

**OBSERVATIONAL
MEDICAL
OUTCOMES
PARTNERSHIP**

**Opportunities for standardized observational
analytics to support causal inference in
comparative effectiveness research**

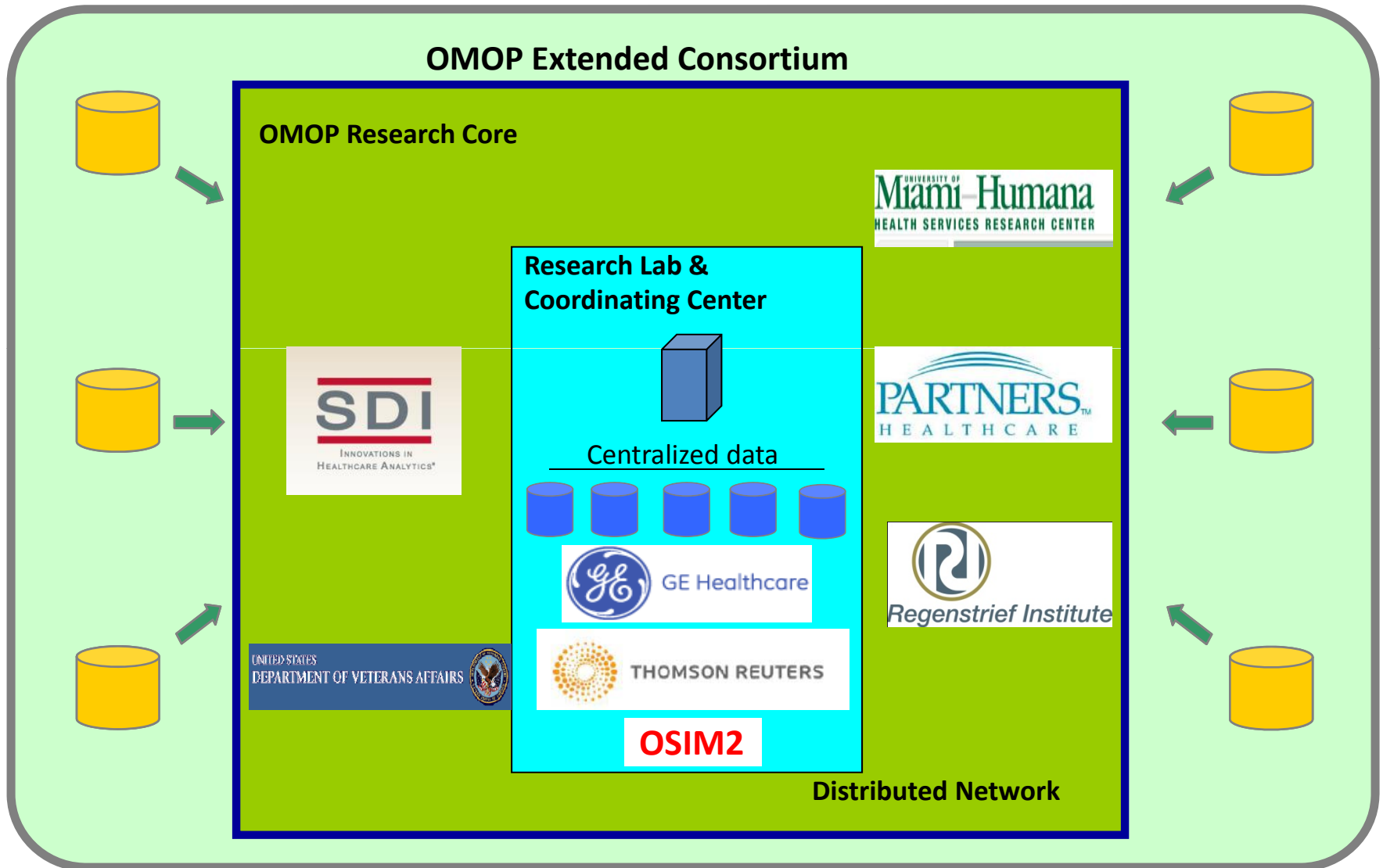
Patrick Ryan
on behalf of OMOP Research Team
October 27, 2011

Observational Medical Outcomes Partnership

Public-Private Research Partnership established to inform the appropriate use of observational healthcare databases for studying the effects of medical products:

- Conducting methodological research to empirically evaluate the performance of alternative methods on their ability to identify true associations
- Developing tools and capabilities for transforming, characterizing, and analyzing disparate data sources across the health care delivery spectrum
- Establishing a shared resource so that the broader research community can collaboratively advance the science

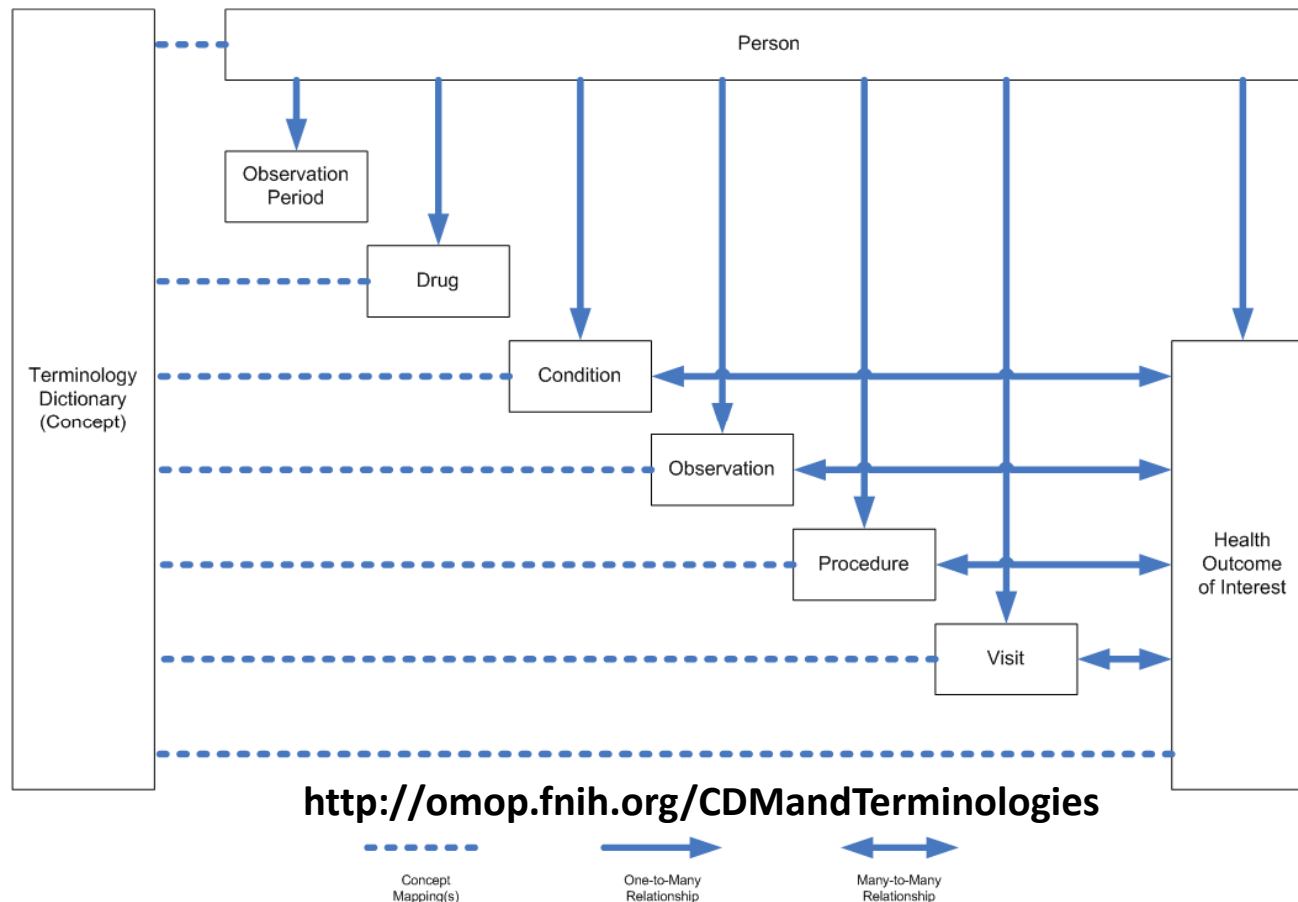
OMOP Data Community – First Two Years



178 million persons with patient-level data

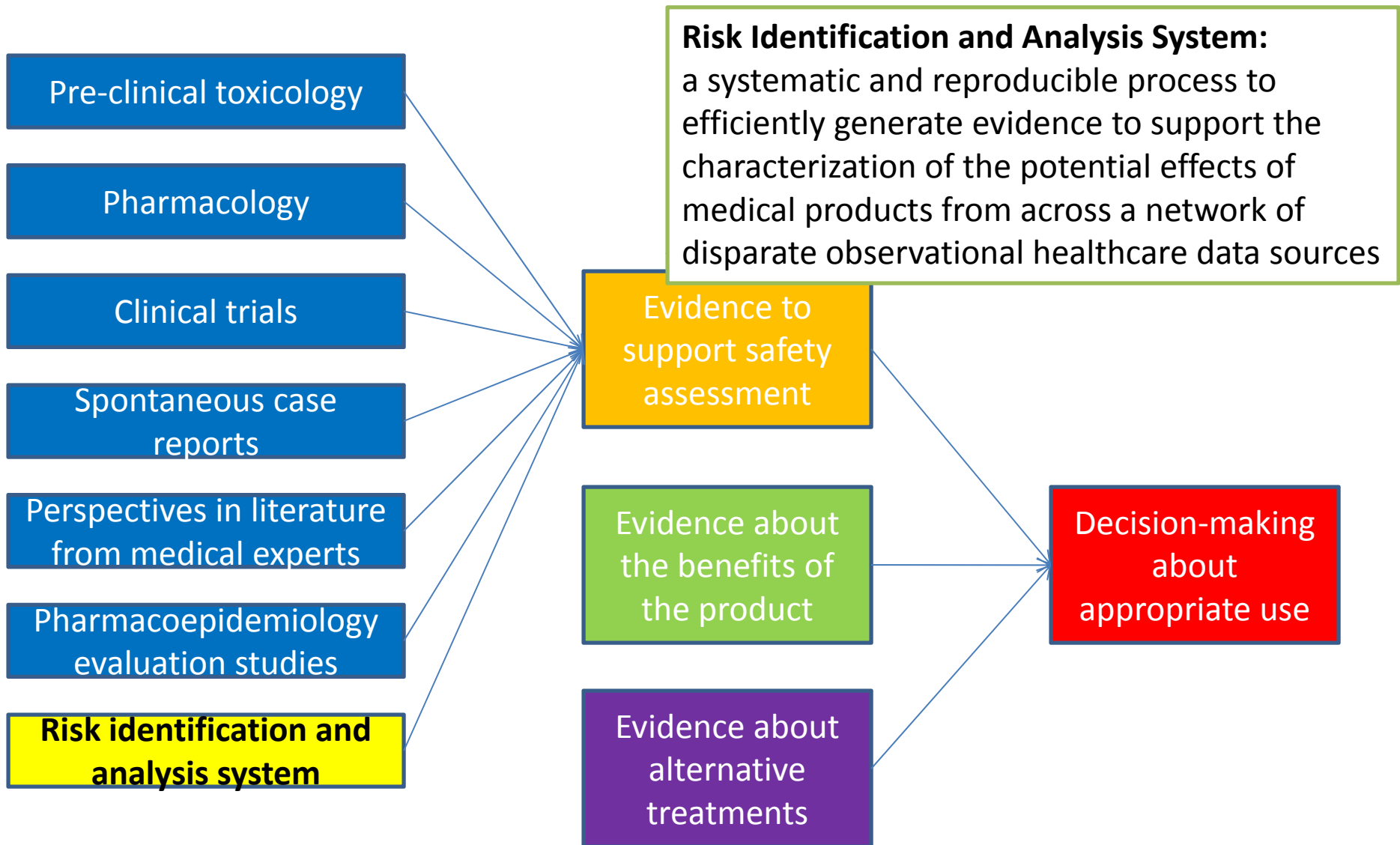
5.4 billion drug exposures, 5.8 billion procedures, 2.3 billion clinical observations

Establishing a common clinical information model

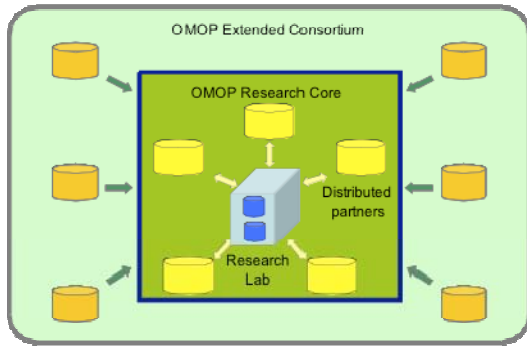


- Developed with broad stakeholder input
- Designed to accommodate disparate types of data (claims and EHRs)
- Optimized to use case of standardized large-scale analytics
- Conceived for active medical product surveillance, but extensible for other use cases
- Applied successfully across OMOP data community

Risk identification and analysis system: One additional piece of evidence to inform medical decision-making

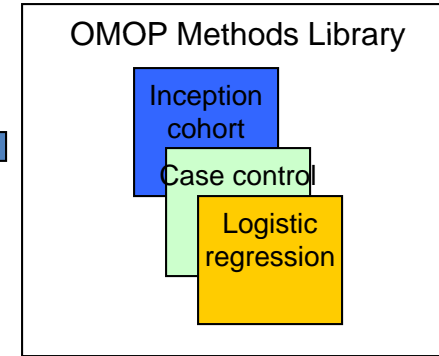
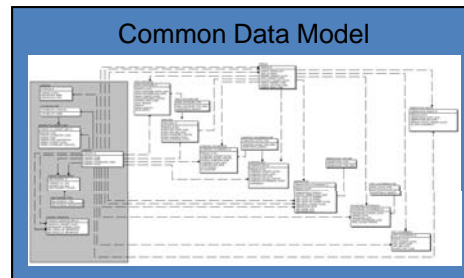


OMOP Research Experiment



- 10 data sources
- Claims and EHRs
- 170M+ lives
- Simulated data (OSIM)

- Open-source
- Standards-based
- Systematic data characterization and quality assurance



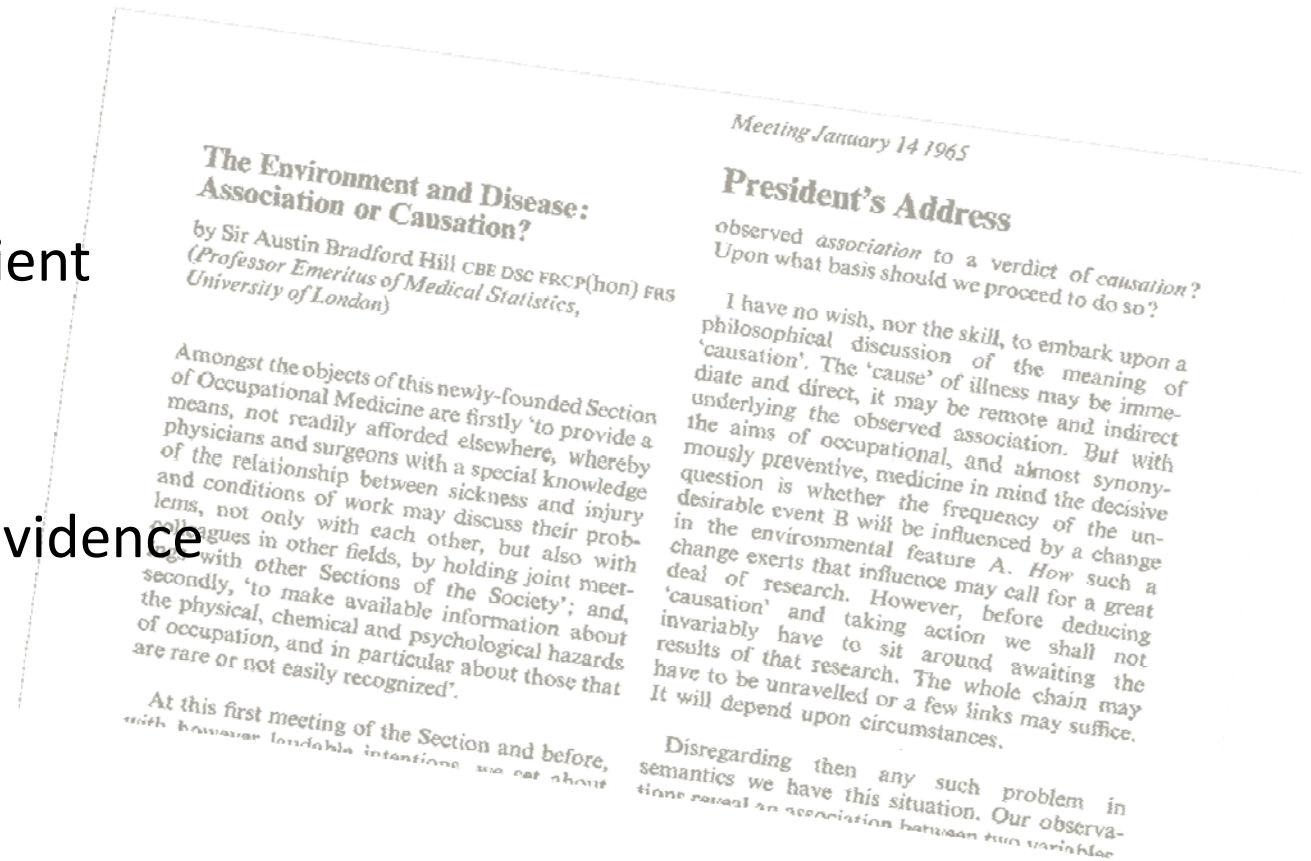
- 14 methods implemented as standardized procedures
- Full transparency with open-source code and documentation
- Epidemiology, statistical and machine learning designs

Drug

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

Hill's causality viewpoints

- Strength of association
- Consistency
- Specificity
- Temporality
- Biological gradient
- Plausibility
- Coherence
- Experimental evidence
- Analogy



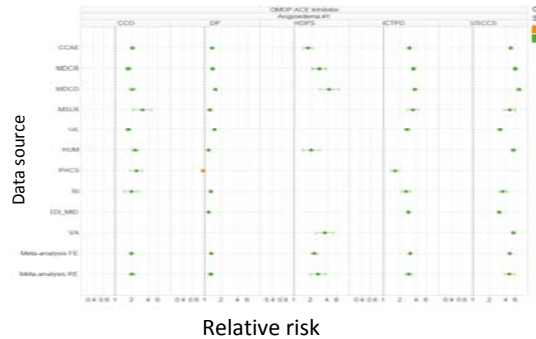
Austin Bradford Hill, "The Environment and Disease: Association or Causation?," *Proceedings of the Royal Society of Medicine*, 58 (1965), 295-300.

Vision for a risk identification and analysis system 'causal dashboard'

Drug Tricyclic antidepressants ▼

Outcome Acute myocardial infarction ▼

Strength of association

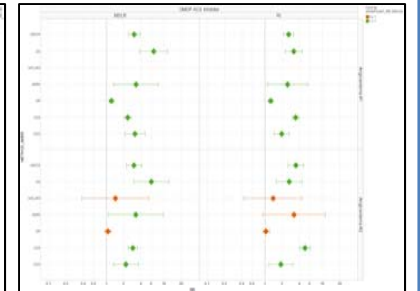
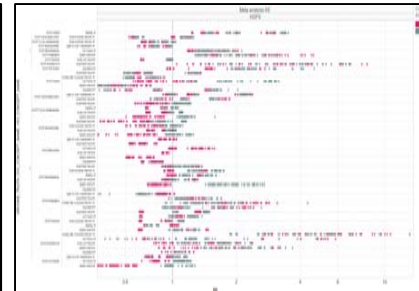
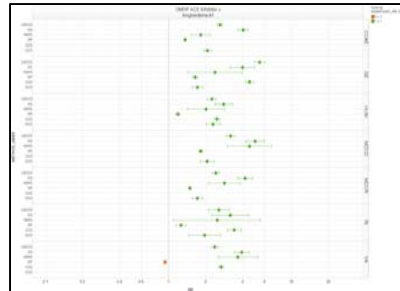


Consistency

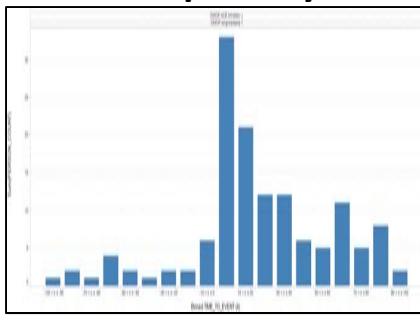
by data source

by method and parameters

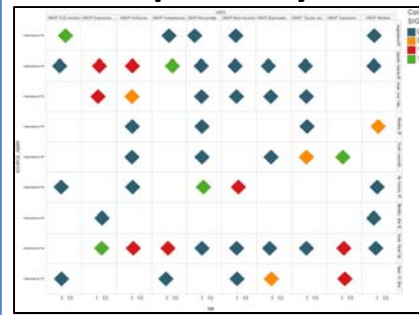
by outcome definition



Temporality

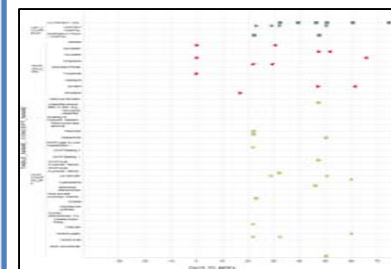


Specificity

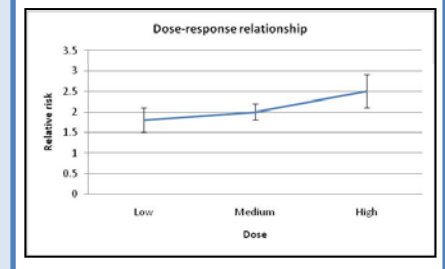


Plausibility

Interactive patient profiles



Biological gradient



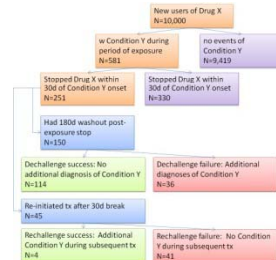
Analogy

Explore related conditions and treatments



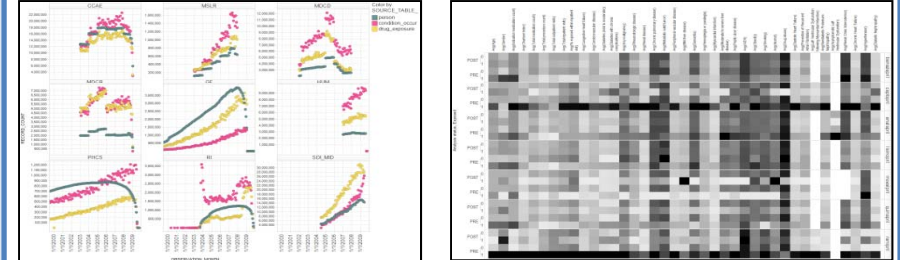
Experimental evidence

Dechallenge/Rechallenge



Coherence

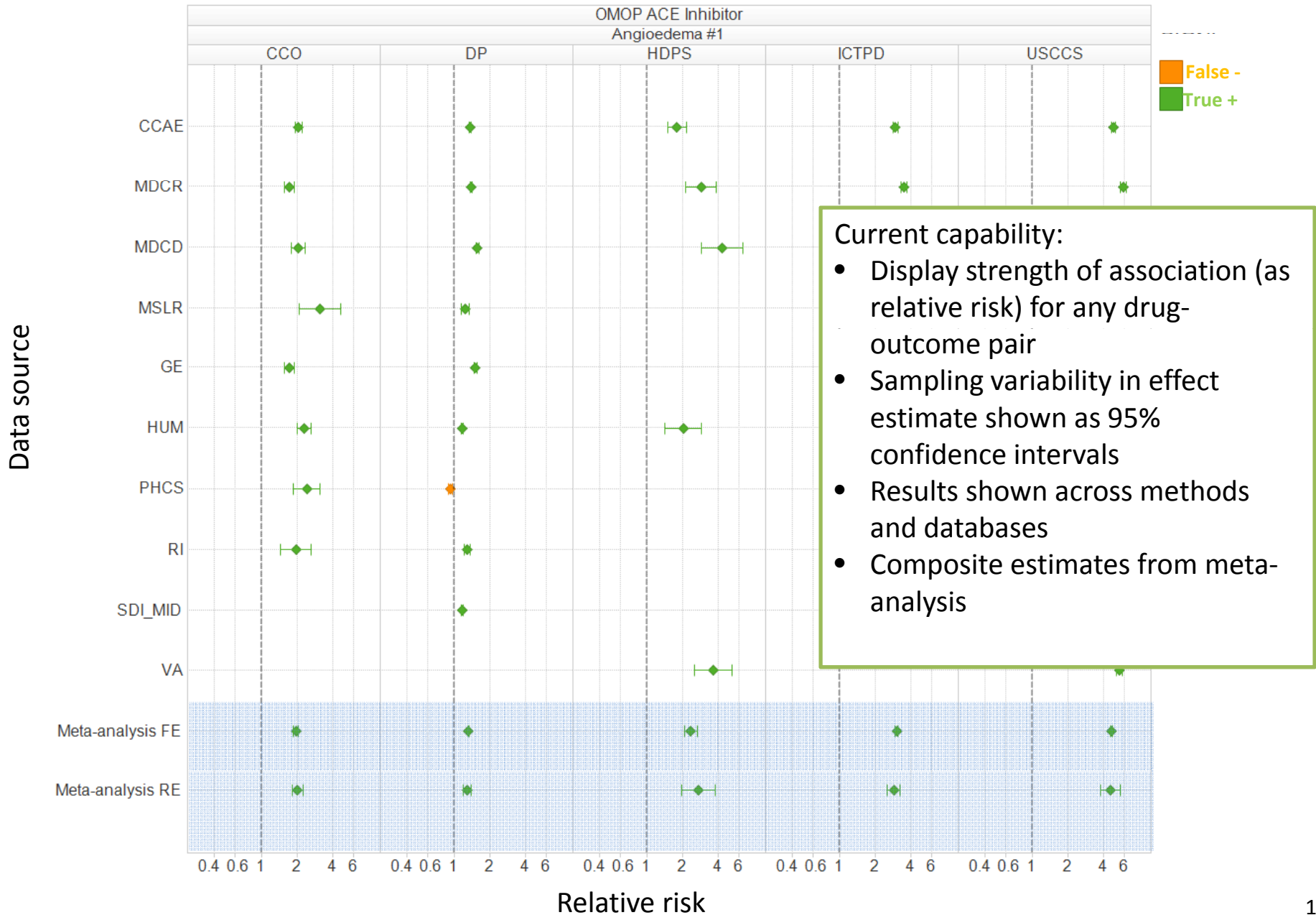
Understand data and cohort to assess potential confounding



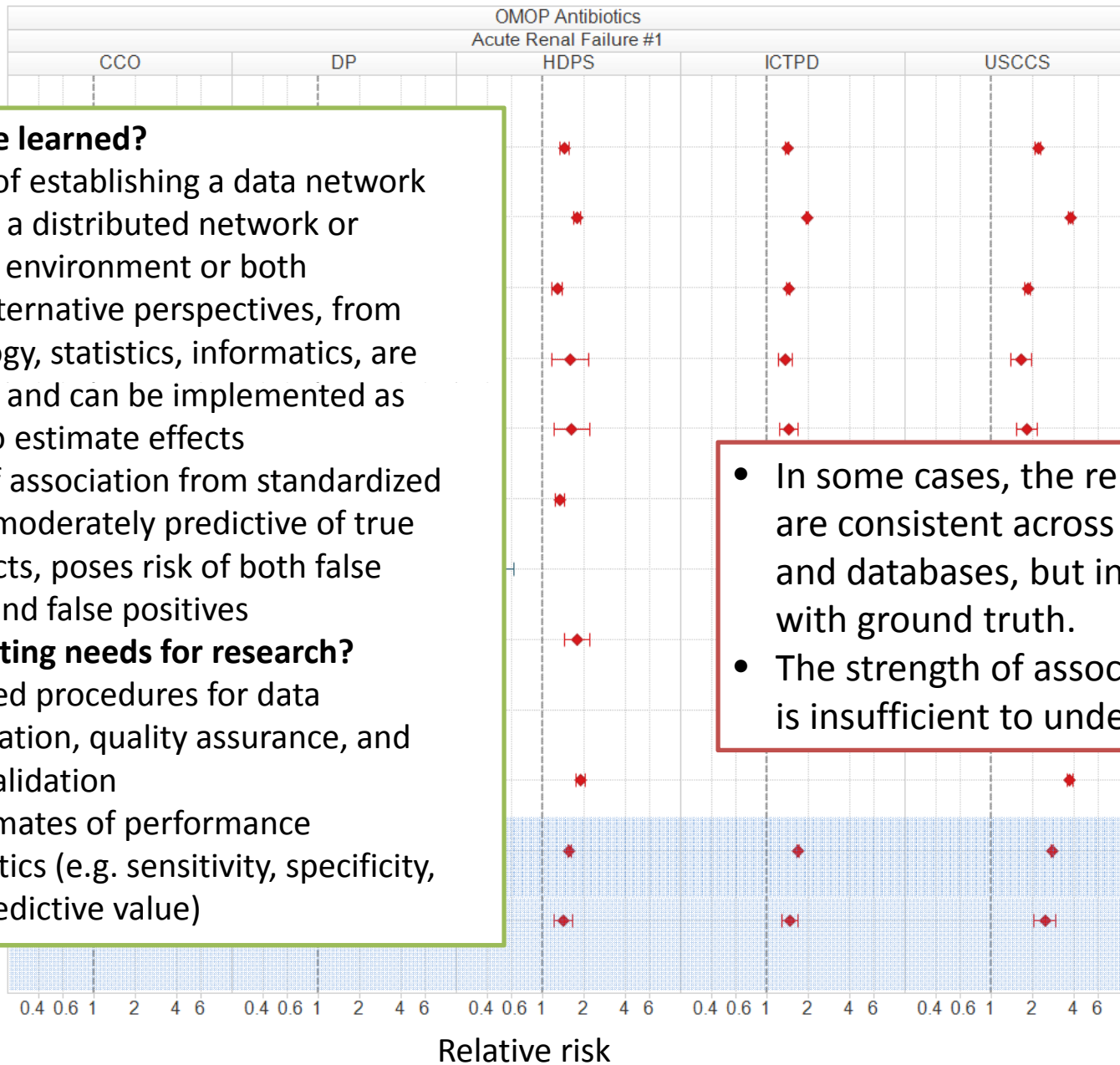
Observational analyses to support each causal consideration

- **Strength of association**
 - Current focus: methods produce effect estimates (RR) of temporal association between exposure and outcome
- Consistency
- Specificity
- Temporality
- Biological gradient
- Plausibility
- Coherence
- Experimental evidence
- Analogy

Strength of association: Ex 1: ACE inhibitor - Angioedema



Strength of association: Ex 2: Antibiotics – Acute Renal Failure



What have we learned?

- Feasibility of establishing a data network with either a distributed network or centralized environment or both
- Multiple alternative perspectives, from epidemiology, statistics, informatics, are considered and can be implemented as methods to estimate effects
- Strength of association from standardized analysis is moderately predictive of true causal effects, poses risk of both false negatives and false positives

What are existing needs for research?

- Standardized procedures for data characterization, quality assurance, and software validation
- Better estimates of performance characteristics (e.g. sensitivity, specificity, positive predictive value)

- In some cases, the relative risks are consistent across methods and databases, but inconsistent with ground truth.
- The strength of association alone is insufficient to understand why

Observational analyses to support each causal consideration

- Strength of association
- **Consistency**
 - We currently consider four types of consistency:
 1. Consistency across different databases (including measures of heterogeneity)
 2. Consistency across different methods
 3. Consistency across parameters within method
 4. Consistency across different HOI definitions
- Specificity
- Temporality
- Biological gradient
- Plausibility
- Coherence
- Experimental evidence
- Analogy

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

What have we learned?

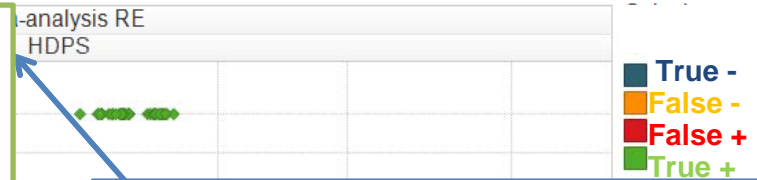
- Effect estimates are highly sensitive to study design decisions
- Substantial heterogeneity in estimates across data sources
- Comparable estimates across alternative standardized vocabularies (ICD9, SNOMED, MedDRA)
- Differential performance by alternative outcome definitions

What are existing needs for research?

- Methods for pooling results across sources
- Systematic process for defining and evaluating HOI definitions
- Explicit rules to map decisions that would be made during custom evaluations into standardized systematic process

bisphosphonates-aplastic anemia when surveillance window is 'all time post-exposure' (RR=1.25)...

- ...but shows no effect when time-at-risk defined by exposure length + 30 days (RR=1)



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)

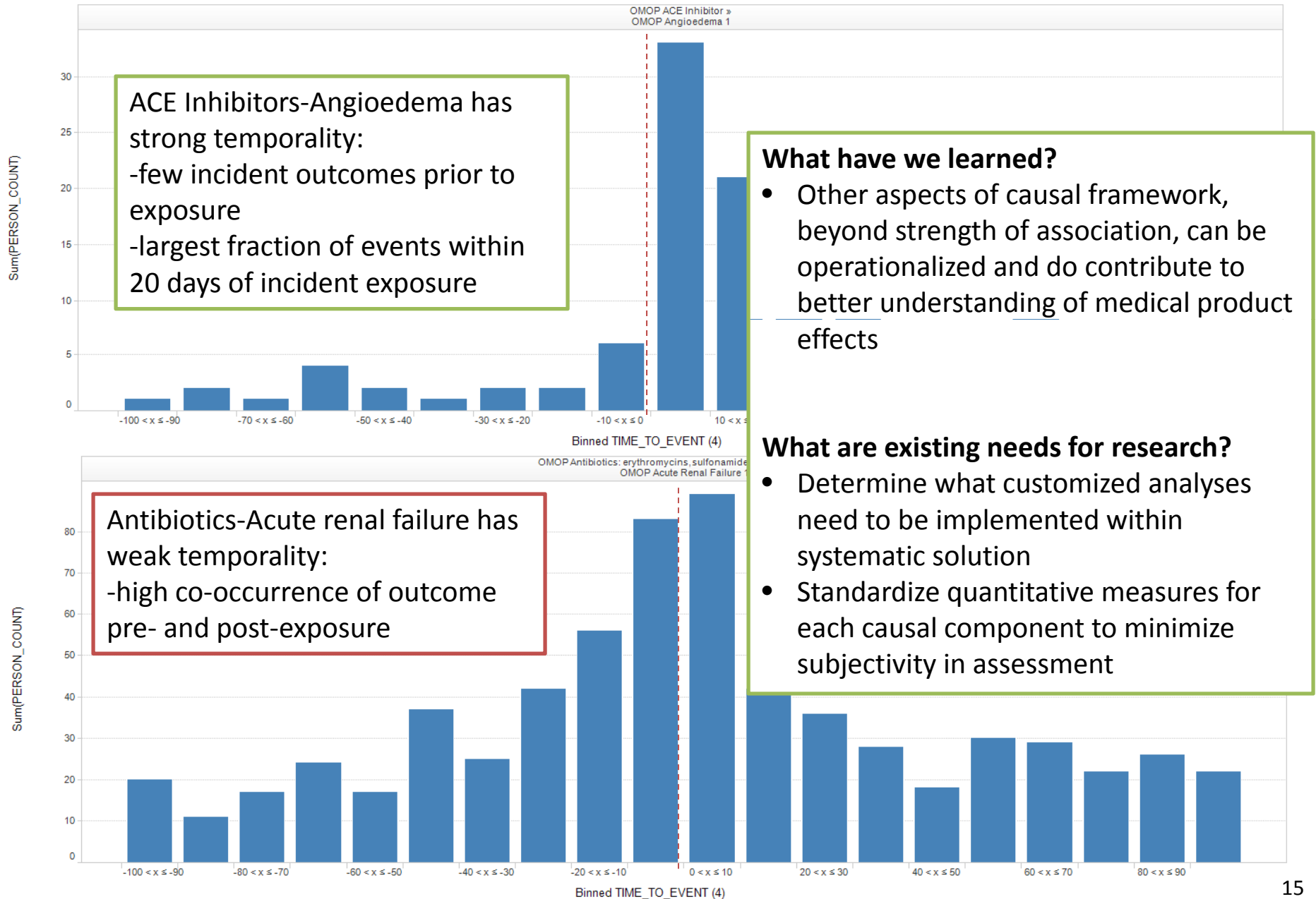
Comparator cohort (2): drugs with same indication, not in same class; most prevalent drug with same indication, not in same class

Relative risk

Observational analyses to support each causal consideration

- Strength of association
- Consistency
- Specificity
- **Temporality**
 - Evaluate time-to-event relationship between exposure and outcome
 - High incidence of events prior to exposure may suggest co-occurrence correlation without causal relationship
- Biological gradient
- Plausibility
- Coherence
- Experimental evidence
- Analogy

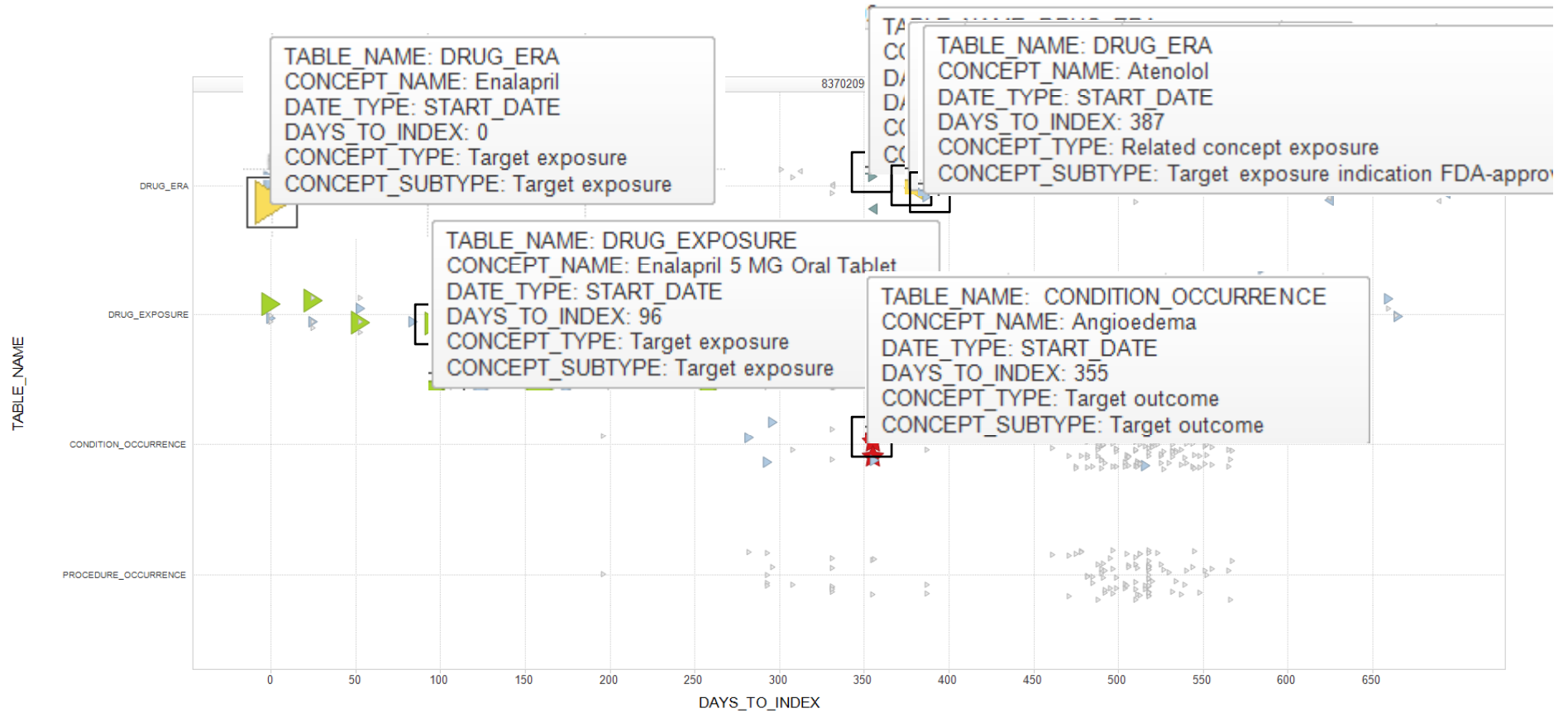
Temporality



Observational analyses to support each causal consideration

- Strength of association
- Consistency
- Specificity
- Temporality
- Biological gradient
- **Plausibility**
 - Explore interactive patient profiles to identify clinically relevant patterns or alternative explanations
 - Extend beyond population-level treatment effects to study patient-centered outcomes
- Coherence
- Experimental evidence
- Analogy

Plausibility



- Target exposure
- Target era
- Target outcome
- Related to exposure
- Related to outcome
- Start Date
- End Date
- ★ Outcome

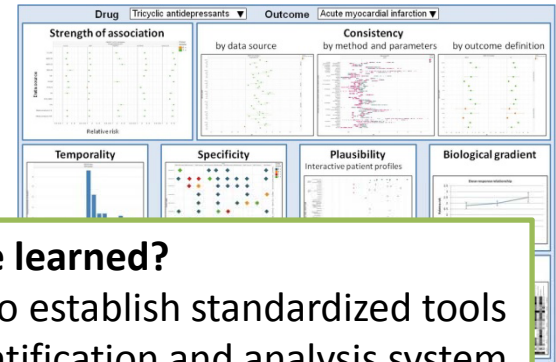
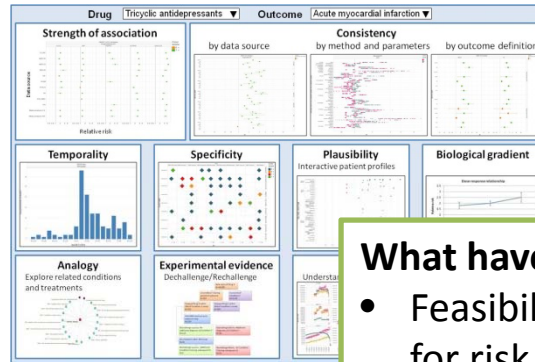
Exploratory framework for studying effects

Unstable angina

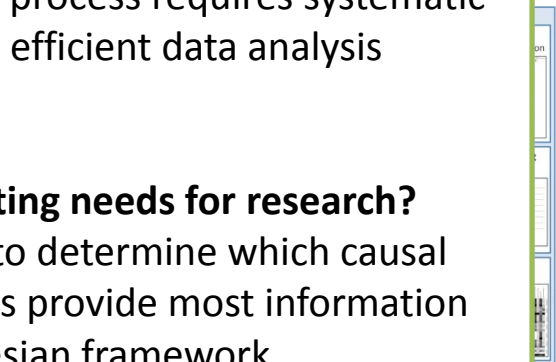
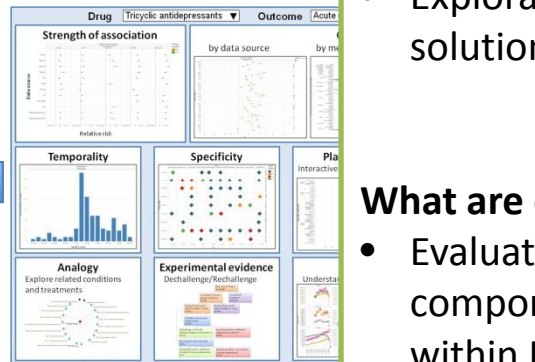
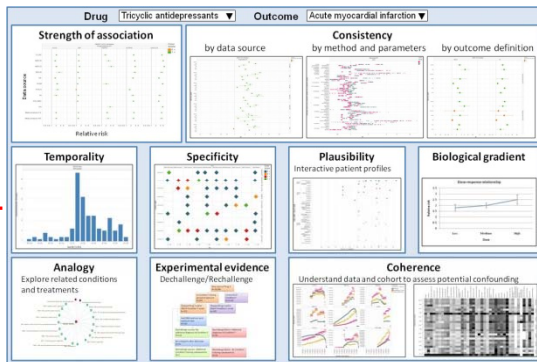
Acute myocardial infarction

Cerebrovascular accident

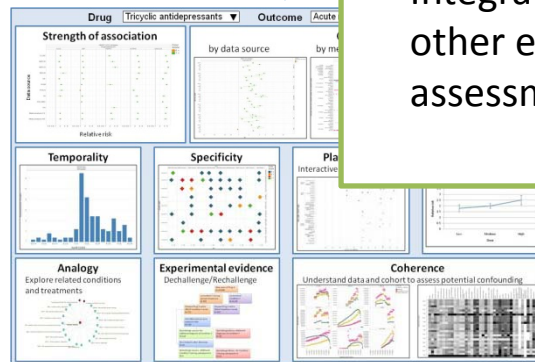
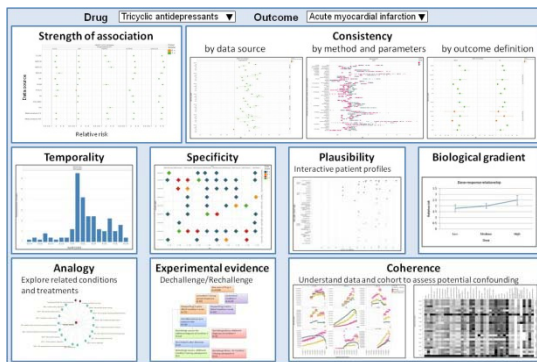
Amitriptyline



Tricyclic antidepressants



SSRIs



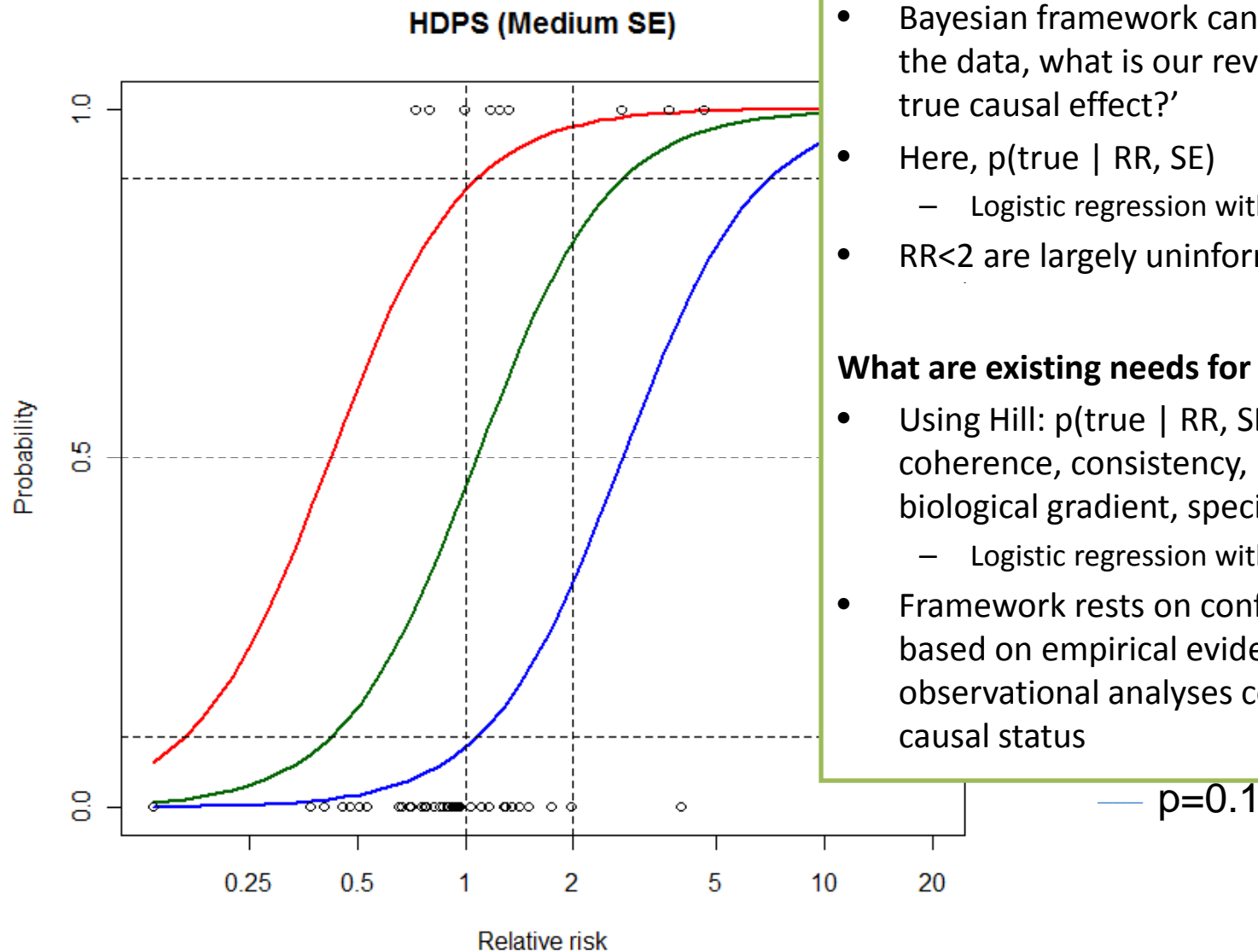
What have we learned?

- Feasibility to establish standardized tools for risk identification and analysis system
- Exploratory process requires systematic solution for efficient data analysis

What are existing needs for research?

- Evaluation to determine which causal components provide most information within Bayesian framework
- Integrating observational analyses with other evidence to support safety assessment

Quantitative framework for studying effects



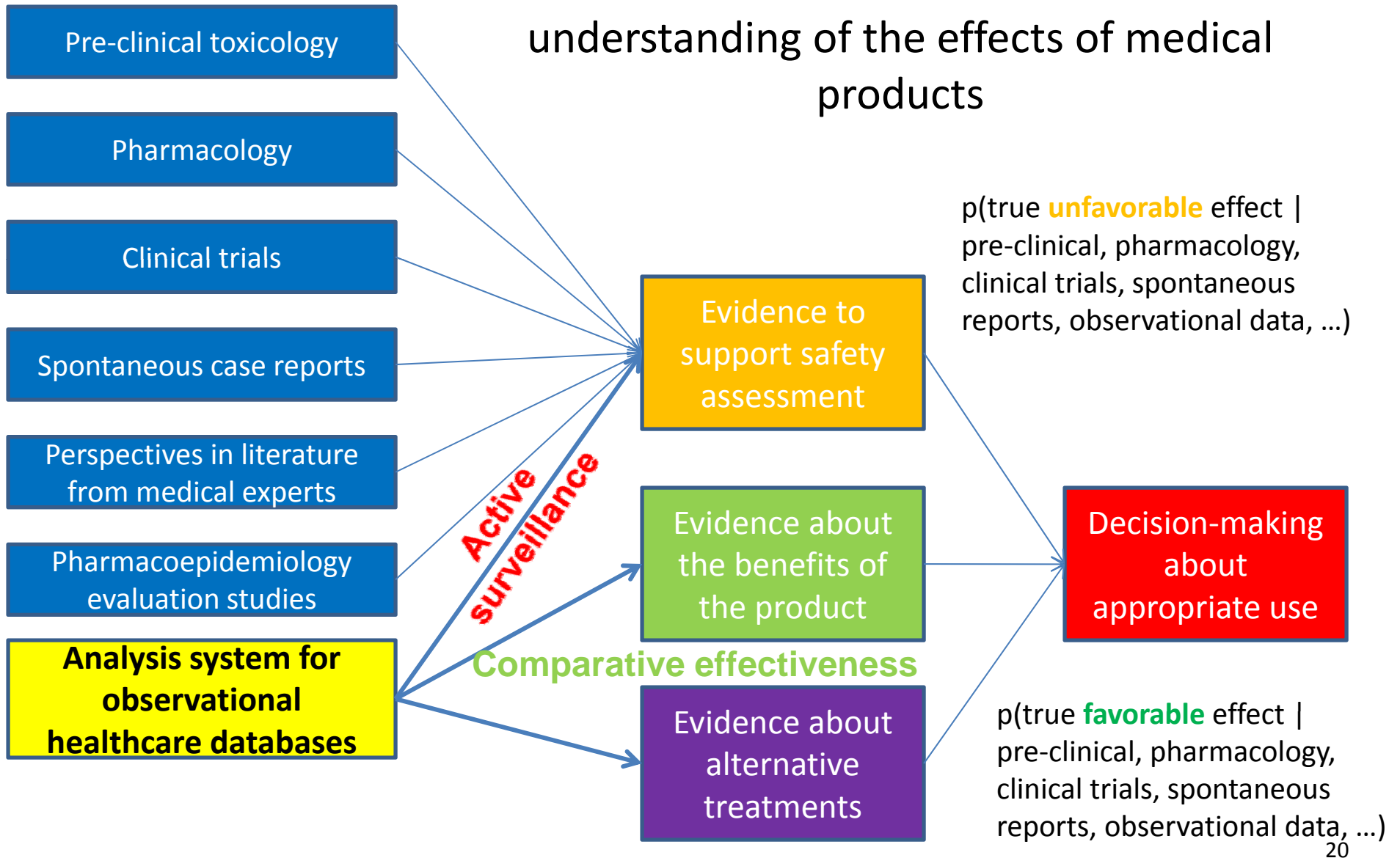
What has been learned?

- Bayesian framework can answer: ‘in light of the data, what is our revised belief of a true causal effect?’
- Here, $p(\text{true} \mid \text{RR}, \text{SE})$
 - Logistic regression with 2 predictors
- $\text{RR} < 2$ are largely uninformative

What are existing needs for research?

- Using Hill: $p(\text{true} \mid \text{RR}, \text{SE}, \text{temporality}, \text{coherence}, \text{consistency}, \text{plausibility}, \text{biological gradient}, \text{specificity}, \text{etc.})$
 - Logistic regression with many predictors
- Framework rests on confidence in model, based on empirical evidence of how observational analyses correspond to true causal status

Opportunities for a coordinated system that leverages a network of observational healthcare databases to enhance our understanding of the effects of medical products



Conclusions

- A standards-based common clinical information model is feasible and can accommodate disparate data sources
- Multiple analytical use cases can be satisfied within one framework, but scope of data needs may vary
- Standardized analytics enable efficient exploration of a large set of research/public health questions in a consistent, transparent, and reproducible process
- Large-scale analytics and interactive visualization can maximize value of EHR data resources by generating clinically meaningful knowledge for all stakeholders

Contact information

Patrick Ryan
Research Investigator
ryan@omop.org

Thomas Scarnecchia
Executive Director
scarnecchia@omop.org

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