surveillance: One additional piece of evidence to inform medical decision-making



A Case Study: Trials for Vioxx - Timeline

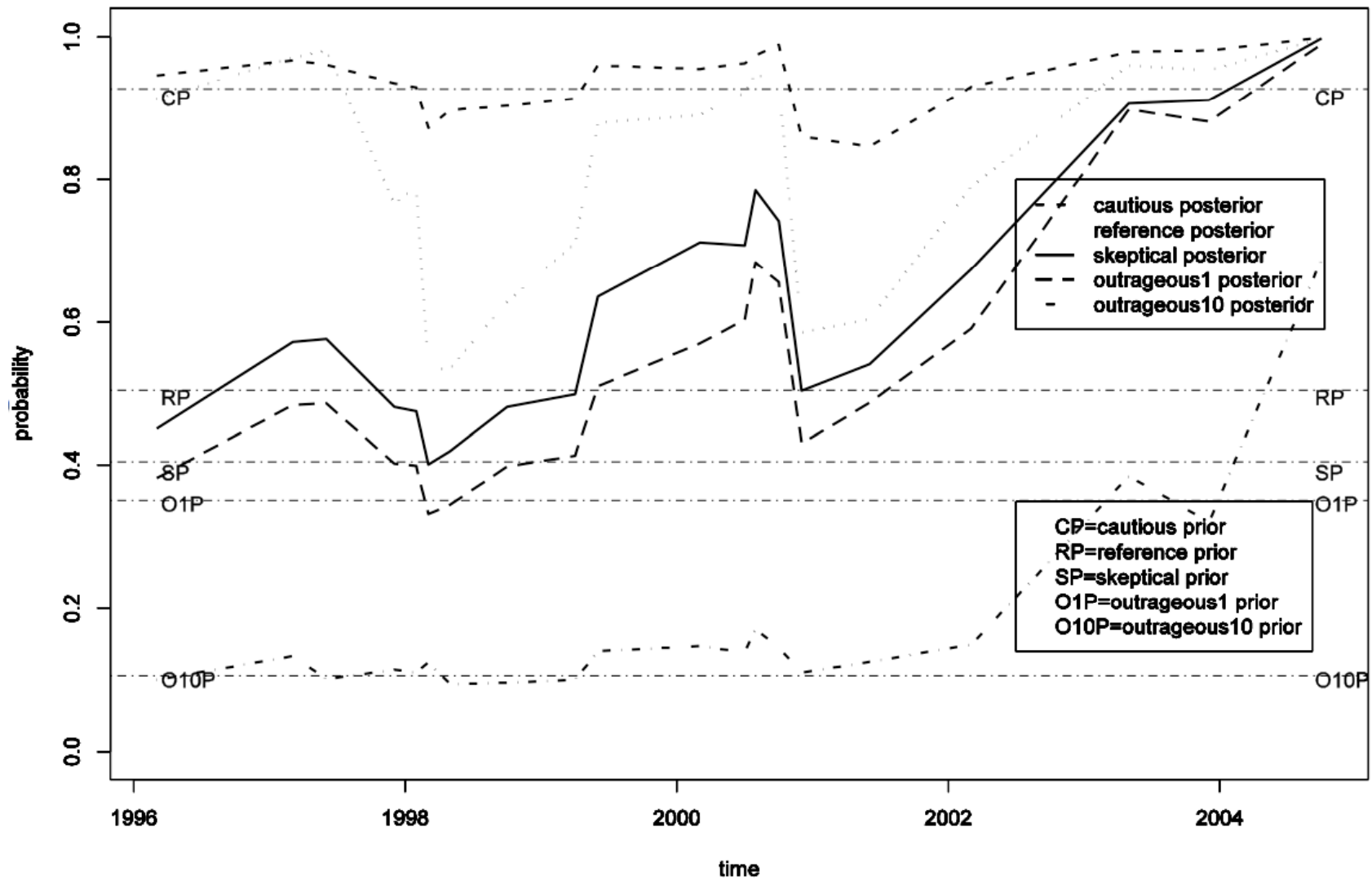


Case Study: CVT Events

			V		Placebo	
Study Block	Study No.	LPO	Events	PYR	Events	PYR
RA	017	5/21/1997	1	8	0	7
	068	9/10/1998	1	49	0	24
	096	7/21/2000	4	97	0	58
	097	6/6/2000	0	137	0	62
	098	7/6/2000	0	11	1	12
	103	7/6/2000	0	44	0	45
	All RA			6	345	1
OA	010	2/8/1996	2	16	0	7
	029	2/5/1997	3	46	0	16
	033	11/18/1997	1	66	1	9
	040	1/1/1998	2	72	0	11
	044	2/18/1998	3	154	0	52
	045	2/18/1998	2	157	3	61
	058	4/1/1998	1	21	0	6
	083	2/9/2000	0	21	0	21
	085	3/3/1999	1	61	0	28
	090	5/17/1999	5	56	0	27
	112	9/8/2000	0	104	0	15
	116	6/22/2000	1	54	0	15
	136	2/5/2002	1	95	1	201
	219	11/28/2003	0	18	1	8
	220	11/24/2003	0	18	0	8
All OA			22	959	6	485
OA/RA	All OA/RA		28	1304	7	692
ALZ	078	4/23/2003	75	1623	52	1762
	091	11/30/2000	13	369	14	381
	126	5/30/2001	11	193	7	197
All ALZ			99	2185	73	2341
APP	122	9/8/2004	75	5700	49	5828

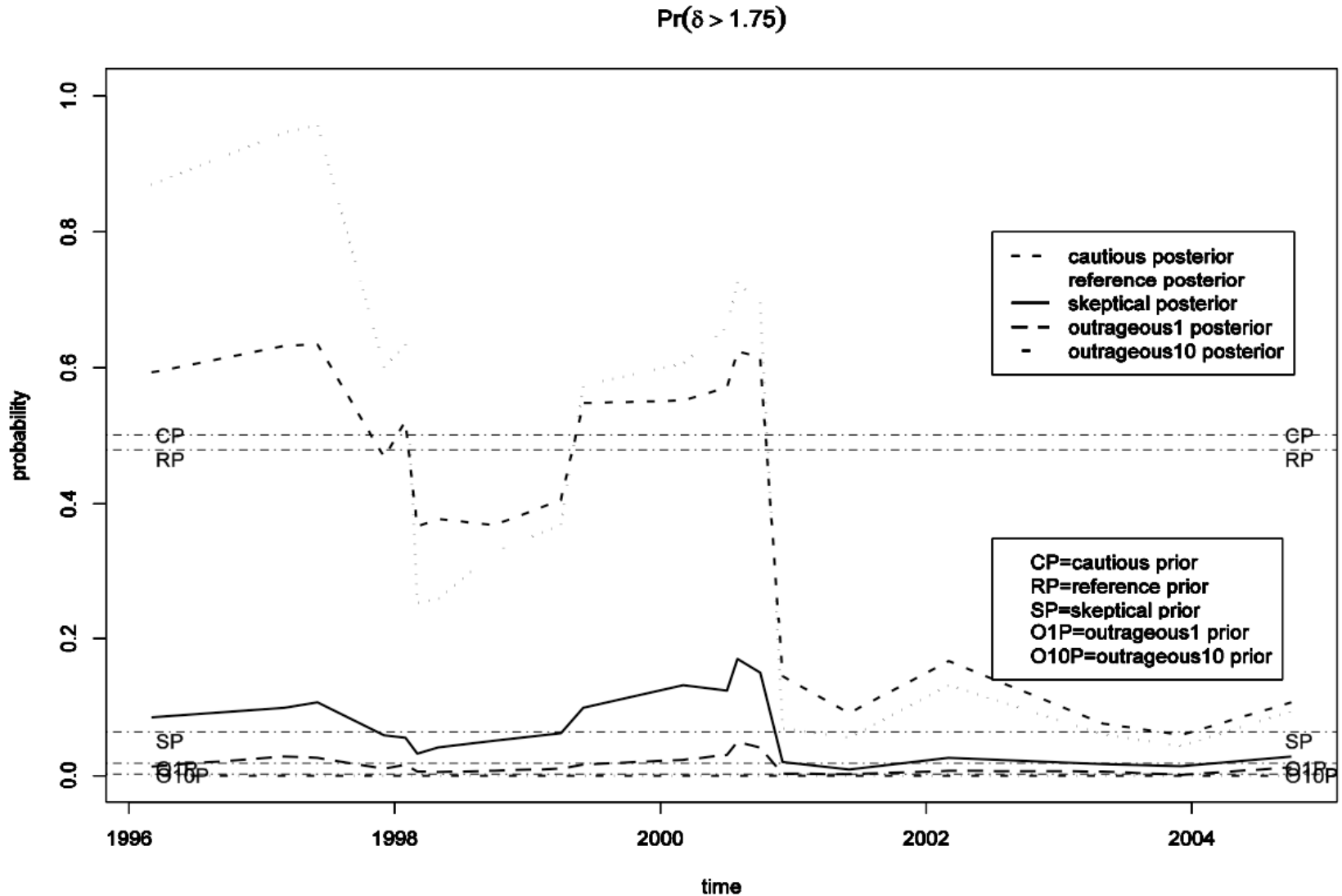
Posterior Probability (1) – block effect model

$$\Pr(\delta > 1.1)$$



Jerry Cheng

Posterior Probability (2) – block effect model



Sequential Uber-Analysis

- Sequential learning
- Silos: RCTs, observational data, spontaneous reports, pharmacology
- Little knowledge transfer
- Bayesian approach provides a snapshot at any point in time that reflects *all* available knowledge and data

Think about this...

	<i>Hospital</i>											
	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

Some Theory (Bayes works!)

Consider observations from a normal distribution:

$$y_1, \dots, y_n \sim N(\mu, 1)$$

Could use the sample mean to estimate μ :

$$\bar{y} = \frac{y_1 + \dots + y_n}{n}$$

"unbiased"



Suppose you want to estimate μ with $\delta(y)$ to minimize the expected squared loss:

$$E_{y|\mu}[(\delta(y) - \mu)^2]$$

Bayesian has to specify a prior distribution for μ , e.g.:

$$N(\theta, 10)$$

Bayesian would typically set $\delta(y)$ to be the posterior mean

$$c\theta + (1 - c)\bar{y}$$

Can easily prove that, e.g., for $n=30$, the Bayesian posterior mean has smaller expected squared loss so long as:

$$|\theta - \mu| < 134.3$$

Back to the hospital data

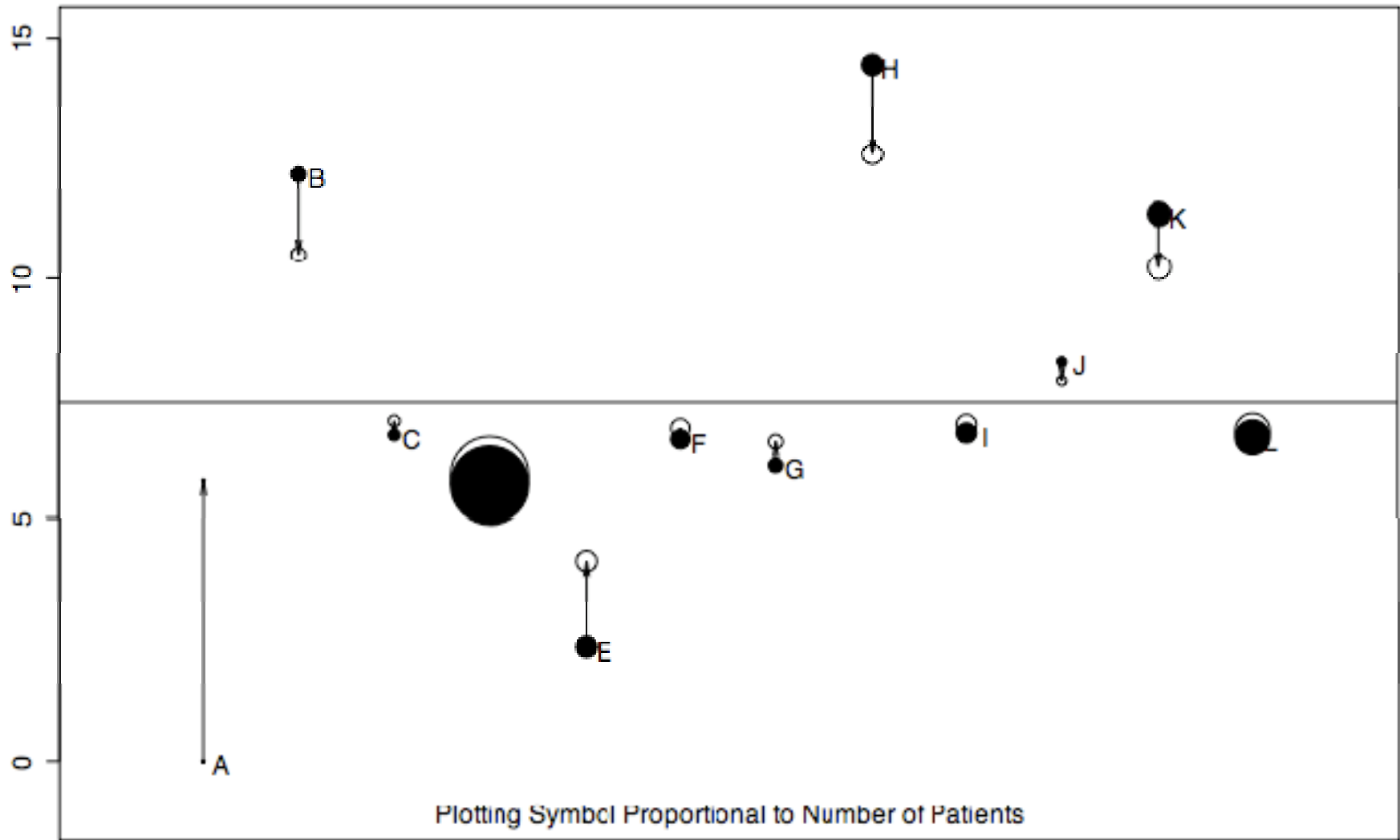
	<i>Hospital</i>											
	A	B	C	D	E	F	G	H	I	J	K	L
No. of ops. n	27	148	119	810	211	196	148	215	207	97	256	360
No. of deaths r	0	18	8	46	8	13	9	31	14	8	29	24

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

Bayesian Shrinkage for the Hospital Example

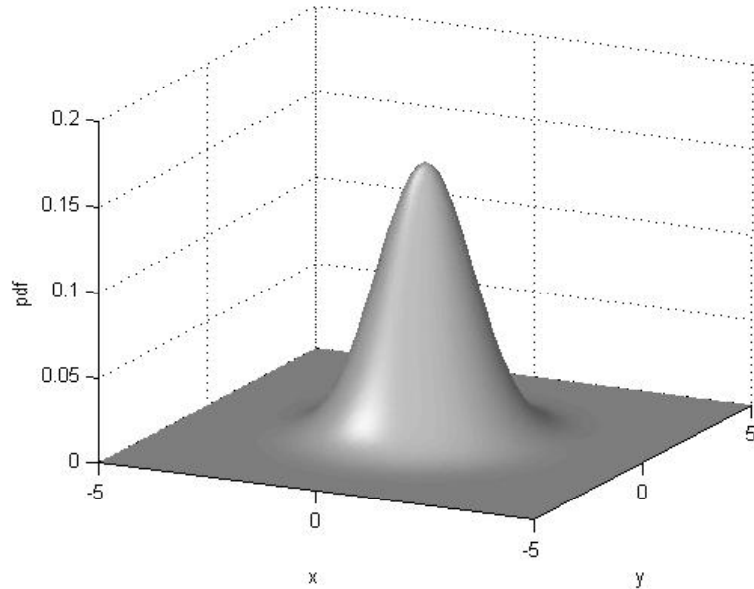


Logistic Regression

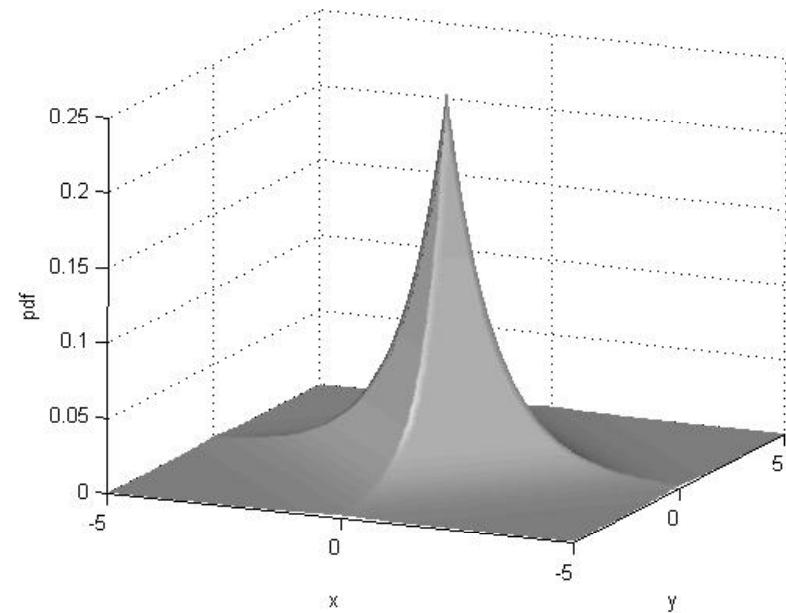
Linear model for log odds of category membership:

$$\log \frac{p(y=1|x_i)}{p(y=-1|x_i)} = \sum b_j x_{ij} = \mathbf{b}x_i$$

Bayesian Perspective



$$\beta_j \sim N(0, \tau^2)$$



$$\beta_j \sim N(0, \tau_j^2)$$

$$\tau_j^2 \sim \exp(\gamma)$$

Family of Priors

- **Two thresholds for rate ratio of an adverse event**
 - δ_U : the drug should not be on market if ratio $>\delta_U$
 - set $\delta_U = 1.75$ expert opinion,
 - δ_L : the drug should be on market if ratio $<\delta_L$
 - set $\delta_L = 1.1$ weigh= the risk against the benefit.
- ✓ **Skeptical Prior**: Gaussian with $\mu=0$ and $P(\delta > \delta_U) = 5\%$
 - Strong prior belief that the drug is not dangerous
- ✓ **Cautious Prior**: Gaussian with $\mu=\delta_U$ and $P(\delta < \delta_L) = 5\%$
 - Strong prior belief that the drug is not safe
- ✓ **Reference Prior**: Gaussian with $\mu=1$, $\sigma=10^3$
- ✓ **Outrageous10 Prior**: Gaussian with $\mu=1$ and $P(\delta > \delta_L) = 10\%$
- ✓ **Outrageous1 Prior**: Gaussian with $\mu=1$ and $P(\delta > \delta_U) = 1\%$