



Standardizing phenotypes for systematic observational analysis: Lessons from OHDSI

Patrick Ryan, PhD

Janssen Research and Development

15 November 2014



A critical need for our community: Open-source phenotype library

Search this site:

Navigation

- 2013 Symposium
- ▷ About Us
- ▷ Research
- ▽ Implementation
 - ▷ Vocabularies
 - Common Data Model
 - Health Outcomes of Interest
 - NATHAN
 - OSCAR
 - GROUCH
 - Methods Library
 - Technology requirements
- ▷ Simulated Data
- Resources
- News & Events
- Community

Health Outcomes of Interest

OMOP established an open-source library of Health Outcomes of Interest (HOI) definitions for use in observational studies. These are a subset of all conditions that are of importance due to their historical associations

with drug toxicities, their medical significance, and/or public health implications.

We have found that there is very little consensus in the literature about outcome definitions. We therefore decided to create a library of HOI definitions that cover a wide representation of the prior work for testing in the OMOP experiment: broad and narrow selections of

diagnosis codes, combination of diagnosis codes with diagnostic or therapeutic procedures and lab values.

Library of HOI Definitions

[Angioedema](#)

[Aplastic Anemia](#)

[Acute Liver Injury](#)

[Acute Kidney Injury](#)

[Bleeding](#)

[Hip Fracture](#)

[Hospitalization](#)

[Acute Myocardial Infarction](#)

[Mortality After Myocardial Infarction](#)

[GI Ulcer Hospitalization](#)

- Background
- HOI Implementation
- OMOP Experiments
- Summary of definitions
- Community

Related Publications

Drug Safety 36(1) Supplement. Published online: 29 Oct. 2013
Alternative outcome definitions and their effect on the performance of methods for observational outcome studies

Christian Reich, Patrick B. Ryan, Martijn J. Schuemie.

Keywords

Health outcomes, definitions, claims database, electronic health record database, surveillance, drug safety surveillance, acute kidney injury, acute liver injury, acute myocardial infarction

29th ICPE, August 25-28, 2013

Alternative outcome definitions and their effect on observational outcome studies

Christian Reich, on behalf of the OMOP research team





Alternative definitions for Acute liver injury

D: Occurrence of at least one diagnosis code

D+P: Occurrence of at least one diagnosis code
AND (diagnostic procedure ≤ 30 d before
OR treatment procedure ≥ 60 d after)

D+P+L: Occurrence of at least one diagnosis code
AND (diagnostic procedure ≤ 30 d before
OR treatment procedure ≥ 60 d after)
AND laboratory results indicative of Hy's law:
(ALT ≥ 3 xULN OR AST ≥ 3 xULN) AND Bilirubin ≥ 2 xULN
within 7 days

L: Laboratory results indicative of Hy's law:
(ALT ≥ 3 xULN OR AST ≥ 3 xULN) AND Bilirubin ≥ 2 xULN
within 7 days

Assessment of Case Definitions for Identifying Acute Liver Injury in Large Observational Databases

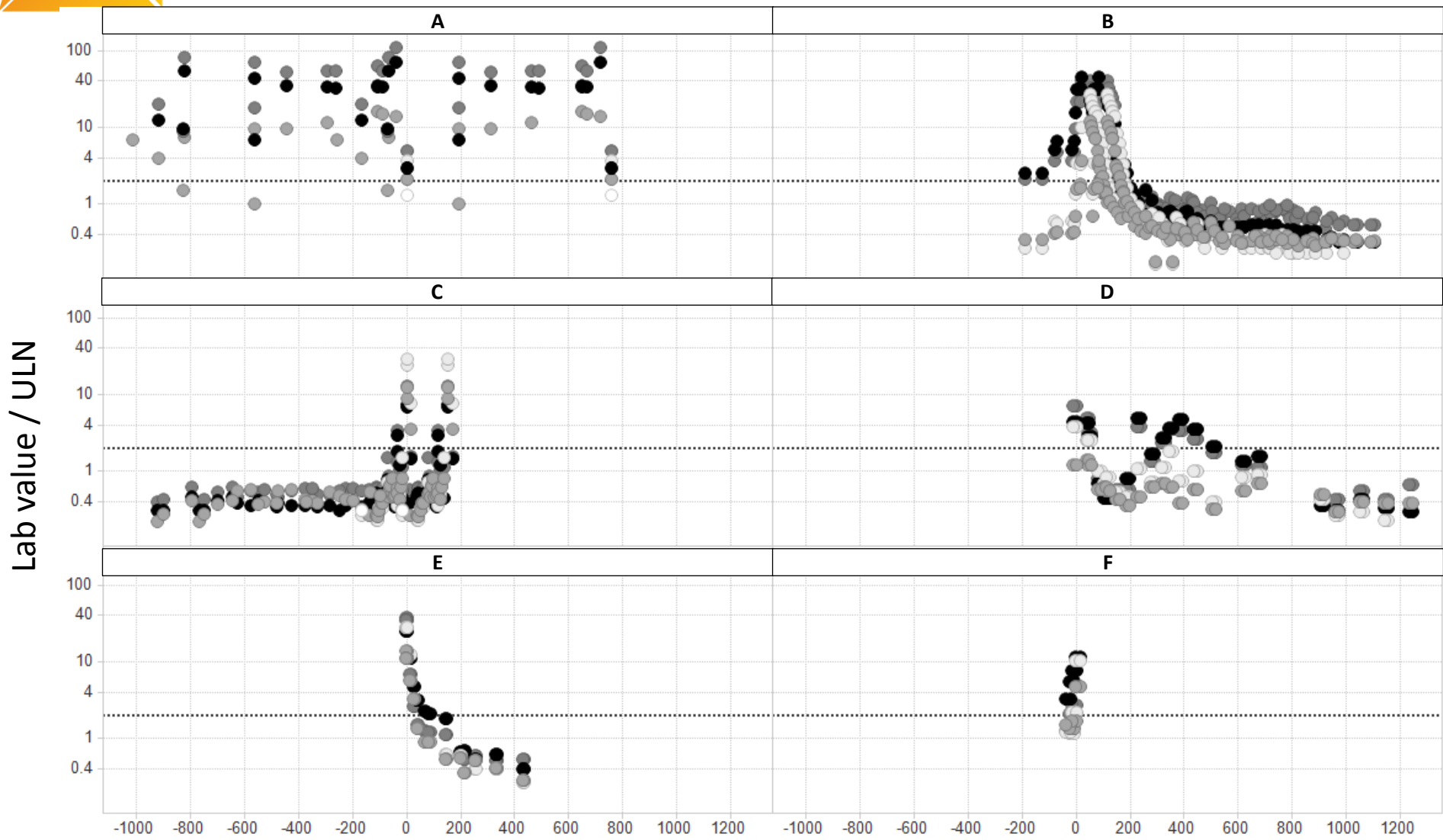
Aaron J. Katz · Patrick B. Ryan ·
Judith A. Racoosin · Paul E. Stang

Table 2 The prevalence and demographic characteristics of patients satisfying four distinct definitions of acute liver injury, by database

Database	Total population [n (000s)]	Patients satisfying acute liver injury definitions			
		D [n (%)]	L [n (%)]	D+P [n (%)]	D+P+L [n (%)]
MarketScan [®] Lab Database	1,229	154,357 (12.6)	191 (0.02)	3,743 (0.3)	28 (0.002)
Regenstrief Institute	2,002	215,930 (10.8)	7,349 (0.4)	2,735 (0.1)	384 (0.02)
Partners HealthCare System	2,942	264,083 (9.0)	36,267 (1.2)	11,032 (0.4)	2,568 (0.09)
National Patient Care Database of the Veterans Health Administration	3,202	345,519 (10.8)	NA	6,453 (0.2)	NA
Humana Inc.	5,197	447,886 (8.6)	NA	8,972 (0.2)	NA
GE Healthcare	11,216	514,118 (4.6)	6,461 (0.1)	75 (0.001)	NA
SDI Health (now IMS Health, Inc.)	90,485	6,491,416 (7.2)	NA	67,102 (0.1)	NA



- Alanine aminotransferase [Enzymatic activity/volume] in Serum or Plasma
- Aspartate aminotransferase [Enzymatic activity/volume] in Serum or Plasma
- Bilirubin [Mass/volume] in Serum or Plasma
- Bilirubin [Presence] in Urine by Test strip
- Direct bilirubin [Mass/volume] in Serum or Plasma
- Indirect bilirubin [Mass/volume] in Serum or Plasma



Lab value / ULN

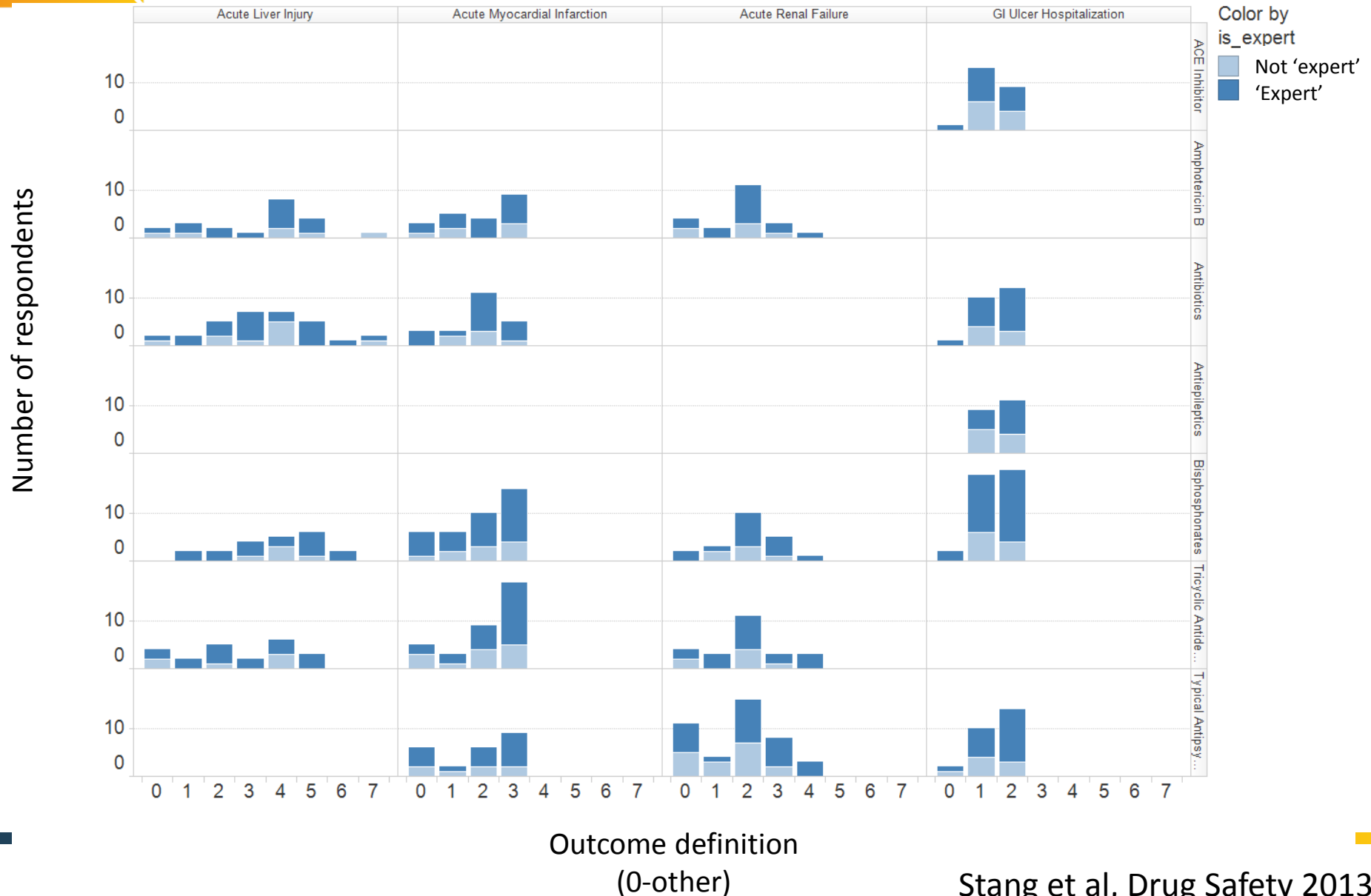
Days from index date

when criteria of (diagnosis + procedure + lab) was satisfied

-- 2*ULN



EDDIE result: little consistency in outcome definition selection





Drug Saf (2013) 36 (Suppl 1):S181–S193

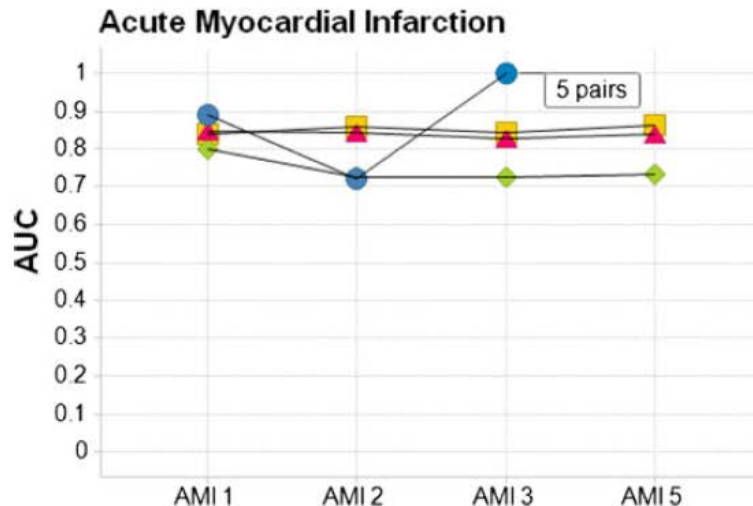
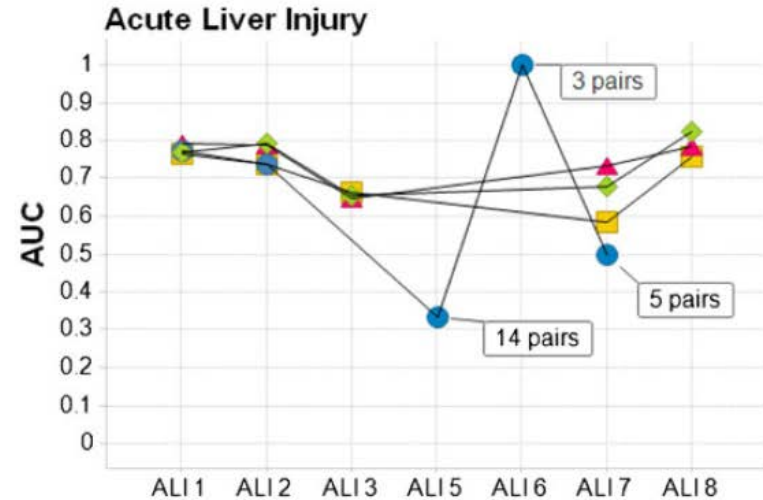
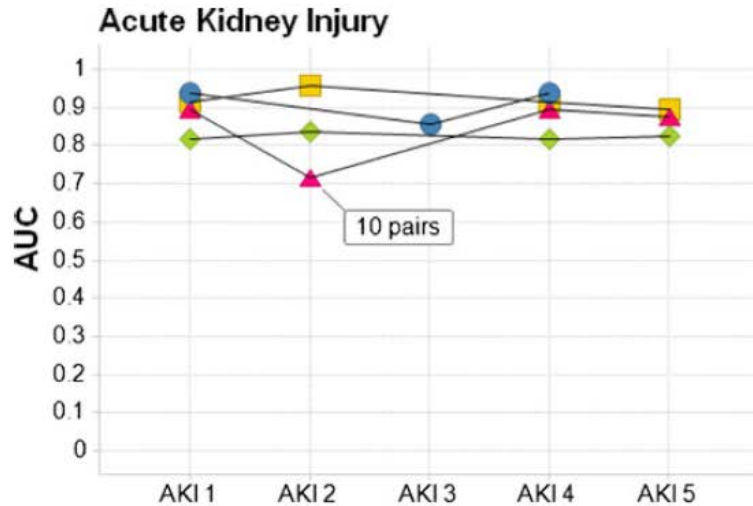
DOI 10.1007/s40264-013-0111-1

ORIGINAL RESEARCH ARTICLE

Alternative Outcome Definitions and Their Effect on the Performance of Methods for Observational Outcome Studies

**Christian G. Reich · Patrick B. Ryan ·
Martijn J. Schuemie**

OMOP finding: Restricting outcome definition did not impact predictive accuracy



▲ CCAE
● GE
◆ MDCD
■ MDCR

- Narrowing definition does not ensure proper trade-off of lower sensitivity vs. higher specificity
- PPV is irrelevant to assessment of precision and accuracy in effect estimation

Fig. 5 AUC values of alternative definitions for each of the three HOI across the four databases. Definitions with only few test cases are labeled. Upper left diagram red markers are equivalent to Fig. 4. HOI

health outcome of interest, *MDCD* MarketScan Multi-state Medicaid, *MDCR* MarketScan Medicare Supplemental Beneficiaries, *CCA* MarketScan Commercial Claims and Encounters, *GE* GE Centricity



Phenotyping is the biggest barrier on the critical path to large-scale real-world evidence generation

