COLUMBIA UNIVERSITY
IN THE CITY OF NEW YORK

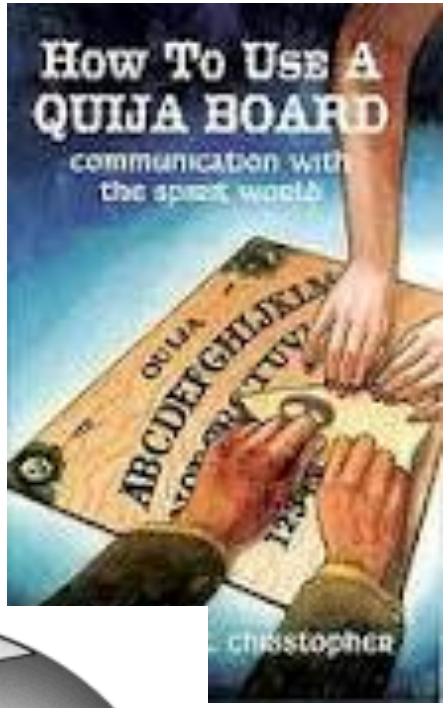
Big-Data-Driven Medicine

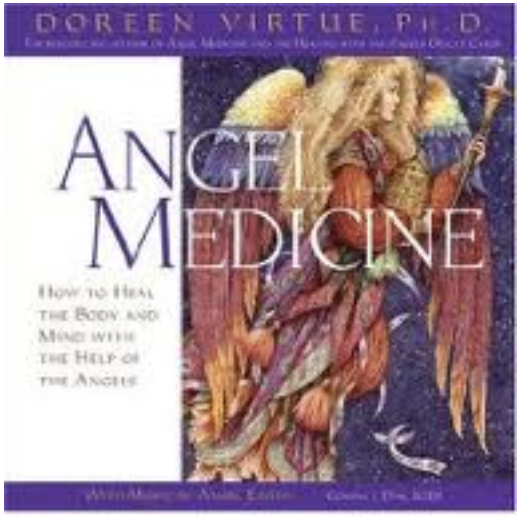
Can big data save medicine?

David Madigan
Department of Statistics
Columbia University

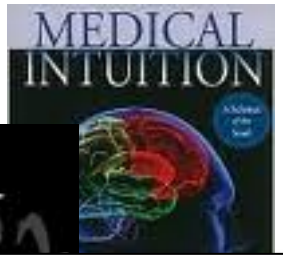
"Evidence-Based Medicine" as against ???


Oh, OUIJA®, can we take the car?
Shall I become a model, or fashion designer?
What college will accept me?
Who's Debbie's date for the Prom?
Are flying saucers for real?


How To Use A
QUIJA BOARD
communication with
the spirit world
christopher


DOREEN VIRTUE, PH.D.
ANGEL
MEDICINE
How to Heal the Body and Mind with the Help of the Angels


IT IS
CERTAIN



MEDICAL INTUITION

clinical judgment,

the application of information based on actual observation of a patient combined with subjective and objective data that lead to a conclusion.

Mosby's Medical Dictionary, 8th edition. © 2009, Elsevier.

Naomi Elliott

Clinical Judgment in Practice Matters:

How advanced practitioners in nursing solve clinical judgment problems in community care contexts

LAMBERT Academic Publishing

How does the combination of Clinical Judgment and Evidence-Based Medicine work in practice?

Coronary Heart Disease (CHD) Score Sheet* FOR MEN

About the CHD Score Sheet

This CHD score sheet can be used to estimate a man's risk of developing CHD over a 10-year period based on age, total cholesterol (TC), HDL cholesterol (HDL-C), blood pressure (BP), and cigarette smoking.

Risk estimates have been derived from the experience of NHLBI's Framingham Heart Study, a predominantly Caucasian population in Massachusetts, USA. The risk algorithm may not fit other populations quite as well.

Step 1

AGE			
Years	Points	Years	Points
20–34	–9	55–59	8
35–39	–4	60–64	10
40–44	0	65–69	11
45–49	3	70–74	12
50–54	6	75–79	13

Step 2

TOTAL CHOLESTEROL					
	Points				
TC (mg/dL)	Age 20–39 y	Age 40–49 y	Age 50–59 y	Age 60–69 y	Age 70–79 y
<160	0	0	0	0	0
160–199	4	3	2	1	0
200–239	7	5	3	1	0
240–279	9	6	4	2	1
≥280	11	8	5	3	1

Step 3

SMOKING					
	Points				
	Age 20–39 y	Age 40–49 y	Age 50–59 y	Age 60–69 y	Age 70–79 y
Nonsmoker	0	0	0	0	0
Smoker	8	5	3	1	1

Step 4

HDL CHOLESTEROL	
HDL-C (mg/dL)	Points
≥60	–1
50–59	0
40–49	1
<40	2

Step 5

BLOOD PRESSURE		
Systolic BP (mm Hg)	Points If Untreated	Points If Treated
<120	0	0
120–129	0	1
130–139	1	2
140–159	1	2
≥160	2	3

Step 6

ADDING UP THE POINTS (Sum from Steps 1–6)	
Age	
TC	
Smoker	
HDL-C	
BP	
Point Total	

CHD RISK

DETERMINE CHD RISK FROM POINT TOTAL	
Point Total	10-year CHD Risk
<0	<1%
0	1%
1	1%
2	1%
3	1%
4	1%
5	2%
6	2%
7	3%
8	4%
9	5%
10	6%
11	8%
12	10%
13	12%
14	16%
15	20%
16	25%
≥17	≥30%

Your chance of developing CHD (angina or heart attack) over the next 10 years is: 2 %

*NCEP Expert Panel. Third report of the Expert Panel on Detection, Evaluation, and Treatment of High Blood Cholesterol in Adults (Adult Treatment Panel III) Executive Summary. Available at http://www.nhlbi.nih.gov/guidelines/cholesterol/atp_iii.htm. Accessed May 31, 2001.

Should John have an angiogram?

Clinical judgment?

Who are we kidding?

48 years old

HDL = 59

CRP normal

triglycerides = 106

LDL = 70

father died of heart dis

university professor

calcium score in 2003 = 19

calcium score in 2010 = 70

BMI = 21.6

stress test normal in 2007

no diabetes

mother died of cancer (83)

EKG unusual in 2009

lipitor

aspirin

exercise

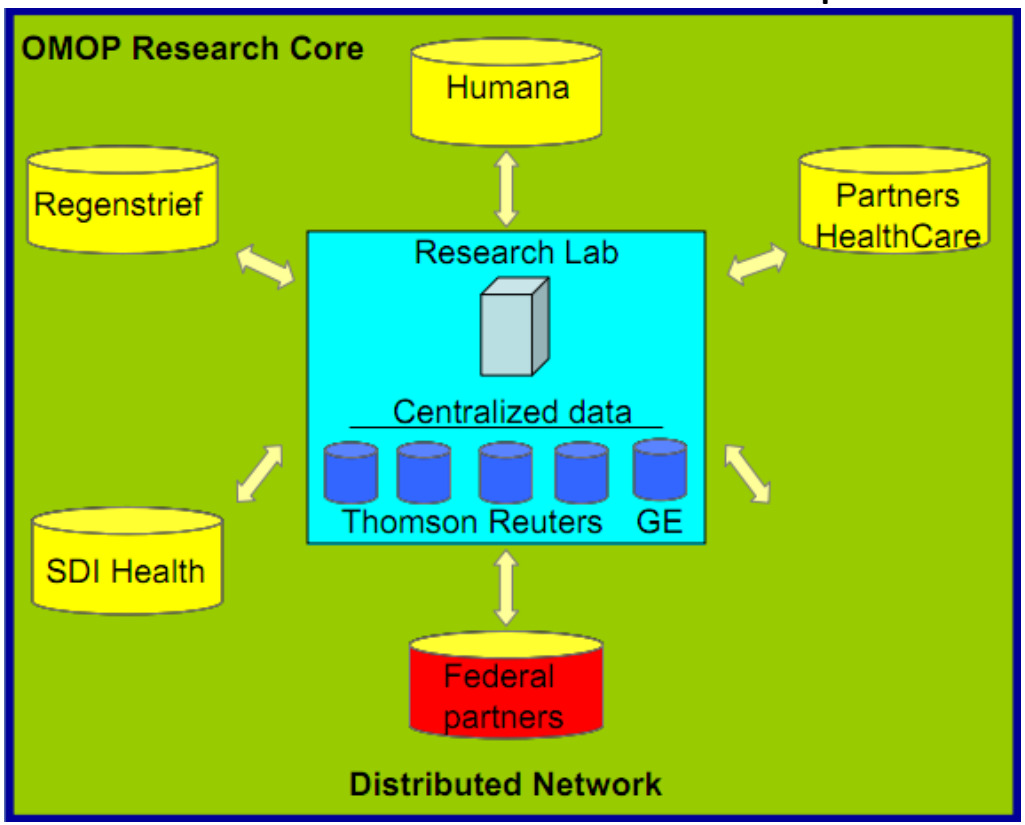
arrhythmia in 2008

genotyping

normal heart ultrasound (2008)

Data-Driven Medicine

Observational Medical Outcomes Partnership

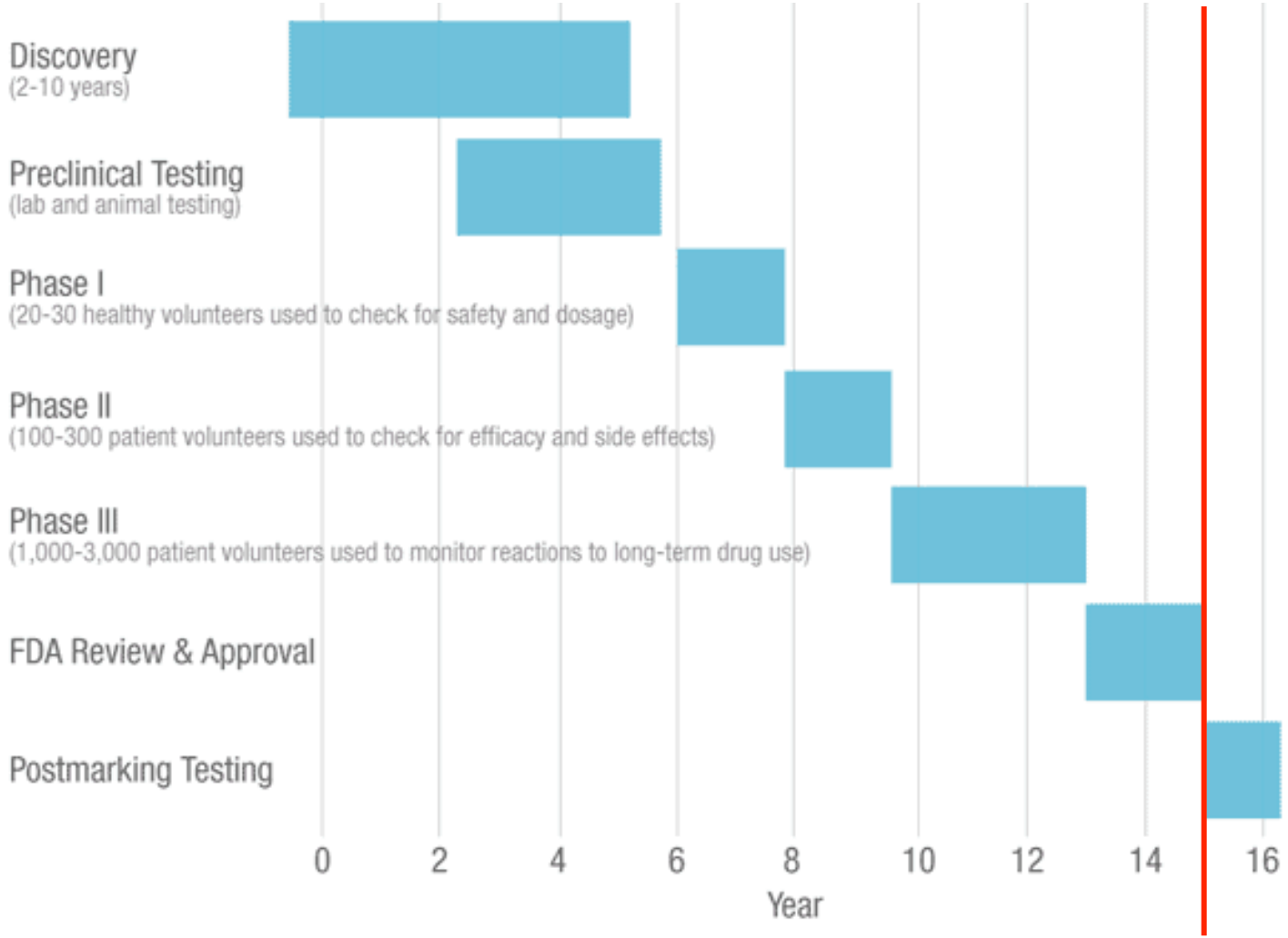


- **Multiple years of medical records for 200+ million people**
- **Largest collection of medical records in the world**
- **32,430 patients just like John**

Many Challenges

- Statistical/Epidemiological
- Computational
- Drug Safety

Safety in Lifecycle of a Drug/Biologic product



Drug Safety

Pre-Approval:

- **High quality data (but small)**
- **No "data mining"**

Post-Approval:

- **Low quality data (but lots of it)**
- **Extensive use of "data mining"**

U.S. Department of Health and Human Services

MEDWATCH

The FDA Safety Information and Adverse Event Reporting Program

For VOLUNTARY reporting of adverse events, product problems and product use errors

Page ____ of ____

Form Approved: OMB No. 0910-0291, Expires: 10/31/08
See OMB statement on reverse.

FDA USE ONLY

Triage unit sequence #

PLEASE TYPE OR USE BLACK INK

A. PATIENT INFORMATION

1. **Patient Identifier** In confidence
2. **Age at Time of Event, or Date of Birth:**
3. **Sex** ☐ Female ☐ Male
4. **Weight** _____ lb or _____ kg

B. ADVERSE EVENT, PRODUCT PROBLEM OR ERROR

Check all that apply:

1. ☐ **Adverse Event** ☐ **Product Problem** *(e.g., defects/malfunctions)*
   ☐ **Product Use Error** ☐ **Problem with Different Manufacturer of Same Medicine**
2. **Outcomes Attributed to Adverse Event** *(Check all that apply)*
   ☐ Death: _____ *(mm/dd/yyyy)*
   ☐ Life-threatening
   ☐ Hospitalization - initial or prolonged
   ☐ Required Intervention to Prevent Permanent Impairment/Damage (Devices)
   ☐ Disability or Permanent Damage
   ☐ Congenital Anomaly/Birth Defect
   ☐ Other Serious (Important Medical Events)
3. **Date of Event** *(mm/dd/yyyy)*
4. **Date of this Report** *(mm/dd/yyyy)*
5. **Describe Event, Problem or Product Use Error**
6. **Relevant Tests/Laboratory Data, Including Dates**
7. **Other Relevant History, Including Preexisting Medical Conditions** *(e.g., allergies, race, pregnancy, smoking and alcohol use, liver/kidney problems, etc.)*

C. PRODUCT AVAILABILITY

Product Available for Evaluation? *(Do not send product to FDA)*

☐ Yes ☐ No ☐ Returned to Manufacturer on: _____ *(mm/dd/yyyy)*

D. SUSPECT PRODUCT(S)

1. **Name, Strength, Manufacturer** *(from product label)*
   #1
   #2
2. **Dose or Amount** / **Frequency** / **Route**
   #1
   #2
3. **Dates of Use** *(If unknown, give duration)* from/to *(or best estimate)*
   #1
   #2
4. **Diagnosis or Reason for Use** *(Indication)*
   #1
   #2
5. **Event Abated After Use Stopped or Dose Reduced?**
   #1 ☐ Yes ☐ No ☐ Doesn't Apply
   #2 ☐ Yes ☐ No ☐ Doesn't Apply
6. **Lot #**
   #1
   #2
7. **Expiration Date**
   #1
   #2
8. **Event Reappeared After Reintroduction?**
   #1 ☐ Yes ☐ No ☐ Doesn't Apply
   #2 ☐ Yes ☐ No ☐ Doesn't Apply
9. **NDC # or Unique ID**

E. SUSPECT MEDICAL DEVICE

1. **Brand Name**
2. **Common Device Name**
3. **Manufacturer Name, City and State**
4. **Model #** / **Lot #** / **Catalog #** / **Expiration Date** *(mm/dd/yyyy)* / **Serial #** / **Other #**
5. **Operator of Device** ☐ Health Professional ☐ Lay User/Patient ☐ Other:
6. **If Implanted, Give Date** *(mm/dd/yyyy)*
7. **If Explanted, Give Date** *(mm/dd/yyyy)*
8. **Is this a Single-use Device that was Reprocessed and Reused on a Patient?** ☐ Yes ☐ No
9. **If Yes to Item No. 8, Enter Name and Address of Reprocessor**

F. OTHER (CONCOMITANT) MEDICAL PRODUCTS

Product names and therapy dates *(exclude treatment of event)*

G. REPORTER *(See confidentiality section on back)*

1. **Name and Address**

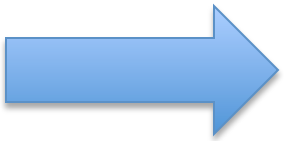
Phone # / **E-mail**

2. **Health Professional?** ☐ Yes ☐ No
3. **Occupation**
4. **Also Reported to:** ☐ Manufacturer ☐ User Facility ☐ Distributor/Importer
5. **If you do NOT want your identity disclosed to the manufacturer, place an "X" in this box:** ☐

FORM FDA 3500 (10/05) **Submission of a report does not constitute an admission that medical personnel or the product caused or contributed to the event.**

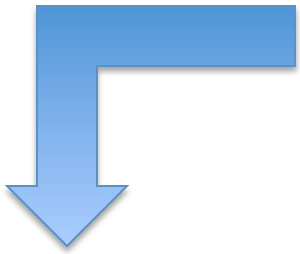
Problems with Spontaneous Reports

- Under-reporting
- Duplicate reports
- No temporal information
- No denominator



Vioxx
Baycol
etc.

IOM
FDAAA



Sentinel
EU-ADR
OMOP

Active surveillance in large, longitudinal databases

(3) Active postmarket risk identification.--

``(A) Definition.--In this paragraph, the term `data' refers to information with respect to a drug approved under this section or under section 351 of the Public Health Service Act, including claims data, patient survey data, standardized analytic files that allow for the pooling and analysis of data from disparate data environments, and any other data deemed appropriate by the Secretary.

``(B) Development of <<NOTE: Deadline.>> postmarket risk identification and analysis methods.--The Secretary shall, not later than 2 years after the date of the enactment of the Food and Drug Administration Amendments Act of 2007, in collaboration with public, academic, and private entities--

``(i) develop methods to obtain access to disparate data sources including the data sources specified in subparagraph (C);

``(ii) develop validated methods for the establishment of a postmarket risk identification and analysis system to link and analyze safety data from multiple sources, with the goals of including, in aggregate--

``(I) at least 25,000,000 patients by July 1, 2010; and

``(II) at least 100,000,000 patients by July 1, 2012; and

BMJ

RESEARCH

Oral bisphosphonates and risk of cancer of oesophagus, stomach, and colorectum: case-control analysis within a UK primary care cohort

Jane Green, clinical epidemiologist,[1] Gabriela Czanner, statistician,[1] Gillian Reeves, statistical epidemiologist,[1] Joanna Watson, epidemiologist,[1] Lesley Wise, manager, Pharmacoepidemiology Research and Intelligence Unit,[2] Valerie Beral, professor of cancer epidemiology[1]

BMJ 2010; 341:c4444

Conclusions The risk of oesophageal cancer increased with 10 or more prescriptions for oral bisphosphonates and with prescriptions over about a five year period.

BMJ study design choices

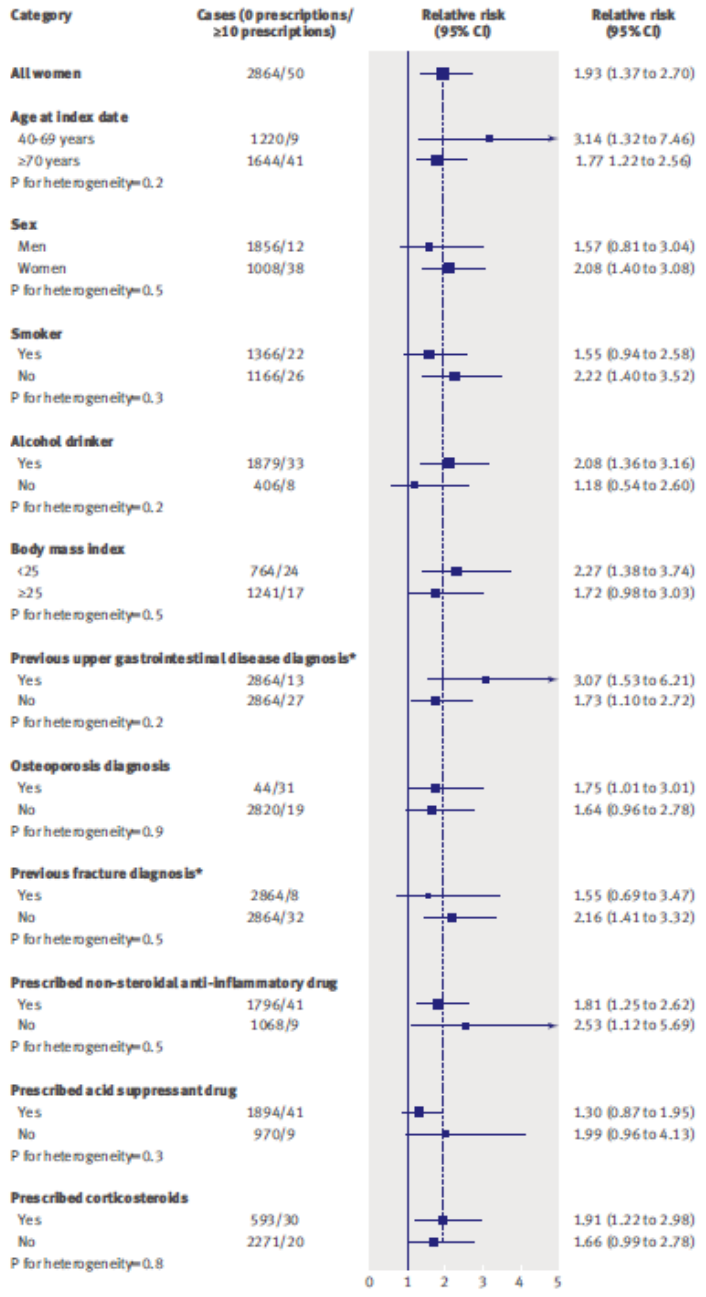
- Data source: General Practice Research Database
- Study design: Nested case-control
- Inclusion criteria: Age > 40
- Case: cancer diagnosis between 1995-2005 with 12-months of follow-up pre-diagnosis
- 5 controls per case
- Matched on age at index date, sex, practice, observation period prior to index
- Exposure definition: >=1 prescription during observation period
- "RR" estimated with conditional logistic regression
- Covariates: smoking, alcohol, BMI before *outcome* index date
- Sensitivity analyses:
  - exposure = 2+ prescriptions
  - covariates not missing
  - time-at-risk = >1 yr post-exposure
- Subgroup analyses:
  - Short vs. long exposure duration
  - Age, Sex, smoking, alcohol, BMI
  - Osteoporosis or osteopenia
  - Fracture pre-exposure
  - Prior diagnosis of Upper GI dx pre-exposure
  - NSAID, corticosteroid, H2blocker, PPI utilization pre-exposure

BMJ Results

Table 2 | Relative risks (RRs) and 95% confidence intervals (CIs) fo bisphosphonates

Oral bisphosphonates	Oesophagus		
	Prescriptions*	Cases/ controls	RR† (95% CI)
Not prescribed	NA	2864/14 376	1.00
Prescribed	13.6/2.4	90/345	1.30 (1.02 to 1.66)
No of prescriptions:			
1-9	3.6/1.0	40/214	0.93 (0.66 to 1.31)
≥10	21.6/3.5	50/131	1.93 (1.37 to 2.70)
Estimated duration of use‡:			
≤1 year	4.9/0.3	31/155	0.98 (0.66 to 1.46)
1-3 years	13.0/2.0	26/114	1.12 (0.73 to 1.73)
≥3 years	22.2/4.6	33/76	2.24 (1.47 to 3.43)

NA=not applicable.
*Prescriptions of bisphosphonates in cases; reported as mean number/mean ye
†All relative risks adjusted for smoking status, alcohol intake, and body mass i
‡Time between first and last prescription.



Relative risks of incident oesophageal cancer in people with ≥10 prescriptions for oral bisphosphonates, compared with those with no prescriptions, by various factors. Relative risks adjusted for smoking status, alcohol intake, and body mass index, as appropriate. *Diagnosis before prescription of bisphosphonates: analyses restricted to those with ≥12 months' observation before first bisphosphonate prescription

BMJ 2010; 341:c4444

ORIGINAL CONTRIBUTION

JAMA®

Exposure to Oral Bisphosphonates and Risk of Esophageal Cancer

Chris R. Cardwell, PhD
Christian C. Abnet, PhD
Marie M. Cantwell, PhD
Liam J. Murray, MD

Context Use of oral bisphosphonates has increased dramatically in the United States and elsewhere. Esophagitis is a known adverse effect of bisphosphonate use, and recent reports suggest a link between bisphosphonate use and esophageal cancer, but this has not been robustly investigated.

Objective To investigate the association between bisphosphonate use and esoph-

JAMA 2010; 304(6): 657-663

Conclusion the use of oral bisphosphonates was not significantly associated with incident esophageal or gastric cancer.

JAMA study design choices

- Data source: General Practice Research Database
- Study design: Cohort
- Inclusion criteria: Age > 40
- Exclusion criteria: Cancer diagnosis in 3 years before index date
- Exposed cohort: Patients with >=1 prescription between 1996-2006
- "Unexposed" cohort: 1-to-1 match with exposed cohort
- Matched on year of birth, sex, practice
- "HR" estimated with Cox proportional hazards model
- Time-at-risk: >6mo from index date
- Covariates:
  - Smoking, alcohol, BMI before *exposure* index date
  - Hormone therapy, NSAIDs, H2blockers, PPIs
- Sensitivity analyses:
  - Excluding people that were in both exposed and unexposed cohorts
  - Exclude patients with missing confounders (not reported)
- Subgroup analyses:
  - Low vs. medium vs. high use, based on defined daily dose
  - Alendronate vs. nitrogen-containing bisphosphonates vs. non-nitrogen-contraining bisphosphonates

1995-2005 in BMJ

Match exposure vs. outcome status; not 5-to-1

Not observation length

Time-at-risk is 'between' two definitions used in BMJ: All time post-exposure and >1yr after index

Different index date

BMJ didn't stratify by hormone therapy

JAMA Results

Table 3. Esophageal (Only) Cancer Incidence in the Bisphosphonate and Matched Control Cohorts

	Bisphosphonate		Control		Risk			
					Unadjusted		Adjusted[a]	
Bisphosphonate Category	Cases	Person-Years	Cases	Person-Years	HR (95% CI)	*P* Value	HR (95% CI)	*P* Value
Any bisphosphonate Prescribed	79	165 400	72	163 480	1.08 (0.79-1.49)	.63	1.07 (0.77-1.49)	.67
Incidence after cumulative prescriptions greater than (in DDDs)[b]								
183	51	104 676	49	104 104	1.04 (0.70-1.53)	.86	1.05 (0.70-1.57)	.82
365	31	73 364	35	73 170	0.88 (0.55-1.43)	.62	0.92 (0.56-1.51)	.74
730	22	40 326	22	40 492	1.00 (0.56-1.81)	.99	0.98 (0.53-1.81)	.95
1095	15	22 813	14	22 891	1.08 (0.52-2.23)	.84	1.01 (0.48-2.12)	.99
Total bisphosphonate intake during follow-up (in DDDs/d)[c]								
Low (0-<0.24)	35	62 922	27	63 648	1.31 (0.80-2.17)	.29	1.24 (0.74-2.09)	.41
Medium (≥0.24-<0.89)	24	58 162	23	55 334	0.98 (0.55-1.74)	.94	1.03 (0.57-1.86)	.92
High (≥0.89)	20	44 316	22	44 497	0.91 (0.50-1.67)	.78	0.90 (0.48-1.68)	.74
Nitrogen-containing bisphosphonates First prescribed	44	106 480	47	106 412	0.94 (0.62-1.41)	.75	0.96 (0.63-1.47)	.86
Incidence after cumulative prescriptions greater than (in DDDs)[b]								
365	30	70 251	34	69 935	0.88 (0.54-1.44)	.61	0.93 (0.56-1.54)	.78
730	22	39 022	22	39 187	1.01 (0.56-1.82)	.99	0.98 (0.53-1.80)	.95
Alendronate First prescribed	33	81 369	42	80 837	0.78 (0.50-1.23)	.29	0.77 (0.48-1.23)	.27
Incidence after cumulative prescriptions greater than (in DDDs)[b]								
365	22	52 308	31	51 741	0.70 (0.41-1.21)	.20	0.68 (0.39-1.19)	.18
730	19	28 898	21	28 904	0.91 (0.49-1.68)	.75	0.85 (0.45-1.61)	.62
Non–nitrogen-containing bisphosphonates First prescribed	35	58 920	25	57 068	1.35 (0.81-2.25)	.25	1.25 (0.73-2.12)	.37

Abbreviations: CI, confidence interval; DDD, defined daily dose; HR, hazard ratio.

[a] Adjusted for body mass index, alcohol, smoking, hormone therapy prescription (before index date), nonsteroidal anti-inflammatory drug prescription (before index date), Barrett esophagus diagnosis (before index date), gastroesophageal reflux disease diagnosis (before index date), H_2 receptor antagonist prescription (before index date), and proton pump inhibitor prescription (before index date).

[b] Person-years and cancer cases occurring after the date of specified prescriptions received for each bisphosphonate cohort member and their matched control. Daily divided dose equivalents: 183 DDDs are equivalent to a 6-month supply; 365 DDDs to a 1-year supply; 730 DDDs to a 2-year supply; and 1095 DDDs to a 3-year supply.

[c] In bisphosphonate cohort (see "Methods" for details of selection of cohorts).

JAMA 2010; 304(6): 657-663

Acne drug tied to a doubled risk of eye problems

Recommend Be the f

Continuing Medical Education
Perelm
School of Med
University of Penns

Curry spice 'lowers risk of heart attack after surgery'

- **Yellow pigment in turmeric known for having anti-inflammatory properties**

Bloomberg News

Bayer Birth-Control Pill May Boost Risks for Blood Clots

By Anna Edney on April 10,

Judge fines J&J $1.1B for antipsychotic drug risks

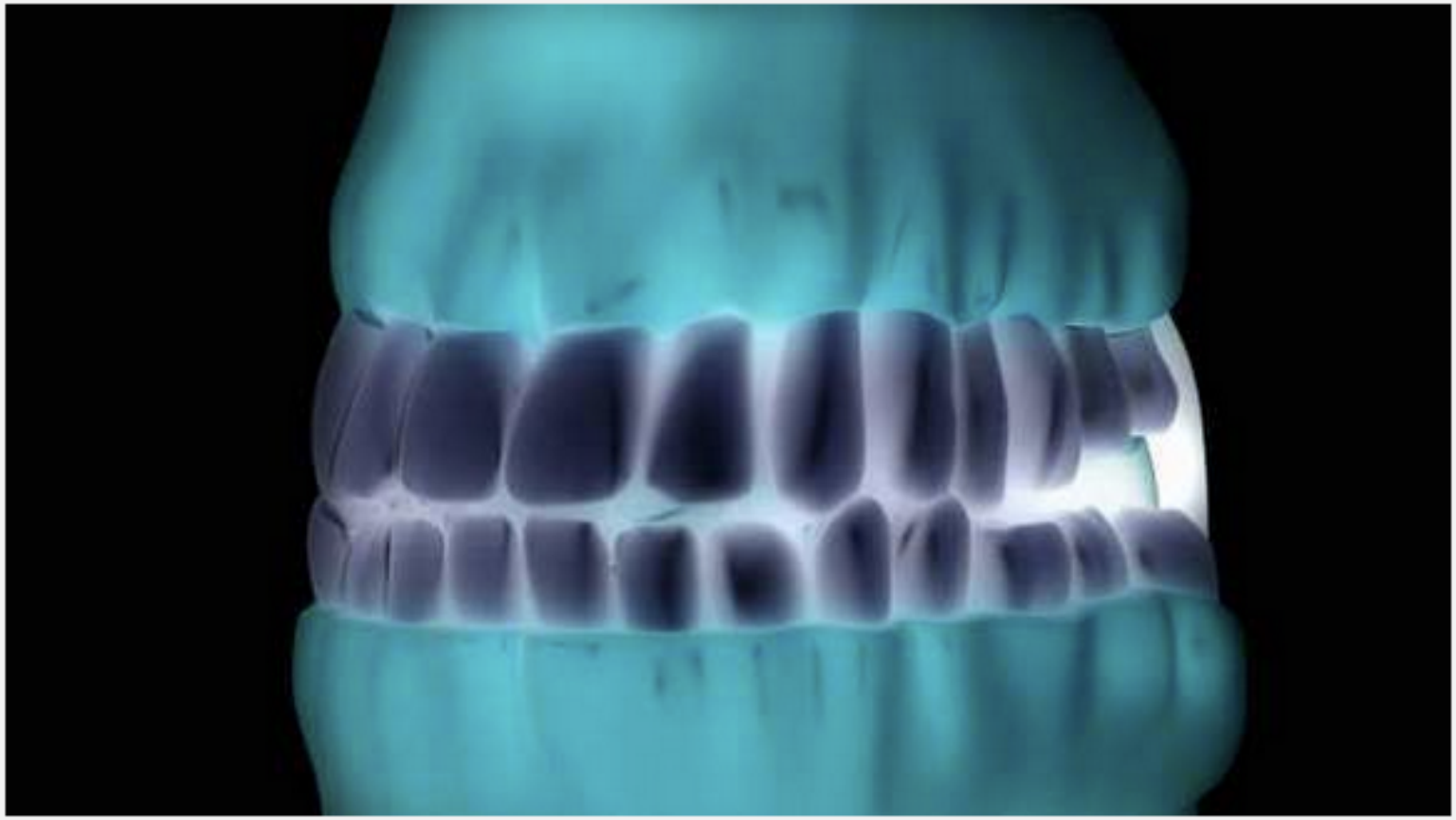
Verdict » Jury says company deceptive about Risperdal.

The Associated Press
First Published Apr 11 2012 04:00 pm • Last Updated Apr 11 2012 08:47 pm

Dental X-ray

By Charles Bankhead, Staff Writer, MedPage Today
Published: April 10, 2012

Gum disease and heart disease -- no link after all?



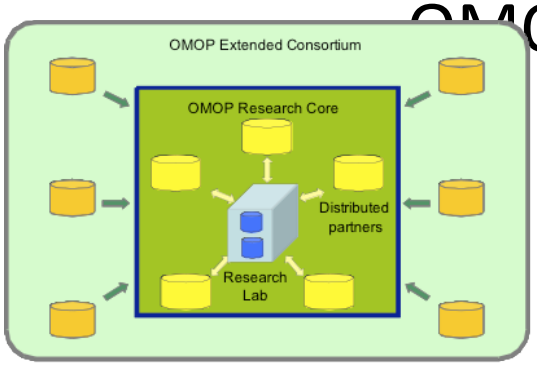
Gum disease, despite what we're told, may not raise the risk of heart disease or strokes. But brush your teeth anyway. (3D4Medical.com / April 18, 2012)

ALSO

By Rosie Mestel, Los Angeles Times / For the Booster Shots blog

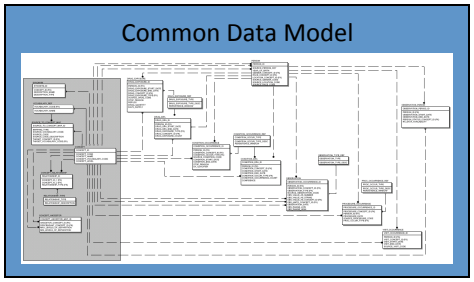
April 18, 2012 | 4:57 p.m.

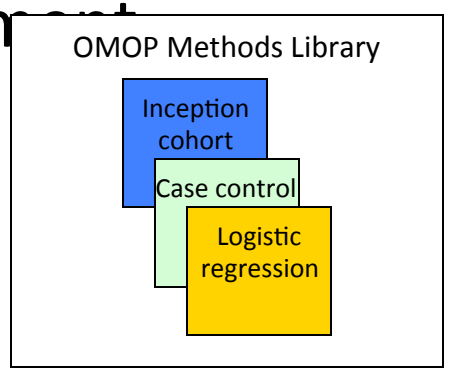
OMOP Research Experiment



- 10 data sources
- Claims and EHRs
- 200M+ lives

- Open-source
- Standards-based



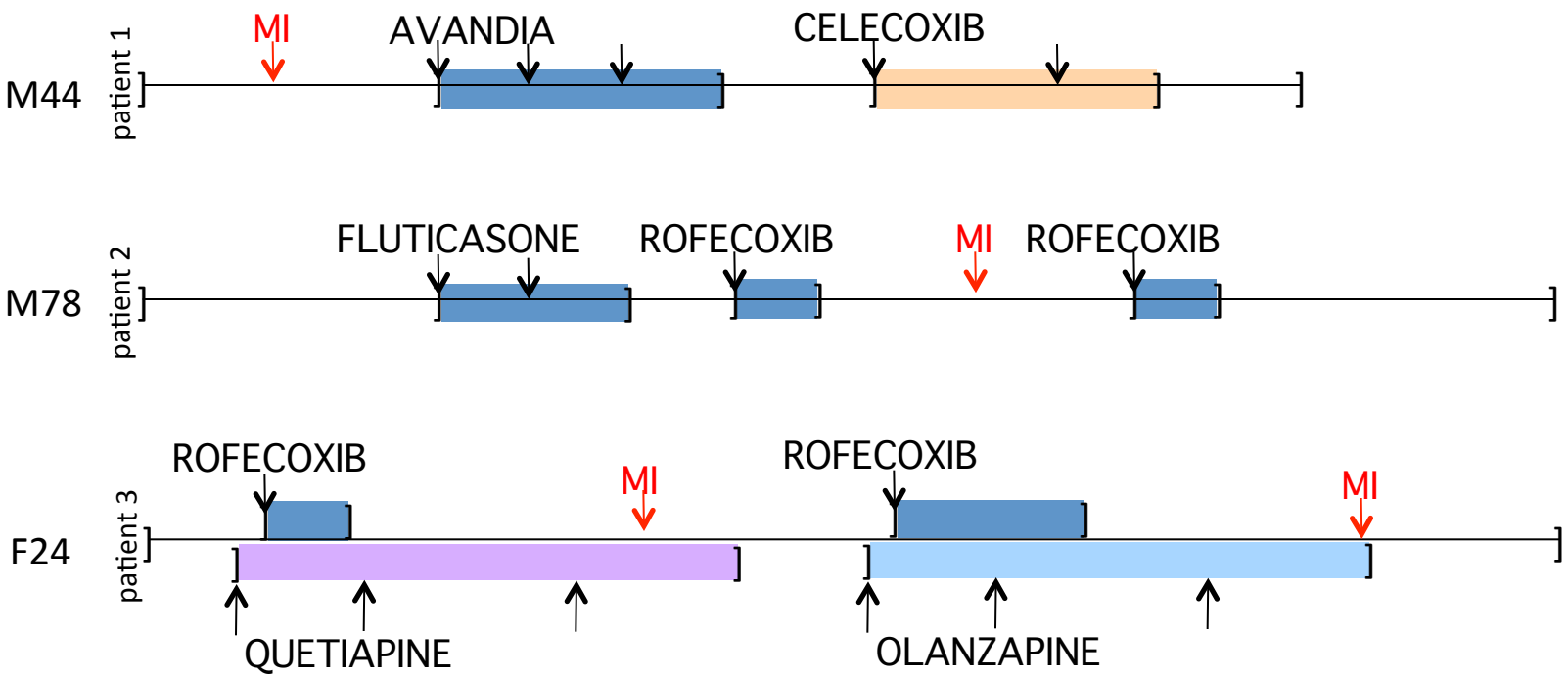


- 14 methods
- Epidemiology designs
- Statistical approaches adapted for longitudinal data

Drug

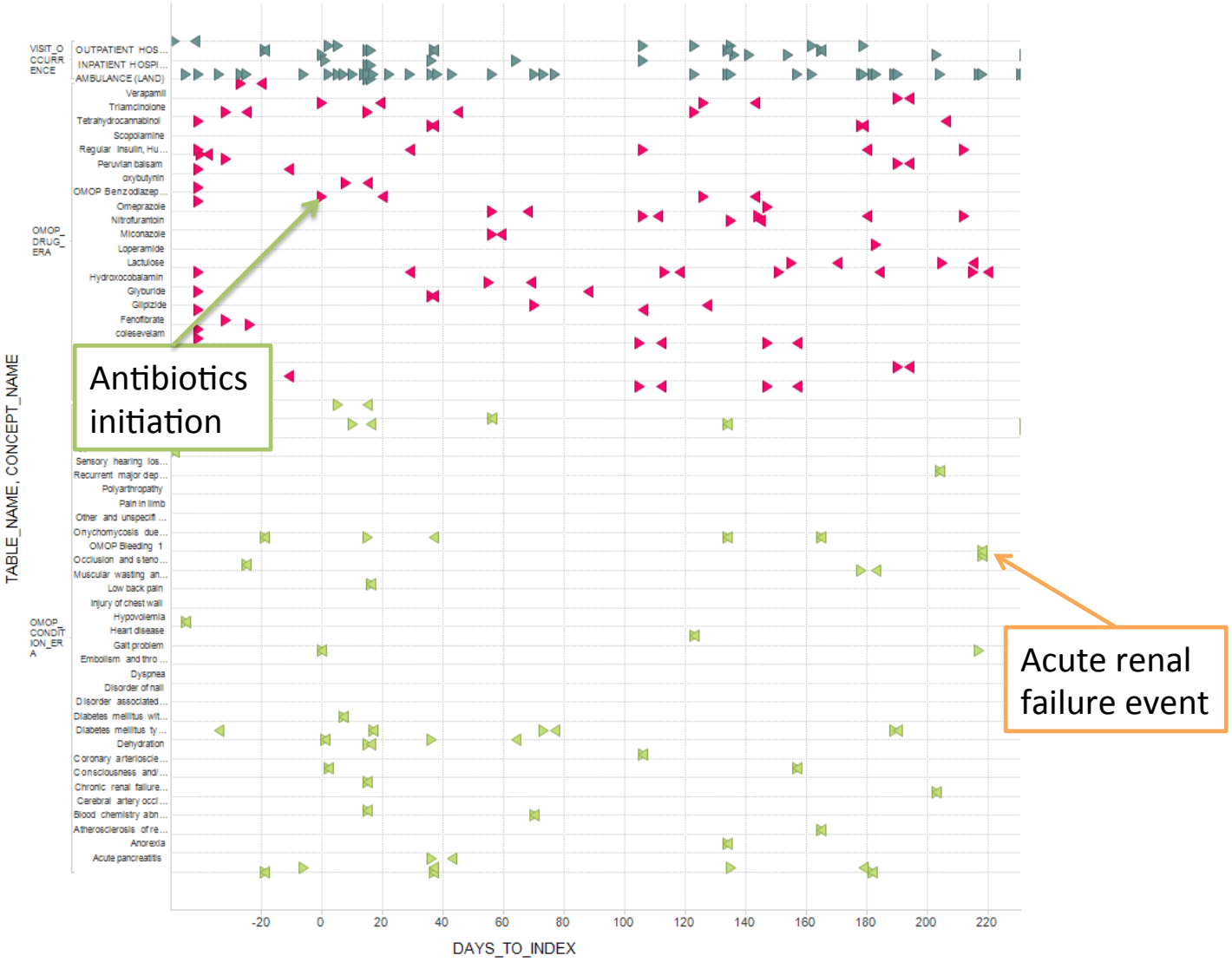
Outcome	**ACE Inhibitors**	**Amphotericin B**	**Antibiotics:** erythromycins, sulfonamides, tetracyclines	**Antiepileptics:** carbamazepine, phenytoin	**Benzodiazepines**	**Beta blockers**	**Bisphosphonates:** alendronate	**Tricyclic antidepressants**	**Typical antipsychotics**	**Warfarin**
Angioedema										
Aplastic Anemia										
Acute Liver Injury										
Bleeding										
Hip Fracture										
Hospitalization										
Myocardial Infarction										
Mortality after MI										
Renal Failure										
GI Ulcer Hospitalization										

What do the data look like?

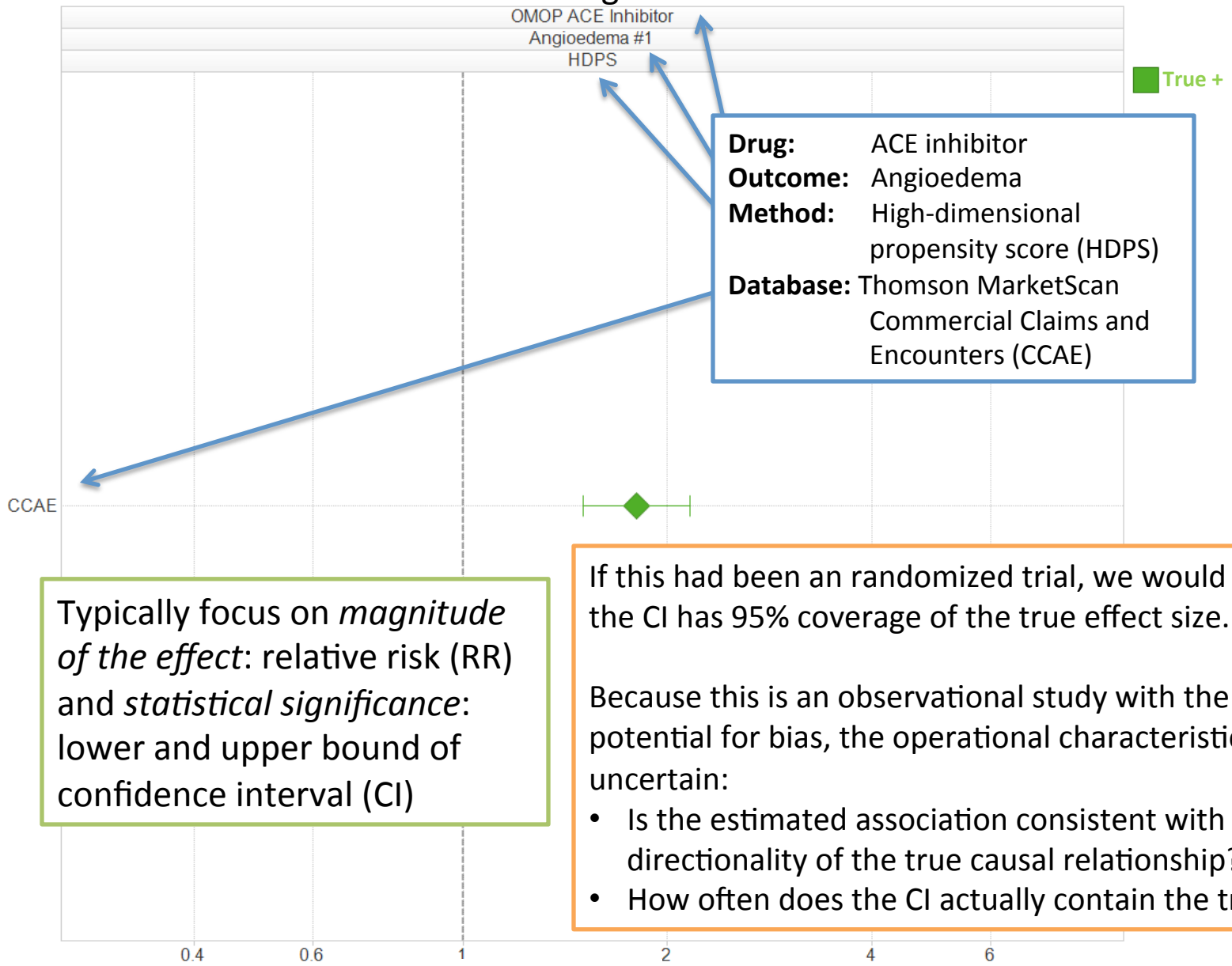


Computational considerations require few passes through the data

OBSERVATIONAL
MEDICAL
OUTCOMES
PARTNERSHIP



Typical scenario: Estimate the effect of one drug on one outcome using one method against one database



Drug: ACE inhibitor
Outcome: Angioedema
Method: High-dimensional propensity score (HDPS)
Database: Thomson MarketScan Commercial Claims and Encounters (CCAE)

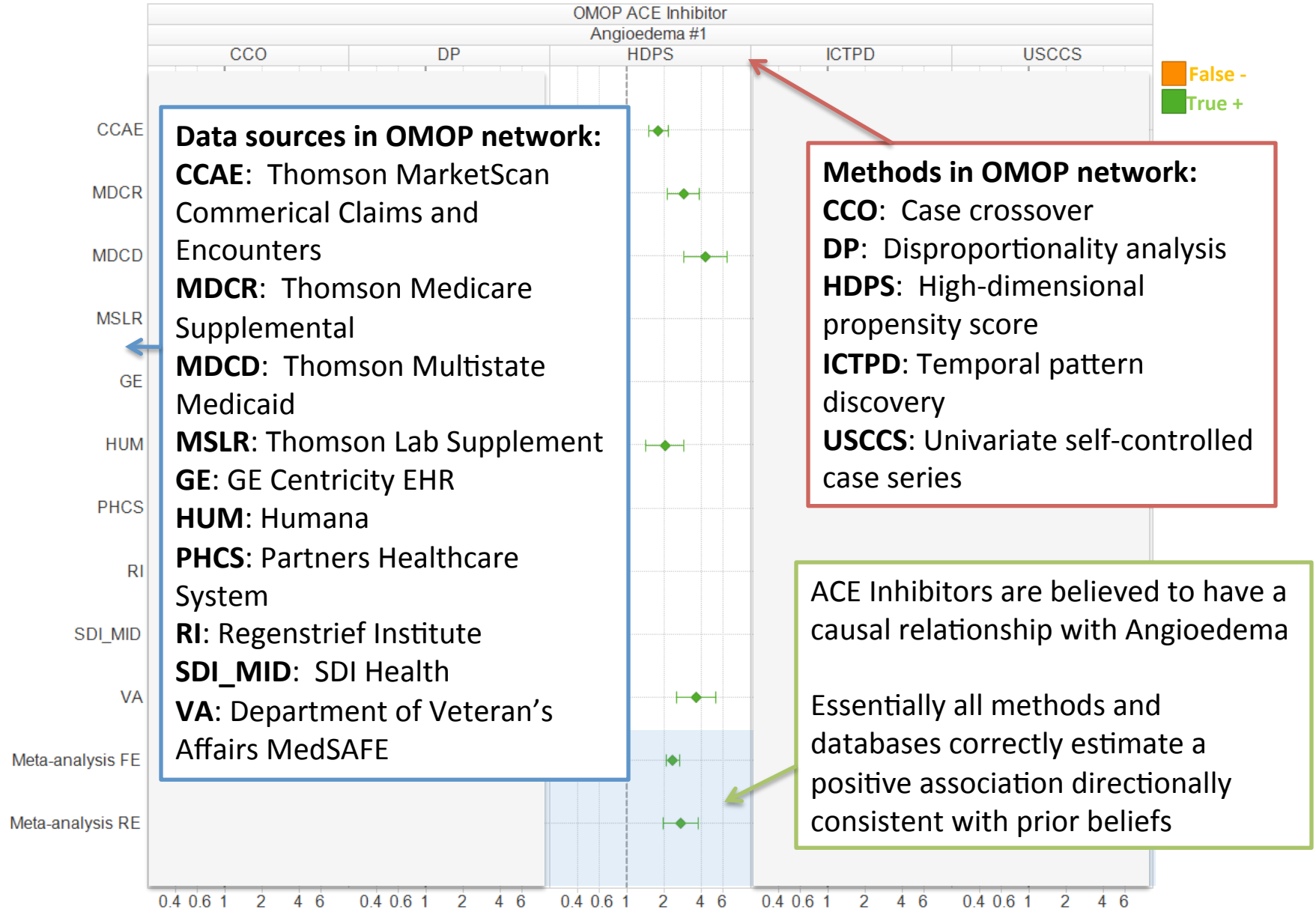
Typically focus on *magnitude of the effect*: relative risk (RR) and *statistical significance*: lower and upper bound of confidence interval (CI)

If this had been an randomized trial, we would know the CI has 95% coverage of the true effect size.

Because this is an observational study with the potential for bias, the operational characteristics are uncertain:

- Is the estimated association consistent with the directionality of the true causal relationship?
- How often does the CI actually contain the truth?

Systematic sensitivity analysis: Estimate the effect using multiple methods across the network of databases



Data sources in OMOP network:

CCAE: Thomson MarketScan Commerical Claims and Encounters

MDCR: Thomson Medicare Supplemental

MDCD: Thomson Multistate Medicaid

MSLR: Thomson Lab Supplement

GE: GE Centricity EHR

HUM: Humana

PHCS: Partners Healthcare System

RI: Regenstrief Institute

SDI_MID: SDI Health

VA: Department of Veteran's Affairs MedSAFE

Methods in OMOP network:

CCO: Case crossover

DP: Disproportionality analysis

HDPS: High-dimensional propensity score

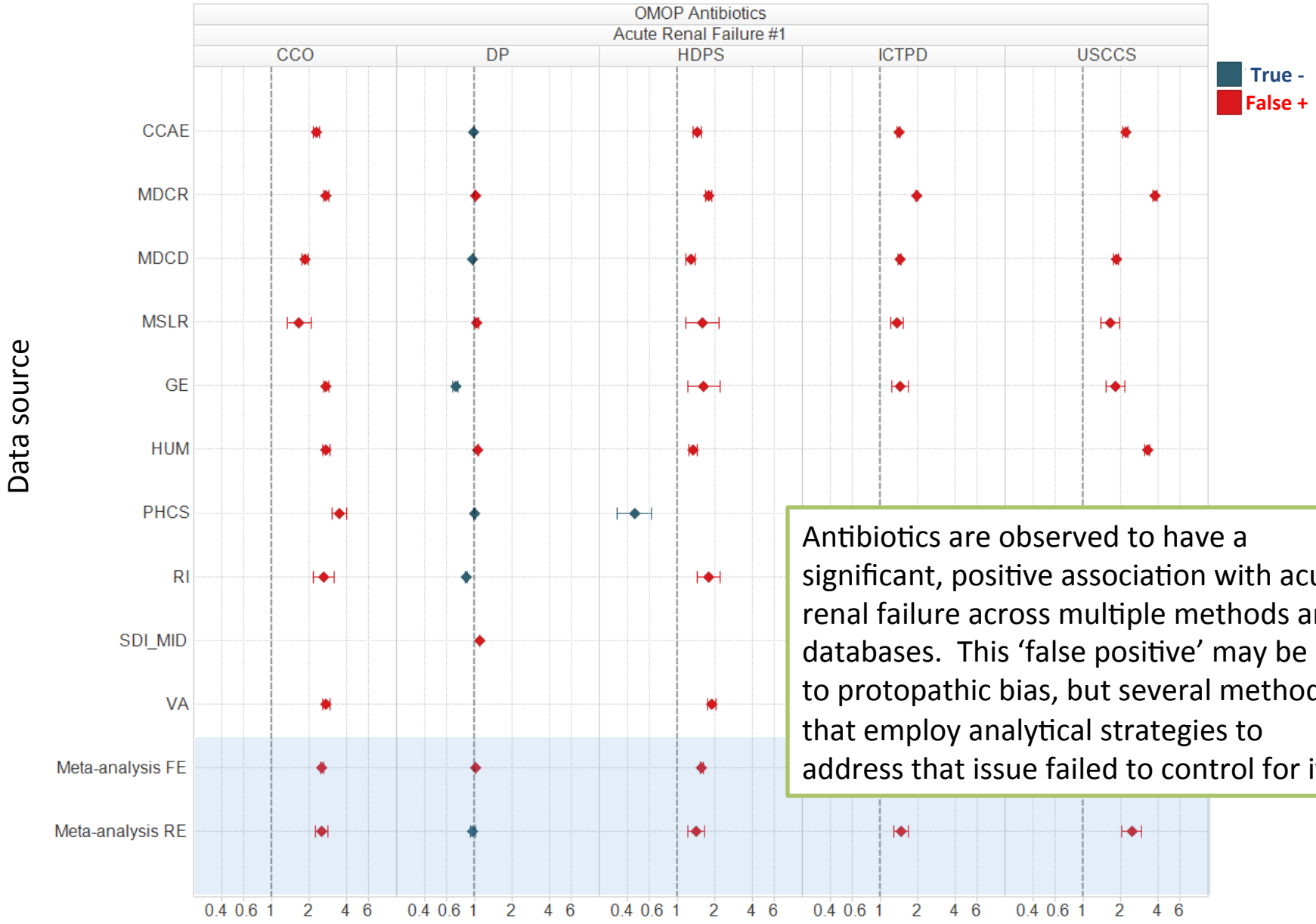
ICTPD: Temporal pattern discovery

USCCS: Univariate self-controlled case series

ACE Inhibitors are believed to have a causal relationship with Angioedema

Essentially all methods and databases correctly estimate a positive association directionally consistent with prior beliefs

Consistent ‘false positive’ observed for ‘negative control’ of Antibiotics and Acute Renal Failure



Antibiotics are observed to have a significant, positive association with acute renal failure across multiple methods and databases. This ‘false positive’ may be due to protopathic bias, but several methods that employ analytical strategies to address that issue failed to control for it.

Measuring method performance

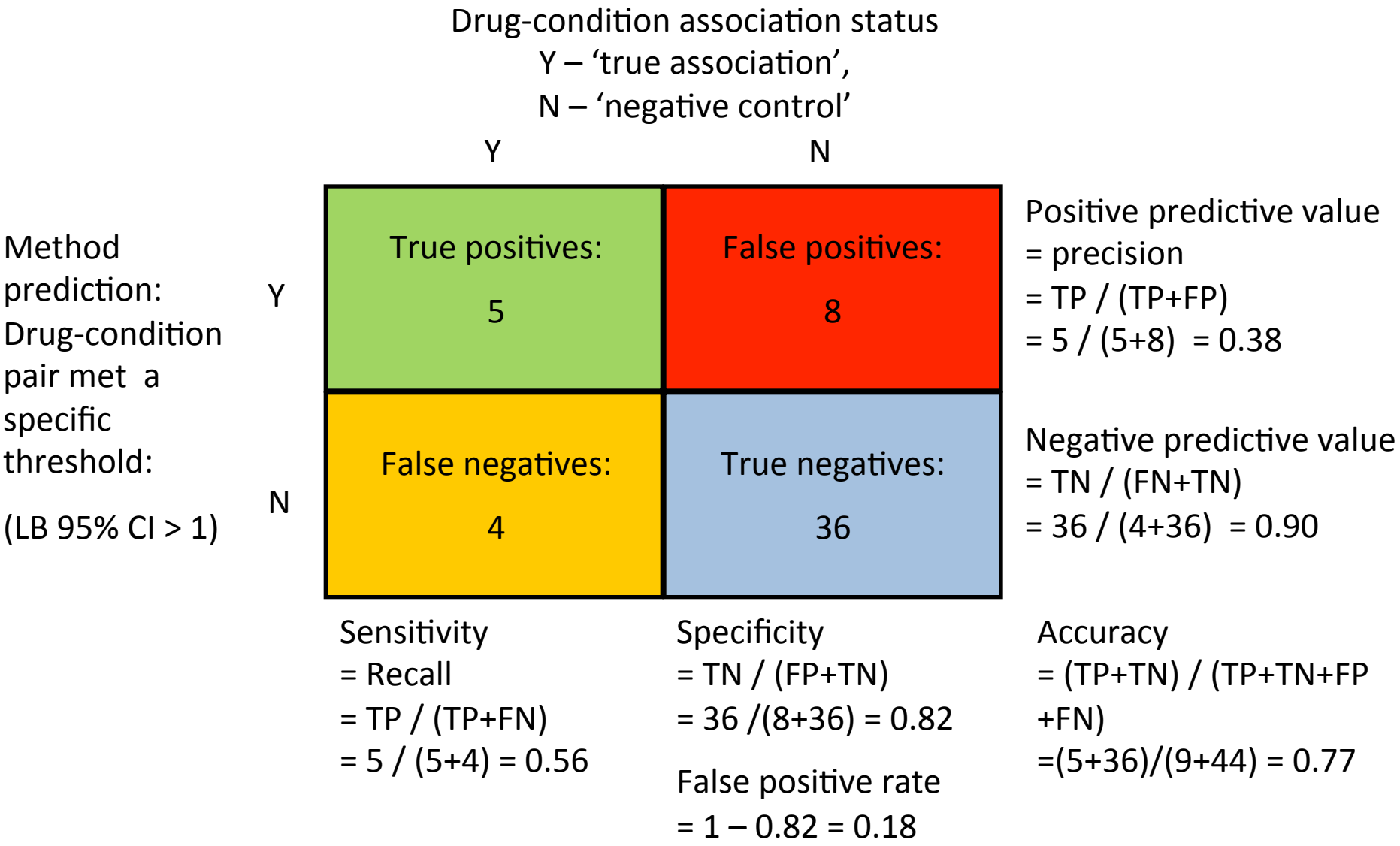
Drug-condition association status
Y – 'true association',
N – 'negative control'

Method prediction: Drug-condition pair met a specific threshold

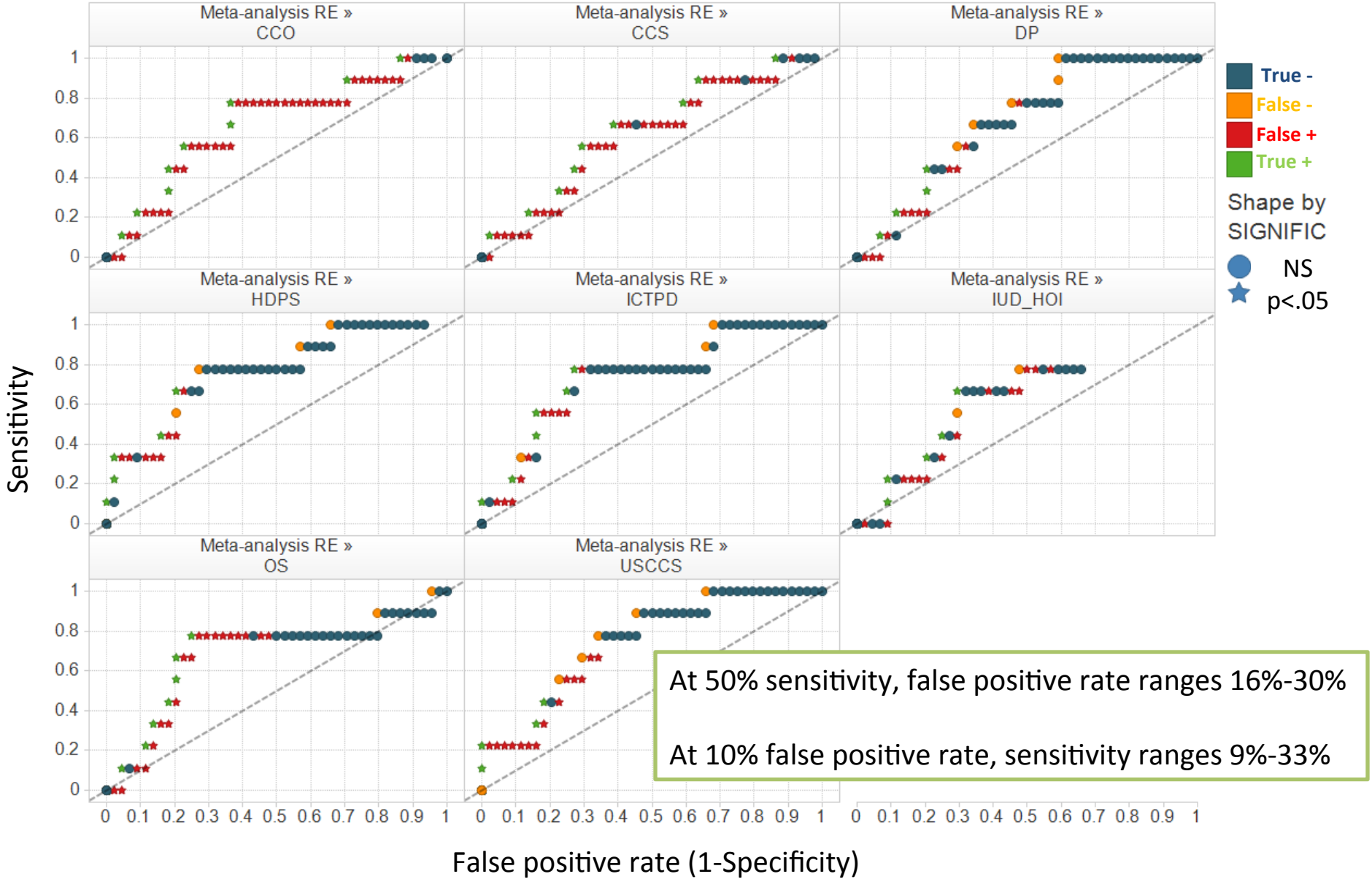
	Y	N
Y	**True positives**	**False positives**
N	**False negatives**	**True negatives**

Question: For any method applied to any data source, what are the expected operating characteristics?

Measuring method performance example: Random-effect meta-analysis of estimates from High-dimensional propensity score



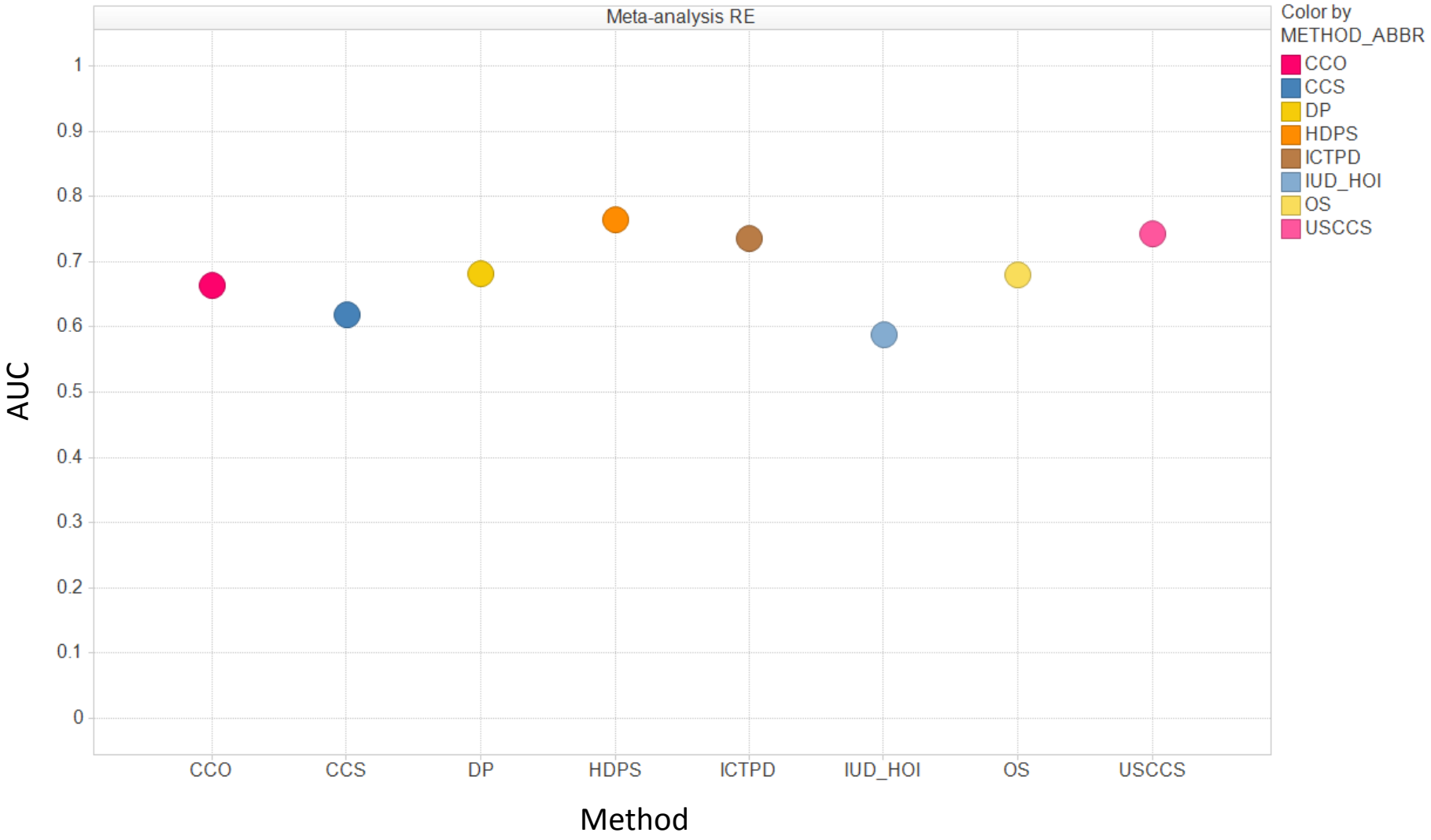
ROC curves of random-effects meta-analysis estimations for all methods



At 50% sensitivity, false positive rate ranges 16%-30%

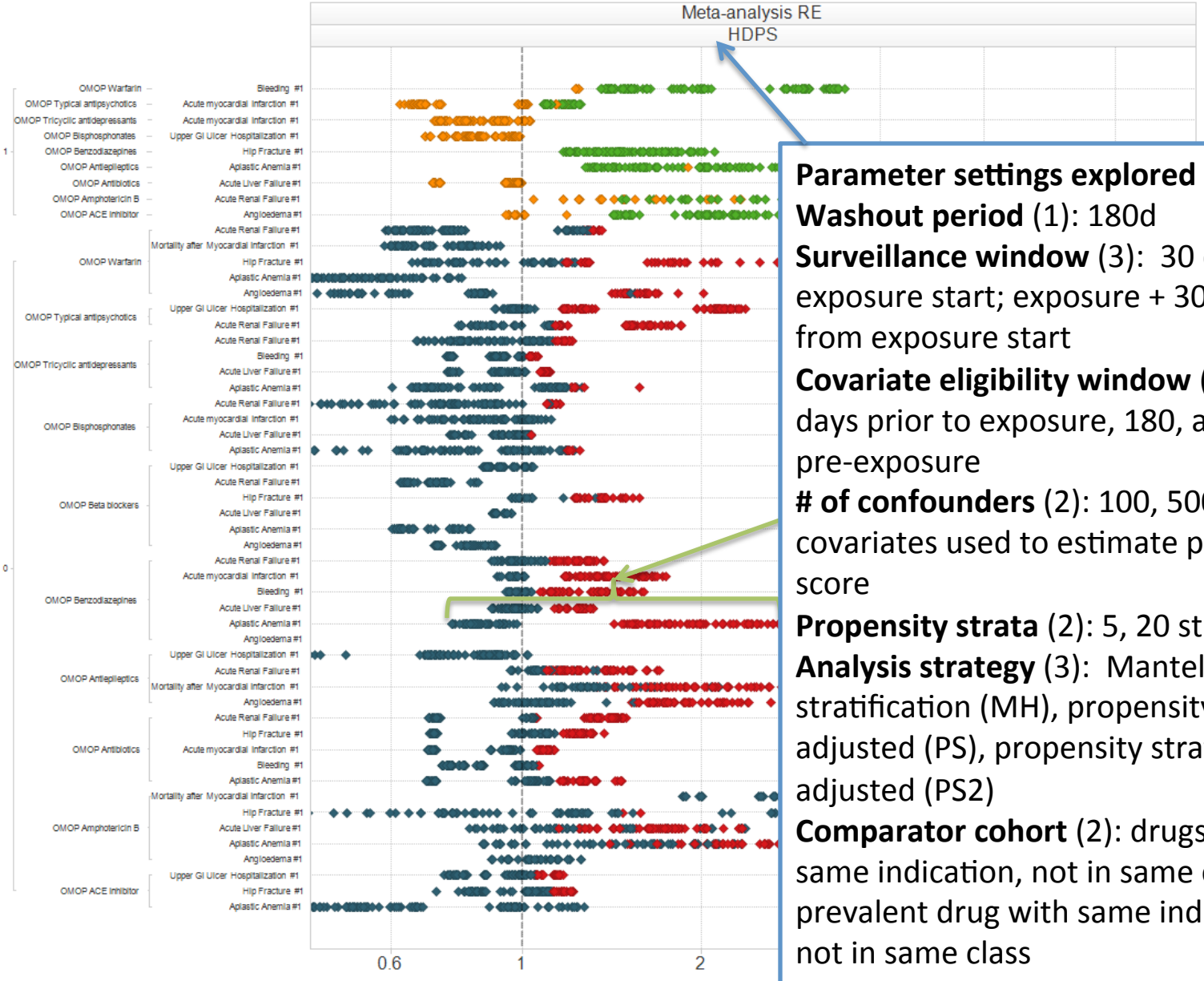
At 10% false positive rate, sensitivity ranges 9%-33%

AUC performance by method



Heterogeneity

Range of estimates across high-dimensional propensity score inception cohort (HDPS) parameter settings



Parameter settings explored in OMOP:

Washout period (1): 180d

Surveillance window (3): 30 days from exposure start; exposure + 30d ; all time from exposure start

Covariate eligibility window (3): 30 days prior to exposure, 180, all-time pre-exposure

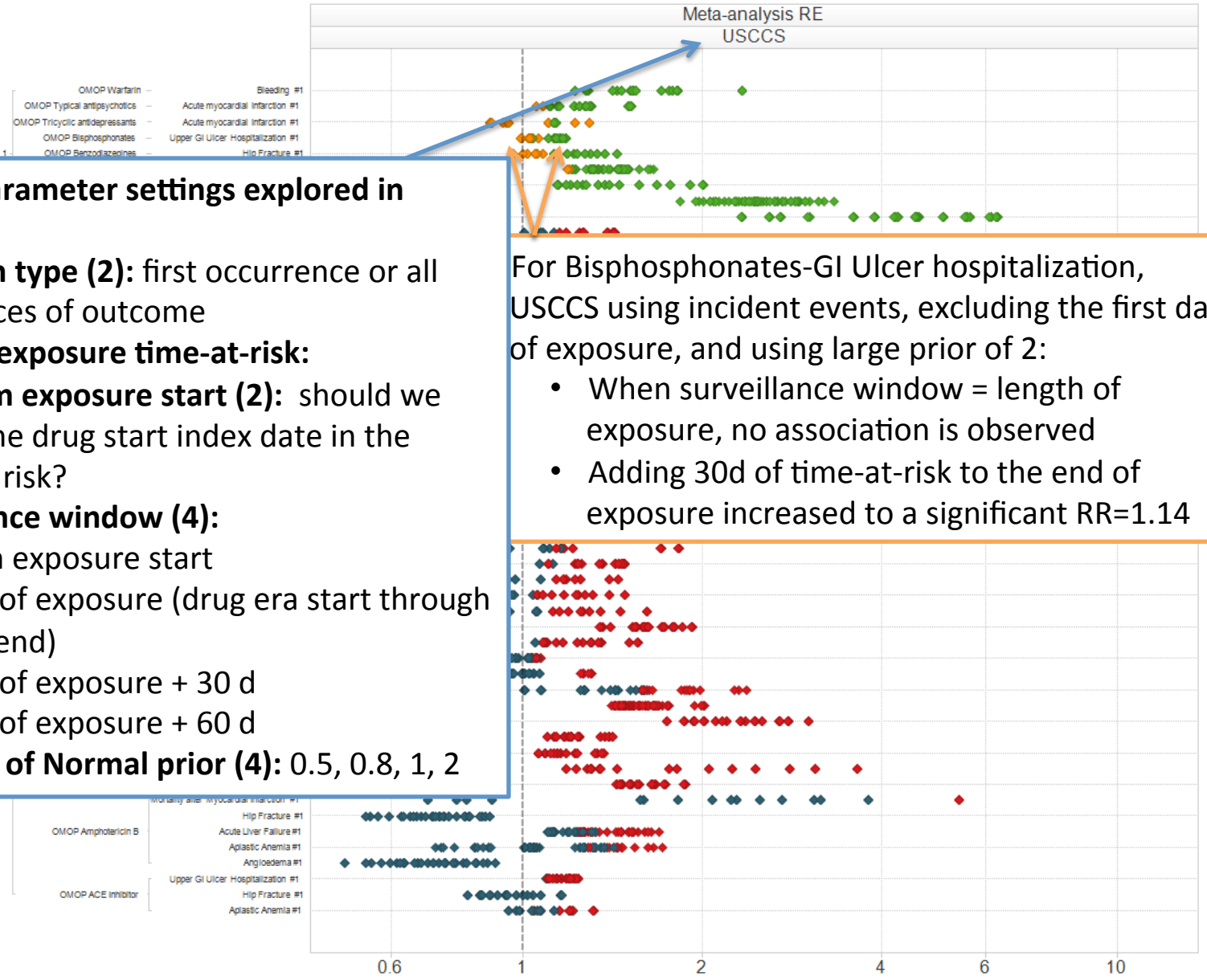
# of confounders (2): 100, 500 covariates used to estimate propensity score

Propensity strata (2): 5, 20 strata

Analysis strategy (3): Mantel-Haenszel stratification (MH), propensity score adjusted (PS), propensity strata adjusted (PS2)

Comparator cohort (2): drugs with same indication, not in same class; most prevalent drug with same indication, not in same class

Range of estimates across univariate self-controlled case series (USCCS) parameter settings



USCCS Parameter settings explored in OMOP:

Condition type (2): first occurrence or all occurrences of outcome

Defining exposure time-at-risk:

Days from exposure start (2): should we include the drug start index date in the period at risk?

Surveillance window (4):

30 d from exposure start

Duration of exposure (drug era start through drug era end)

Duration of exposure + 30 d

Duration of exposure + 60 d

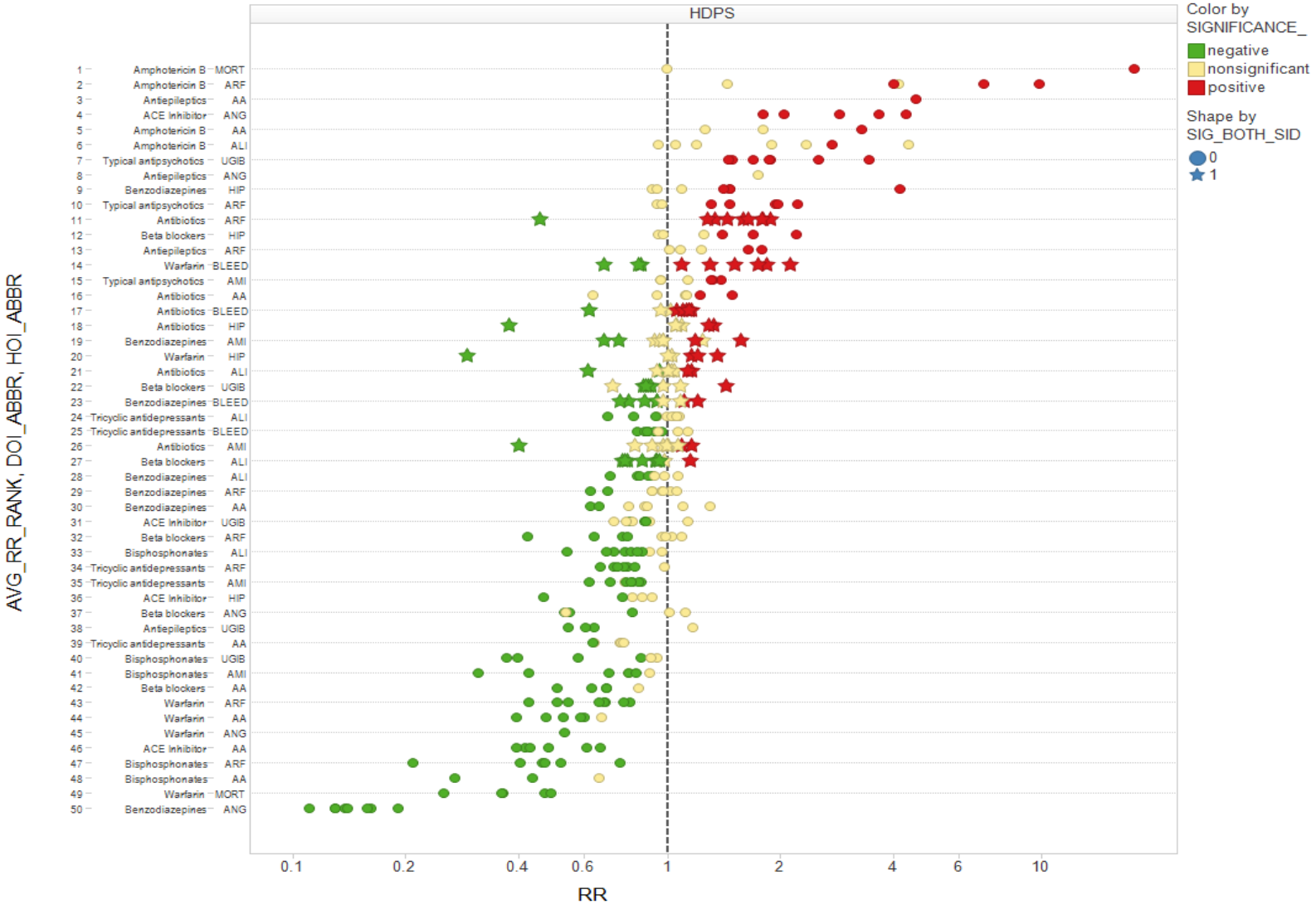
Precision of Normal prior (4): 0.5, 0.8, 1, 2

For Bisphosphonates-GI Ulcer hospitalization, USCCS using incident events, excluding the first day of exposure, and using large prior of 2:

- When surveillance window = length of exposure, no association is observed
- Adding 30d of time-at-risk to the end of exposure increased to a significant RR=1.14

Fix everything *except* the database…

Cohort

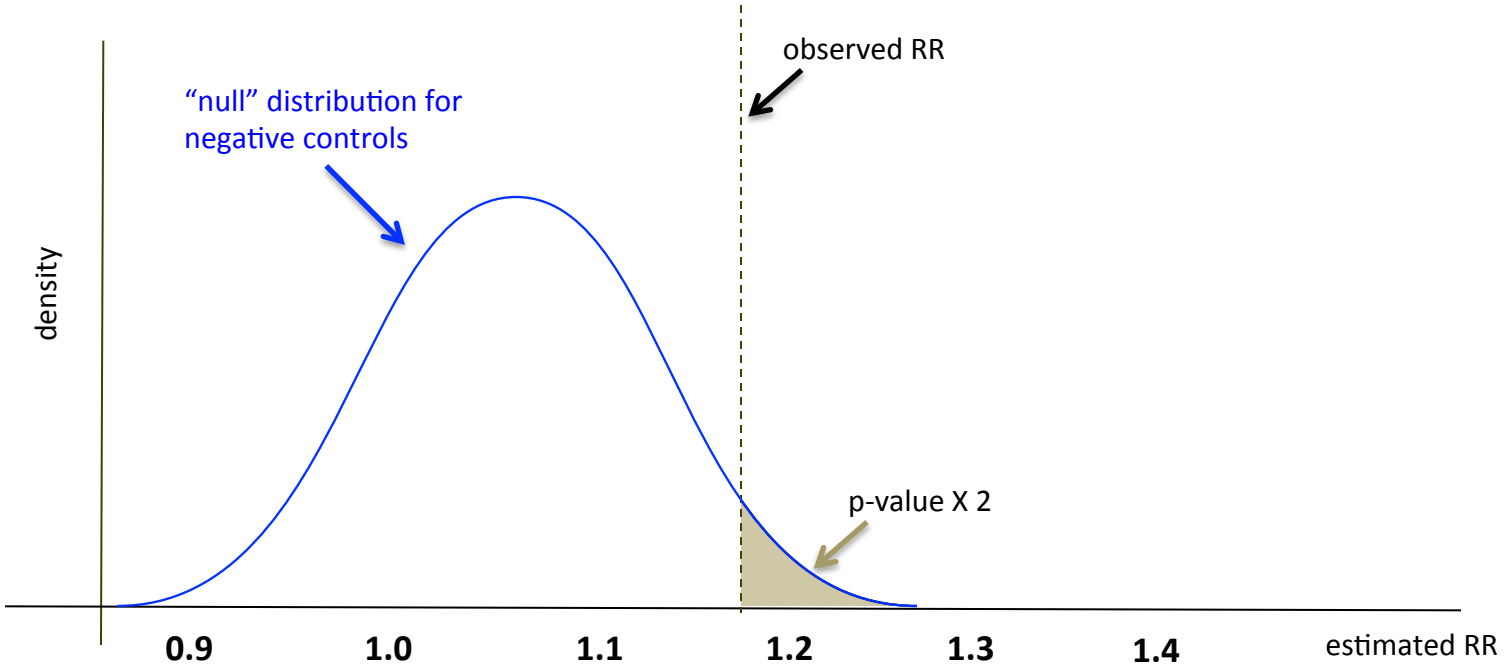

HDPS
AVG_RR_RANK, DOI_ABBR, HOI_ABBR
1 Amphotericin B MORT
2 Amphotericin B ARF
3 Antiepileptics AA
4 ACE Inhibitor ANG
5 Amphotericin B AA
6 Amphotericin B ALI
7 Typical antipsychotics UGIB
8 Antiepileptics ANG
9 Benzodiazepines HIP
10 Typical antipsychotics ARF
11 Antibiotics ARF
12 Beta blockers HIP
13 Antiepileptics ARF
14 Warfarin BLEED
15 Typical antipsychotics AMI
16 Antibiotics AA
17 Antibiotics BLEED
18 Antibiotics HIP
19 Benzodiazepines AMI
20 Warfarin HIP
21 Antibiotics ALI
22 Beta blockers UGIB
23 Benzodiazepines BLEED
24 Tricyclic antidepressants ALI
25 Tricyclic antidepressants BLEED
26 Antibiotics AMI
27 Beta blockers ALI
28 Benzodiazepines ALI
29 Benzodiazepines ARF
30 Benzodiazepines AA
31 ACE Inhibitor UGIB
32 Beta blockers ARF
33 Bisphosphonates ALI
34 Tricyclic antidepressants ARF
35 Tricyclic antidepressants AMI
36 ACE Inhibitor HIP
37 Beta blockers ANG
38 Antiepileptics UGIB
39 Tricyclic antidepressants AA
40 Bisphosphonates UGIB
41 Bisphosphonates AMI
42 Beta blockers AA
43 Warfarin ARF
44 Warfarin AA
45 Warfarin ANG
46 ACE Inhibitor AA
47 Bisphosphonates ARF
48 Bisphosphonates AA
49 Warfarin MORT
50 Benzodiazepines ANG
0.1
0.2
0.4
0.6
1
2
4
6
10
RR
Color by SIGNIFICANCE_
negative
nonsignificant
positive
Shape by SIG_BOTH_SID
0
1

SCCS

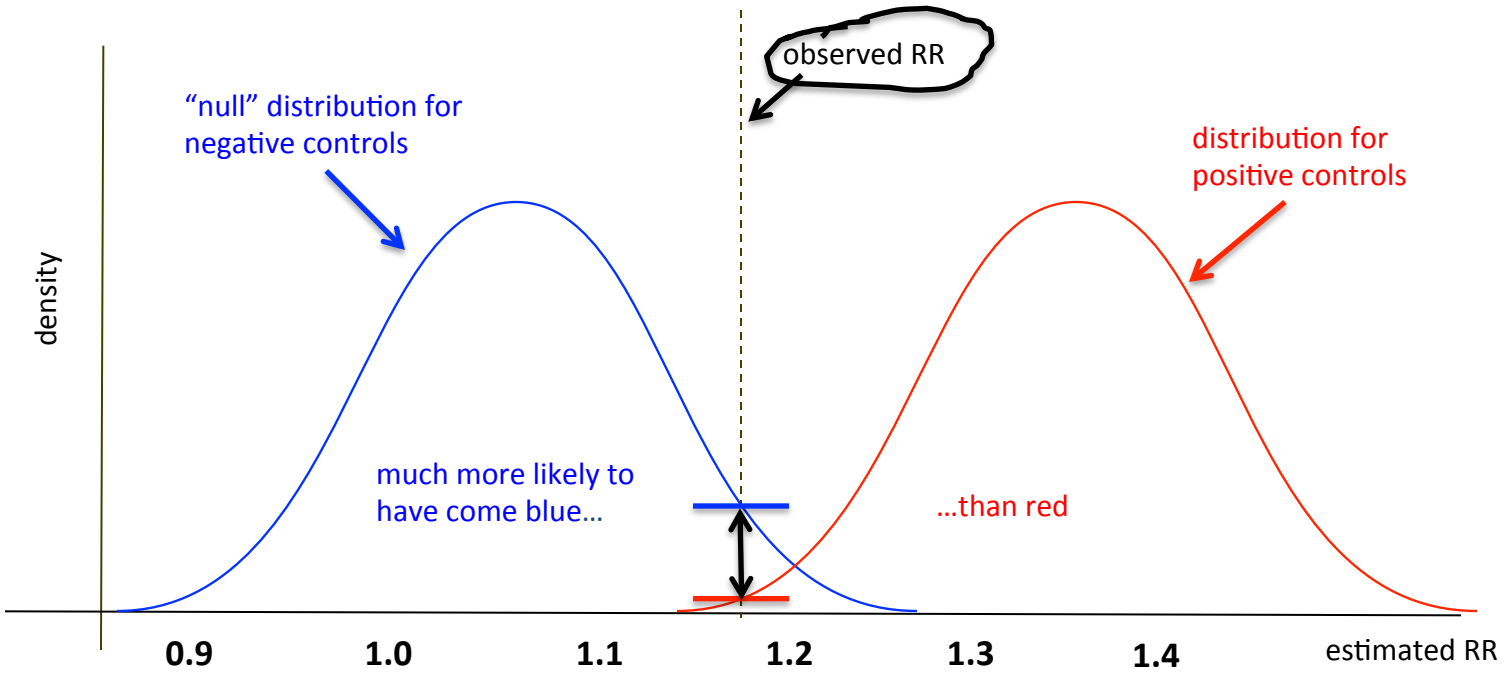

USCCS
AVG_RR_RANK, DOI_ABBR, HOI_ABBR
1 ACE Inhibitor ANG
2 Antibiotics ARF
3 Typical antipsychotics ARF
4 Amphotericin B ARF
5 Typical antipsychotics UGIB
6 Antibiotics AA
7 Warfarin AA
8 Benzodiazepines AA
9 Antiepileptics ANG
10 Amphotericin B ALI
11 Antiepileptics AA
12 Antibiotics BLEED
13 Warfarin BLEED
14 Antibiotics ALI
15 ACE Inhibitor UGIB
16 Bisphosphonates UGIB
17 Benzodiazepines ARF
18 Antibiotics HIP
19 Tricyclic antidepressants ARF
20 Beta blockers UGIB
21 Warfarin ANG
22 Antibiotics AMI
23 Beta blockers ARF
24 Benzodiazepines ANG
25 Benzodiazepines HIP
26 Beta blockers AA
27 Bisphosphonates ARF
28 ACE Inhibitor AA
29 Warfarin ARF
30 Typical antipsychotics AMI
31 Antiepileptics UGIB
32 Tricyclic antidepressants AA
33 Benzodiazepines ALI
34 Benzodiazepines BLEED
35 Beta blockers ALI
36 Amphotericin B ANG
37 Beta blockers ANG
38 Warfarin HIP
39 Warfarin MORT
40 Tricyclic antidepressants BLEED
41 Bisphosphonates AMI
42 Bisphosphonates AA
43 Tricyclic antidepressants ALI
44 Tricyclic antidepressants AMI
45 ACE Inhibitor HIP
46 Bisphosphonates ALI
47 Benzodiazepines AMI
48 Amphotericin B MORT
49 Beta blockers HIP
50 Amphotericin B AA
51 Antiepileptics ARF
52 Antiepileptics MORT
53 Amphotericin B HIP
0.1
0.2
0.4
0.6
1
2
4
6
RR
Color by SIGNIFICANCE_
negative
nonsignificant
positive
Shape by SIG_BOTH_SID
0
1

Interpreting Biased Estimates

p-values are tail probabilities

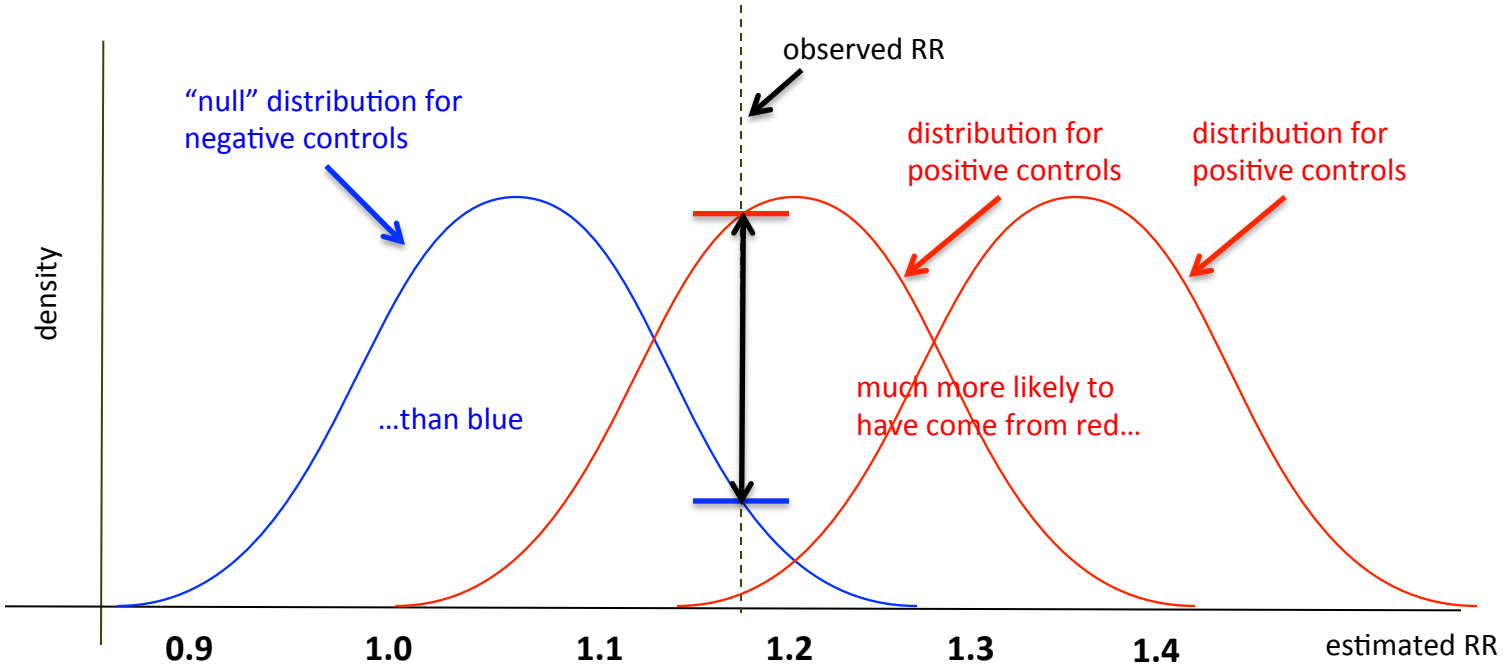


We also have positive controls



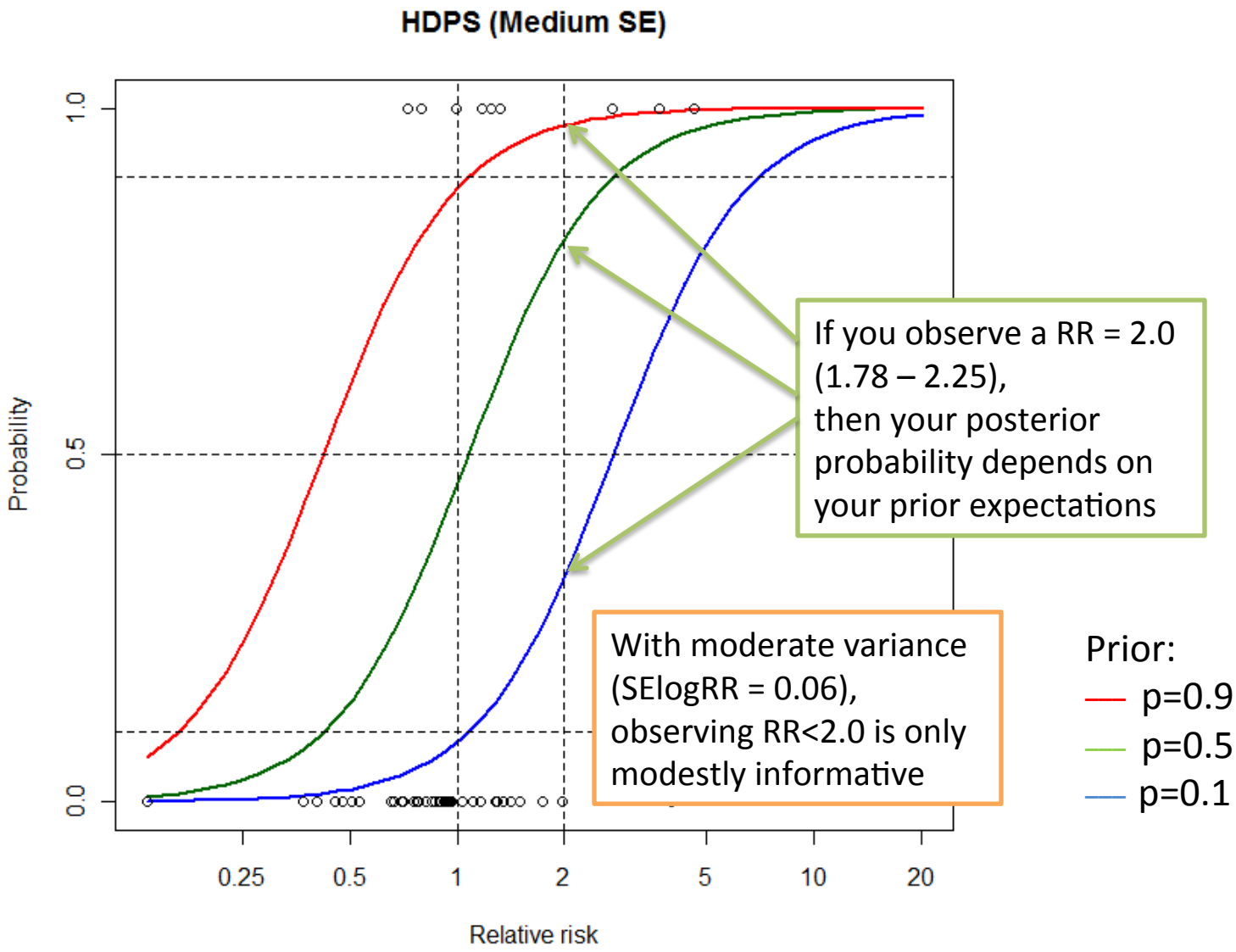
- moderate p-value, small Pr(positive control)

But if AUC is small…

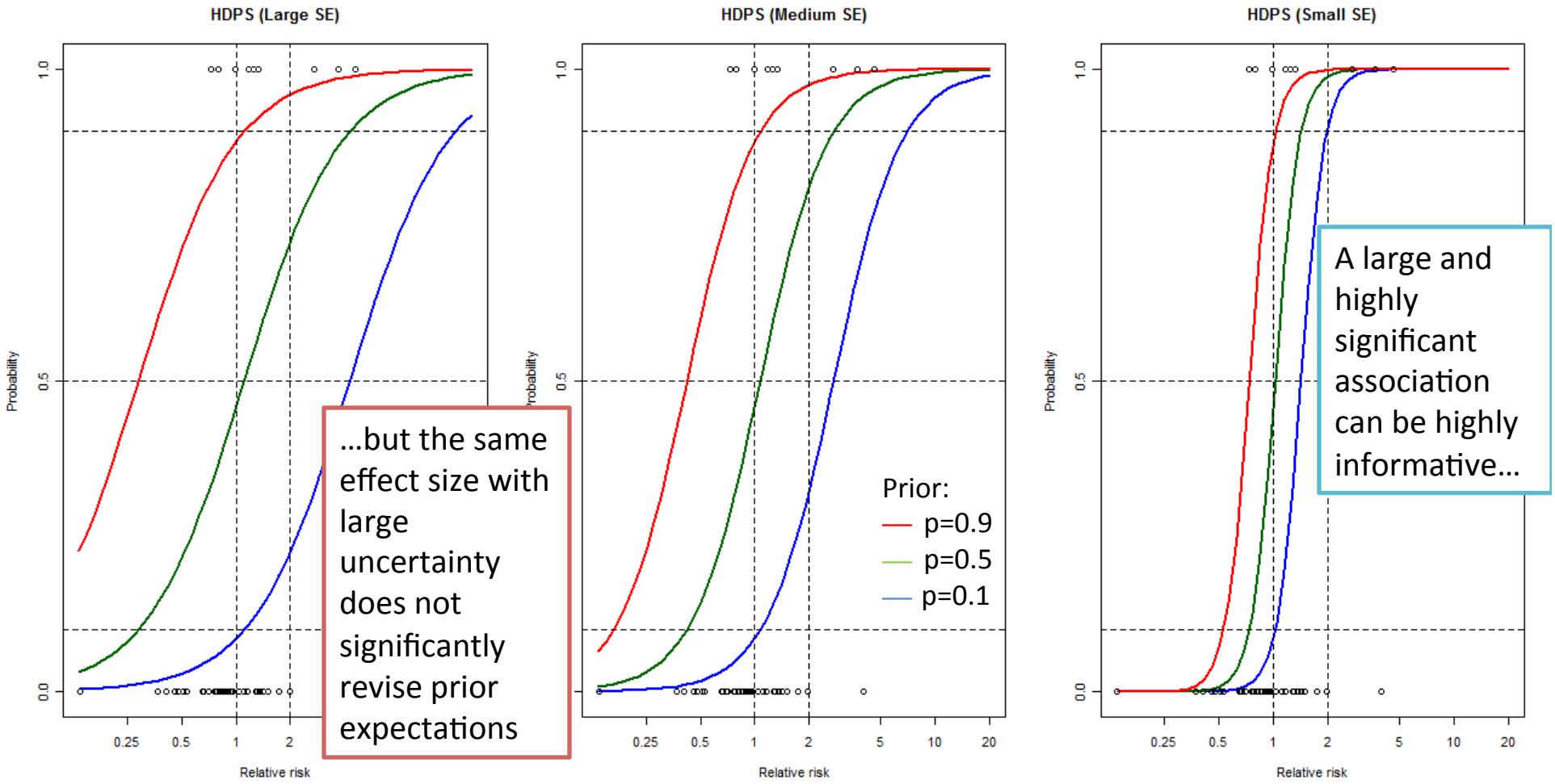


- despite the moderate p-value, Pr(positive control) is large!
- the null distribution doesn't tell the whole story

Revising prior expectations in light of new evidence



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)

Exploring clopidogrel and upper

BJCP British Journal of Clinical Pharmacology

DOI:10.1111/j.1365-2125.2008.03154.x

Gastro-intestinal haemorrhage risks of selecti antag look

Lucie Opatrny,[1,2] J. A. 'Chris' Delaney[1,3] & Samy Suissa[1,3]

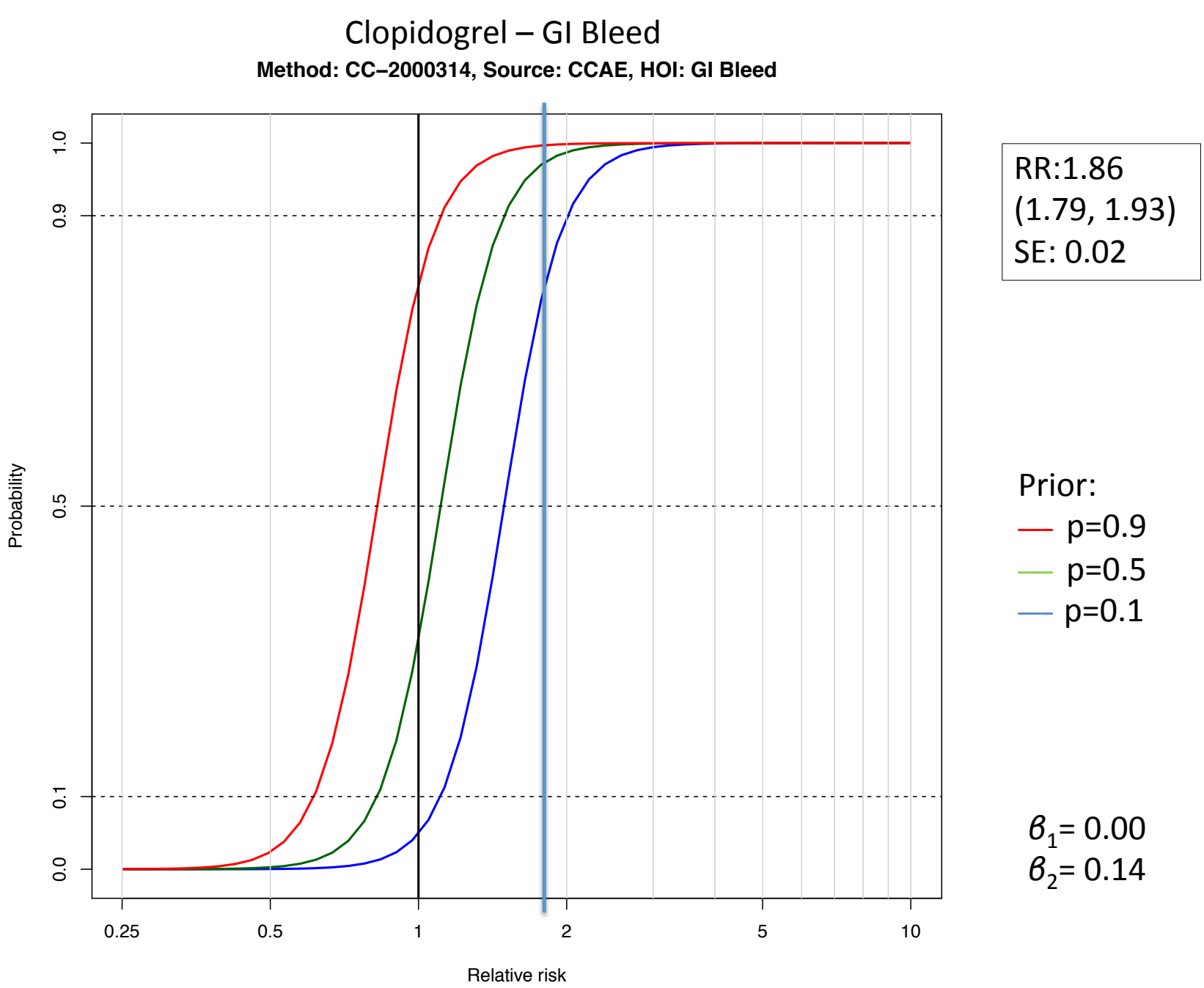
[1]Division of Clinical Epidemiology, [2]Division of Internal Medicine, McGill University Health Center and [3]Department of Epidemiology, Biostatistics and Occupational Health, McGill University, Montreal, Canada

Correspondence
Dr Samy Suissa PhD, Division of Clinical Epidemiology, Royal Victoria Hospital, 687 Pine Avenue West, Ross 4.29, Montreal, Quebec, Canada, H3A 1A1.
Tel.: + 1 51 4843 1564

2.07 1.66, 2.58

Received
21 July 2007
Accepted
24 January 2008
Published *OnlineEarly*
6 May 2008

Br J Clin Pharmacol / **66**:1 / 76–81


Clopidogrel – GI Bleed
Method: CC–2000314, Source: CCAE, HOI: GI Bleed
Probability
0.0
0.1
0.5
0.9
1.0
0.25
0.5
1
2
5
10
Relative risk
RR:1.86
(1.79, 1.93)
SE: 0.02
Prior:
p=0.9
p=0.5
p=0.1
θ_1= 0.00
θ_2= 0.14

Exploring isoniazid and acute liver

CMAJ RESEARCH

Adverse events associated with treatment of latent tuberculosis in the general population

Benjamin M. Smith MD, Kevin Schwartzman MD MPH, Gillian Bartlett PhD, Dick Menzies MD MSc

ABSTRACT

Background: Guidelines recommend tr
of latent tuberculosis in patients at ir
risk for active tuberculosis. Studies inv
ing the association of therapy with
adverse events have not included th
treated population nor accounted for
bidities or occurrence of similar event
untreated general population. Our o
was to estimate the risk of adverse events requiring hospital admission that were associated with therapy for latent tuberculosis infection in the general population.

Methods: Using administrative health data from the province of Quebec, we created a historical cohort of all residents dispensed therapy for latent tuberculosis between 1998 and 2003. Each patient was matched on age, sex and postal region with two untreated residents. The observation period was 18 months (from 6 months before to 12 months after initiation of therapy). The primary outcome was hospital admission for therapy-associated adverse events.

Results: During the period of observation, therapy for latent tuberculosis was dispensed to 9145 residents, of whom 95% started isoni-

6.4 (2.2–18.3)

years, the odds of hospital admission for a hepatic event among patients treated for latent tuberculosis infection was significantly greater than among matched untreated people after adjustment for comorbidities (odds ratio [OR] 6.4, 95% CI 2.2–18.3). Excluding patients with comorbid illness, there were two excess admissions to hospital for hepatic events per 100 patients initiating therapy compared with the rate among untreated people over 65 years (95% CI 0.1–3.87).

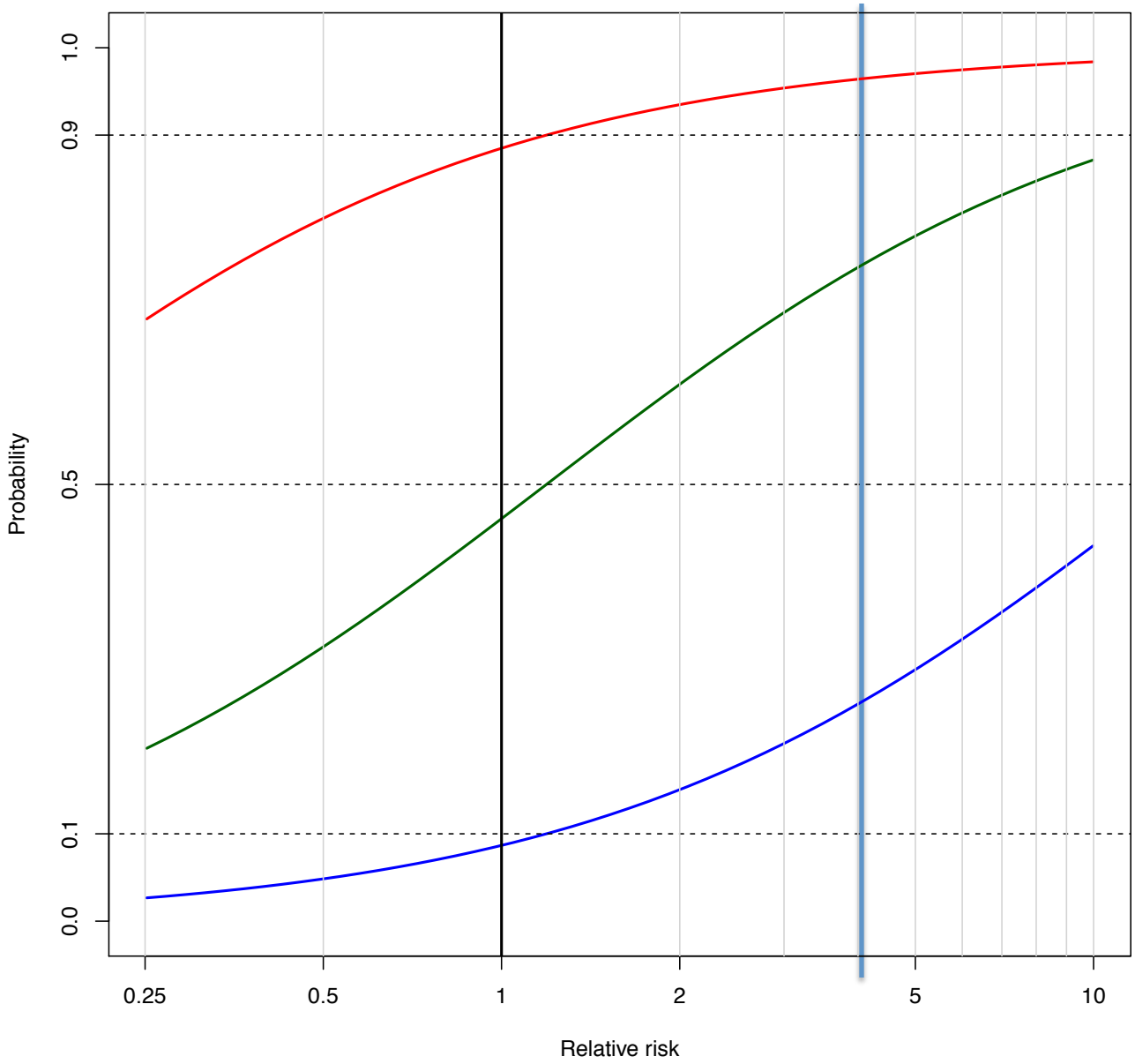
Interpretation: The risk of adverse events requiring hospital admission increased significantly among patients over 65 years receiving treatment for latent tuberculosis infection. The decision to treat latent tuberculosis infection in elderly patients should be made after careful consideration of risks and benefits.

/cmaj.091824

CMAJ, February 22, 2011, 183(3)

Isoniazid - Acute Liver Injury

Method: CM–21000211, Source: MDCR, HOI: Acute Liver Injury



RR:4.03
(2.69, 6.03)
SE: 0.21

Prior:
p=0.9
p=0.5
p=0.1

β_1= 0.79
β_2= 0.02

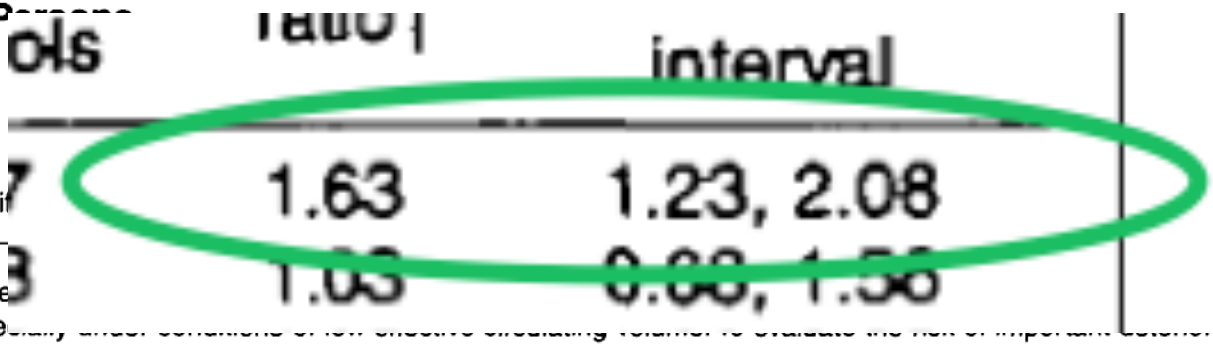
Exploring ibuprofen and acute kidney

American Journal of Epidemiology
Copyright © 2000 by The Johns Hopkins University School of Hygiene and Public Health
All rights reserved

Vol. 151, No. 5
Printed in U.S.A.

Nonsteroidal Antiinflammatory Drugs and Acute Renal Failure in Elderly P

Marie R. Gri

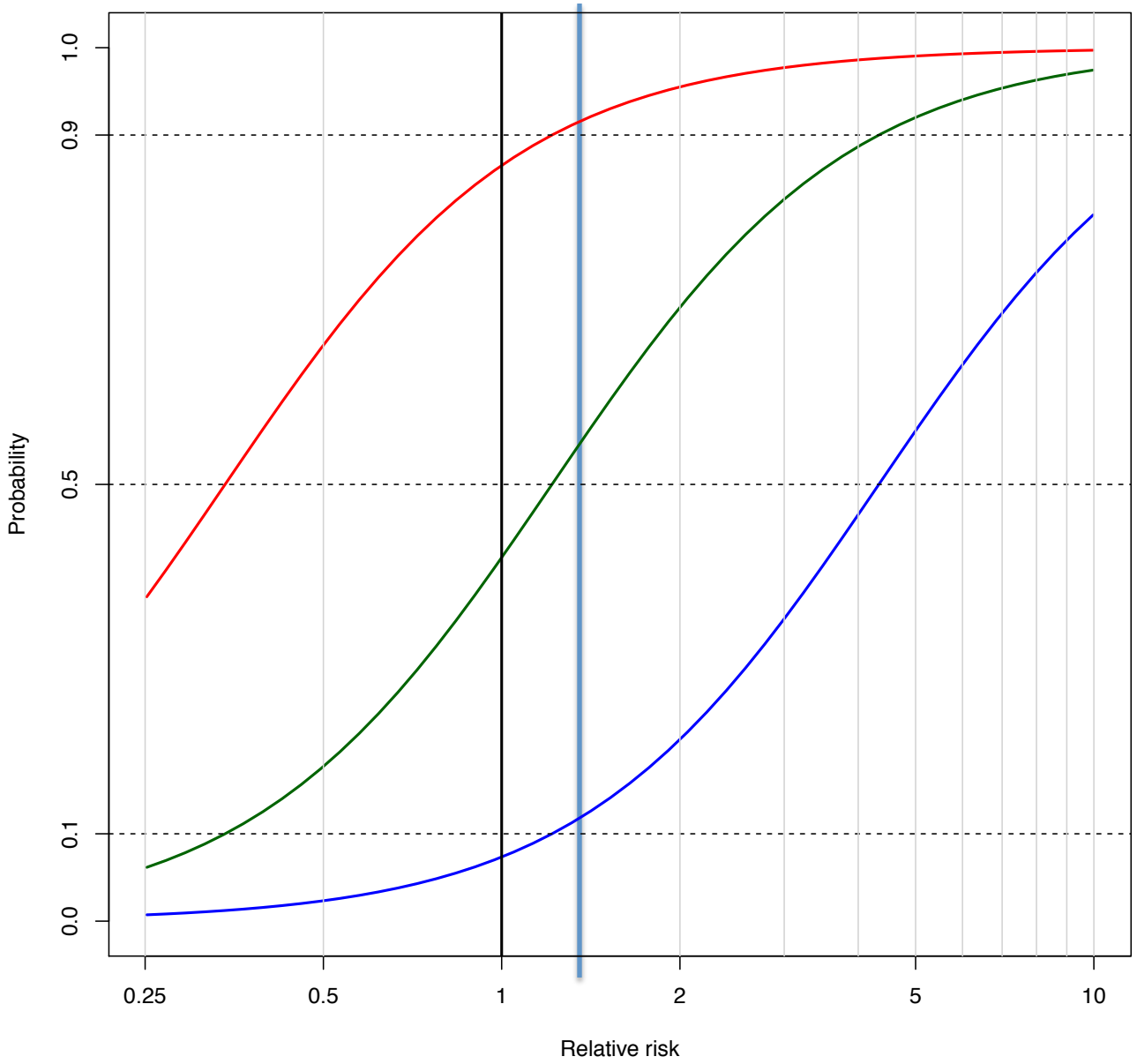


Re tion, espe n of renal function due to this effect, the authors performed a nested case-control study using Tennessee Medicaid enrollees aged ≥65 years in 1987–1991. Cases were patients who had been hospitalized with community-acquired acute renal failure; they were selected on the basis of medical record review of Medicaid enrollees with selected discharge diagnoses. Information on the timing, duration, and dose of prescription NSAIDs used, demographic factors, and comorbidity was gathered from computerized Medicaid-Medicare data files. Of the 1,799 patients with acute renal failure (4.51 hospitalizations per 1,000 person-years), 18.1% were current users of prescription NSAIDs as compared with 11.3% of 9,899 randomly selected population controls. After control for demographic factors and comorbidity, use of NSAIDs increased the risk of acute renal failure 58% (adjusted odds ratio = 1.58; 95% confidence interval (CI): 1.34, 1.86). For ibuprofen, which accounted for 35% of NSAID use, odds ratios associated with dosages of ≤1,200 mg/day, >1,200–<2,400 mg/day, and ≥2,400 mg/day were 0.94 (95% CI: 0.58, 1.51), 1.89 (95% CI: 1.34, 2.67), and 2.32 (95% CI: 1.45, 3.71), respectively (test for linear trend: $p = 0.009$). Prescription NSAID use resulted in an estimated 25 excess hospitalizations associated with renal failure per 10,000 years of use. Thus, NSAIDs represent a relatively uncommon but avoidable cause of acute renal failure in frail elderly persons. *Am J Epidemiol* 2000;151:488–96.

Am J Epidemiol 2000;151:488–96.

Ibuprofen – Acute Kidney Injury

Method: CC–2000241, Source: MDCD, HOI: Acute Kidney Injury



RR:1.26
(1.16, 1.36)
SE: 0.04

Prior:
p=0.9
p=0.5
p=0.1

β_1= 0.00
β_2= 0.07

Exploring indomethacin and acute myocardial infarction

Current Use of Nonsteroidal Antiinflammatory Drugs and the Risk of Acute Myocardial Infarction

Lorenz M. Fischer, M.Sc., Raymond G. Schlienger, Ph.D., M.P.H.,
Christian M. Matter, M.D., Hershel Jick, M.D., and Christoph R. Meier, Ph.D., M.Sc.

Study Ob n during
current .
Design. I
Setting. (
0.96 (0.66–1.38)
1.36 (0.82–2.25)
0.95 (0.53–1.69)

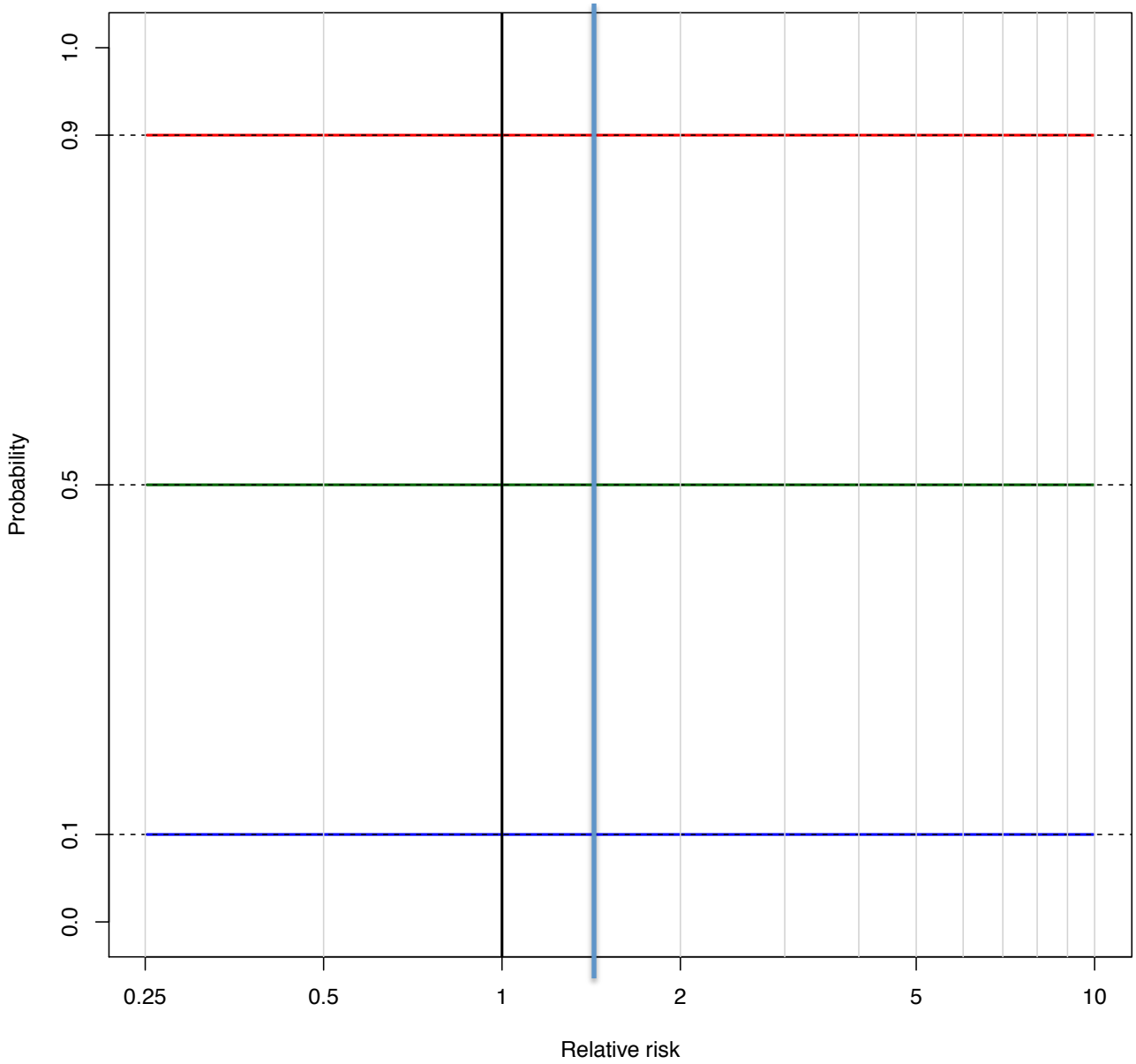
Subjects. A total of 8688 case patients, aged 89 years or younger, with a first-time acute myocardial infarction and 33,923 control subjects matched on age, sex, calendar time, and general practice attended.

Intervention. The United Kingdom General Practice Research Database was searched for potential cases of first-time acute myocardial infarction between January 1995 and April 2001. Control subjects without acute myocardial infarction were identified at random.

(Pharmacotherapy 2005;25(4):503–510)

Indomethacin – Acute Myocardial Infarction

Method: CC–2000241, Source: GE, HOI: Acute Myocardial Infarction



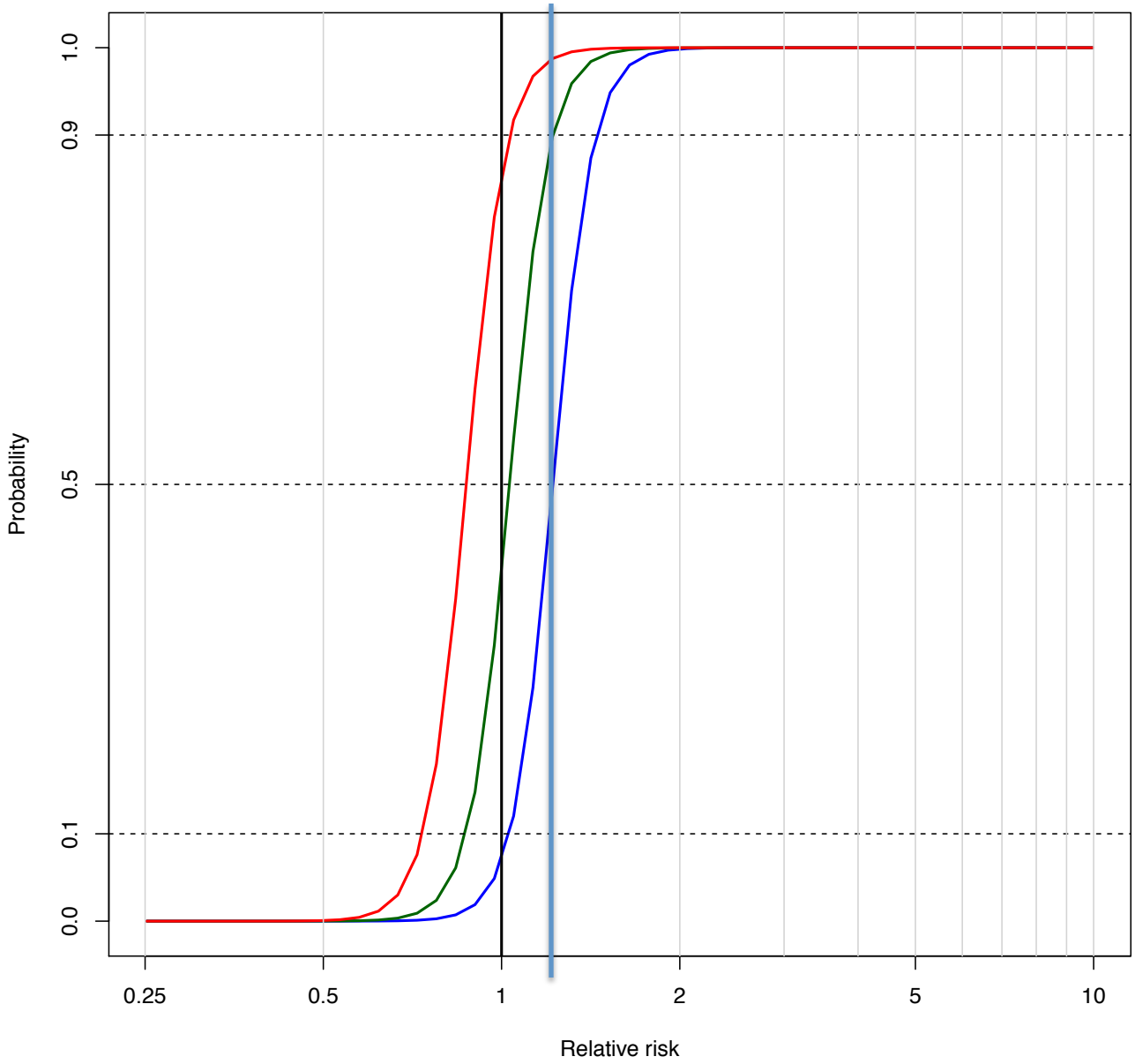
RR:1.44
(1.20, 1.72)
SE: 0.09

Prior:
p=0.9
p=0.5
p=0.1

β_1= 0.00
β_2= 0.00

Indomethacin – Acute Myocardial Infarction

Method: SCCS–1953010, Source: GE, HOI: Acute Myocardial Infarction

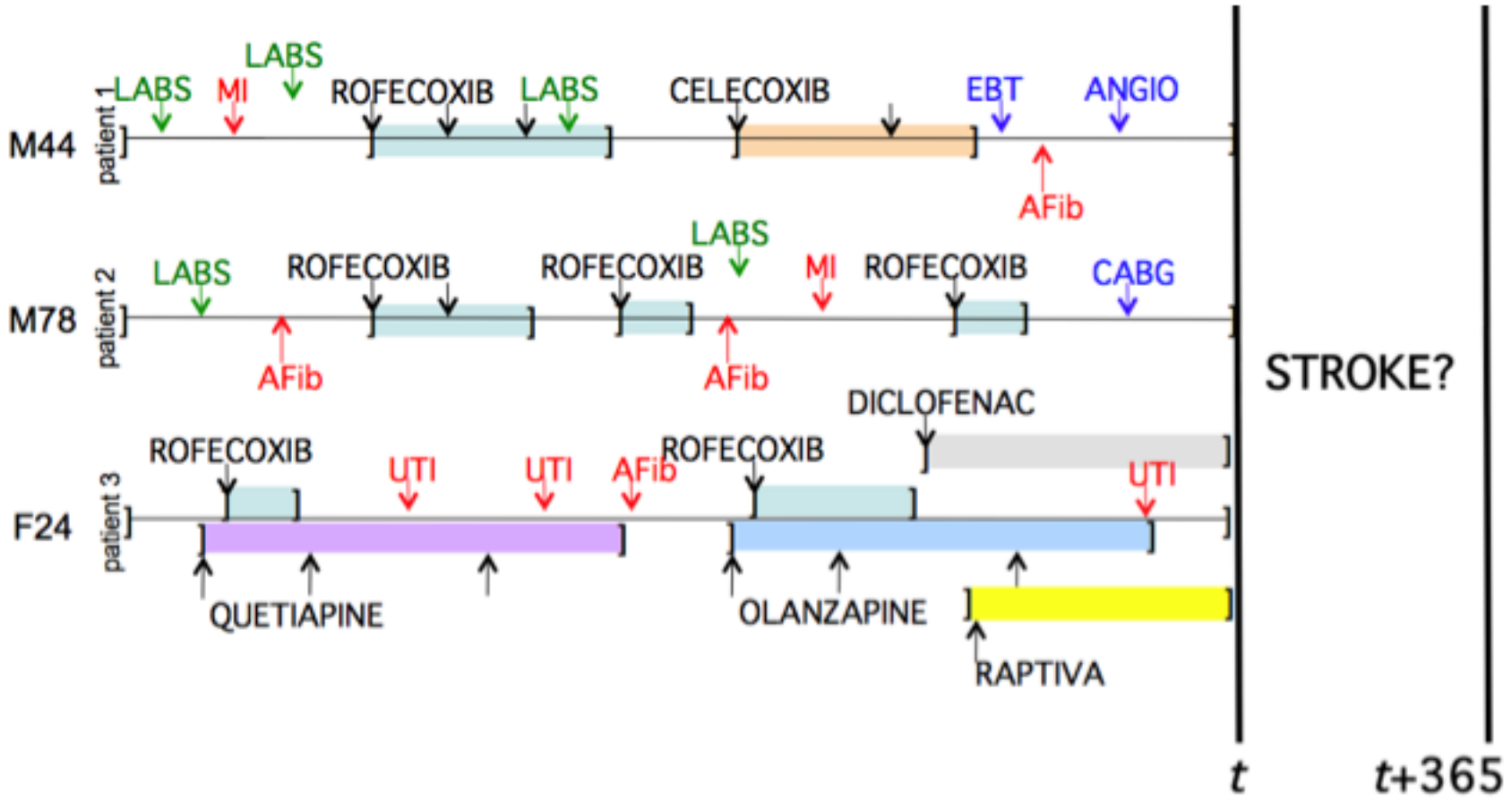


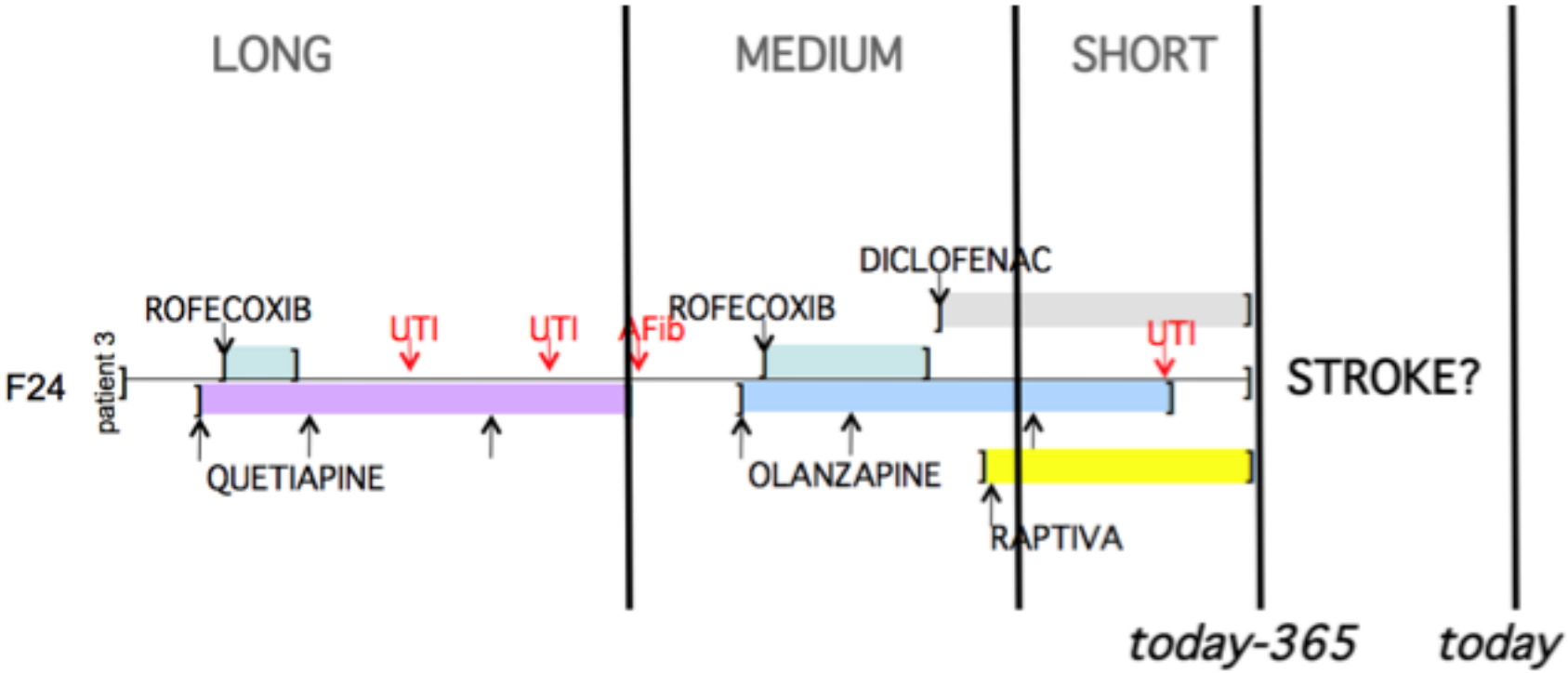
RR:1.16
(1.06, 1.72)
SE: 0.05

Prior:
p=0.9
p=0.5
p=0.1

θ_1= 4.97
θ_2= 0.37

Predictive Modeling


M44
patient 1
LABS
MI
LABS
ROFECOXIB
LABS
CELECOXIB
EBT
ANGIO
AFib
M78
patient 2
LABS
ROFECOXIB
ROFECOXIB
LABS
MI
ROFECOXIB
CABG
AFib
AFib
F24
patient 3
ROFECOXIB
UTI
UTI
AFib
ROFECOXIB
DICLOFENAC
UTI
QUETIAPINE
OLANZAPINE
RAPTIVA
STROKE?
t
t+365


LONG
MEDIUM
SHORT
F24
patient 3
ROFECOXIB
UTI
UTI
AFib
ROFECOXIB
DICLOFENAC
UTI
STROKE?
QUETIAPINE
OLANZAPINE
RAPTIVA
today-365
today

Discussion

- Current approaches to reporting observational studies are broken
- Big data can transform medicine but we have to stop kidding ourselves about
  - the role of “clinical” judgment in a data-rich world, and
  - the reliability of observational studies

Three Pieces of the Solution

1. Sensitivity Analysis

Greenland S. Multiple-bias modelling for analysis of observational data. Journal of the Royal Statistical Society Series a-Statistics in Society 2005;168:267-291.

2. Establish Operating Characteristics

- How well do observational studies work?
- How big is the bias in practice?
- How often do 95% intervals contain the true RR?
- What does this mean?

1.67 (95% CI, 1.33-2.09).

Three Pieces of the Solution

3. Prediction of Future Observables

4. Avoid reliance on human art

Collaborators

- Patrick Ryan, J&J
- Martijn Schuemie, Erasmus University
- Marc Suchard, UCLA
- Bill DuMouchel, Oracle
- Marc Overhage, Siemens
- Paul Stang, J&J
- Ivan Zorych, Columbia
- Tyler McCormick, U. Washington
- Cynthia Rudin, MIT
- Ben Latham, MIT

Let’s not do this!