Statistical Methods for Drug Safety Surveillance

Big Data to the Rescue?

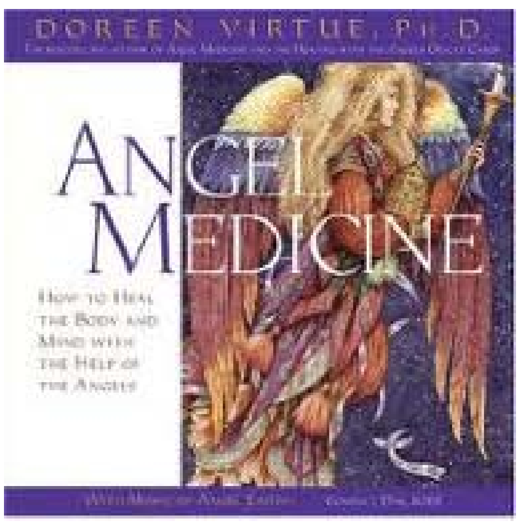
David Madigan
Department of Statistics
Columbia University

December 7, 2011 at The Wharton School,
University of Pennsylvania

“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?
Should
steady?


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

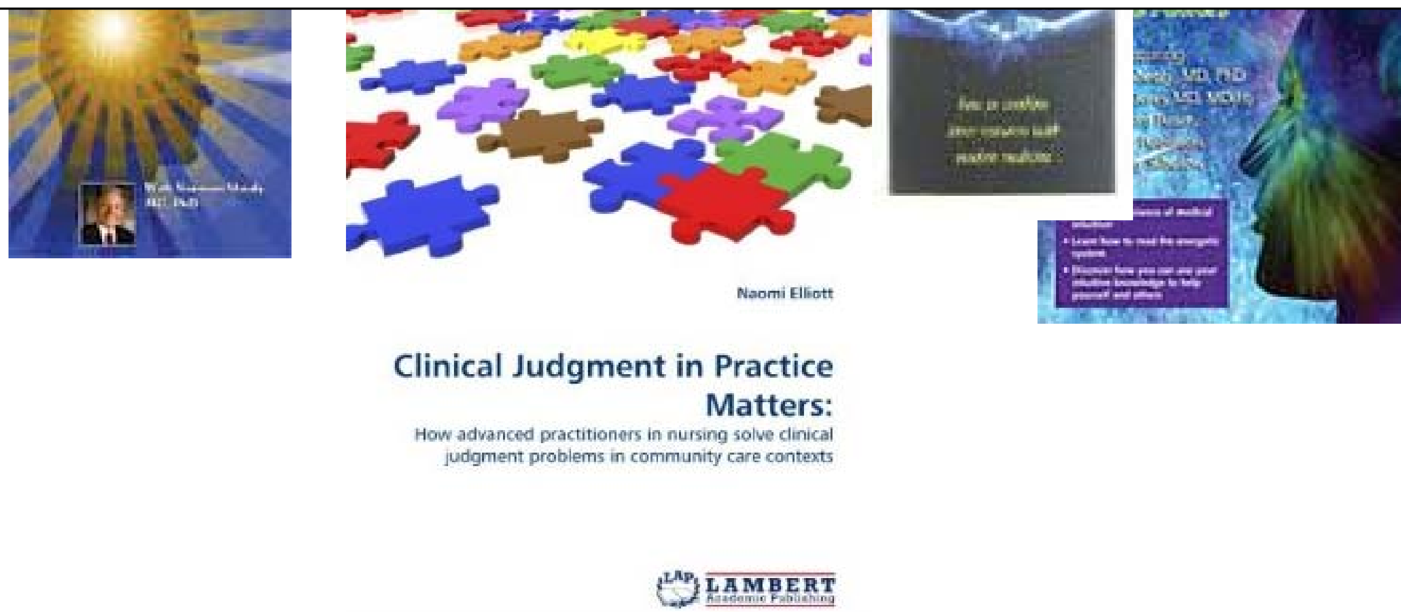

DOREEN VIRTUE, Ph.D.
ANGEL
MEDICINE
How to Heal
the Body and
Mind with
the Help of
the Angels


IT IS
CERTAIN

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.



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.

MERCK

Should John have an angiogram?

48 yea… o…

HDL = 5…

…RP …ormal

triglycerides = 106

LDL = 70

father died of heart di…

university professor

Clinical judgment?

calcium score in 2003 = 19

ca…cium score in 20…8 = 4…

calcium score in 2010 = 70

BMI = 21.6

stress test normal in 2007

no diabetes

mother died of cancer (83)

EKG unusual in 2009

Who are we kidding?

…spirin

…exercise

lipit…

…ed w…

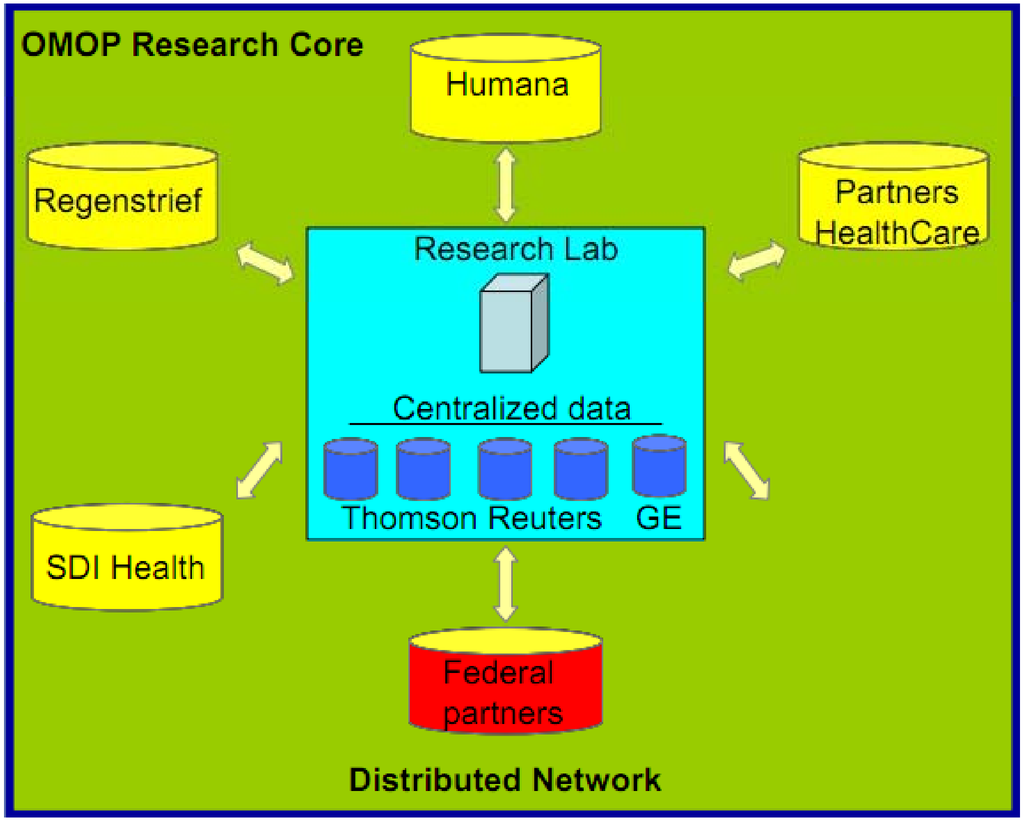
arrhythmia in 2008

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

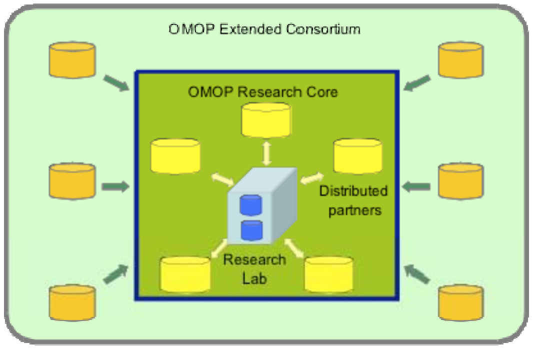
- Statistical/Epidemiological
- Computational
- Drug Safety

OMOP

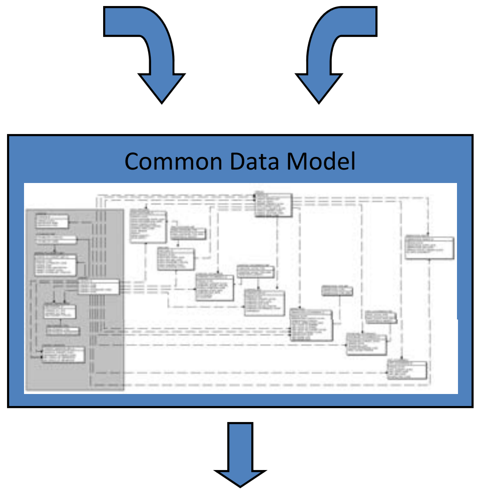
- How well can statistical methods applied to the OMOP data detect known drug safety issues?
- FDA is building Sentinel

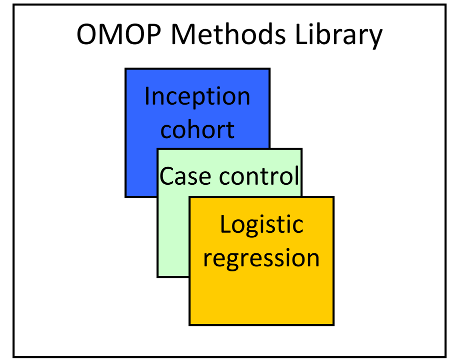
OMOP Research Experiment

- Open-source
- Standards-based

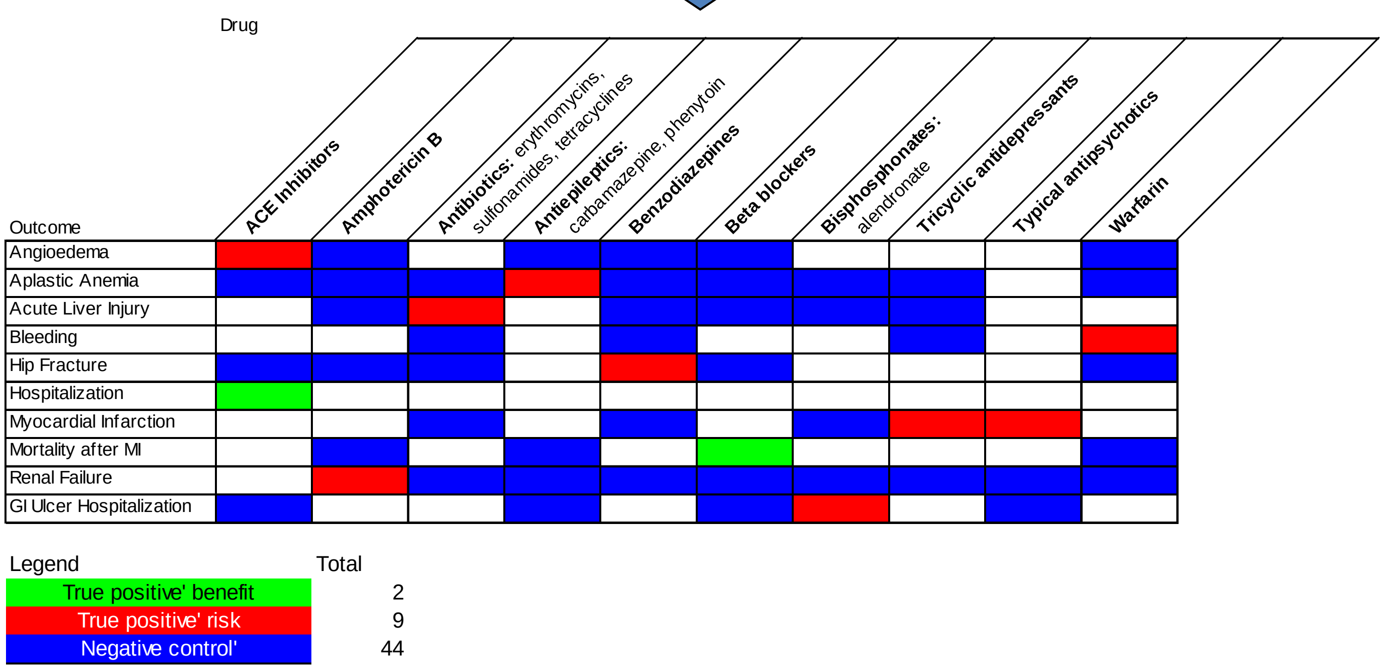


- 10 data sources
- Claims and EHRs
- 200M+ lives

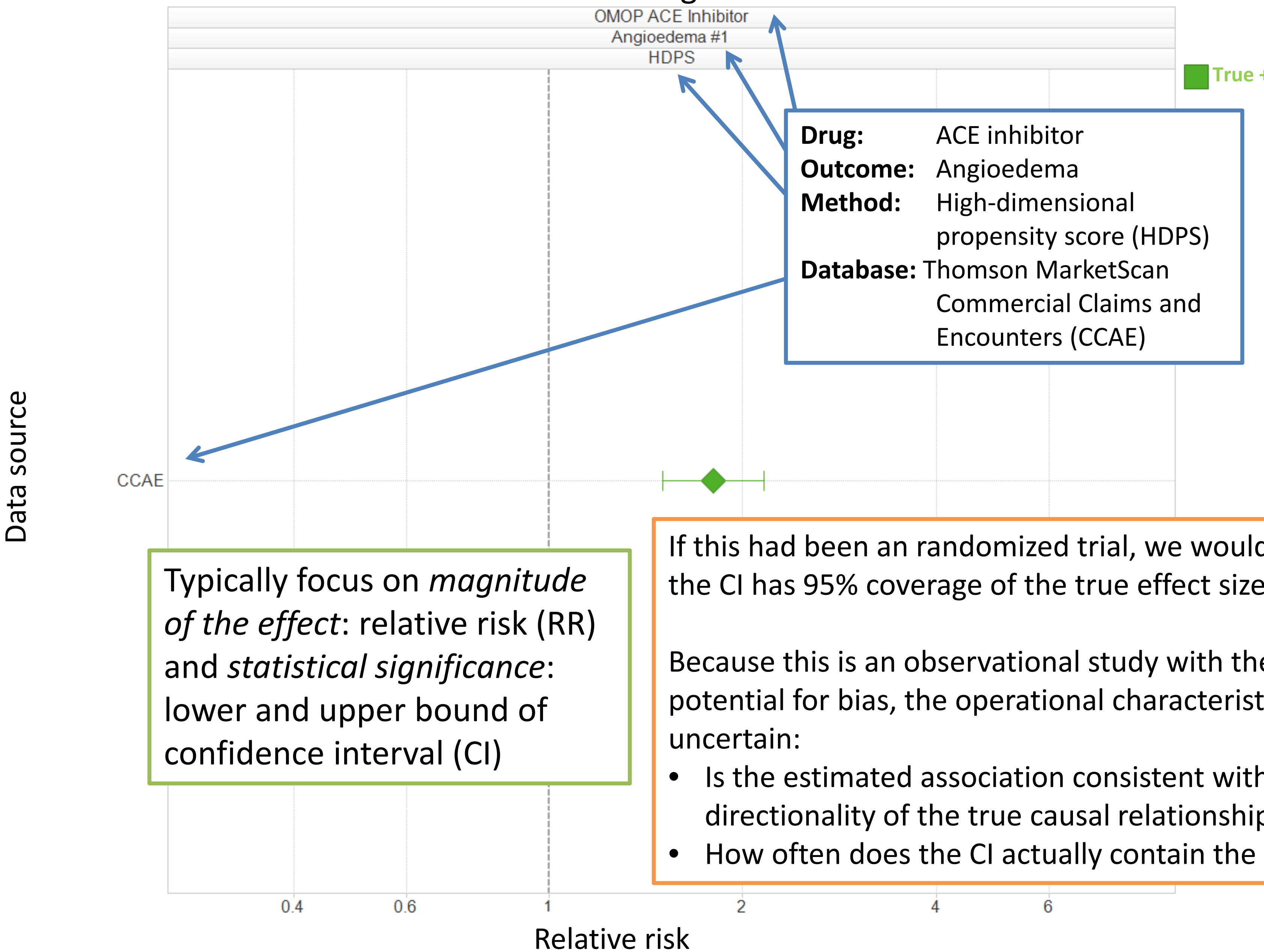




- 14 methods
- Epidemiology designs
- Statistical approaches adapted for longitudinal data



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



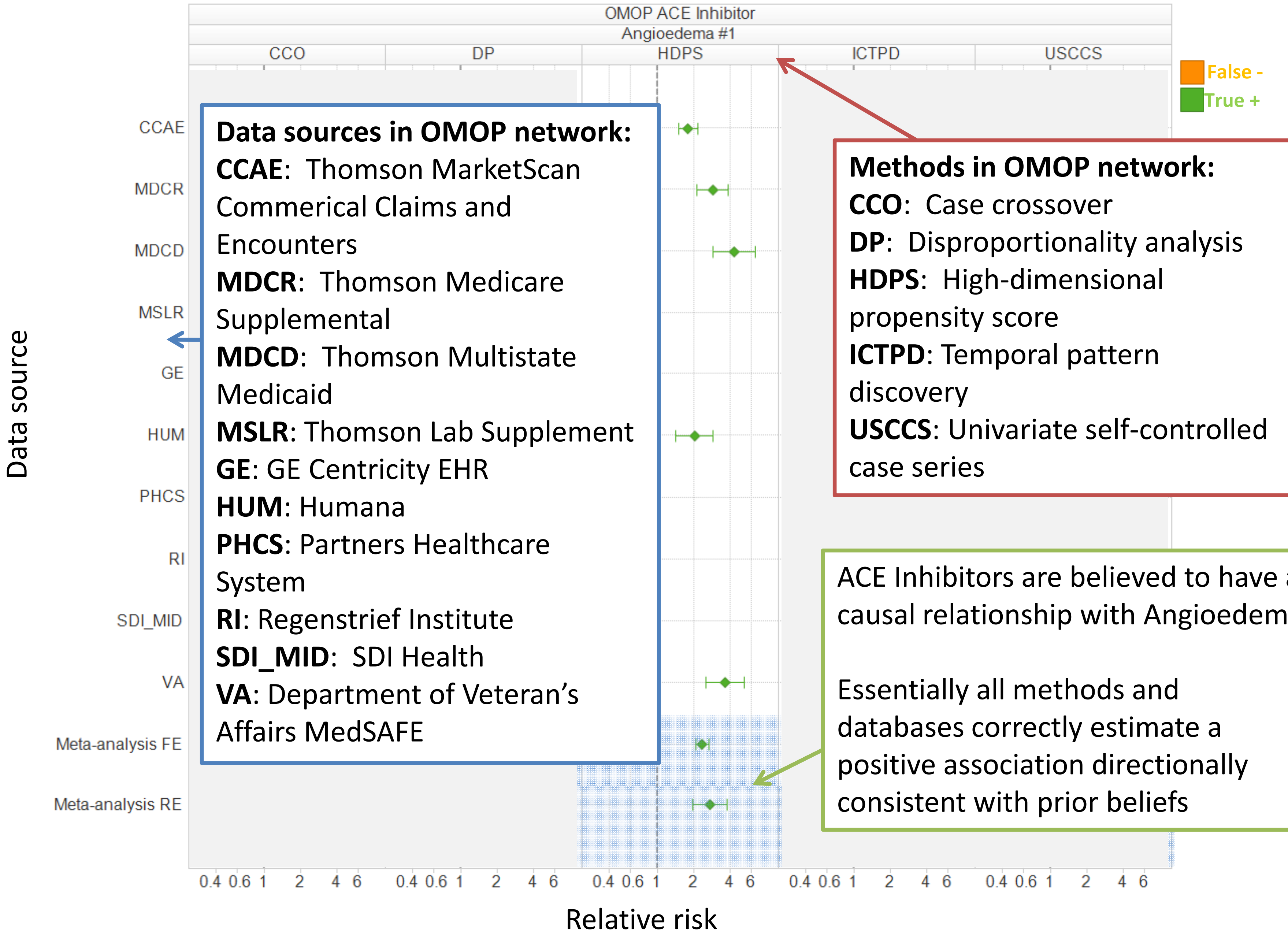
Typically focus on *magnitude of the effect*: relative risk (RR) and *statistical significance*: lower and upper bound of confidence interval (CI)

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



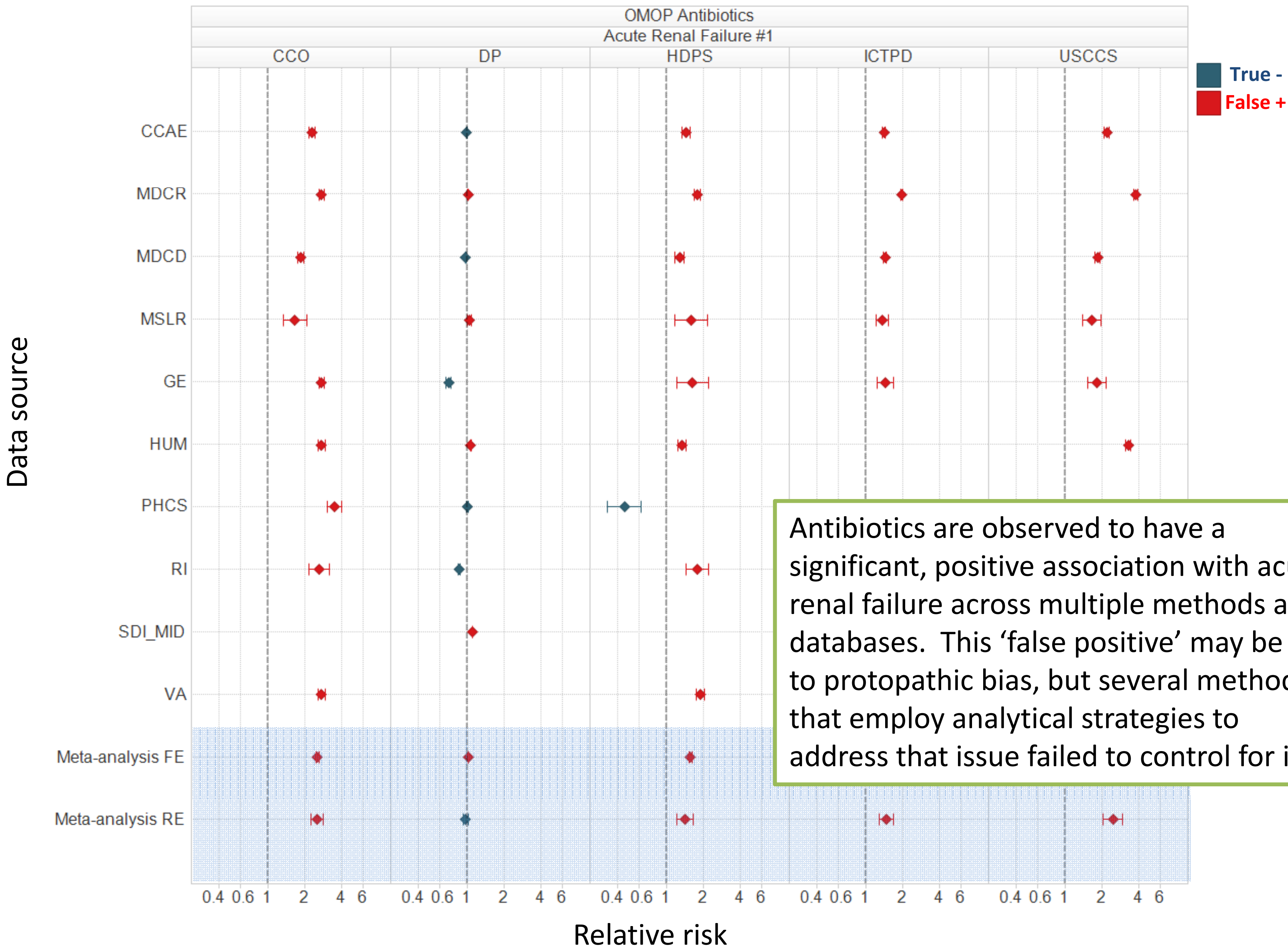
Data sources in OMOP network:
CCAE: Thomson MarketScan Commerical Claims and Encounters
MDCR: Thomson Medicare Supplemental
MDCD: Thomson Multistate Medicaid
MSLR: Thomson Lab Supplement
GE: GE Centricity EHR
HUM: Humana
PHCS: Partners Healthcare System
RI: Regenstrief Institute
SDI_MID: SDI Health
VA: 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?

‘Ground truth’ assumed for Monitoring Health Outcomes of Interest

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

Legend	Total
True positive' benefit	2
True positive' risk	9
Negative control'	44

http://omop.fnih.org/OMOPWhitePapers

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

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

Method prediction: Drug-condition pair met a specific threshold:

(LB 95% CI > 1)

	Y	N
Y	True positives: 5	False positives: 8
N	False negatives: 4	True negatives: 36

Positive predictive value
= precision
= TP / (TP+FP)
= 5 / (5+8) = 0.38

Negative predictive value
= TN / (FN+TN)
= 36 / (4+36) = 0.90

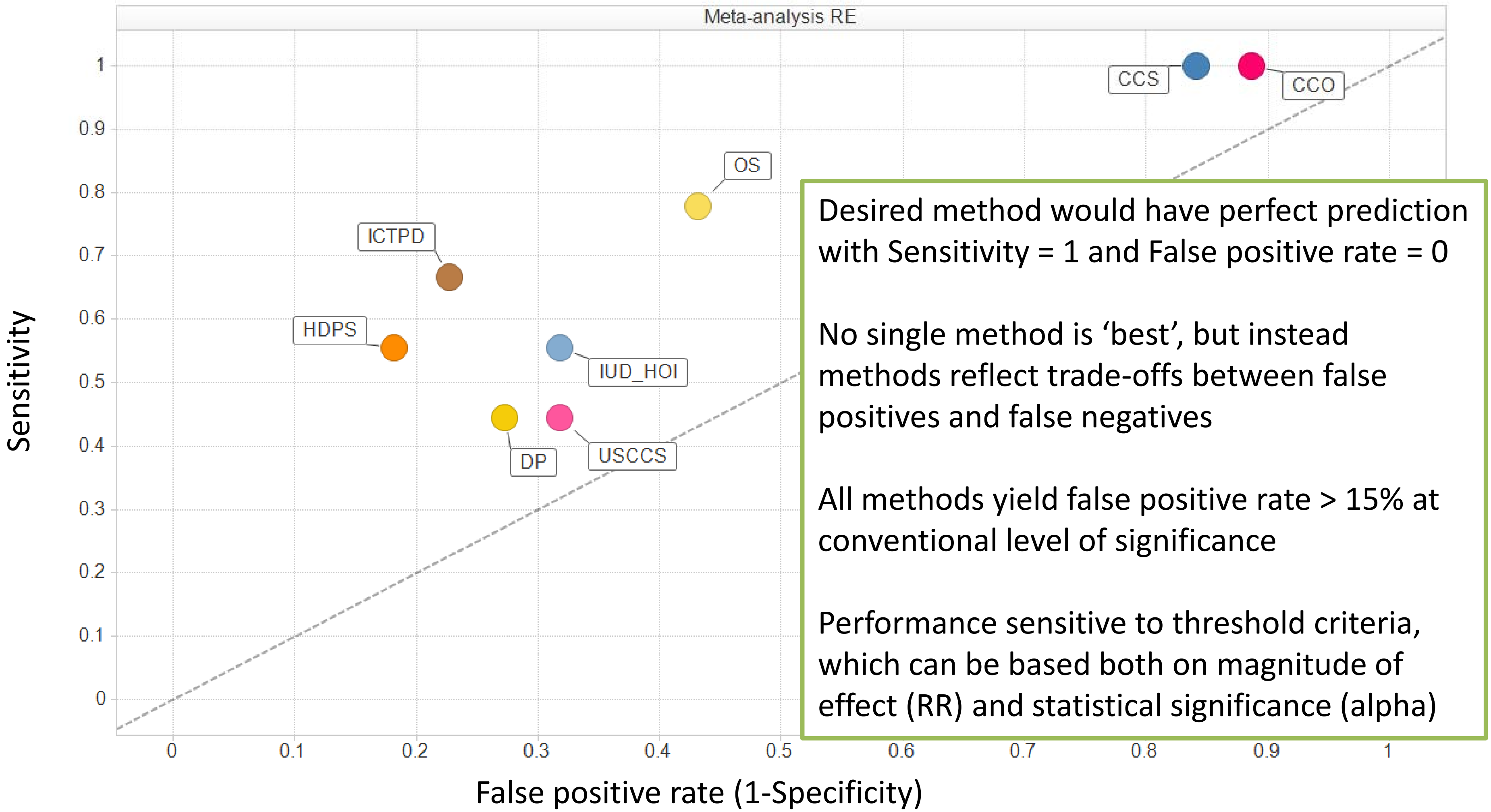
Sensitivity
= Recall
= TP / (TP+FN)
= 5 / (5+4) = 0.56

Specificity
= TN / (FP+TN)
= 36 /(8+36) = 0.82

False positive rate
= 1 – 0.82 = 0.18

Accuracy
= (TP+TN) /
(TP+TN+FP+FN)
=(5+36)/(9+44) = 0.77

Comparing methods by sensitivity and specificity at alpha=0.05



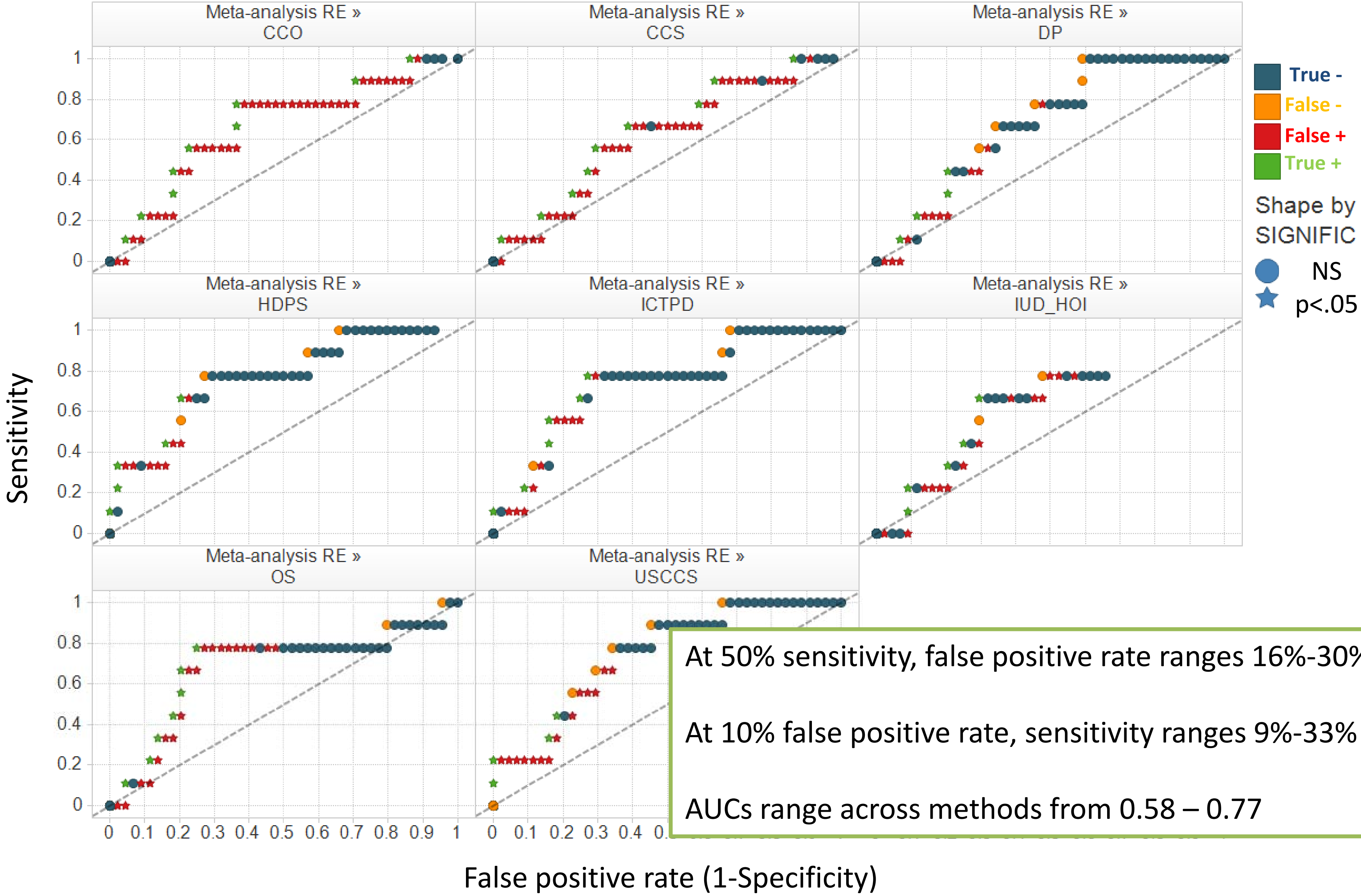
Desired method would have perfect prediction with Sensitivity = 1 and False positive rate = 0

No single method is 'best', but instead methods reflect trade-offs between false positives and false negatives

All methods yield false positive rate > 15% at conventional level of significance

Performance sensitive to threshold criteria, which can be based both on magnitude of effect (RR) and statistical significance (alpha)

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%

AUCs range across methods from 0.58 – 0.77

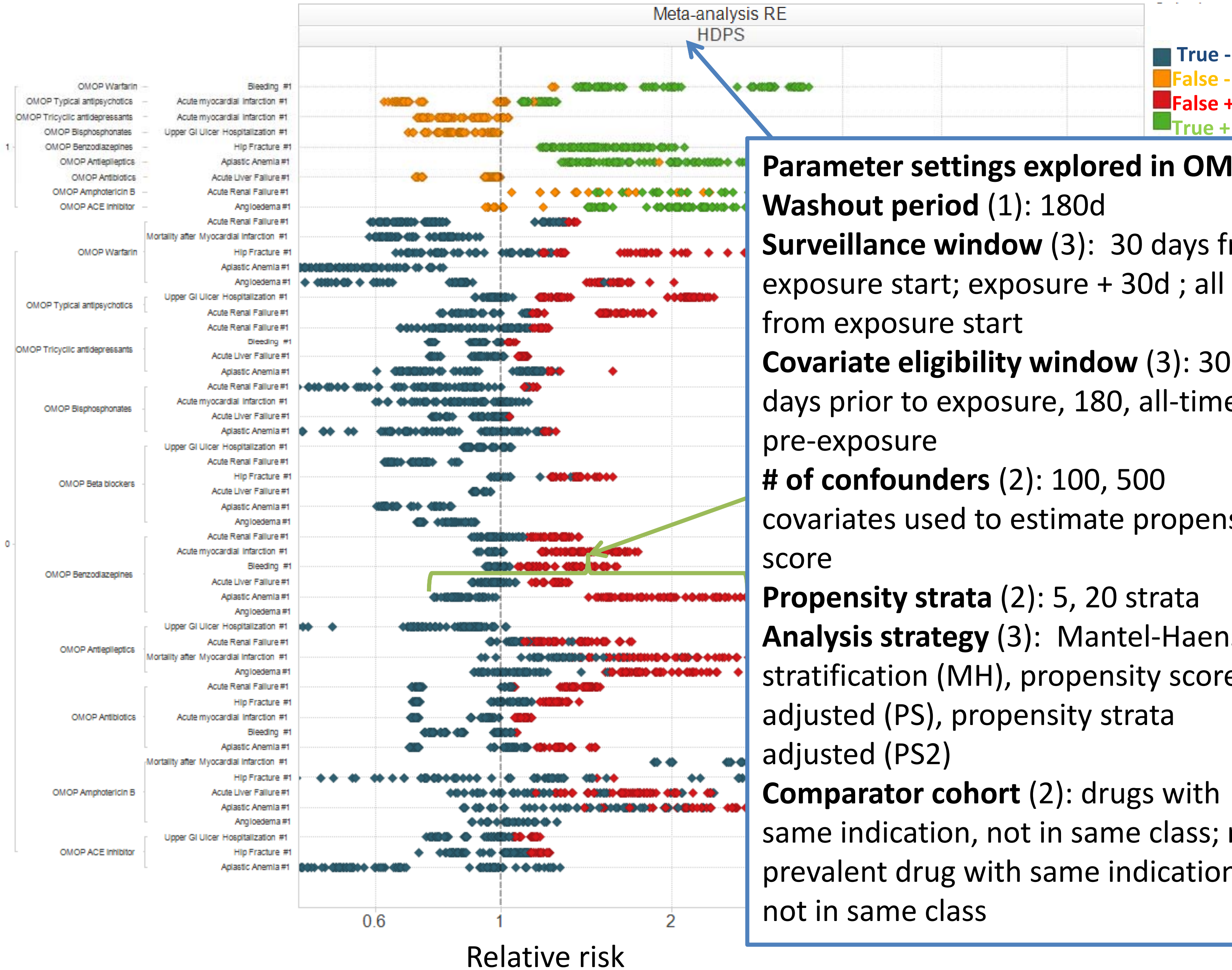
Risk identification methods under

Method name	Contributor	Release date
Disproportionality analysis		
Disproportionality analysis (DP)	Columbia / Merck	15-Mar-10
IC Temporal Pattern Discovery (ICTPD)	Uppsala Monitoring Centre	23-May-10
HSIU cohort method (HSIU)	Regenstrief / Indiana University	8-Jun-10
Case-based methods		
Univariate self-controlled case series (USCCS)	Columbia	2-Apr-10
Multi-set case control estimation (MSCCE)	Columbia / GlaxoSmithKline	16-Apr-10
Bayesian logistic regression (BLR)	Rutgers / Columbia	21-Apr-10
Case-control surveillance (CCS)	Lilly	2-May-10
Case-crossover (CCO)	University of Utah	1-Jun-10
Exposure-based methods		
Observational screening (OS)	ProSanos / GlaxoSmithKline	8-Apr-10
High-dimensional propensity score (HDPS)	Columbia	6-Aug-10
Incident user design (IUD-HOI)	University of North Carolina	26-Oct-10
Sequential testing methods		
Maximized Sequential Probability Ratio Test (MSPRT)	Harvard Pilgrim / Group Health	25-Jul-10
Conditional sequential sampling procedure (CSSP)	Harvard Pilgrim / Group Health	30-Aug-10

In what follows, we have chosen one parameter combination for each method that performs best for the meta-analysis estimates

http://omop.fnih.org/MethodsLibrary

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



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

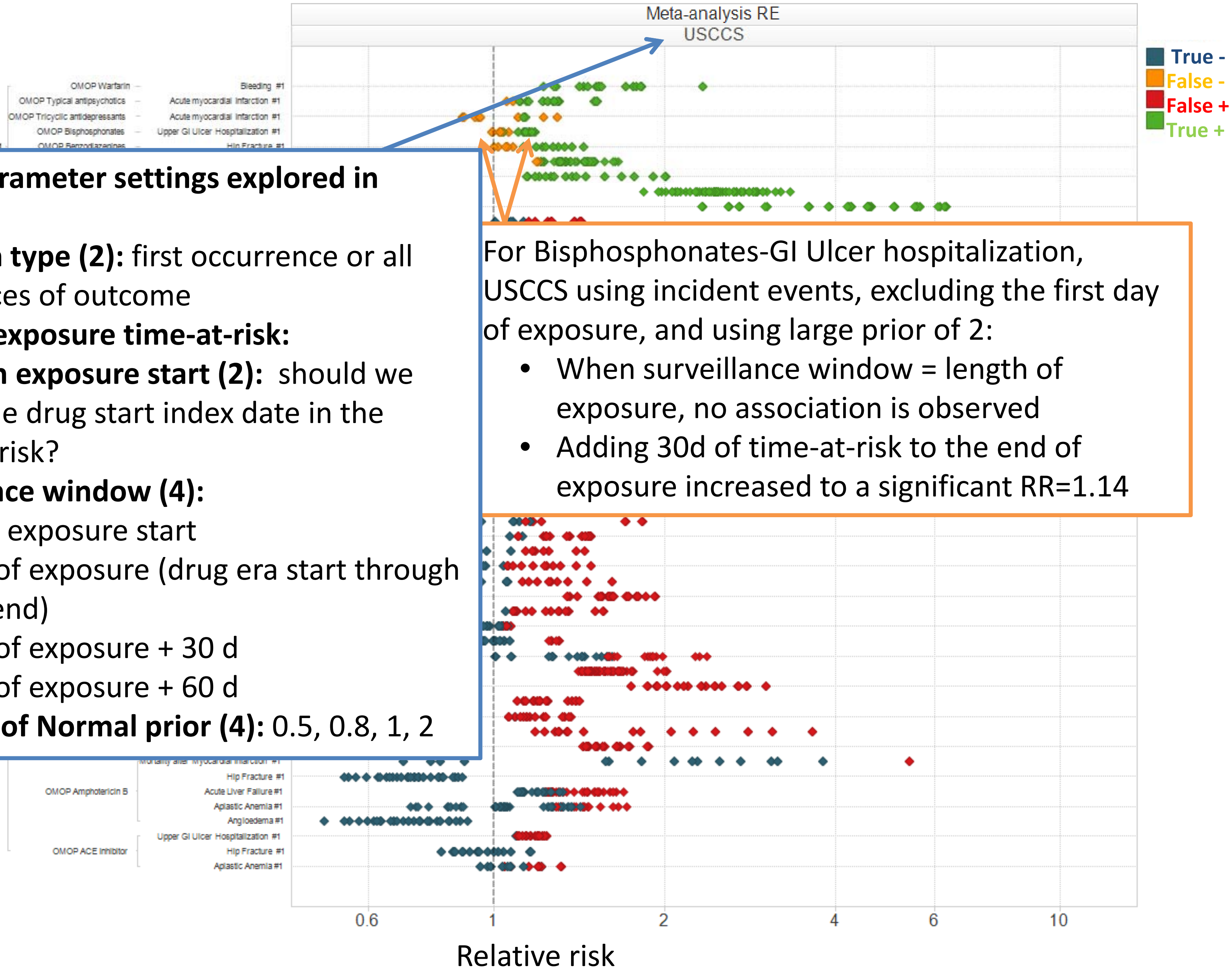
# of confounders (2): 100, 500 covariates used to estimate propensity score

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

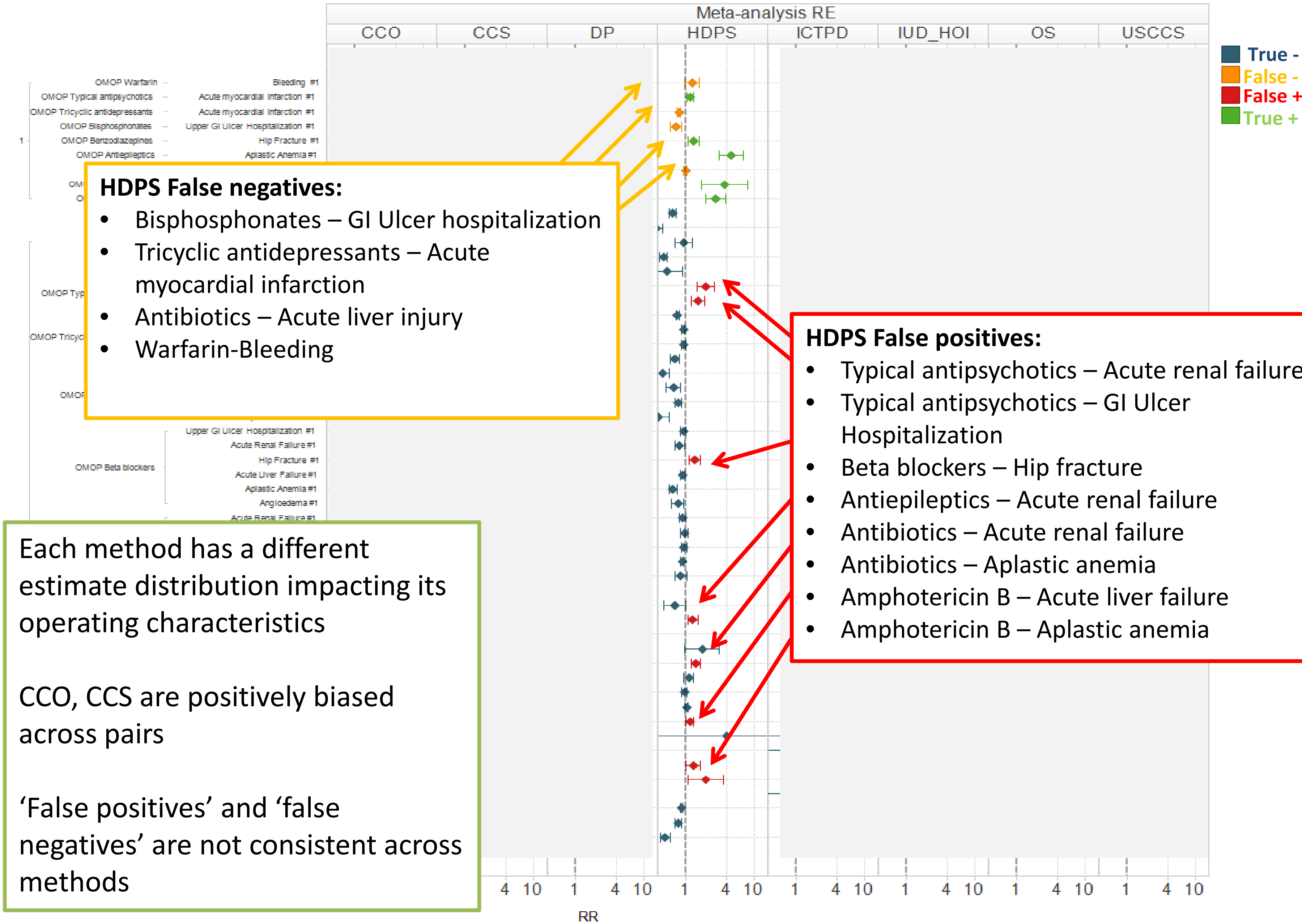


USCCS Parameter settings explored in OMOP:
Condition type (2): first occurrence or all occurrences of outcome
Defining exposure time-at-risk:
Days from exposure start (2): should we include the drug start index date in the period at risk?
Surveillance window (4):
30 d from exposure start
Duration of 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

Distribution of estimates across all drug-outcome pairs



HDPS False negatives:

- Bisphosphonates – GI Ulcer hospitalization
- Tricyclic antidepressants – Acute myocardial infarction
- Antibiotics – Acute liver injury
- Warfarin-Bleeding

HDPS False positives:

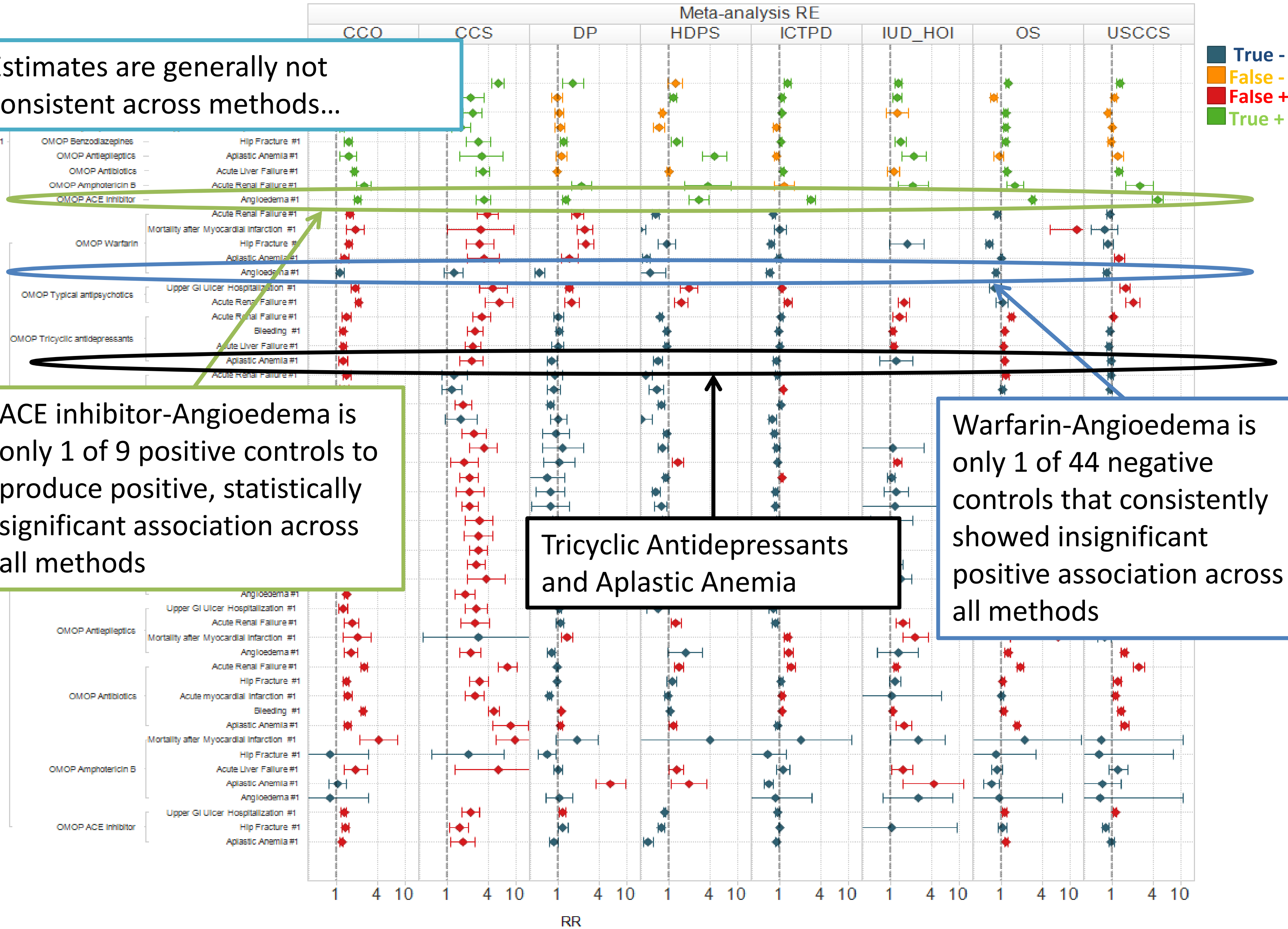
- Typical antipsychotics – Acute renal failure
- Typical antipsychotics – GI Ulcer Hospitalization
- Beta blockers – Hip fracture
- Antiepileptics – Acute renal failure
- Antibiotics – Acute renal failure
- Antibiotics – Aplastic anemia
- Amphotericin B – Acute liver failure
- Amphotericin B – Aplastic anemia

Each method has a different estimate distribution impacting its operating characteristics

CCO, CCS are positively biased across pairs

'False positives' and 'false negatives' are not consistent across methods

Distribution of estimates across all drug-outcome pairs



Estimates are generally not consistent across methods…

ACE inhibitor-Angioedema is only 1 of 9 positive controls to produce positive, statistically significant association across all methods

Tricyclic Antidepressants and Aplastic Anemia

Warfarin-Angioedema is only 1 of 44 negative controls that consistently showed insignificant positive association across all methods

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.

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

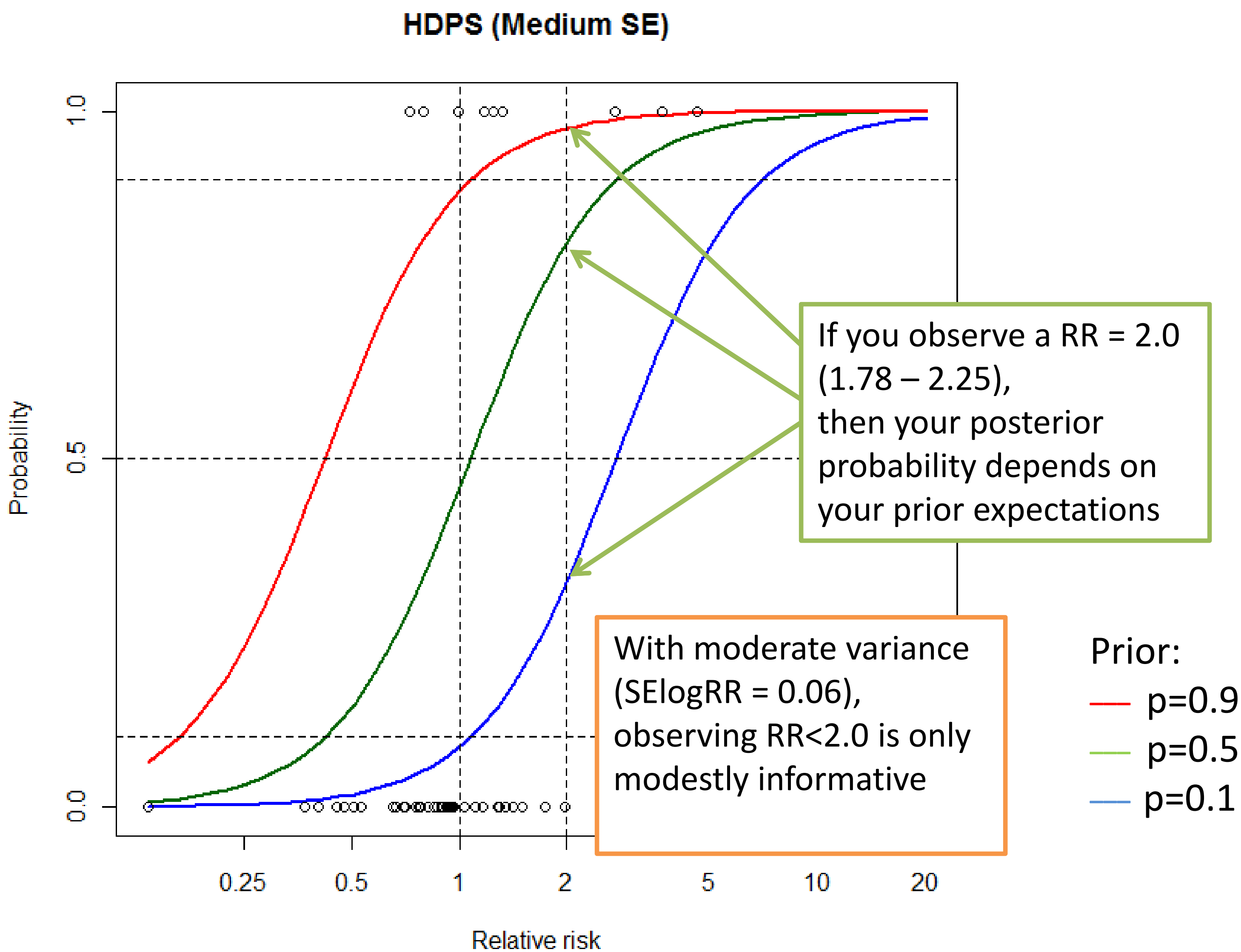
Conclusion the use of oral bisphosphonates was not significantly associated with incident esophageal or gastric cancer.

“Data”:

Effect estimates from one method against one database across an array of drug-outcome pairs

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 antidepressa		0.63 (0.07)	1.02 (0.02)	0.96 (0.01)	0.80 (0.04)			0.82 (0.06)	
OMOP Typical antipsychotic					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 a risk identification system



Conclusion

- Reliance on clinical judgment is scary
- Massive observational data can help
- Nontrivial challenges remain