OBSERVATIONAL MEDICAL OUTCOMES PARTNERSHIP

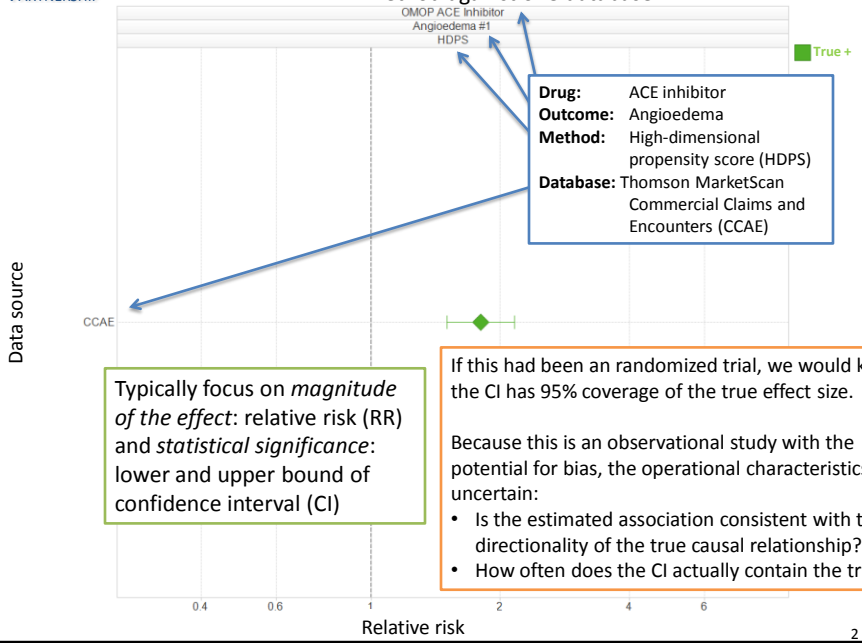
Establishing Operating Characteristics of Active Surveillance Approaches

Patrick Ryan
on behalf of OMOP Research Team
February 16, 2011

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

OBSERVATIONAL MEDICAL OUTCOMES PARTNERSHIP

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.

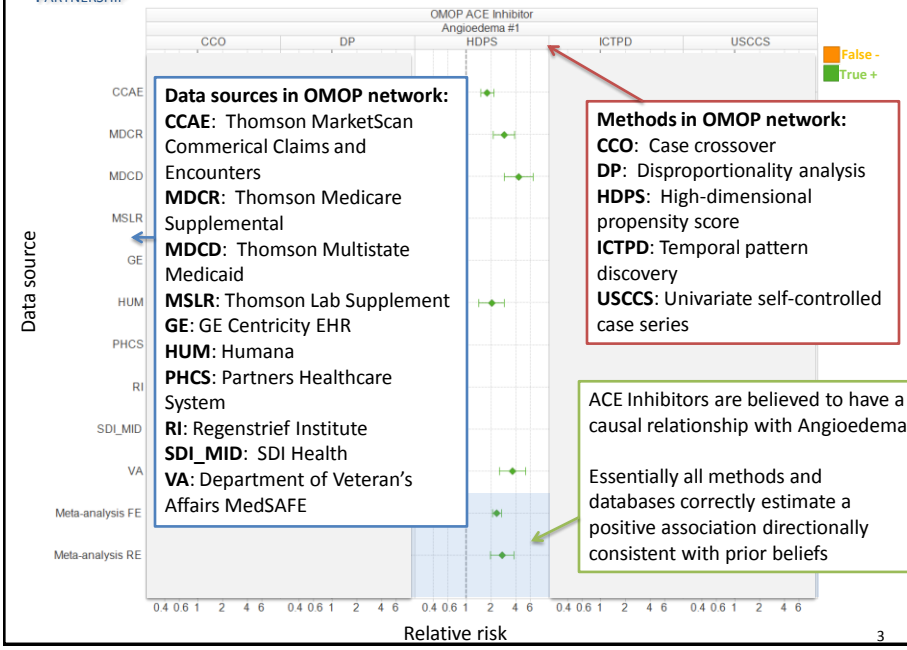
Because this is an observational study with the potential for bias, the operational characteristics are uncertain:

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

2

OBSERVATIONAL MEDICAL OUTCOMES PARTNERSHIP

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

3

OBSERVATIONAL MEDICAL OUTCOMES PARTNERSHIP

Measuring method performance

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

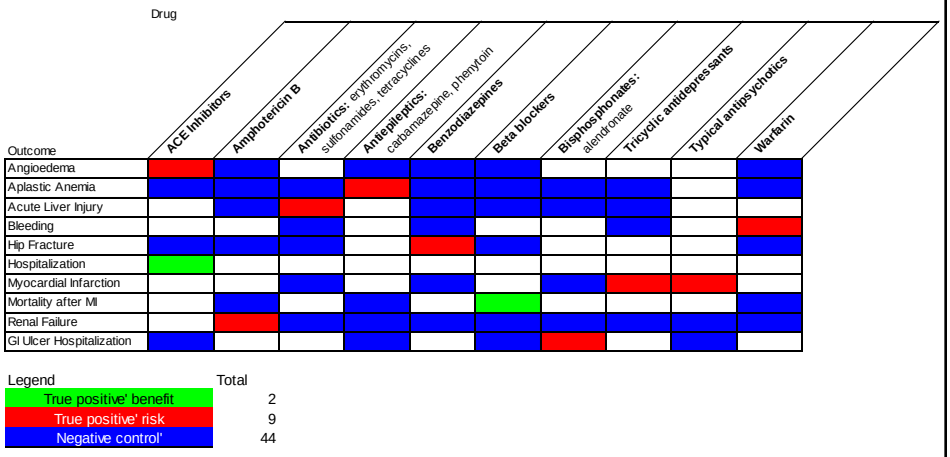
Method prediction: Drug-condition pair met a specific threshold	Y	N
Y	**True positives**	**False positives**
N	**False negatives**	**True negatives**

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

4

OBSERVATIONAL MEDICAL OUTCOMES PARTNERSHIP

'Ground truth' assumed for Monitoring Health Outcomes of Interest



http://omop.fnih.org/OMOPWhitePapers

5

OBSERVATIONAL MEDICAL OUTCOMES PARTNERSHIP

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

6

OBSERVATIONAL MEDICAL OUTCOMES PARTNERSHIP

Active surveillance methods under evaluation in OMOP experiment

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)	Harvard Medical School / 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

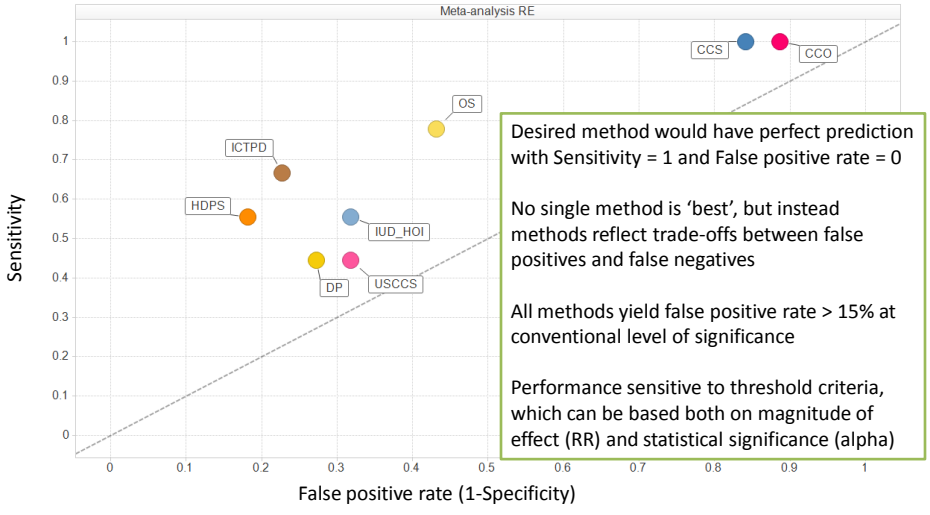
In what follows, we have chosen one parameter combination for each method that performs best for the meta-analysis estimates

http://omop.fnih.org/MethodsLibrary

7

OBSERVATIONAL MEDICAL OUTCOMES PARTNERSHIP

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

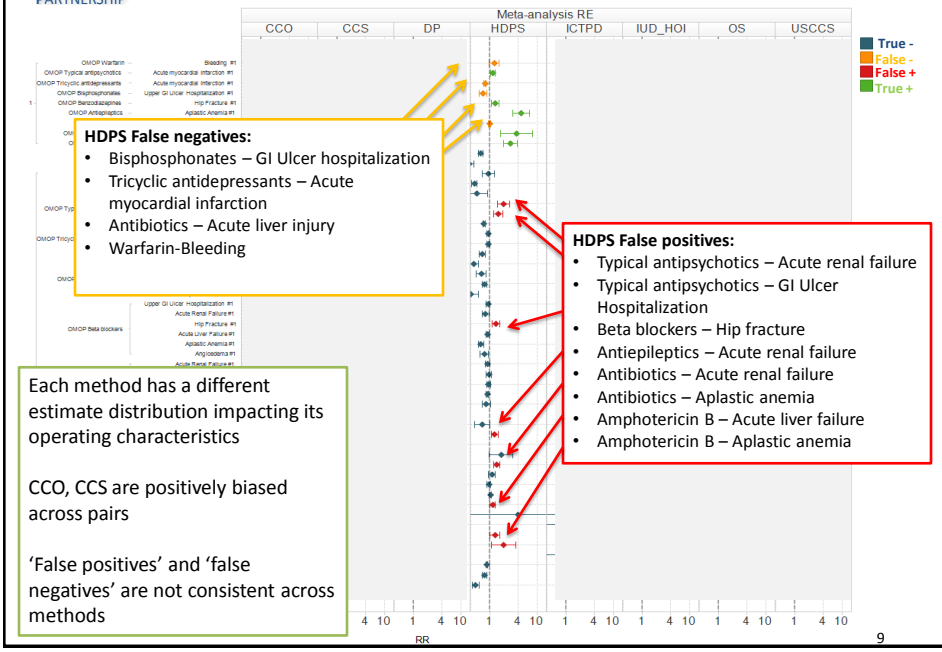
All methods yield false positive rate > 15% at conventional level of significance

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

8

OBSERVATIONAL MEDICAL OUTCOMES PARTNERSHIP

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

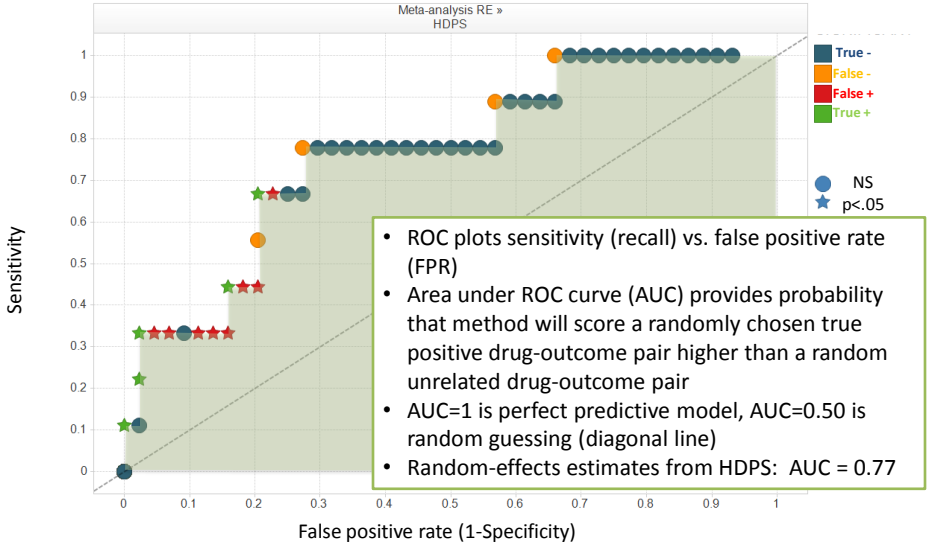
CCO, CCS are positively biased across pairs

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

9

OBSERVATIONAL MEDICAL OUTCOMES PARTNERSHIP

Receiver Operating Characteristic (ROC) curve

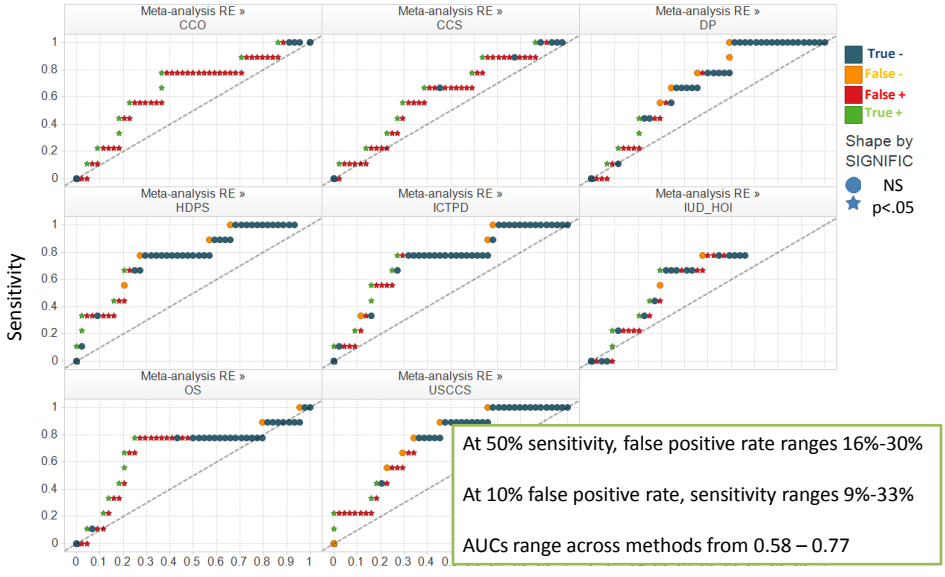


- ROC plots sensitivity (recall) vs. false positive rate (FPR)
- Area under ROC curve (AUC) provides probability that method will score a randomly chosen true positive drug-outcome pair higher than a random unrelated drug-outcome pair
- AUC=1 is perfect predictive model, AUC=0.50 is random guessing (diagonal line)
- Random-effects estimates from HDPS: AUC = 0.77

10

OBSERVATIONAL MEDICAL OUTCOMES PARTNERSHIP

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

11

OBSERVATIONAL MEDICAL OUTCOMES PARTNERSHIP

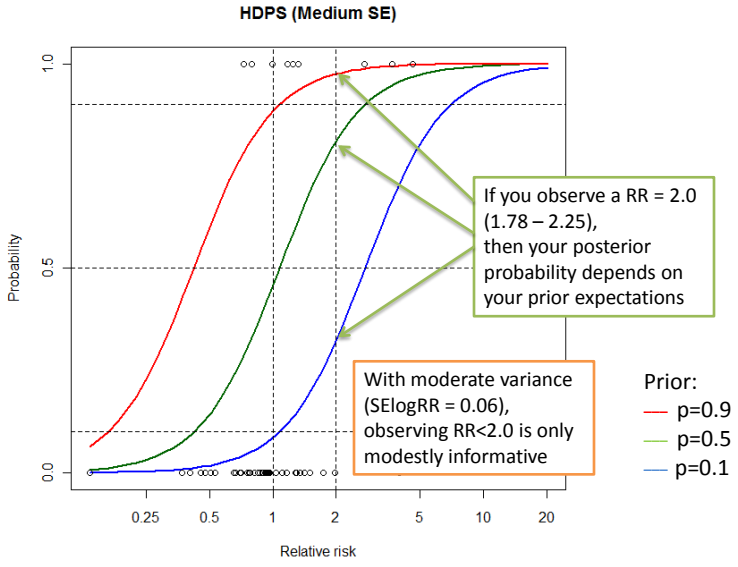
So given these operating characteristics, what can we expect to do in practice….

- 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(true|RR,SE) ~ p(RR,SE|true)*p(true)
  - p(true) is the prior probability of true association; consider a family of priors: skeptical (0.1), indifferent (0.5), enthusiastic (0.9)
  - p(RR,SE|true) can be estimated from empirical data (OMOP experimental results)

12

OBSERVATIONAL MEDICAL OUTCOMES PARTNERSHIP

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



13

OBSERVATIONAL MEDICAL OUTCOMES PARTNERSHIP

Concluding thoughts

- Operating characteristics are fundamental to understanding how an active surveillance system can complement current practice to support a comprehensive safety assessment
- No one clear 'best' method, as it depends on tolerance for false positives vs. false negatives, but current evidence suggests system can achieve:
  - At 50% sensitivity, false positive rate ranges 16%-30%
  - At 10% false positive rate, sensitivity ranges 9%-33%
- Accuracy of a method when applied to all available data is only the start and may be optimistic
  - Newly marketed medical products accumulate exposures over time and the population exposed may shift
  - Timeliness of detection is an additional desired performance metric, but first depends on accuracy to know that timely findings are correct and stable
- Further empirical research needed to have more complete understanding of operating characteristics before widespread adoption
  - More test cases needed to describe overall performance and identify specific limitations of methods
  - Experiments need to be conducted in both real data and simulated data to provide a complete summary of method behavior

14