

**OBSERVATIONAL
MEDICAL
OUTCOMES
PARTNERSHIP**

**The Role of Statistics and
Opportunities for Statisticians in
Active Drug Safety Surveillance**

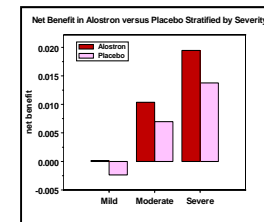
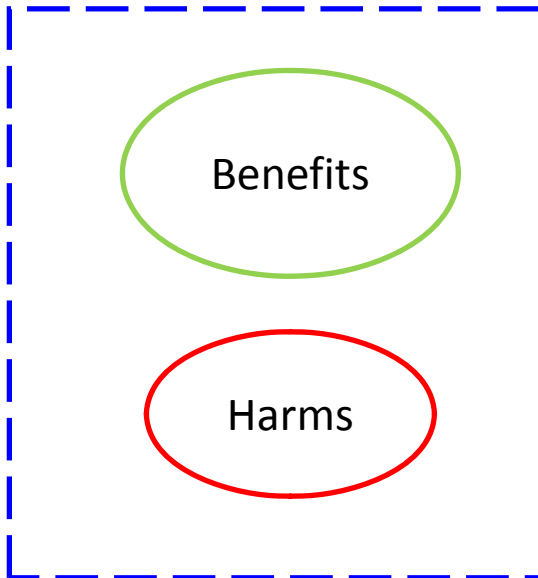
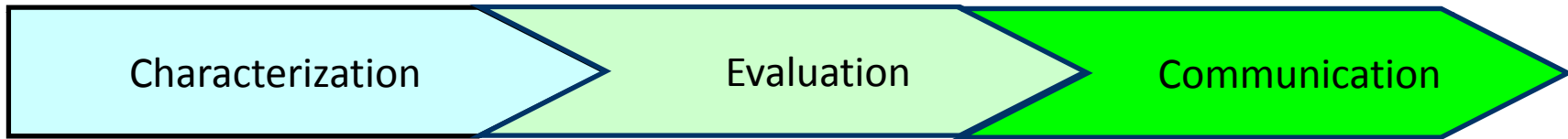
Patrick Ryan

GlaxoSmithKline

Observational Medical Outcomes Partnership (OMOP)

August 01, 2010

Steps within the process of understanding of medicines



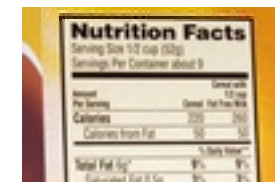
The NEW ENGLAND JOURNAL of MEDICINE

Effect of Rosiglitazone on the Risk of Myocardial Infarction and Death from Cardiovascular Causes

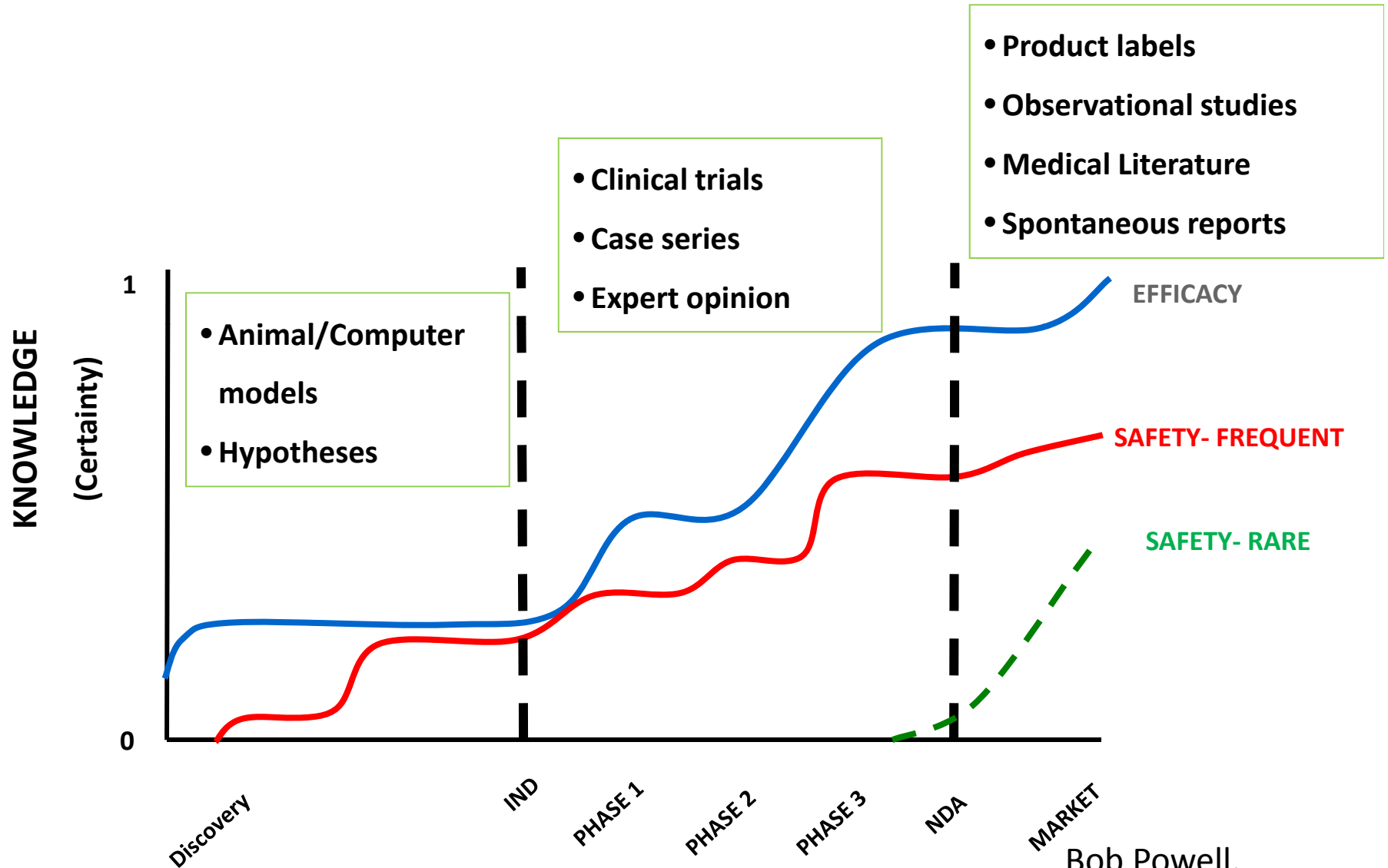
Steven E. Nissen, M.D., and Kathy Walsh, M.P.H.

ABSTRACT

BACKGROUND Rosiglitazone is widely used to treat patients with type 2 diabetes mellitus, but the effect on cardiovascular morbidity and mortality has not been determined.



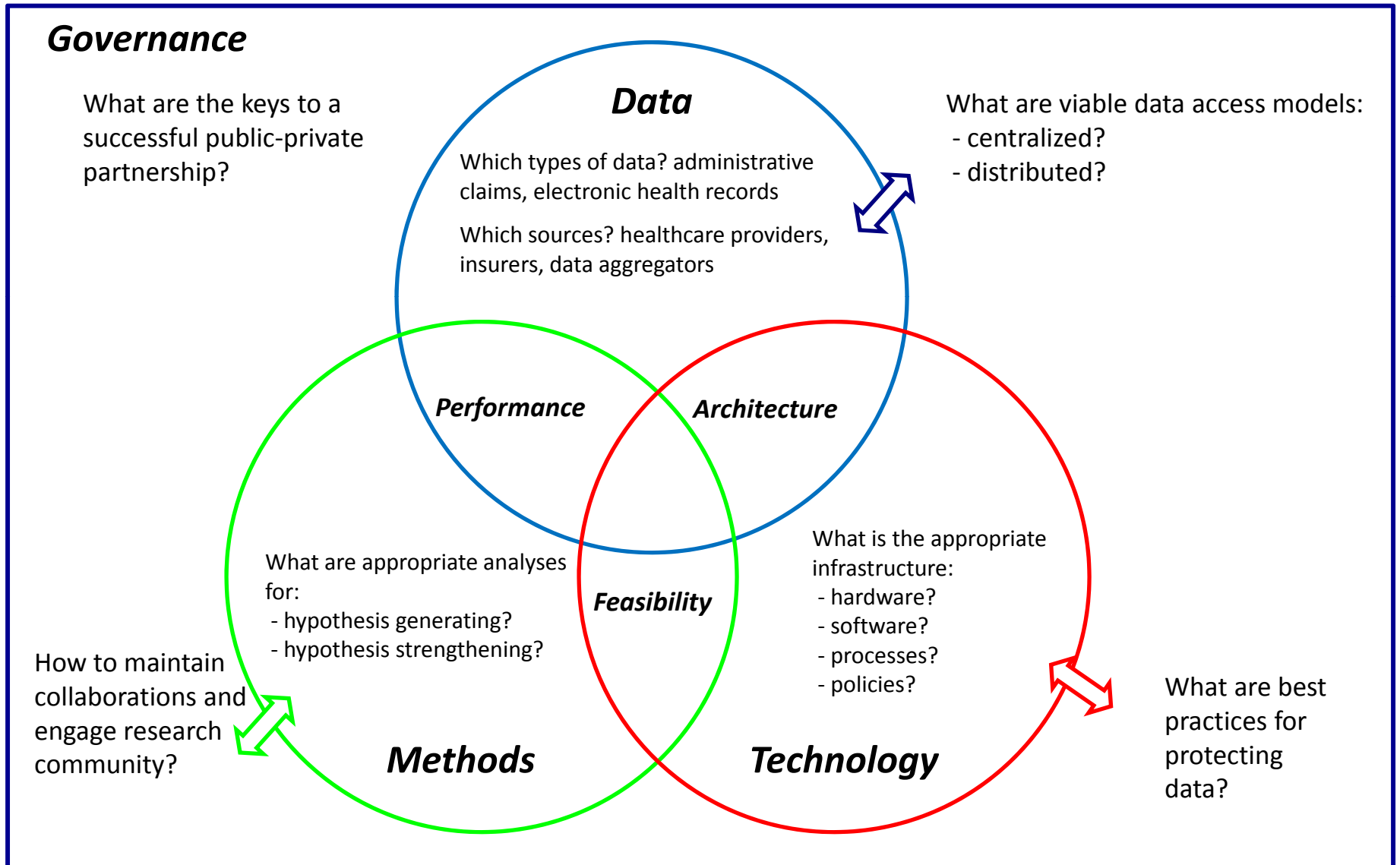
Diversity in data and analyses throughout development



Post-approval opportunities for statisticians

- Identify and evaluate emerging safety concerns of medical products using observational healthcare databases (***active surveillance***)
- Explore patient subgroups that have differentiated response (***personalized medicine***)
- Study long-term outcomes of alternative treatments in real-world populations (***comparative effectiveness***)
- Integrate disparate data sources to provide composite evaluation that weighs the evidence of the all potential effects of a medicine (***benefit-risk analysis***)

Outstanding questions for active surveillance



Observational Medical Outcomes Partnership

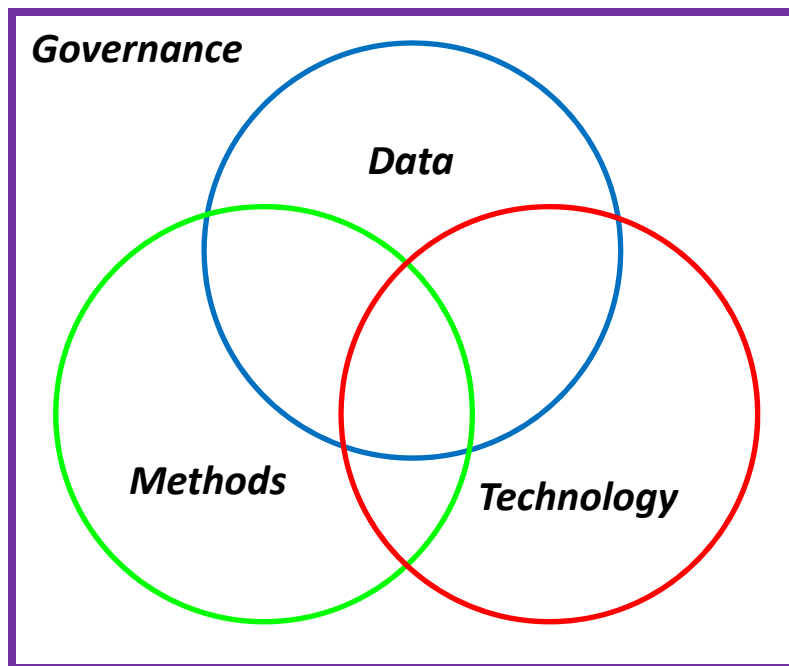
A public-private partnership to serve the public health by testing whether multi-source observational data can improve our ability to assess drug safety and benefits.

- Assess the appropriate technology and data infrastructure required for systematic monitoring of observational data
- Develop and test the feasibility and performance of the analysis methods
- Evaluate required governance structures

OMOP: <http://omop.fnih.org>

Breadth and diversity of OMOP research community

OMOP's research community requires active participation from all key stakeholders, including government, academia, industry, health care organizations, and patient groups.



Over 100 collaborating partners

Governance

- 10 Executive Board members, chaired by FDA and managed by Foundation for NIH
- 21 Advisory Board members
- Led by 6 research investigators and PMO

Methods

- 17 methods collaborators

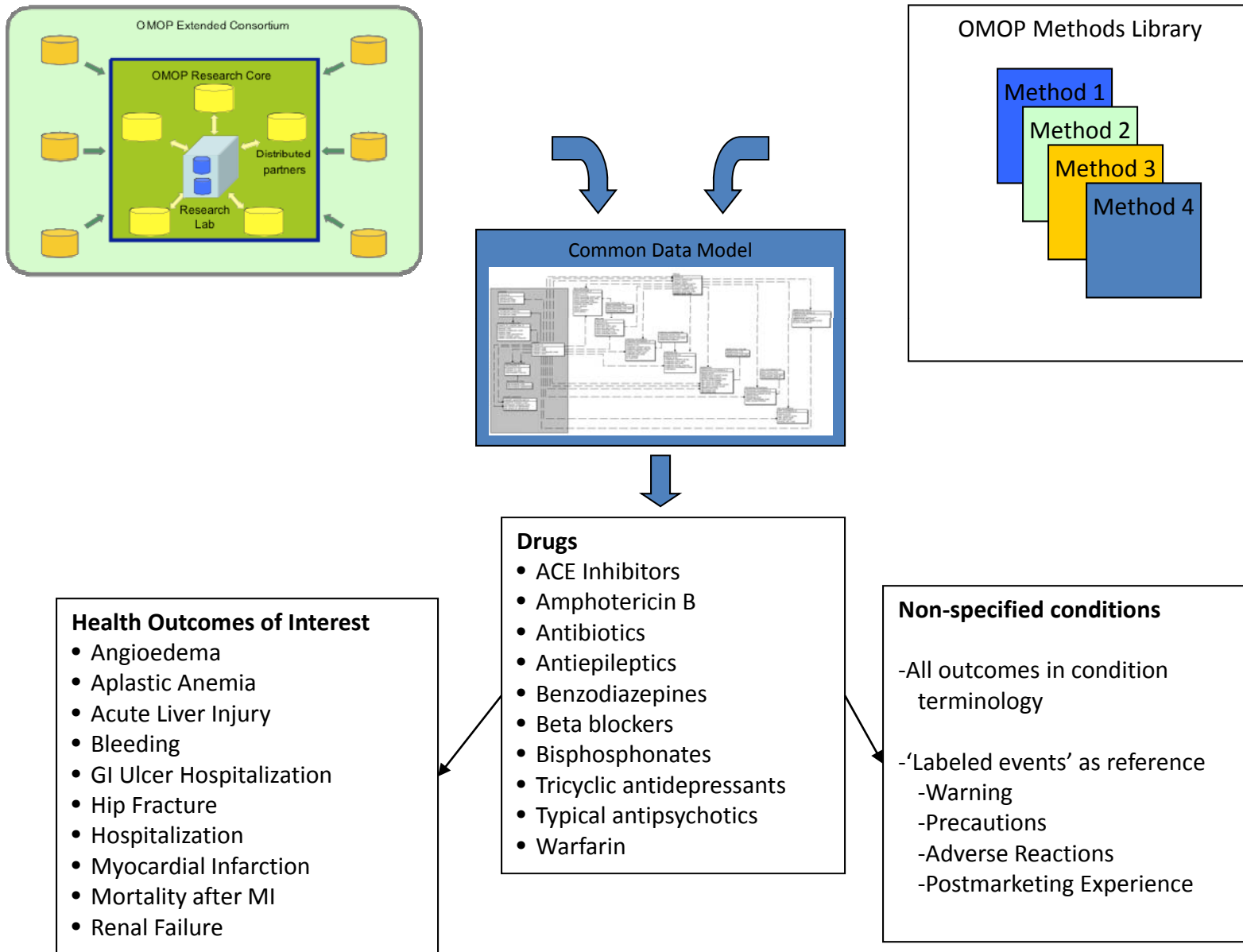
Data

- 6 distributed research partners
- 5 central databases included in the OMOP Research Lab

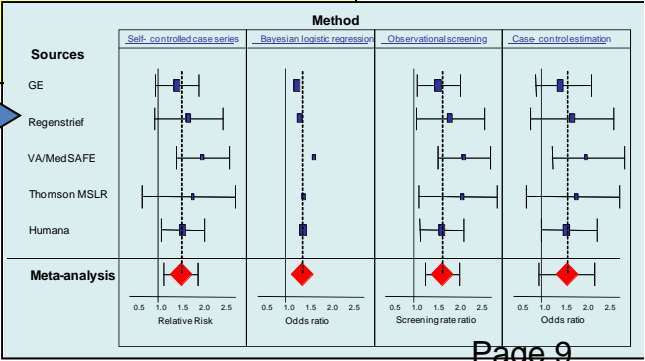
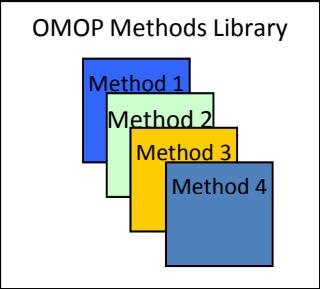
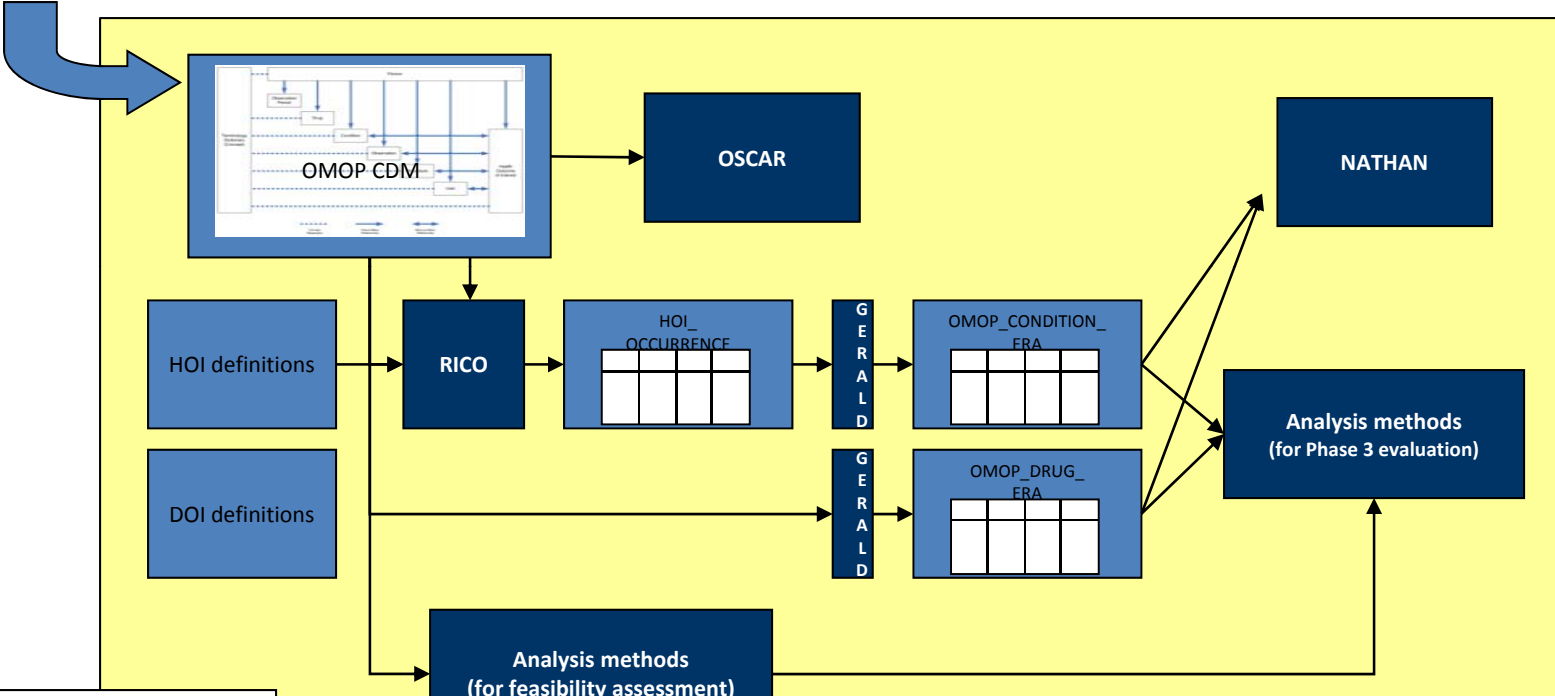
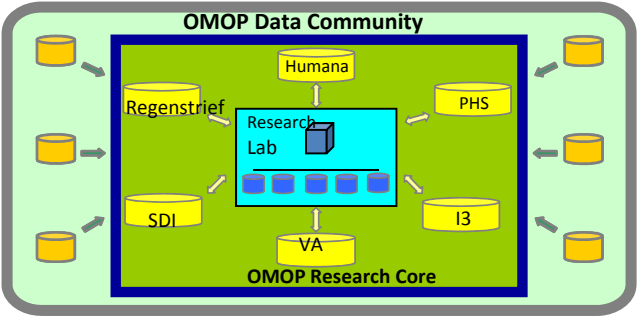
Technology

- 2 data access models, 7 different systems architectures

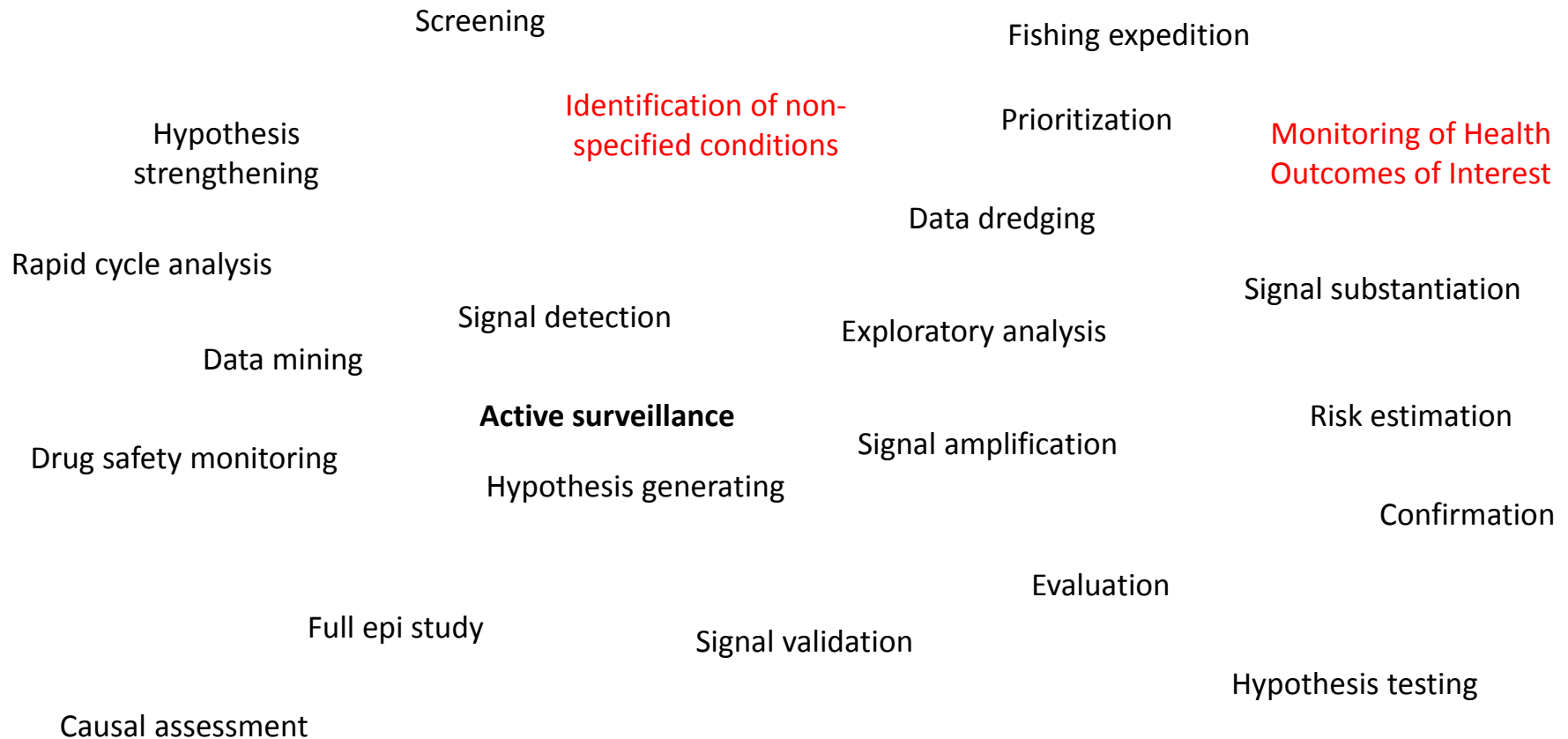
OMOP research experiment workflow



OMOP Progression

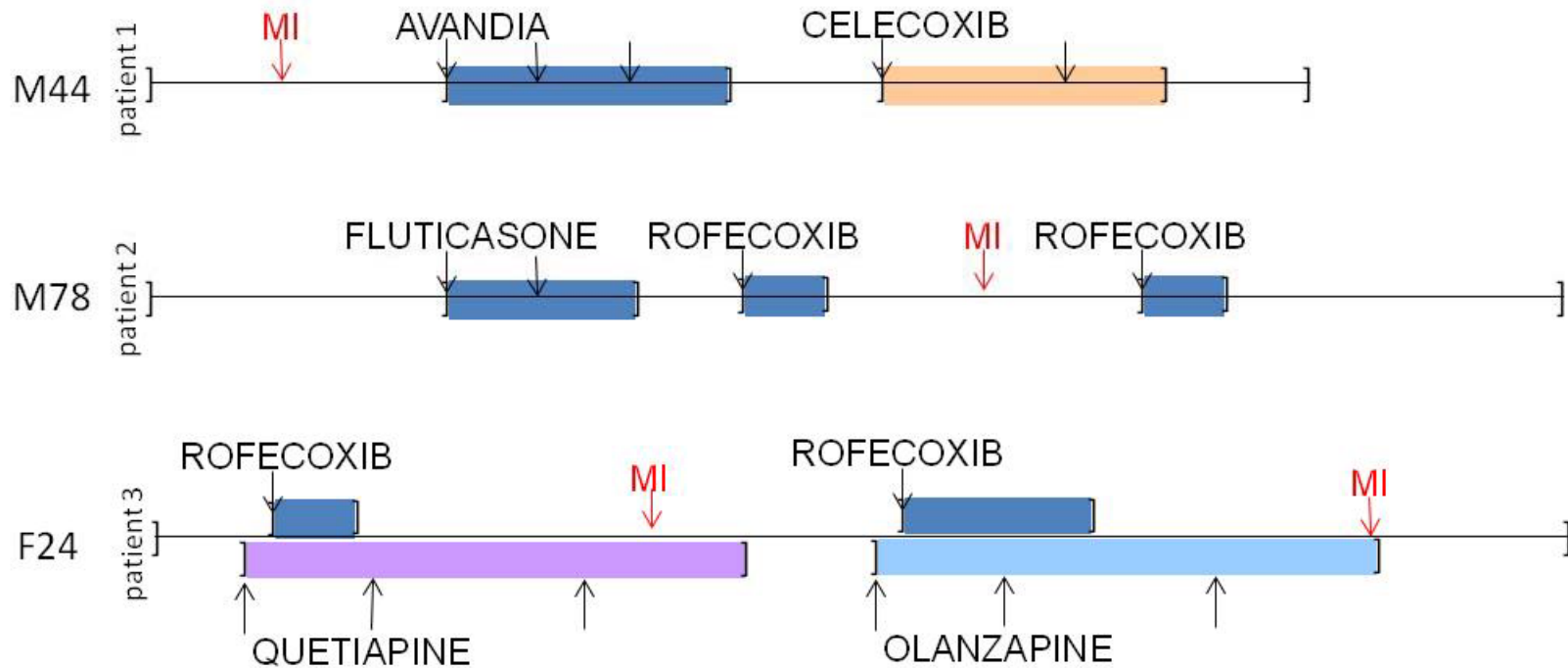


Characterizing Drug-Outcome Associations



Fundamental task: Estimate the strength of the drug-outcome relationship

What do the data look like?



Database contain millions of persons with years of (incomplete) longitudinal data
Computational considerations require efficient use of data in analyses

OMOP's Methods Landscape

Disproportionality Analysis

	<i>AE j = Yes</i>	<i>AE j = No</i>
Drug <i>i</i> = Yes	<i>a=20</i>	<i>b=100</i>
Drug <i>i</i> = No	<i>c=100</i>	<i>d=1080</i>

- Distinct Patients
 - SRS
 - Modified SRS
- X
- MGPS
BCPNN
PRR X Stratified
Chi
etc.

- Temporal Pattern Discovery (WHO)

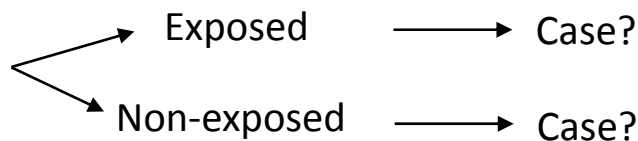
Sequential Methods

	<i>AE j = Yes</i>	<i>AE j = No</i>
Drug <i>i</i> = Yes	<i>a=20</i>	
Drug <i>i</i> = No		

← *Compare to baseline Poisson*

- Maximized Sequential Probability Ratio Test (MaxSPRT)
- Conditional Sequential Sampling Procedure (CSSP)

Exposure Based Methods

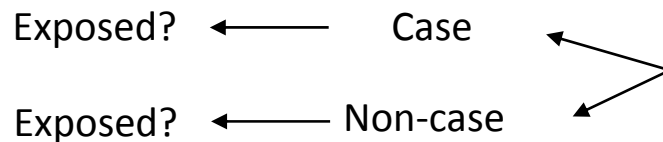


- Observational screening
- HSIU
- Incident User Designs
- High-Dimensional Propensity Scoring
- Local control

OMOP Methods Library at: <http://omop.fnih.org/MethodsLibrary>

OMOP's Methods Landscape

Case Based Methods



- Case control surveillance
- Multiset case control estimation
- Self-controlled case series
- Case crossover

Other Methods

- Hi-Dimensional logistic regression
- Statistical relational learning

Future Methods

- Multivariate self-controlled case series
- Case-time control
- Lasso propensity scoring
- Online algorithms
- OMOP Cup (60+ submissions)

OMOP Methods Library at: <http://omop.fnih.org/MethodsLibrary>

Methodological considerations common across multiple approaches

- Exposure definition
 - Incident vs. prevalent exposure
 - Source of data capture
- Outcome definition
 - Incident vs. prevalent events
 - Diagnosis codes vs. HOI
- Defining temporal relationship
 - Time from exposure start
 - Time after exposure end
- Comparator selection
- Inclusion/exclusion criteria
 - Baseline history
 - Follow-up time
- Covariate selection and adjustment
 - Matching
 - Stratification
 - Multivariate modeling
- Output metric/statistic
 - Estimation vs. testing
 - Relative vs. attributable risk
 - Measure of uncertainty

Each method has input parameters that encode these choices

OMOP Methods Library

The screenshot shows the OMOP Methods Library website. The header includes the logo for the Foundation for the National Institutes of Health and the text "Observational Medical Outcomes Partnership". The main content area is titled "OMOP Methods Library - Download Methods" and features a list of available downloads and guidelines. The list includes:

- OMOP Methods development guidelines**
- Disproportionality Analysis Method - OMOP Research Team**
 - Disproportionality Analysis Method specification 7Dec 2009
 - Disproportionality Analysis Method Source Code and Examples 2Feb 2010
 - Disproportionality Analysis Feasibility Test #1 17Jan 2010
 - Disproportionality Analysis Feasibility Test #2 17Jan 2010
- Multi-Set Case-Control Estimation - OMOP Research Team**
 - Multi-set case-control Method specification 7Dec 2009
 - Multi-set case-control Method Source Code and Examples 2Feb 2010
 - Multi-set case-control Feasibility Test #1 17Jan 2010
 - Multi-set case-control Feasibility Test #2 17Jan 2010
- Bayesian Logistic Regression - OMOP Research Team**
 - Bayesian logistic regression specification 2Feb 2010

- Standardized procedures are being developed to analyze *any* drug and *any* condition
- All programs being made publicly available to promote transparency and consistency in research
- Methods will be evaluated in OMOP research against specific test case drugs and Health Outcomes of Interest

'Ground truth' for Monitoring Health Outcomes of Interest

Test cases to be used for evaluating method performance for 'Monitoring of Health Outcomes of Interest'

Outcome	Drug									
	ACE Inhibitors	Amphotericin B	Antibiotics	Antiepileptics	Benzodiazepines	Beta blockers	Bisphosphonates	Tricyclic antidepressants	Typical antipsychotics	Warfarin
Angioedema	R	N		N	N	N				N
Aplastic Anemia	N	N	N	R	N	N	N	N		N
Acute Liver Injury		N	R		N	N	N	N		
Bleeding			N		N			N		R
GI Ulcer Hospitalization	N			N		N	R		N	
Hip Fracture	N	N	N		R	N				N
Hospitalization	B									
Myocardial Infarction			N		N		N	R	R	
Mortality after MI		N		N		B				N
Renal Failure		R	N	N	N	N	N	N	N	N

Legend	Total
B- 'True positive' benefit	2
R- 'True positive' risk	9
N- 'Negative control'	44
Avoid selection due to labeling	
Not selected due to correlation with HOI	

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

Positive predictive value
 = precision =
 $TP / (TP + FP)$

Negative predictive value
 = $TN / (FN + TN)$

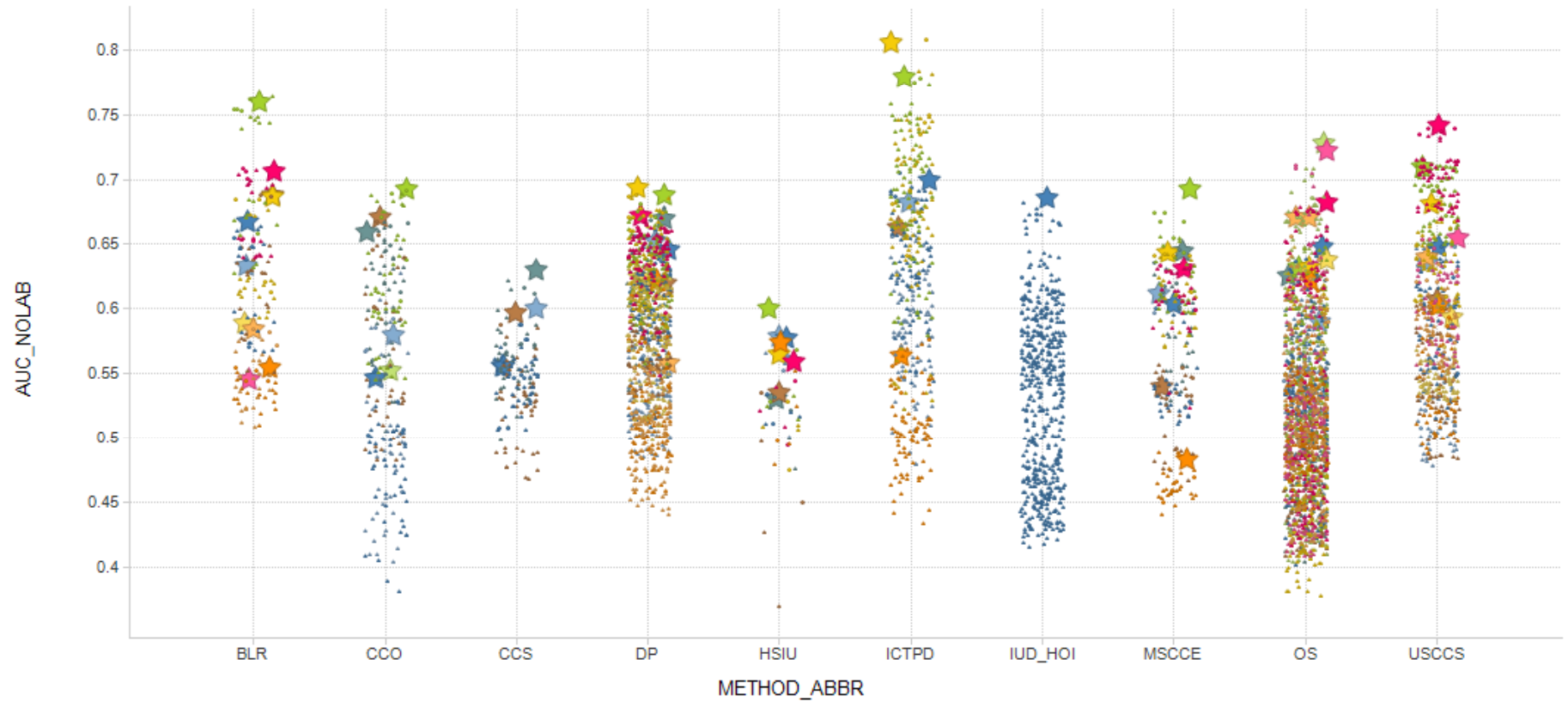
Sensitivity
 = Recall =
 $TP / (TP + FN)$

Specificity
 = $TN / (FP + TN)$

Accuracy measures from different perspectives

- **Area under ROC curve (AUC)** – what is the composite tradeoff between sensitivity and specificity at all possible threshold values?
- **Mean average precision (MAP)** – on average, what proportion of predictions are true at different thresholds?
- **p@k** - among the top k (e.g. 100) scores, what percentage are true?
- **recall@FPR** - if we can tolerate a particular false positive rate (e.g. 5%), what fraction of the true positives will we identify?
- **Average false positive rate** – what is the average false positive rate observed for each of the true positives?

Visualizing performance of alternative methods across a network of databases



Data is for illustrative purposes only

Summary

- Increased focus on post-approval analyses presents significant opportunities for statisticians
- Post-approval analyses can meaningfully inform health care policy, but appropriate use of data and methods needs to be established
- Further research and application require statisticians to play a more active role in the design, conduct, and evaluation of analyses

Contact information

Patrick Ryan

Research Investigator

Patrick.b.ryan@gsk.com

Thomas Scarnecchia

Executive Director

tscarnecchia@fnih.org

Emily Welebob

Senior Program Manager, Research

ewelebob@fnih.org

OMOP website: <http://omop.fnih.org>