

# OBSERVATIONAL MEDICAL OUTCOMES PARTNERSHIP

## **The Observational Medical Outcomes Partnership: Results and Future Perspectives**

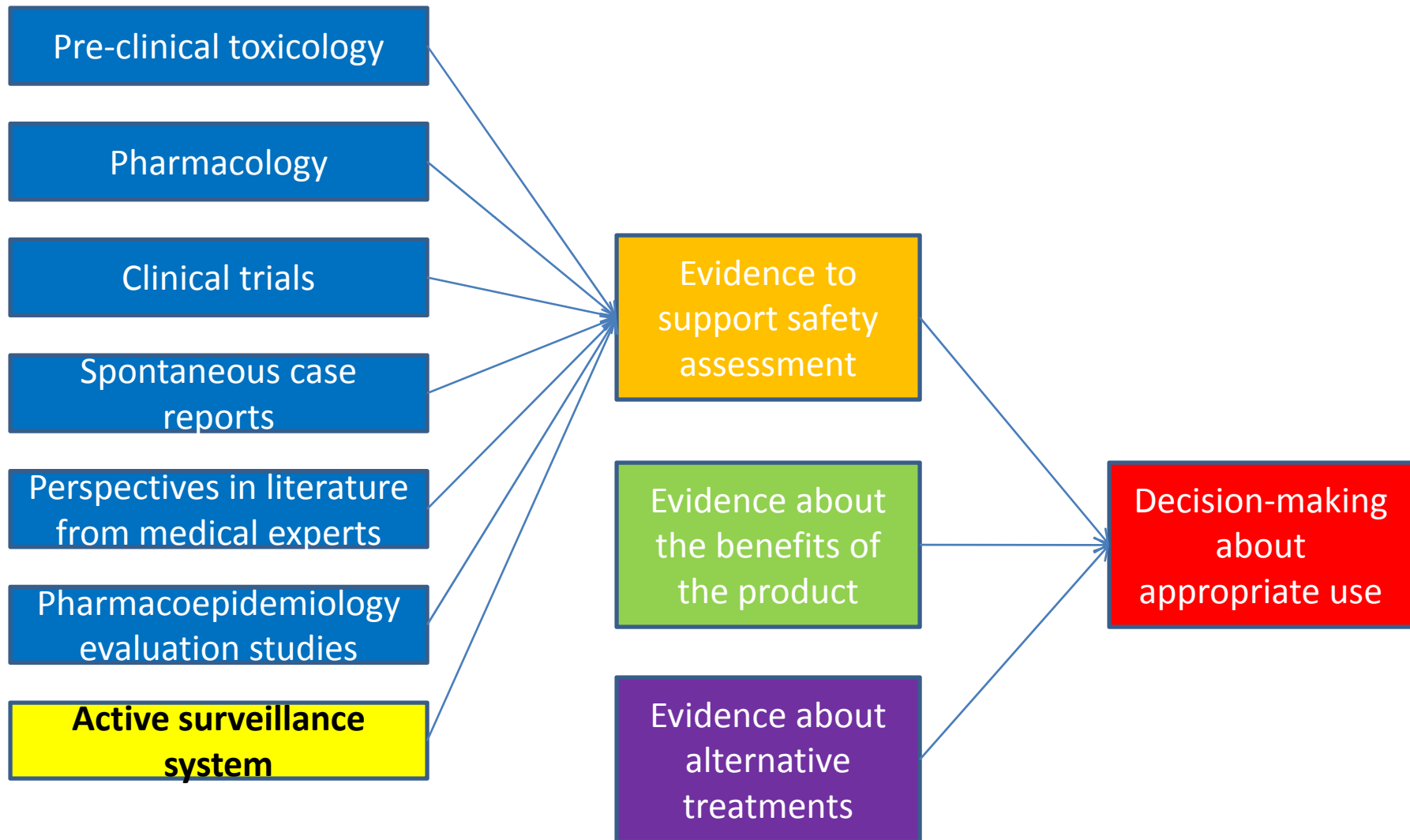
Patrick Ryan  
on behalf of OMOP Research Team  
April 9, 2011

Full results and audio presentations from OMOP Symposium available at:  
<http://omop.fnih.org/OMOP2011Symposium>

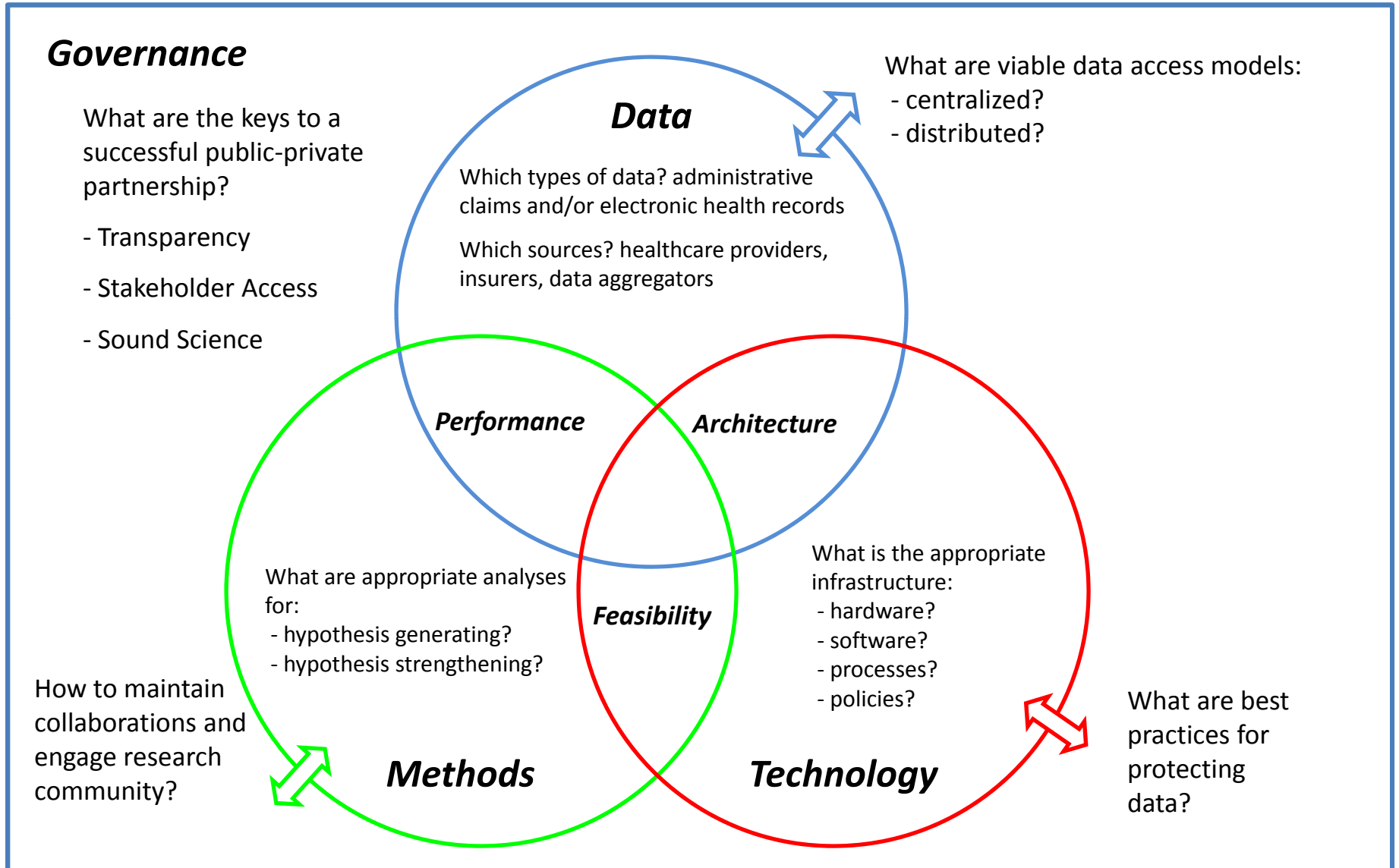
## Disclosures

- Since I began working with OMOP:
  - I was an employee of GSK and am currently an employee of J&J
  - I studied at UNC and have participated in funded/unfunded projects with >10 academic centers
  - I have actively collaborated with regulatory agencies, scientific industry and health systems, in US and internationally
- BIAS ALERT: I am enthusiastic about the opportunities for systematic observational analyses, and confident that we can develop a reliable active surveillance system if we all work together...
- ....so the empirical findings are from OMOP, but any editorials I make can be considered my own

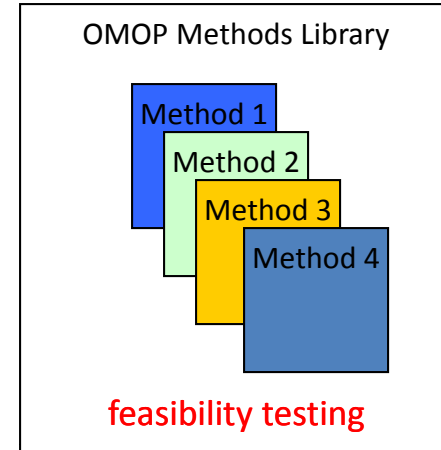
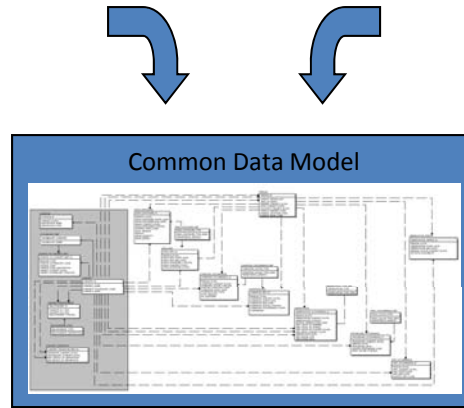
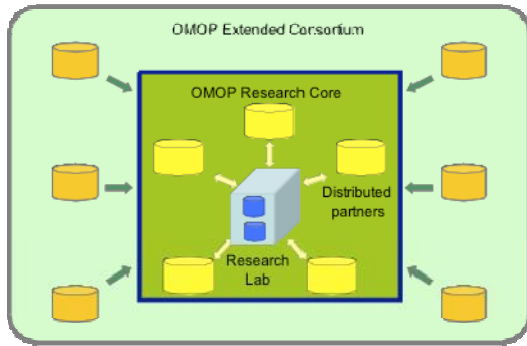
# Active surveillance: One additional piece of evidence to inform medical decision-making



# Outstanding Questions For Active Surveillance



# OMOP research experiment workflow



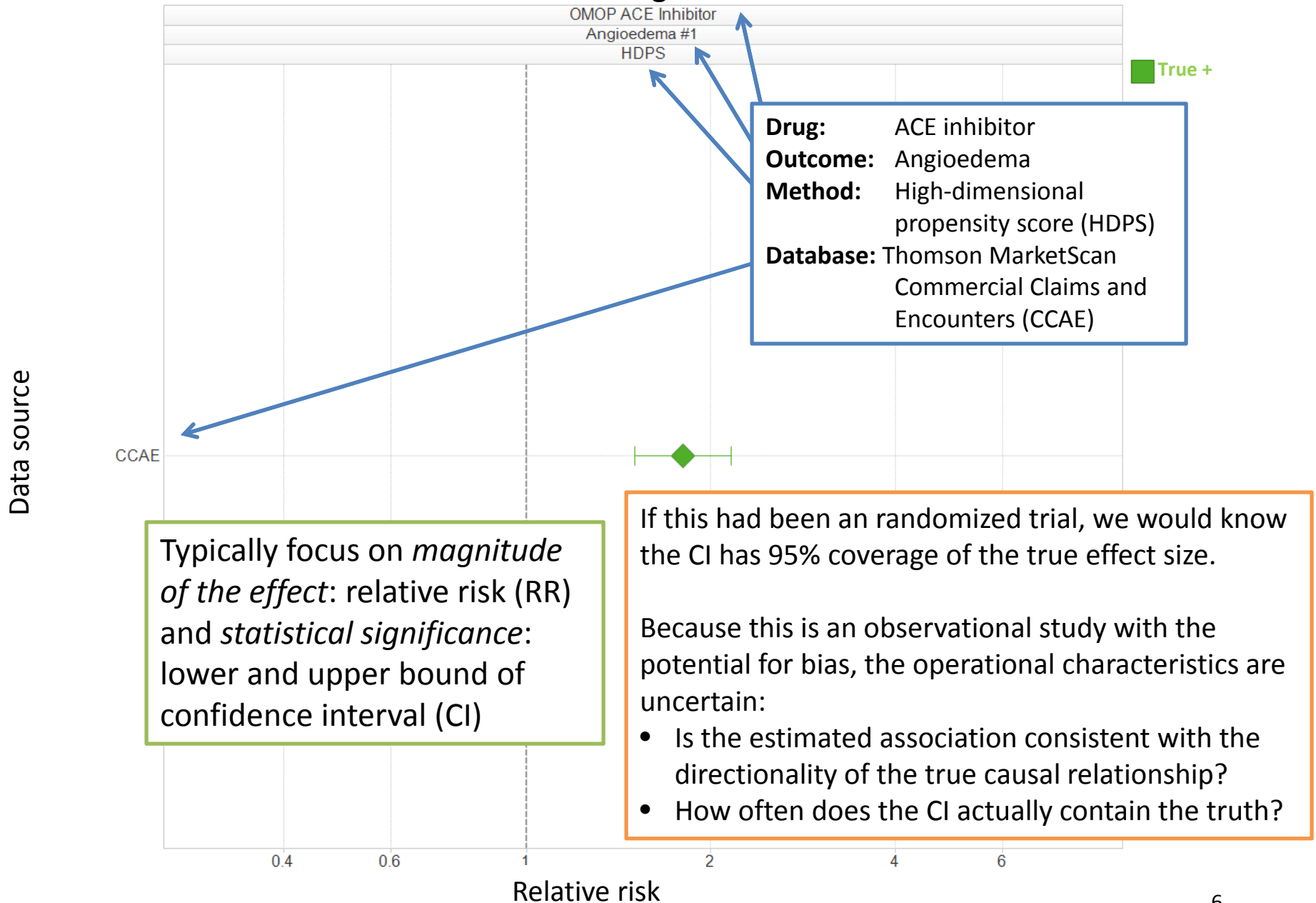
Drug

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

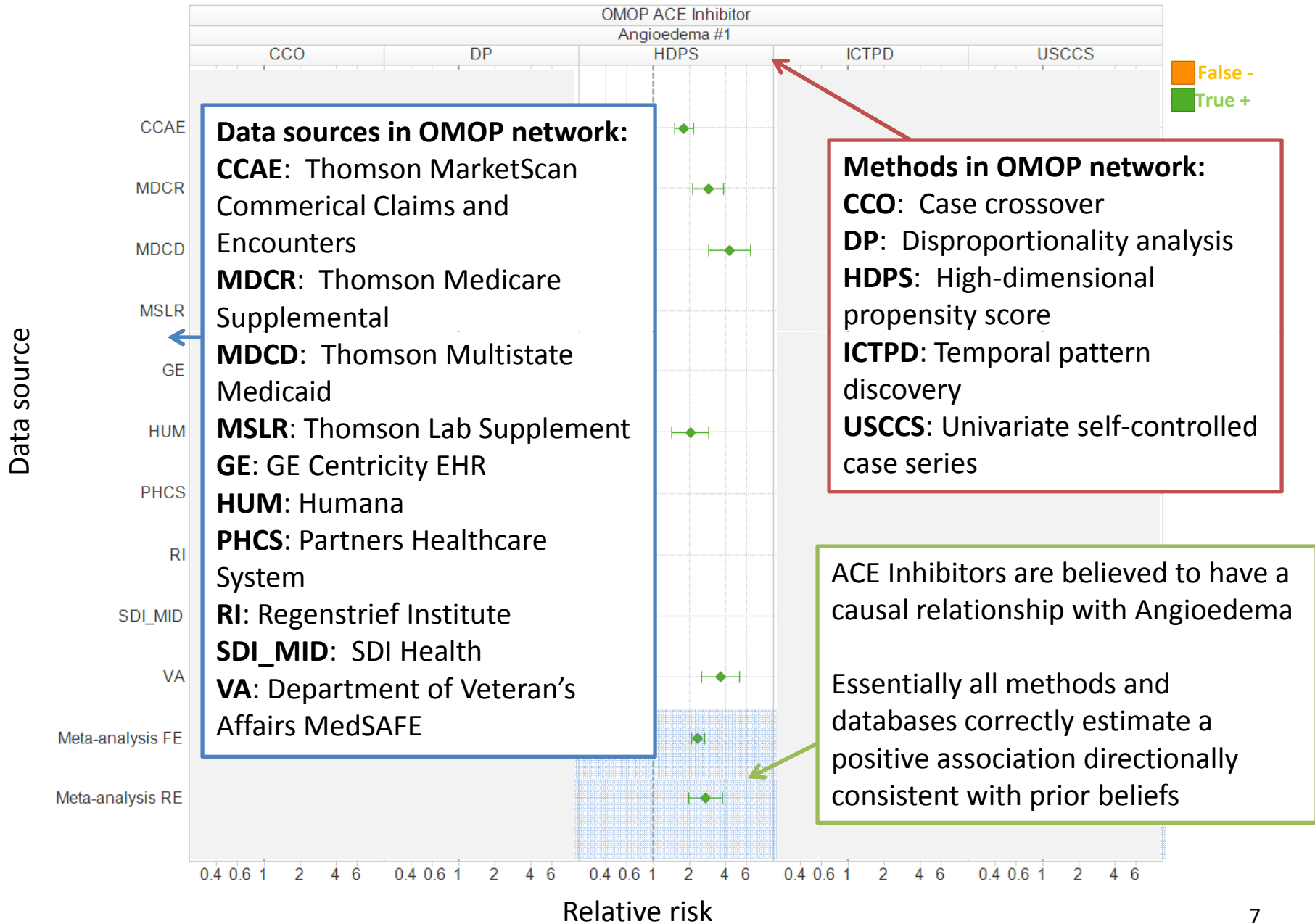
Legend

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

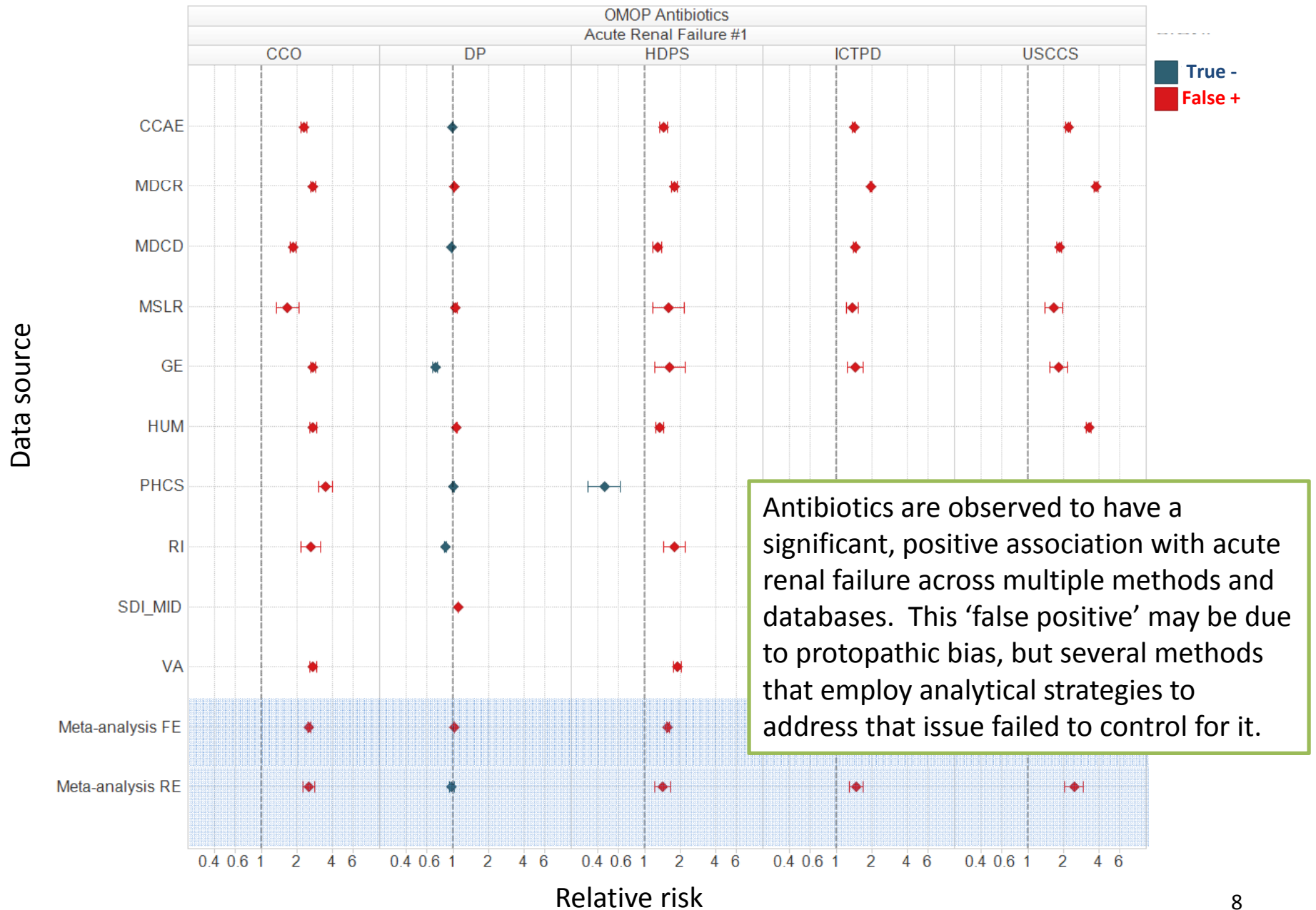
## Typical scenario: Estimate the effect of one drug on one outcome using one method against one database



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



## Consistent 'false positive' observed for 'negative control' of Antibiotics and Acute Renal Failure



# Measuring method performance

Drug-condition association status

Y – ‘true association’,

N – ‘negative control’

Y

N

Method prediction:  
Drug-condition  
pair met a  
specific  
threshold

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	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			
Bleeding			Negative control				Negative control			True positive risk
Hip Fracture	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

# 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’

Y

N

Method prediction:  
Drug-condition pair met a specific threshold:  
(LB 95% CI > 1)

Y

N

True positives: 5	False positives: 8
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  
=  $(TP+TN) / (TP+TN+FP+FN)$   
=  $(5+36) / (9+44) = 0.77$

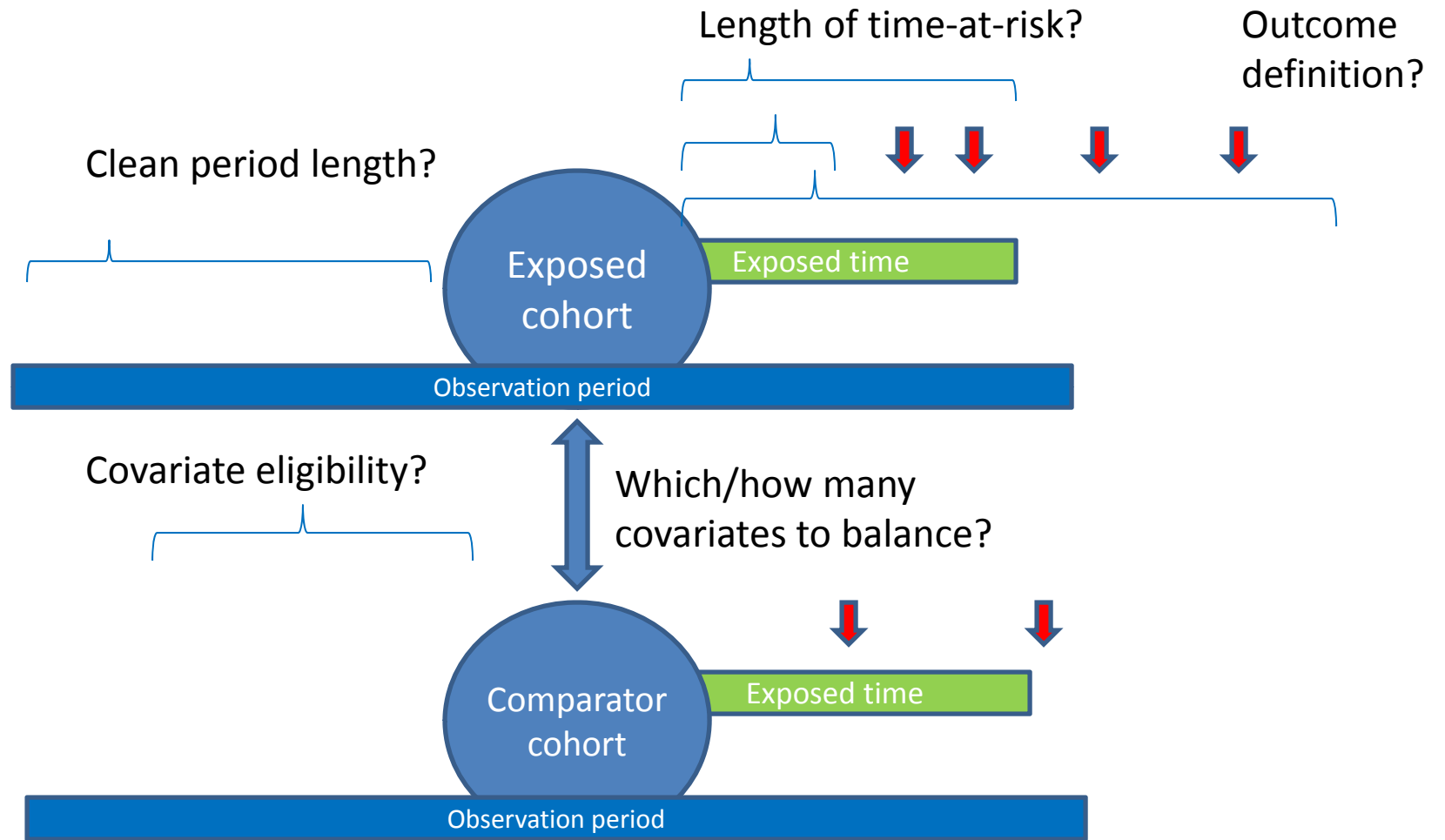
# Active surveillance methods under evaluation in OMOP experiment

Method name	Contributor	Release date
<b>Disproportionality analysis</b>		
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
<b>Case-based methods</b>		
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
<b>Exposure-based methods</b>		
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
<b>Sequential testing methods</b>		
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>

# Exploration of test cases within inception cohort design

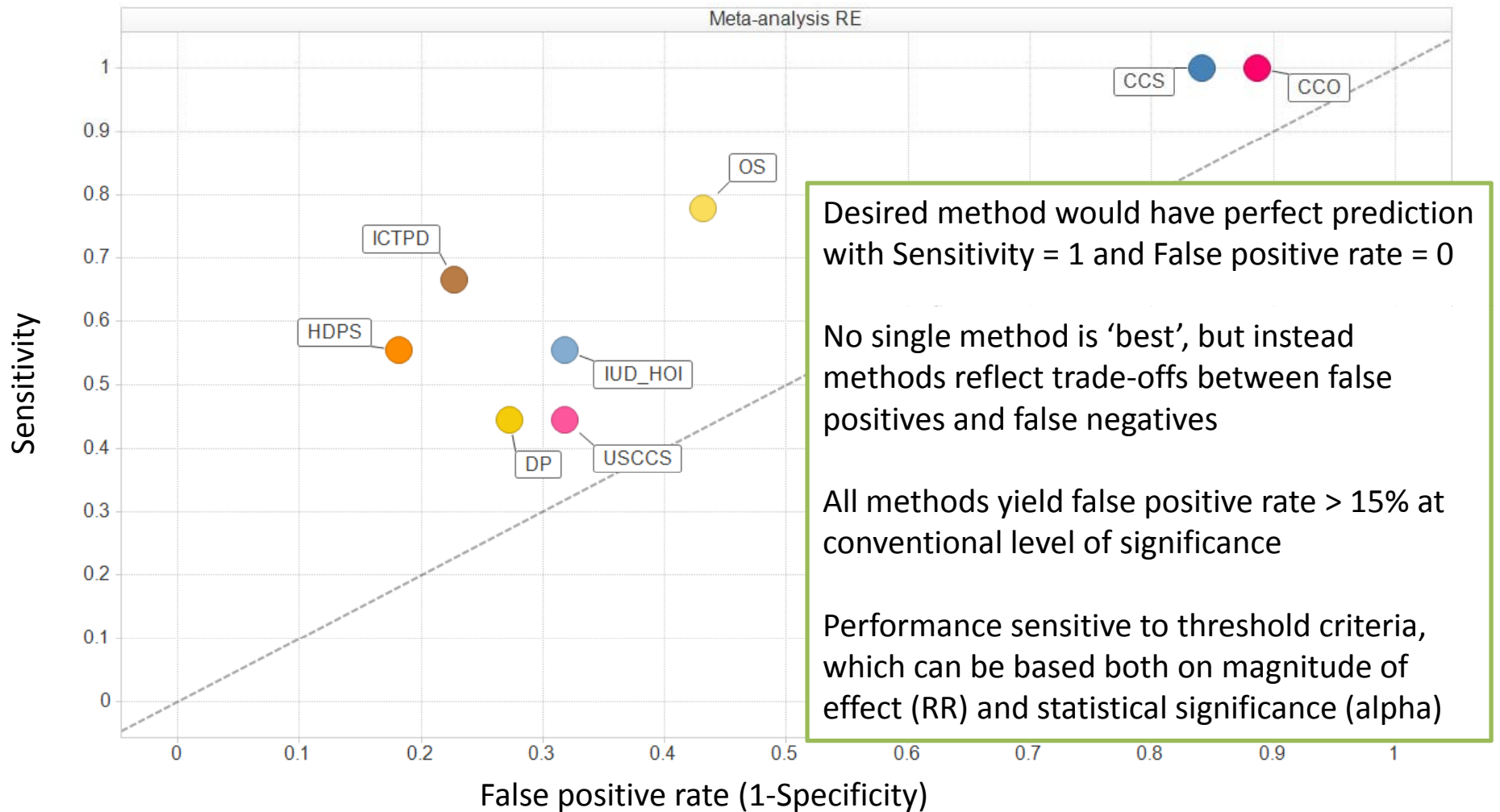


Exclusion criteria:  
Indications  
Contraindications

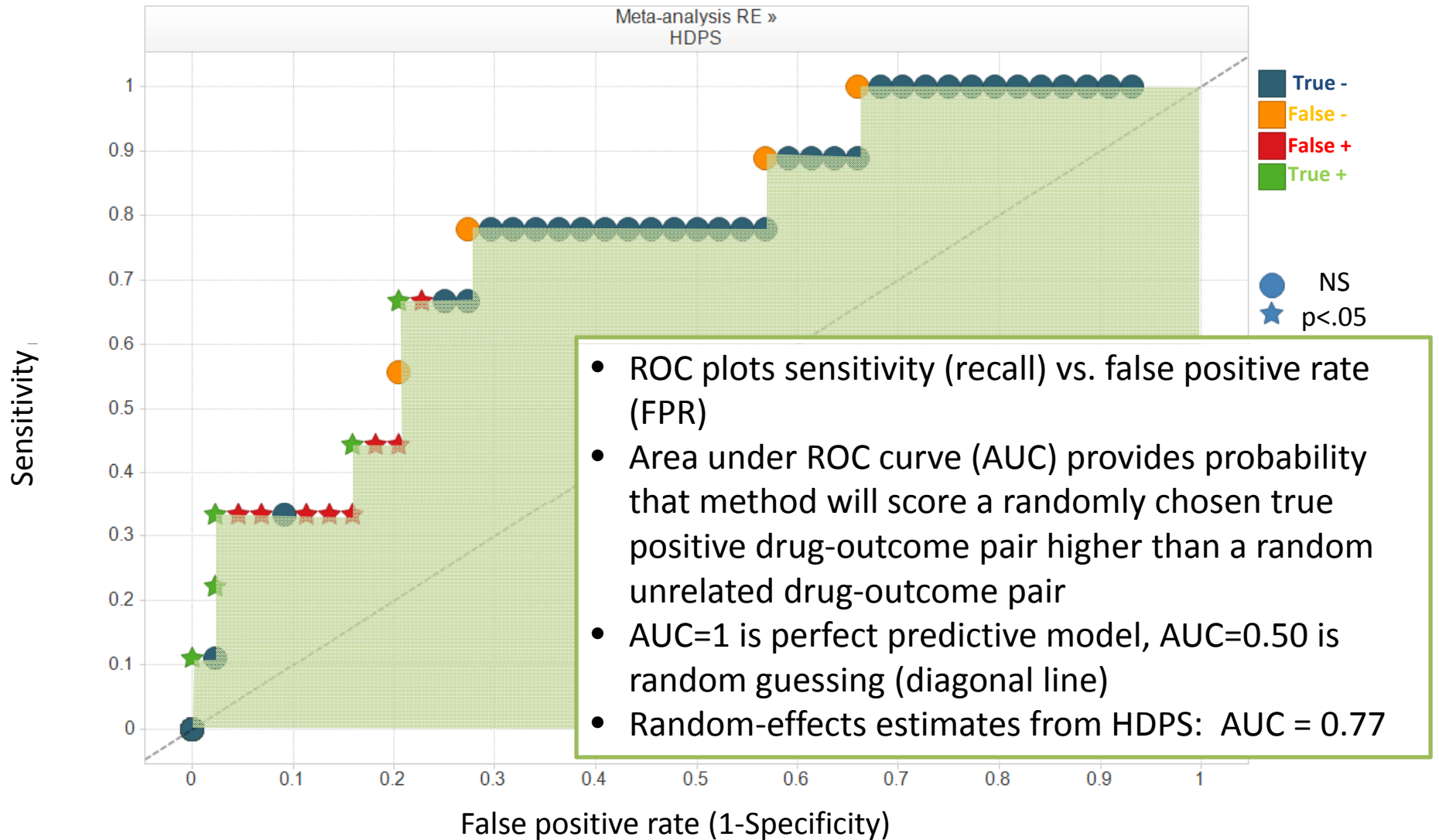
Which active  
comparator?

Propensity score adjustment strategy?  
Stratification (# of strata?)  
Multivariate adjustment

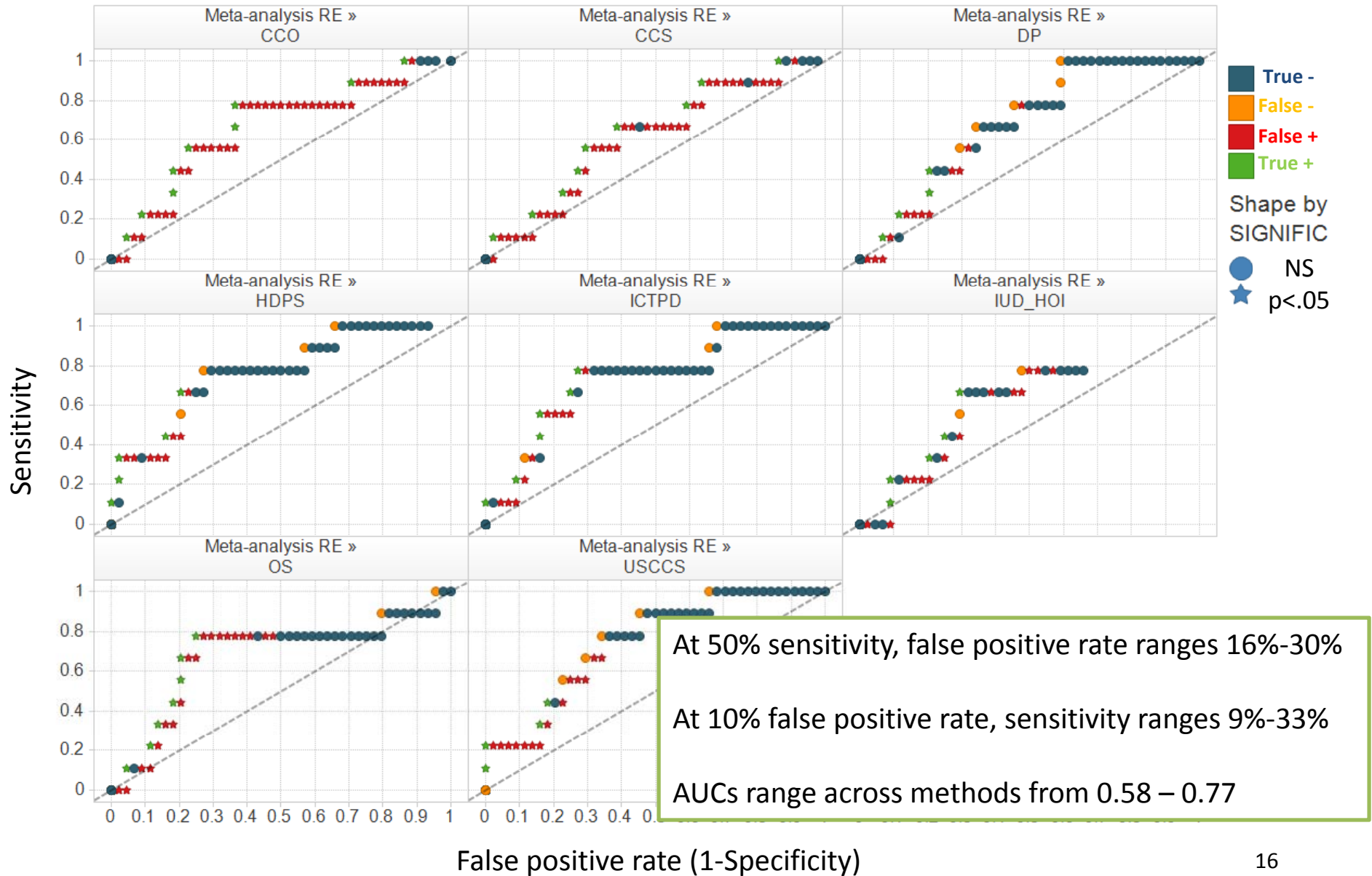
## Comparing methods by sensitivity and specificity at alpha=0.05



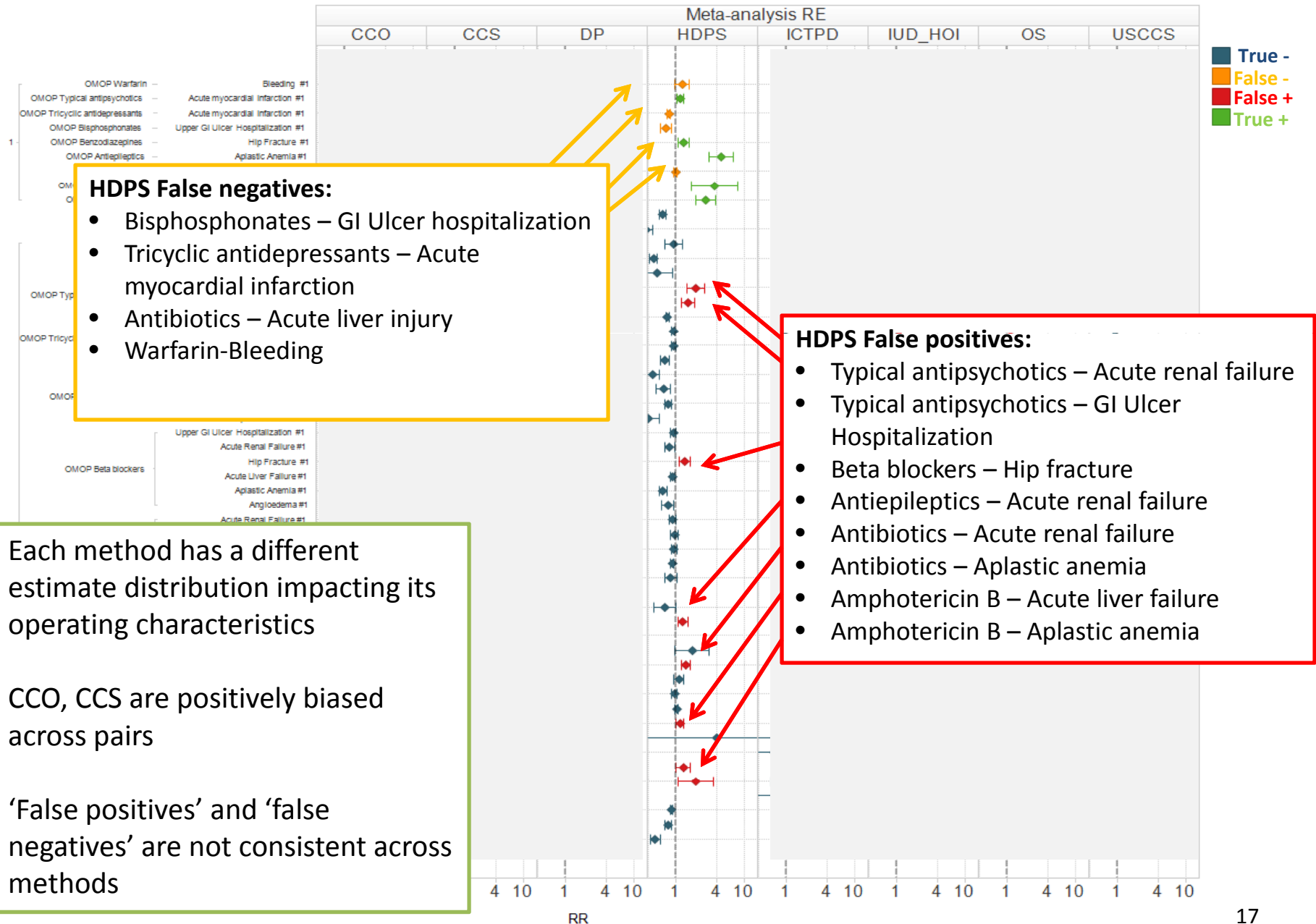
# Receiver Operating Characteristic (ROC) curve



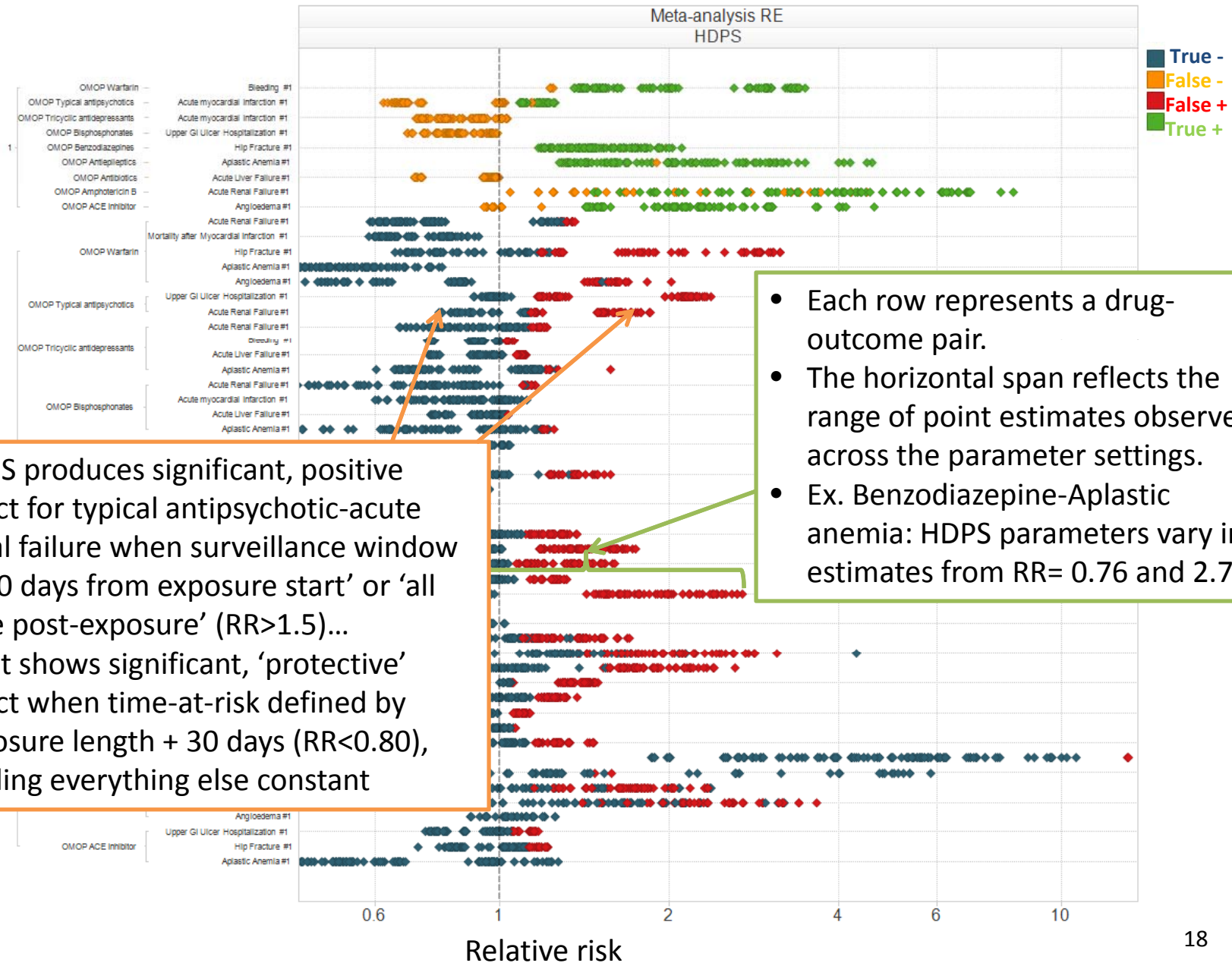
# ROC curves of random-effects meta-analysis estimations for all methods



# Distribution of estimates across all drug-outcome pairs



# Range of estimates across HDPS parameter settings



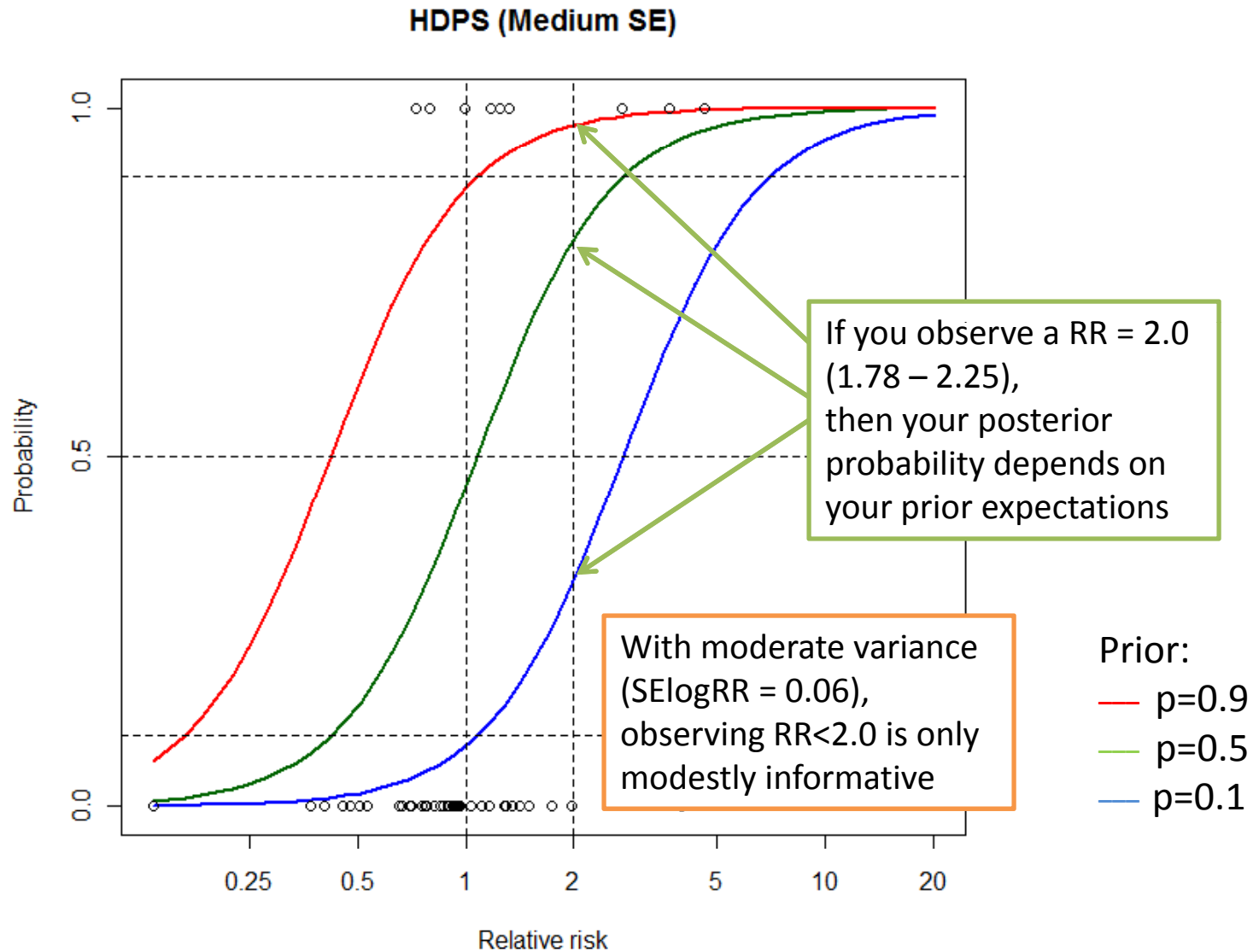
- Each row represents a drug-outcome pair.
- The horizontal span reflects the range of point estimates observed across the parameter settings.
- Ex. Benzodiazepine-Aplastic anemia: HDPS parameters vary in estimates from RR= 0.76 and 2.70

- HDPS produces significant, positive effect for typical antipsychotic-acute renal failure when surveillance window is '30 days from exposure start' or 'all time post-exposure' (RR>1.5)...
- ...but shows significant, 'protective' effect when time-at-risk defined by exposure length + 30 days (RR<0.80), holding everything else constant

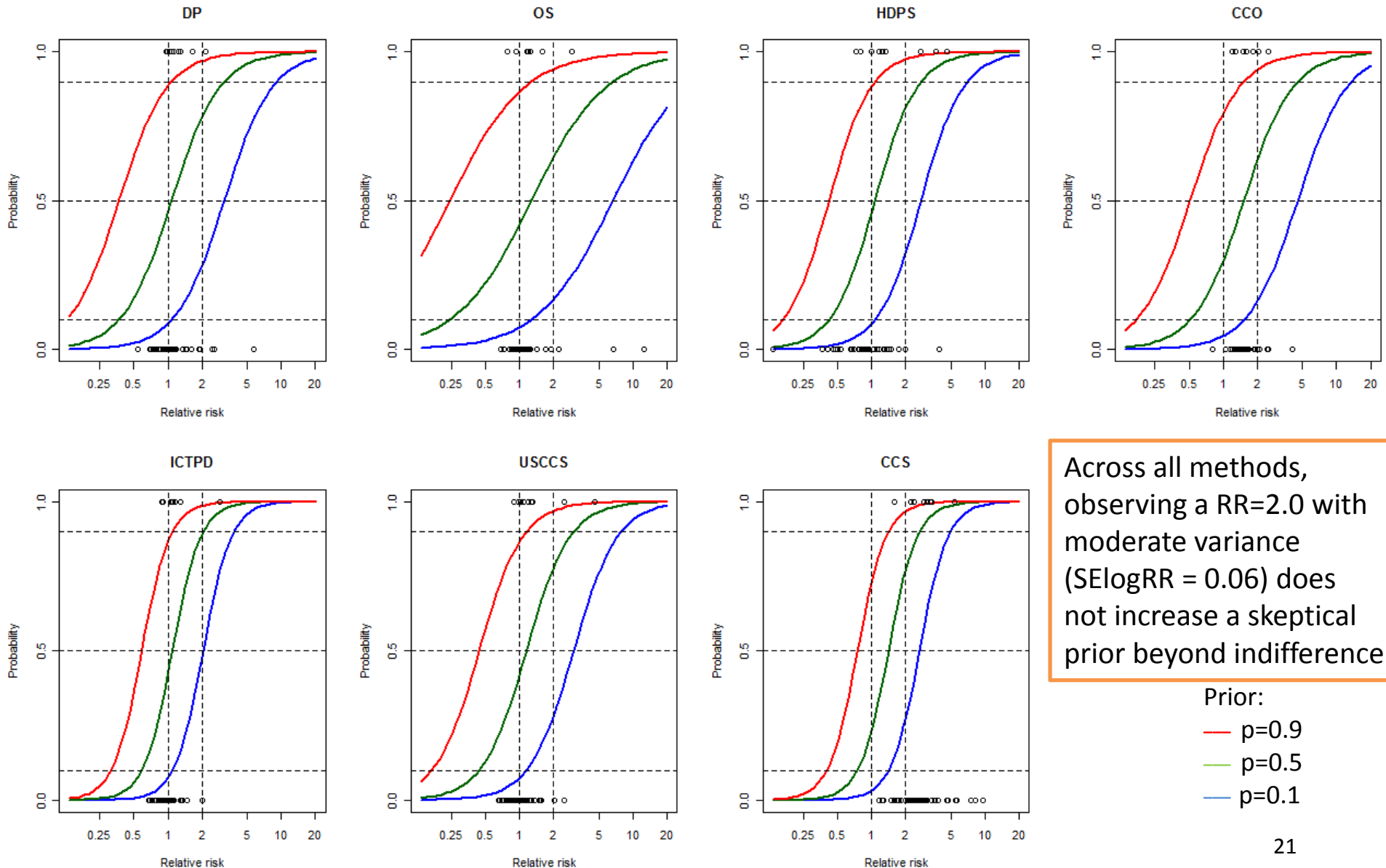
## So given these operating characteristics, what can we expect to do in practice....

- Use case: An emerging safety concern is raised for a new medical product. The association between the drug and outcome could be estimated by running an OMOP active surveillance method across the network of observational databases
  - The method will produce a relative risk and standard error from each participating data source, which can then be pooled together in a meta-analytic framework
  - Hypothetical scenario: The random-effects meta-analysis yields an RR=2.0 with SE=0.06.
  - Question: what is the probability that there is a true causal relationship given this observed association?
- Bayes rule enables such calculation...
  - $p(\text{true} | \text{RR}, \text{SE}) \sim p(\text{RR}, \text{SE} | \text{true}) * p(\text{true})$
  - $p(\text{true})$  is the prior probability of true association; consider a family of priors: skeptical (0.1), indifferent (0.5), enthusiastic (0.9)
  - $p(\text{RR}, \text{SE} | \text{true})$  can be estimated from empirical data (OMOP experimental results)

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



# Revising prior expectations in light of new evidence from an active surveillance system: Impact of using estimates from different methods



## Concluding thoughts

- An active surveillance system can complement current practice by providing evidence to support a comprehensive safety assessment
- No one clear 'best' method, as it depends on tolerance for false positives vs. false negatives
- In this experiment, active surveillance methods achieved:
  - At 50% sensitivity, false positive rate ranges 16%-30%
  - At 10% false positive rate, sensitivity ranges 9%-33%
- Need to be cautious in interpreting results from single method in single database
  - Replication does not necessarily provide complete confidence
- Further empirical research needed to have more complete understanding of operating characteristics before widespread adoption