



Observational Health Data Sciences and Informatics

Nigam Shah

nigam@stanford.edu

Stanford University



OHDSI

OBSERVATIONAL HEALTH DATA SCIENCES AND INFORMATICS

1. Tools
2. Methods
3. Data network



Tools – Achilles, Atlas

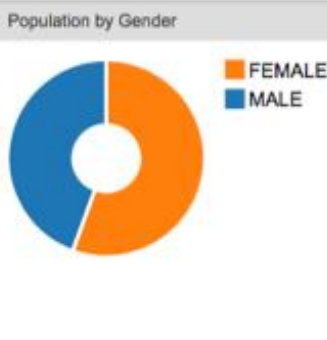
Person Summary

Source name: CMS-SYNPUF synthetic data
Number of persons: 2.33M

Year of Birth



Population by Gender



Cohort

Ischemic stroke incident inpatient events (replication of Lee et al, J Clin Psychiatry 2016)

Save Close Copy Delete

Definition Concept Sets Generation Reporting Explore Export

Cohort definition: A cohort is defined as the set of persons satisfying one or more inclusion criteria for a duration of time. One person may qualify for one cohort multiple times during non-overlapping time intervals. Cohorts are constructed in ATLAS by specifying cohort entry criteria and cohort exit criteria. Cohort entry criteria involve selecting one or more initial events, which determine the start date for cohort entry, and optionally specifying additional inclusion criteria which filter to the qualifying events. Cohort exit criteria are applied to each cohort entry record to determine the end date when the person's episode no longer qualifies for the cohort.

All Cohort Entry Criteria Cohort Exit Criteria

Initial event cohort: Events are recorded time-stamped observations for the persons, such as drug exposures, conditions, procedures, measurements and visits. All events have a start date and end date, though some events may have a start date and end date with the same value (such as procedures or measurements). The event index date is set to be equal to the event start date.

People having any of the following: [Add Initial Event...](#)

a condition occurrence of Ischemic stroke (replication of Lee et al, J C ▾) [Add](#) [Add criteria attribute...](#) [Delete Criteria](#)

✗ for the first time in the person's history

✗ occurrence start is: Between ▾ 2005-01-01 and 2010-12-31

✗ with a Visit occurrence of: ☒ Inpatient Visit [Add](#) [Import](#)

with continuous observation of at least 0 ▾ days before and 0 ▾ days after event index date

Limit initial events to: earliest event ▾ per person.

Initial event inclusion criteria: From among the initial events, include:

People having of the following criteria: [Add New Criteria...](#)

with 1 ▾ using all occurrences of:

a procedure occurrence of computed tomography (CT) or magnetic re ▾ [Add](#) [Add criteria attribute...](#) [Delete Criteria](#)

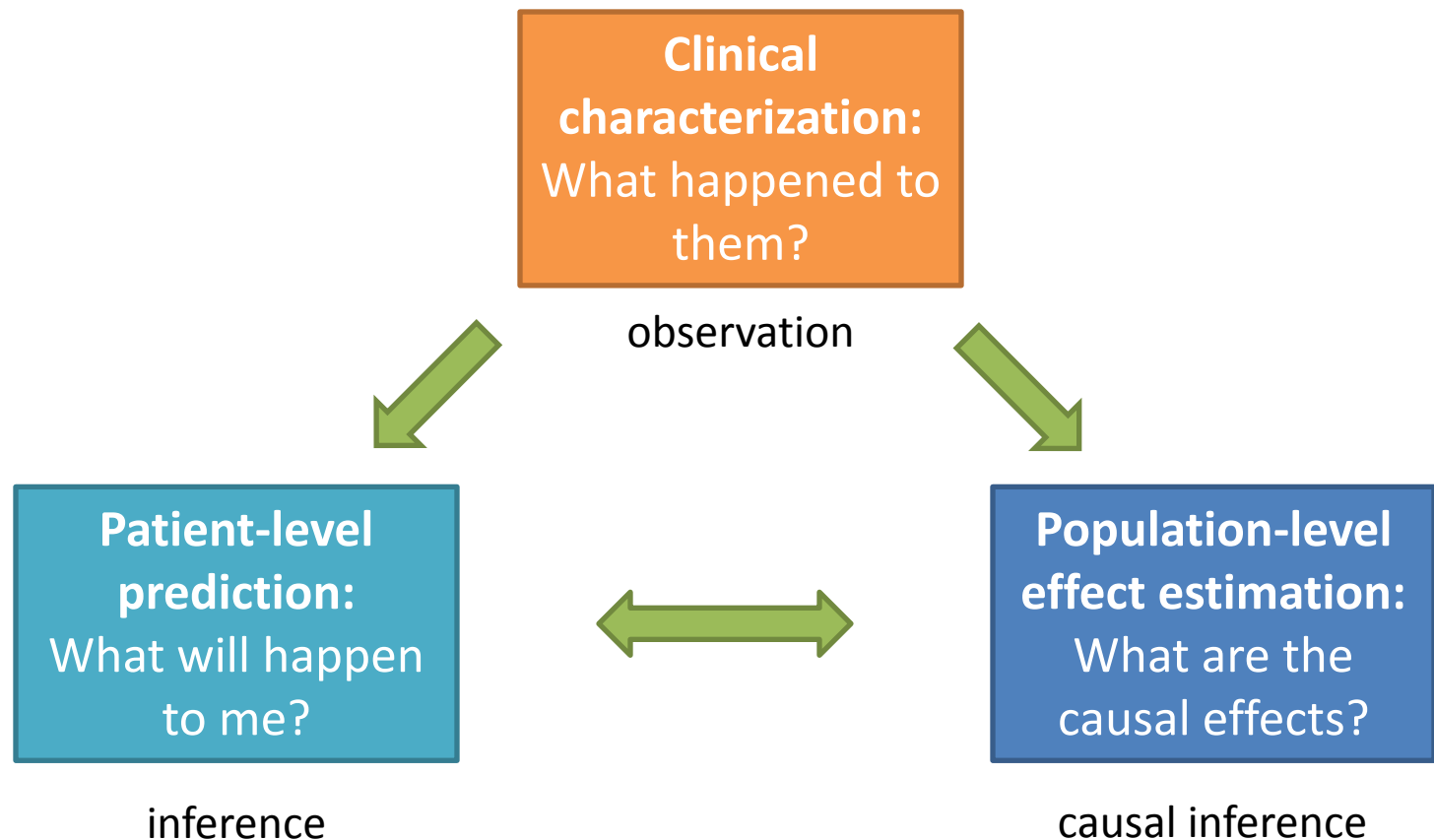
✗ with a Visit occurrence of: ☒ Inpatient Visit [Add](#) [Import](#)

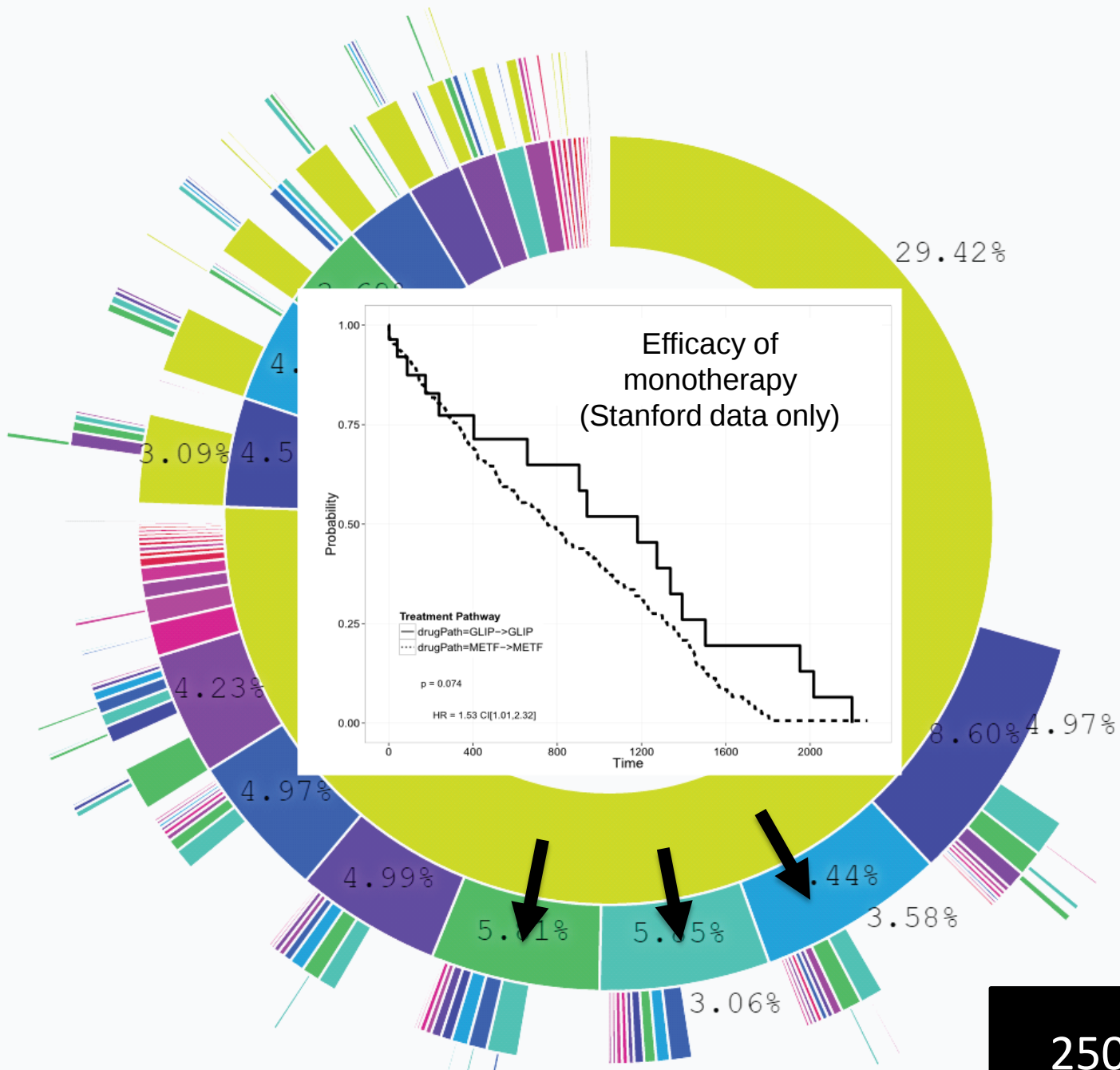
starting between 7 ▾ days Before ▾ and 7 ▾ days After ▾ event index date [and ending any time.](#)

Limit cohort of initial events to: earliest event ▾ per person.



Complementary evidence to inform the patient journey





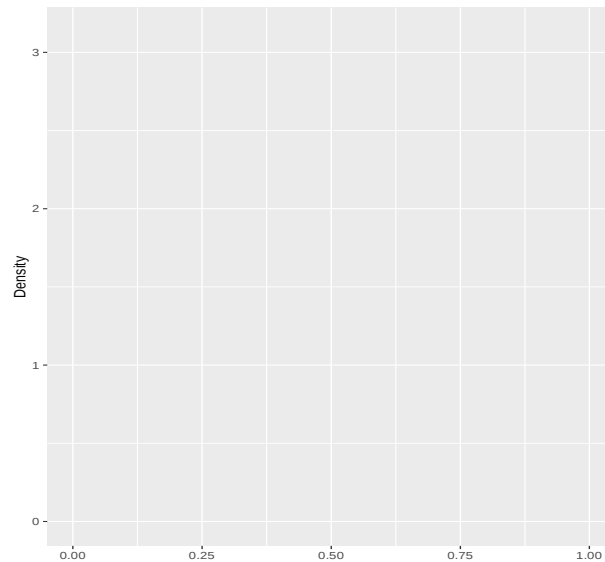
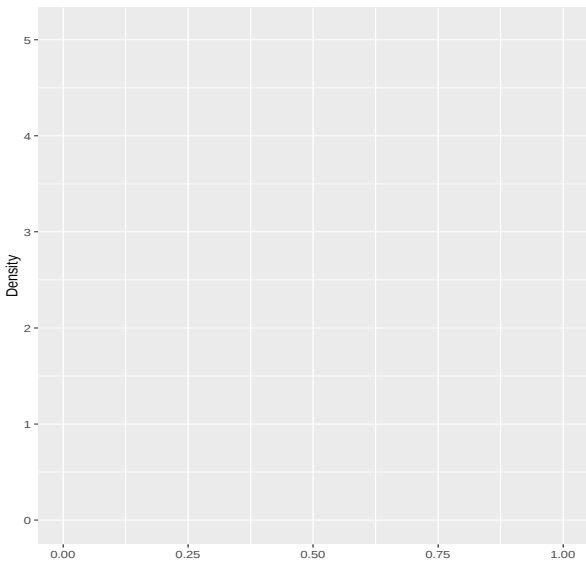
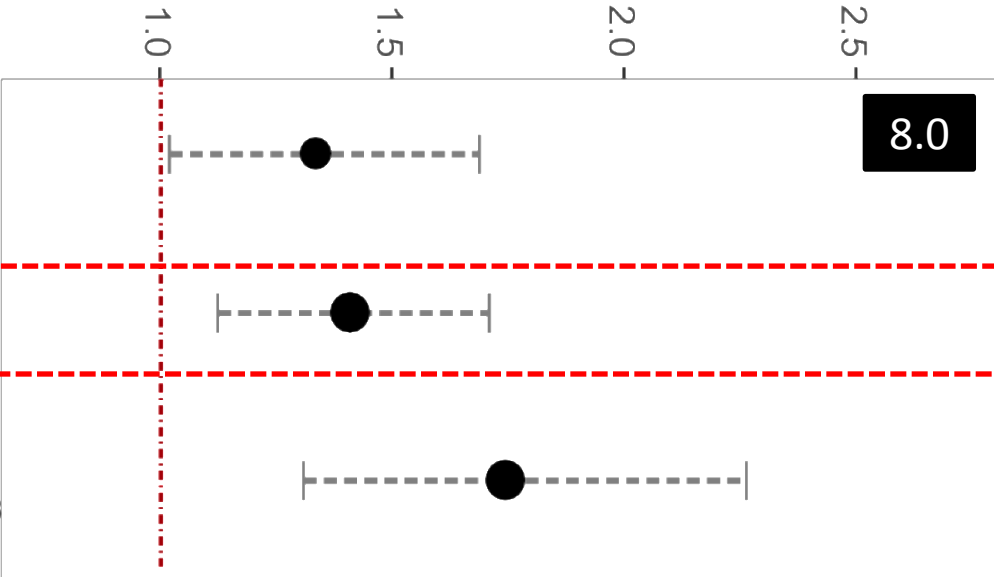
- Metformin
- pioglitazone
- sitagliptin
- Glipizide
- glimepiride
- Gliclazide
- Glyburide
- rosiglitazone
- Insulin, Glargine, Human
- exenatide
- Insulin, Aspart, Human
- liraglutide
- saxagliptin
- Insulin, Lispro, Human
- Glucose
- Insulin, Isophane, Human

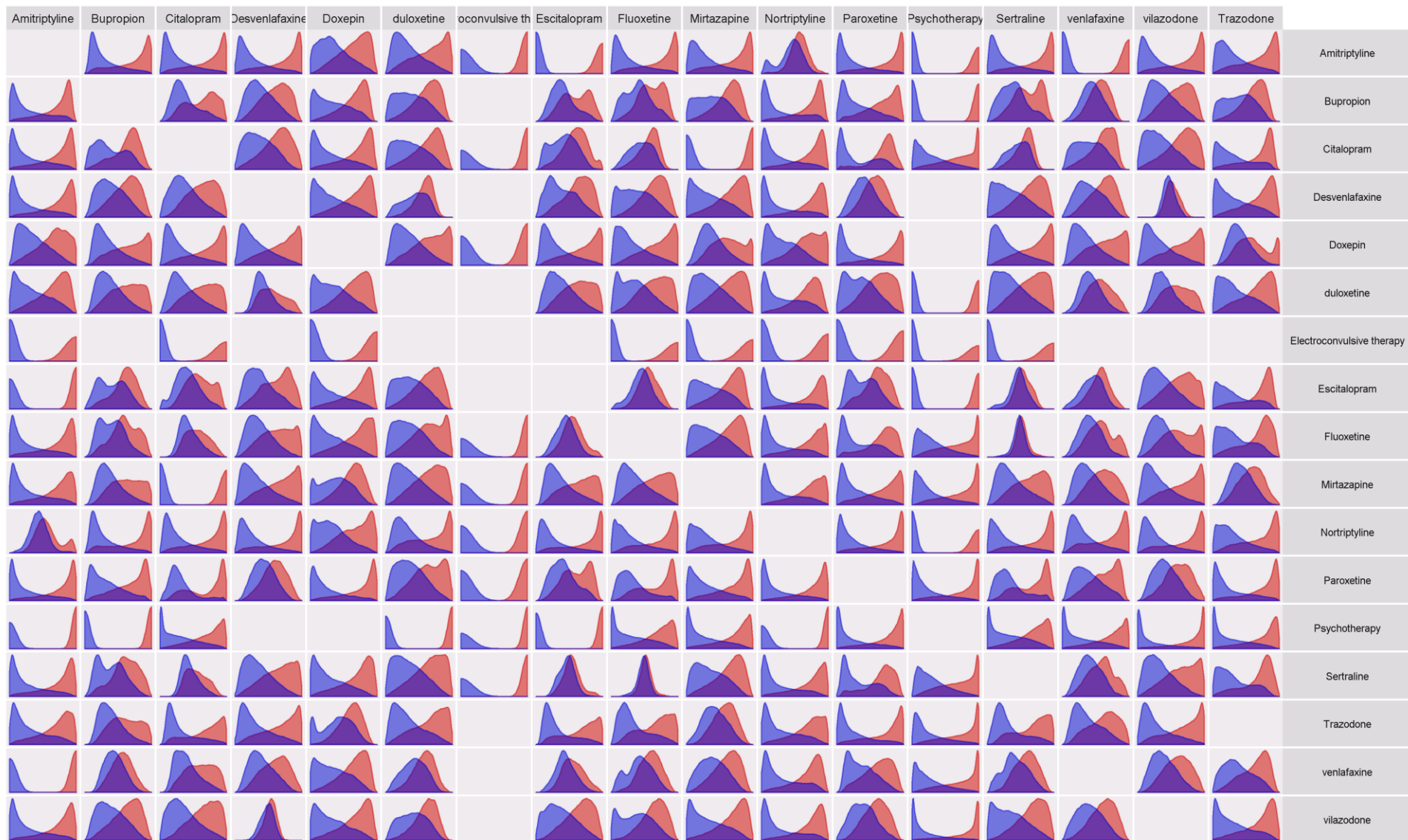
Effective treatment pathways

Metformin → Glipizide vs. Pioglitazone

Metformin → Sitagliptin vs. Glipizide

Metformin → Sitagliptin vs. Pioglitazone

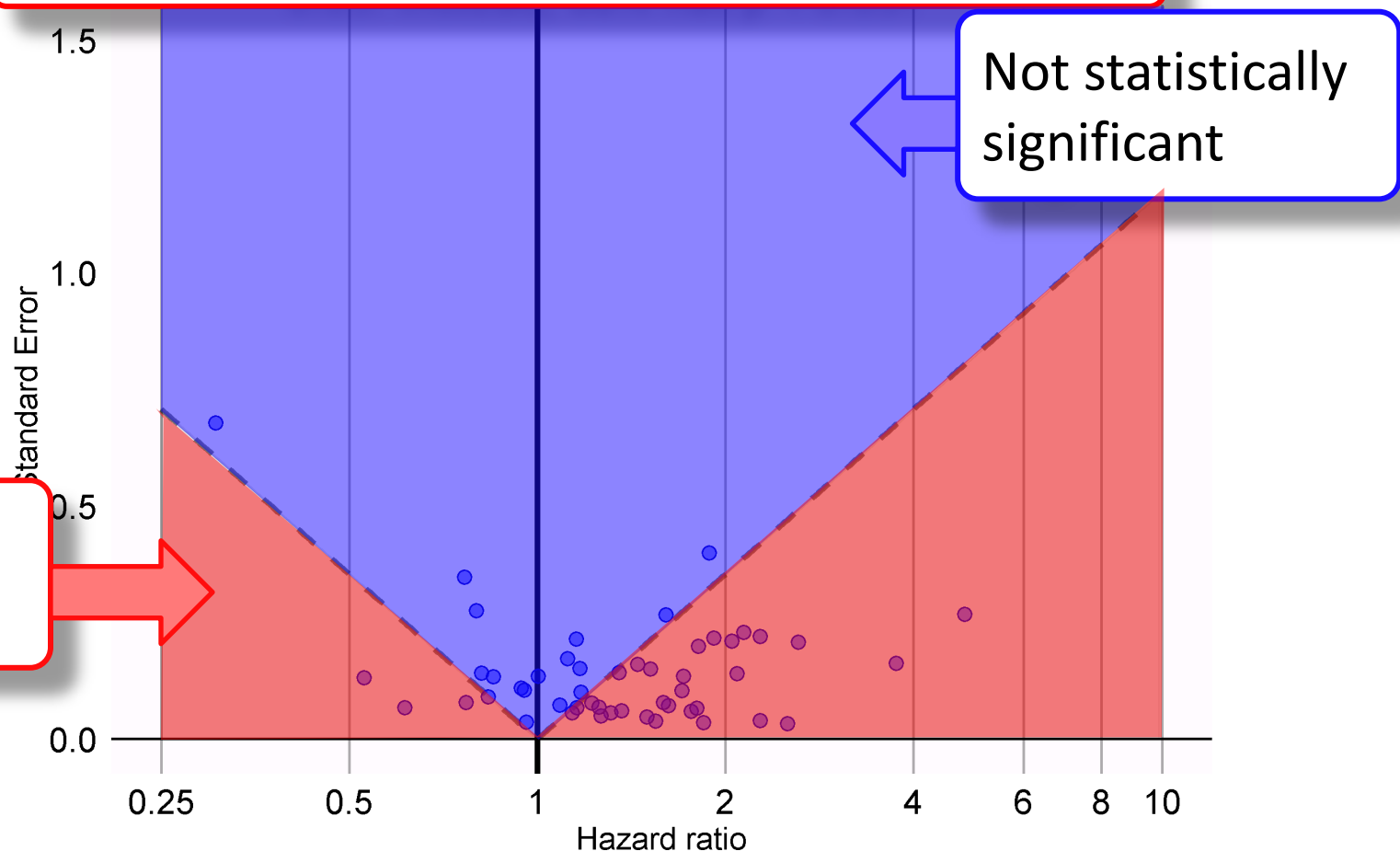






Calibration using negative controls

We would expect 5% of negative controls to have $p < 0.05$. Instead, 68% has $p < 0.05$!



Statistically significant

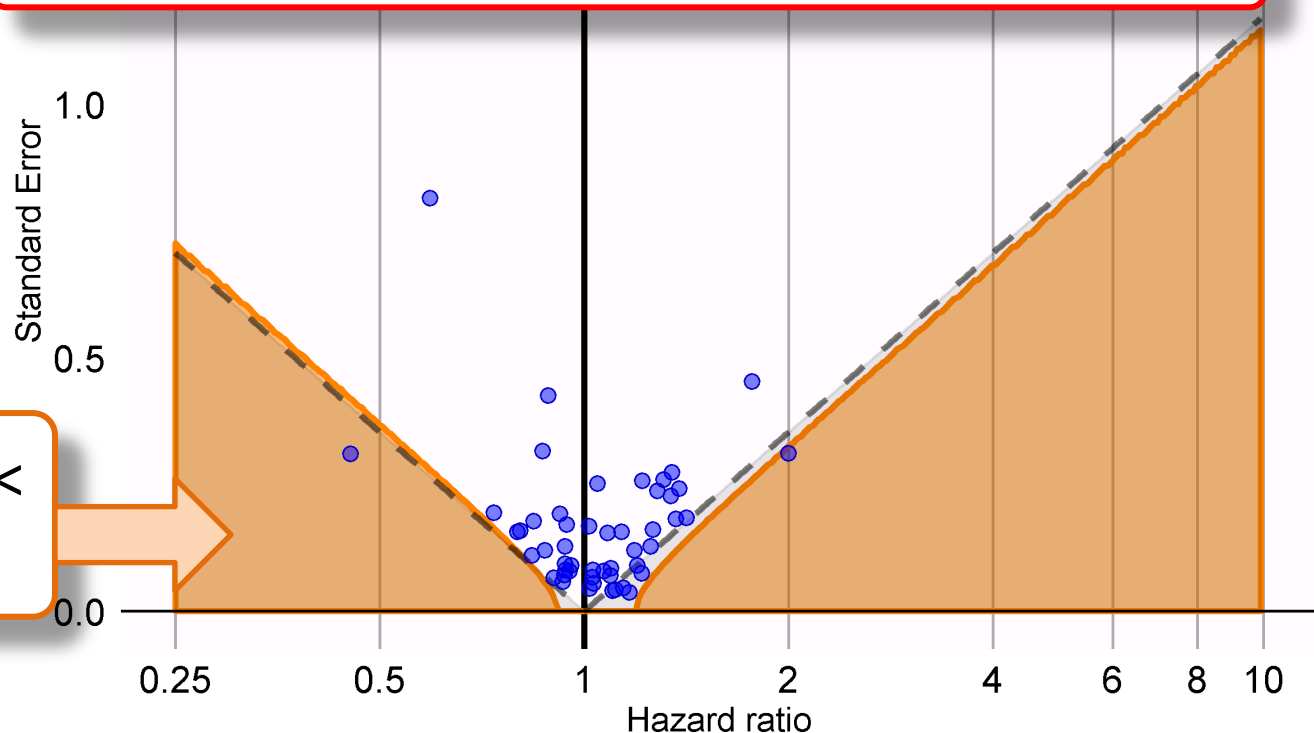
Not statistically significant



Calibration using negative controls

When using the propensity score, 16% have $p < 0.05$

After calibration, 4% have $p < 0.05$ (was 16%)



Calibrated $p < 0.05$

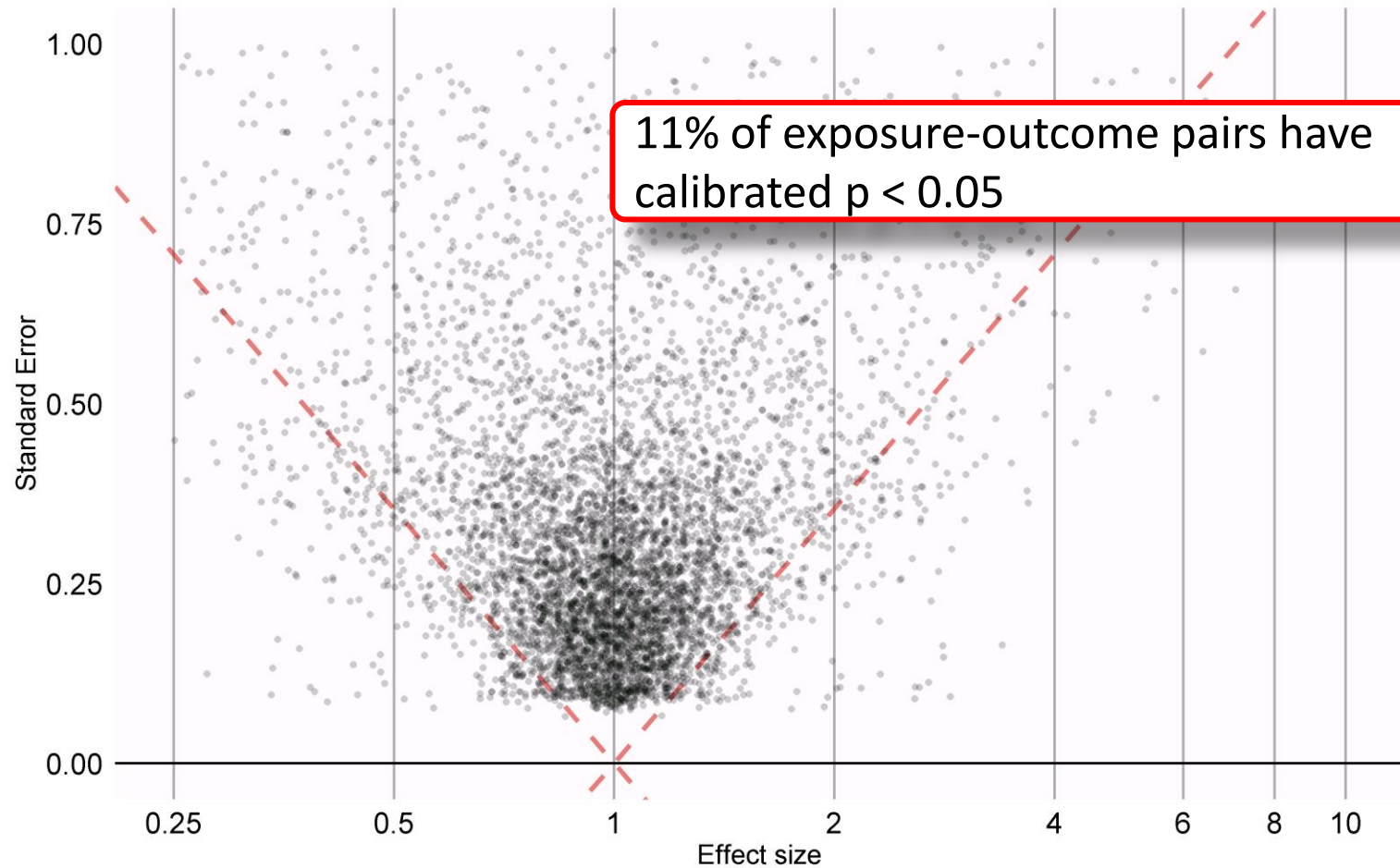


OHDSI recommendations (and tool support) for evidence generation

- ✓ Produce standard diagnostics
 - E.g. for cohort studies diagnose the propensity score distribution, covariate balance, etc.
- ✓ Include negative controls
 - Estimate the error when the null is true
- ✓ Create positive controls
 - Estimate the error when $RR > 1$
- ✓ Calibrate p-value and confidence intervals
 - Restoring nominal characteristics

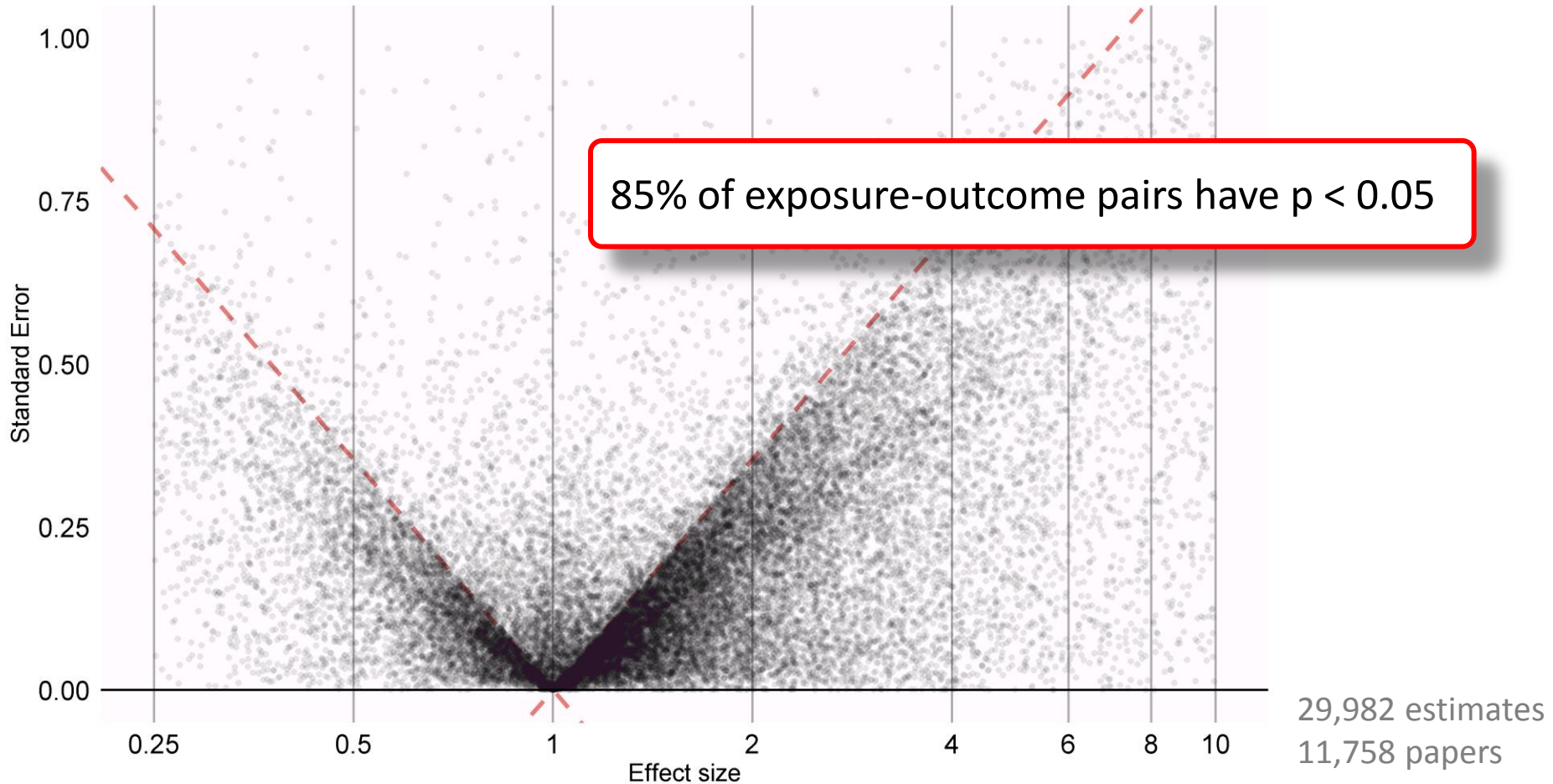


Results from all x all comparison





Observational research in literature (done one at a time)



Point of care randomization /
large simple trial



Queue / Consider
for randomization at
point of care

Clinical situation



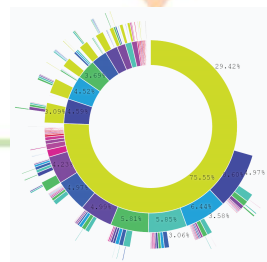
Guideline available?

Yes

**Use level A
guideline**

No

Useful byproduct



**High
priority**

Priority list of
clinical situations

Increment
priority

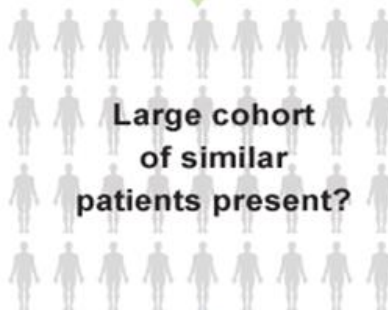
**Large cohort
of similar
patients present?**

Yes

**Use
practice-based
evidence**

No

**Use professional
judgment**





George Hripcsak
David Madigan
Patrick Ryan
Martijn Shuemie

Christian Reich
Marc Suchard
Jon Duke