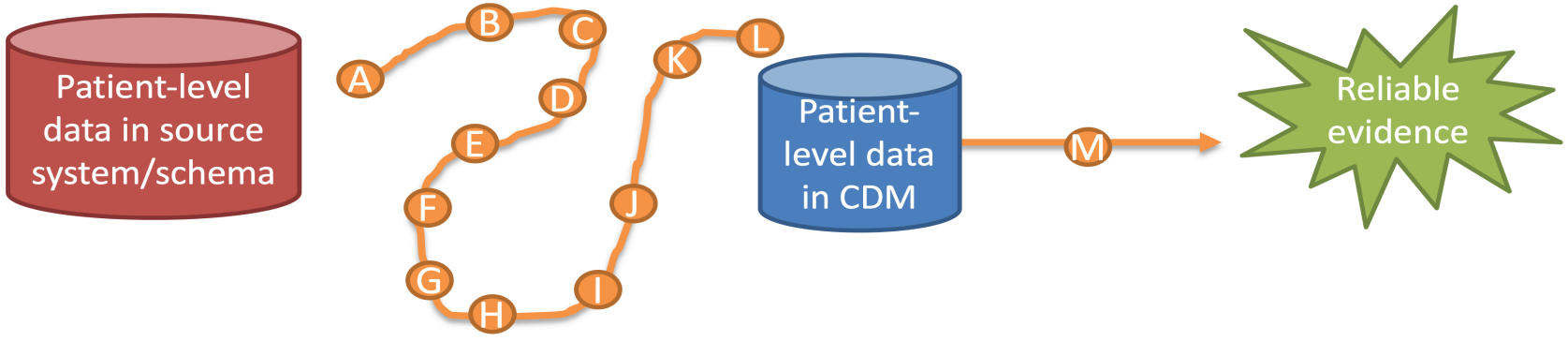


Standardized Analytics | What Evidence?

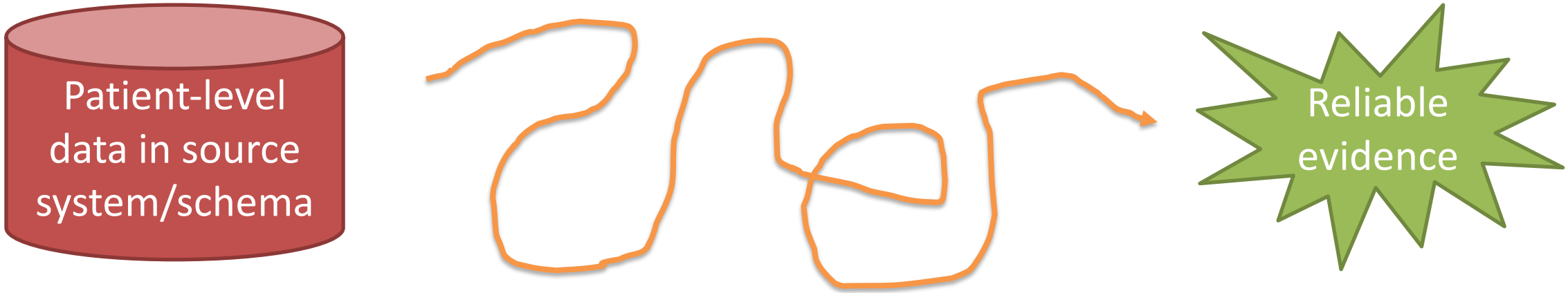
OHDSI
OBSERVATIONAL HEALTH DATA SCIENCES AND INFORMATICS

OHDSI Standardized Analytics for Evidence Generation

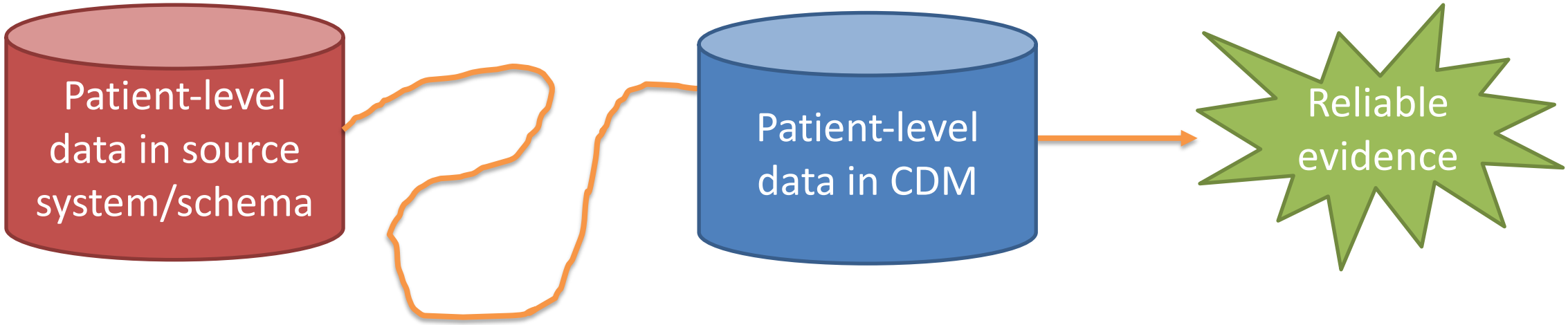
Anthony Sena, Jenna Reps, Martijn Schuemie

The journey through OHDSI's standardized analytics

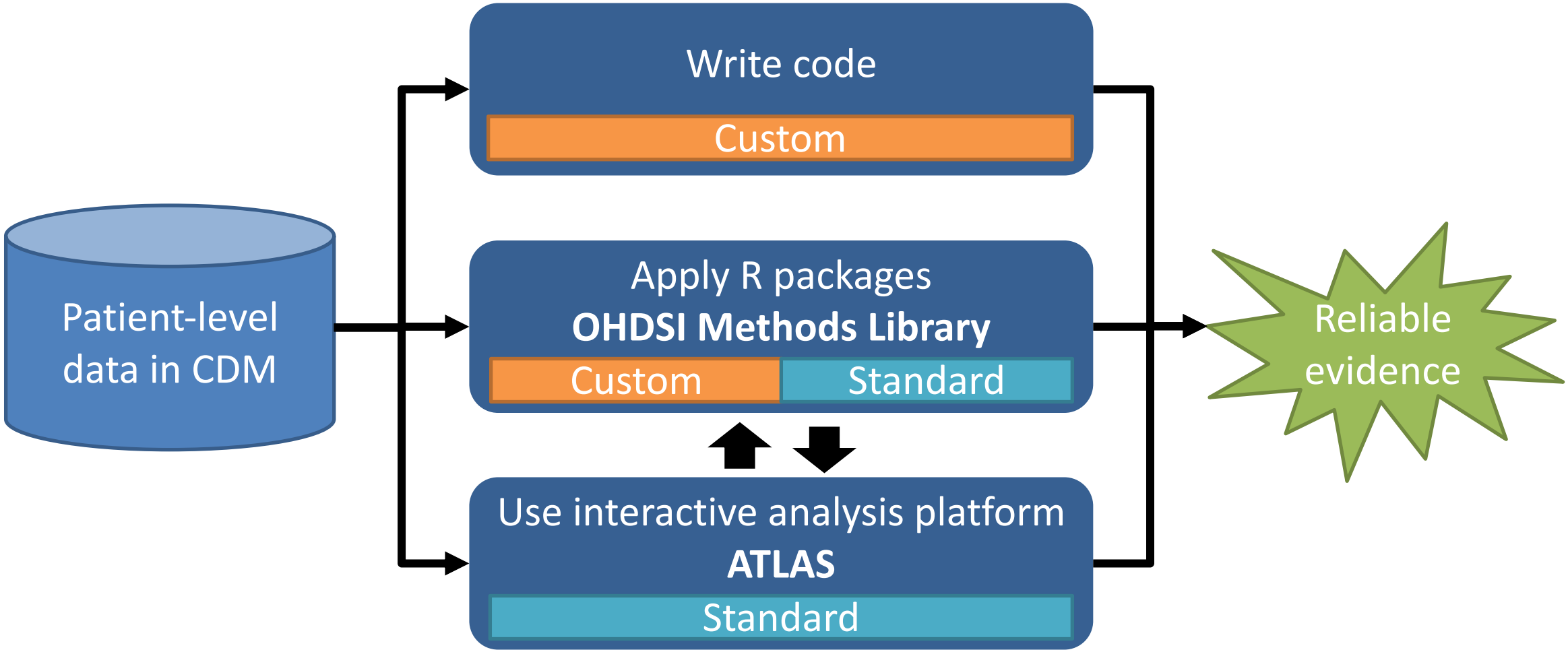
The journey to real-world evidence



The journey to real-world evidence



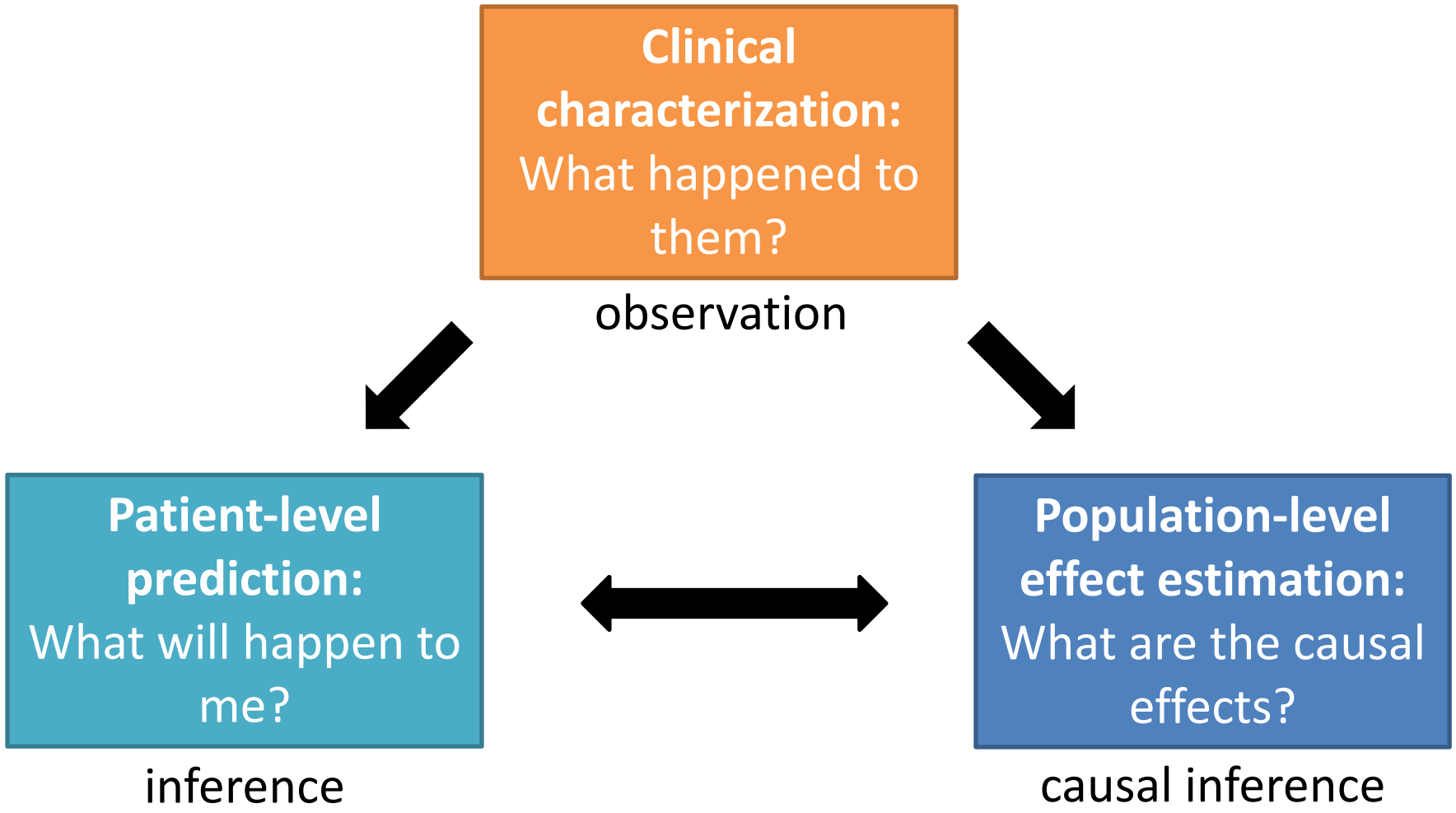
The journey to real-world evidence



Generating reliable evidence for hypertension

- Hypertension is **prevalent** and a **leading risk factor** for many health outcomes
- There are **many alternative treatments**
  - ACE, Thiazide diuretics, ARB, CCB...
  - Mono vs. combo therapy
- **Multiple potential effects**
  - Intended outcomes (e.g. reduced **acute myocardial infarction**)
  - Adverse drug reactions (e.g. increased **angioedema**)
- **Extensively studied** in clinical trials
  - What can we learn from real-world data?

Complementary evidence to inform the patient journey



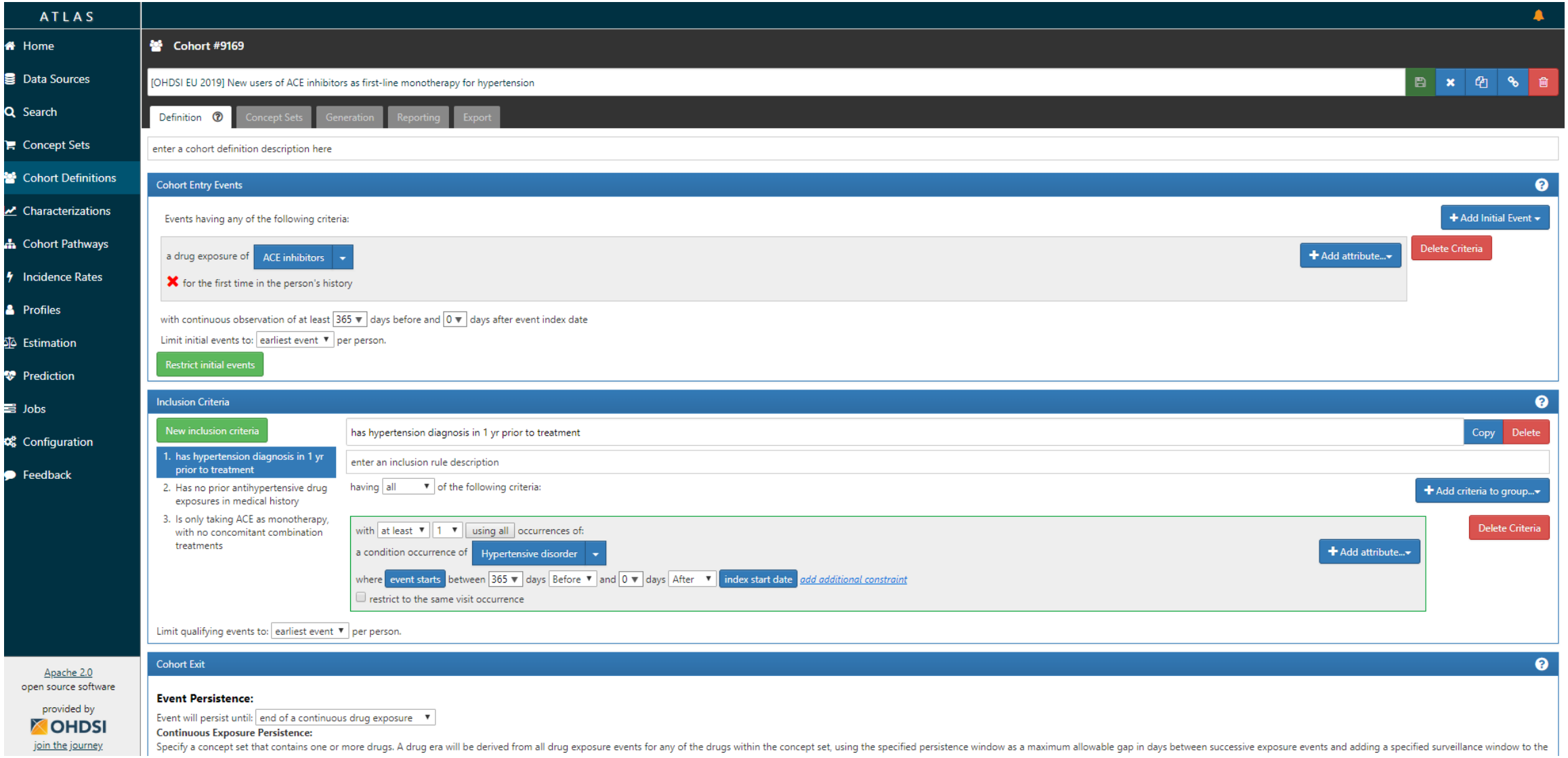
The journey through OHDSI's standardized analytics: Clinical Characterization

Anthony Sena

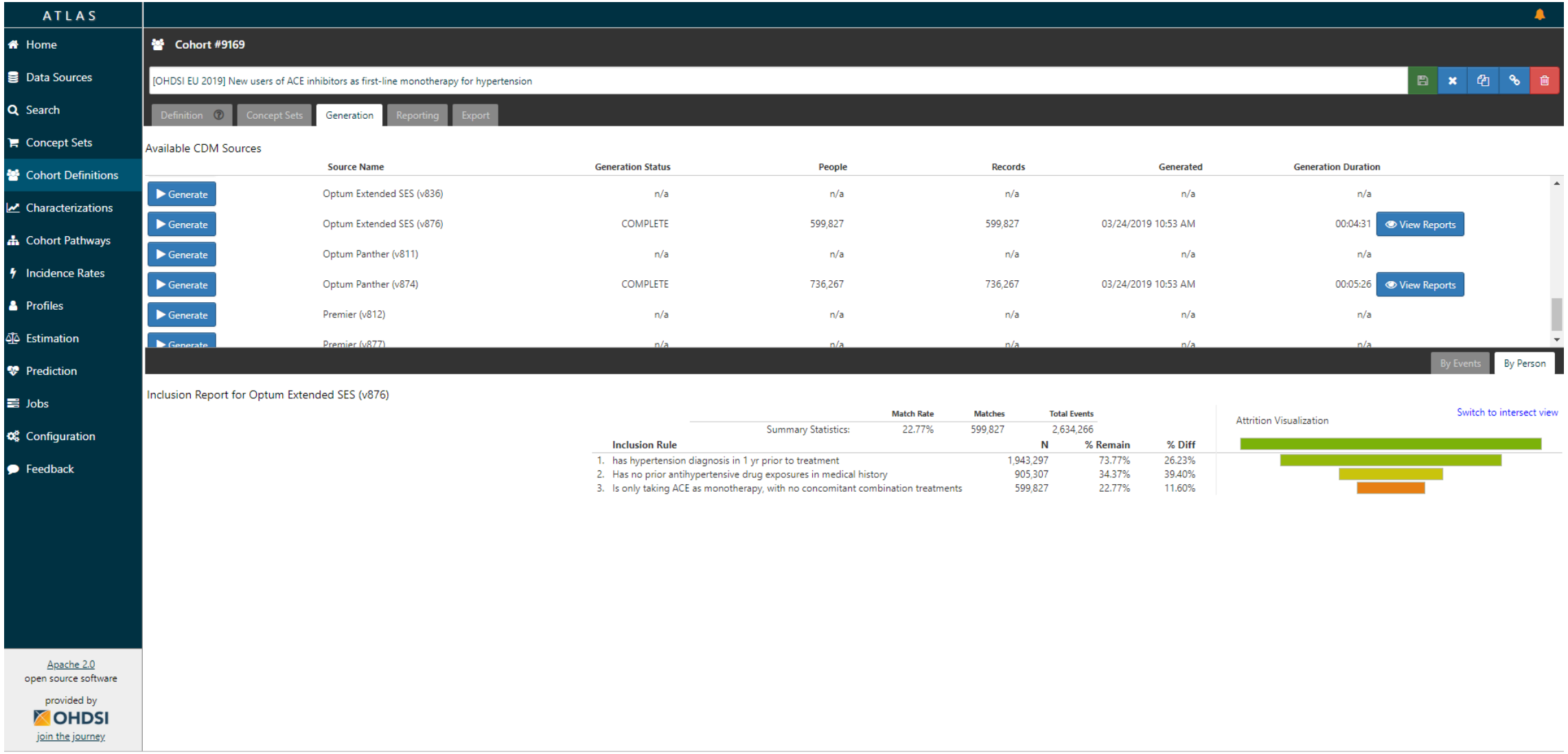
Janssen Research and Development

Erasmus MC

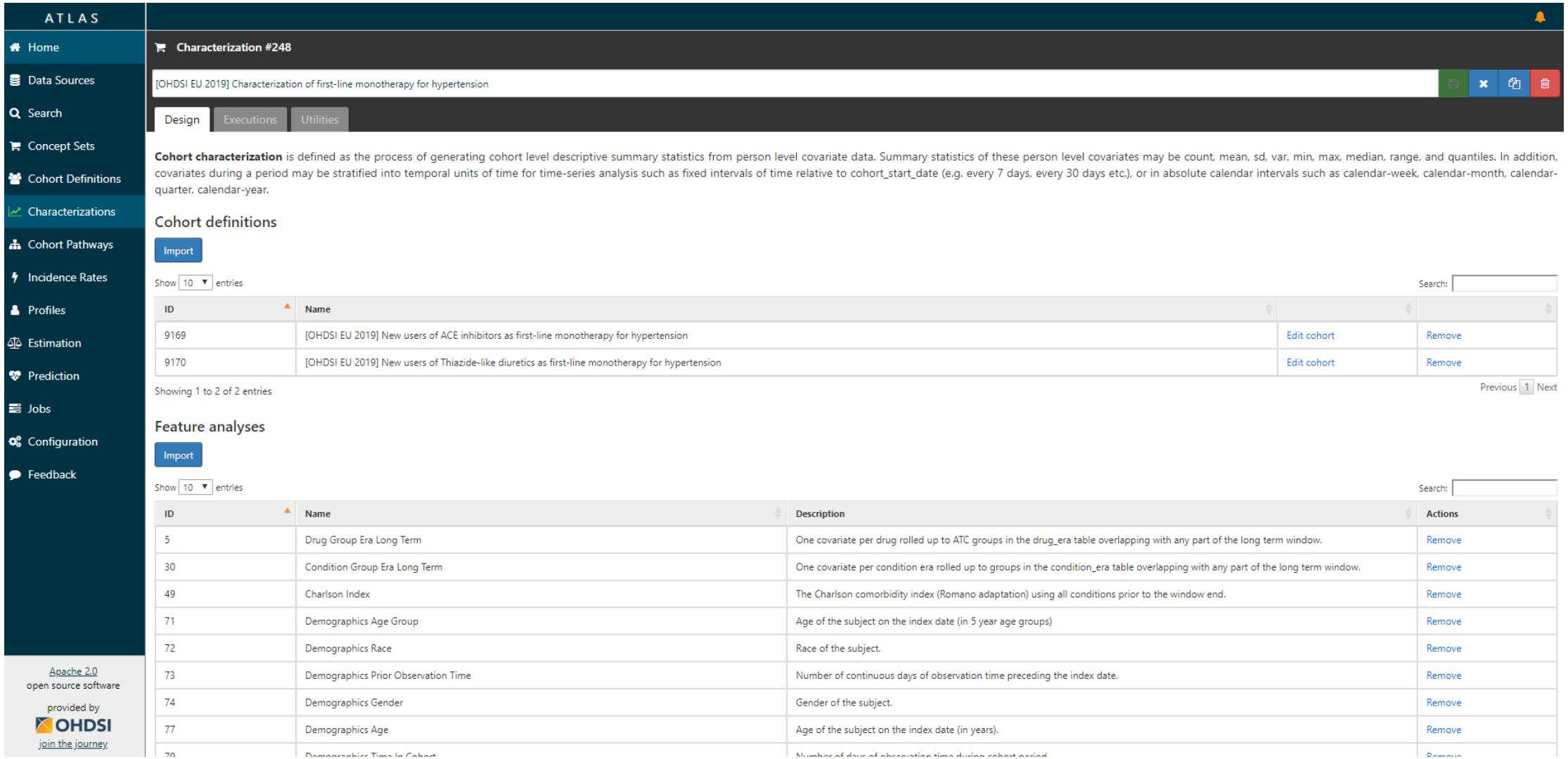
Cohort definition - design



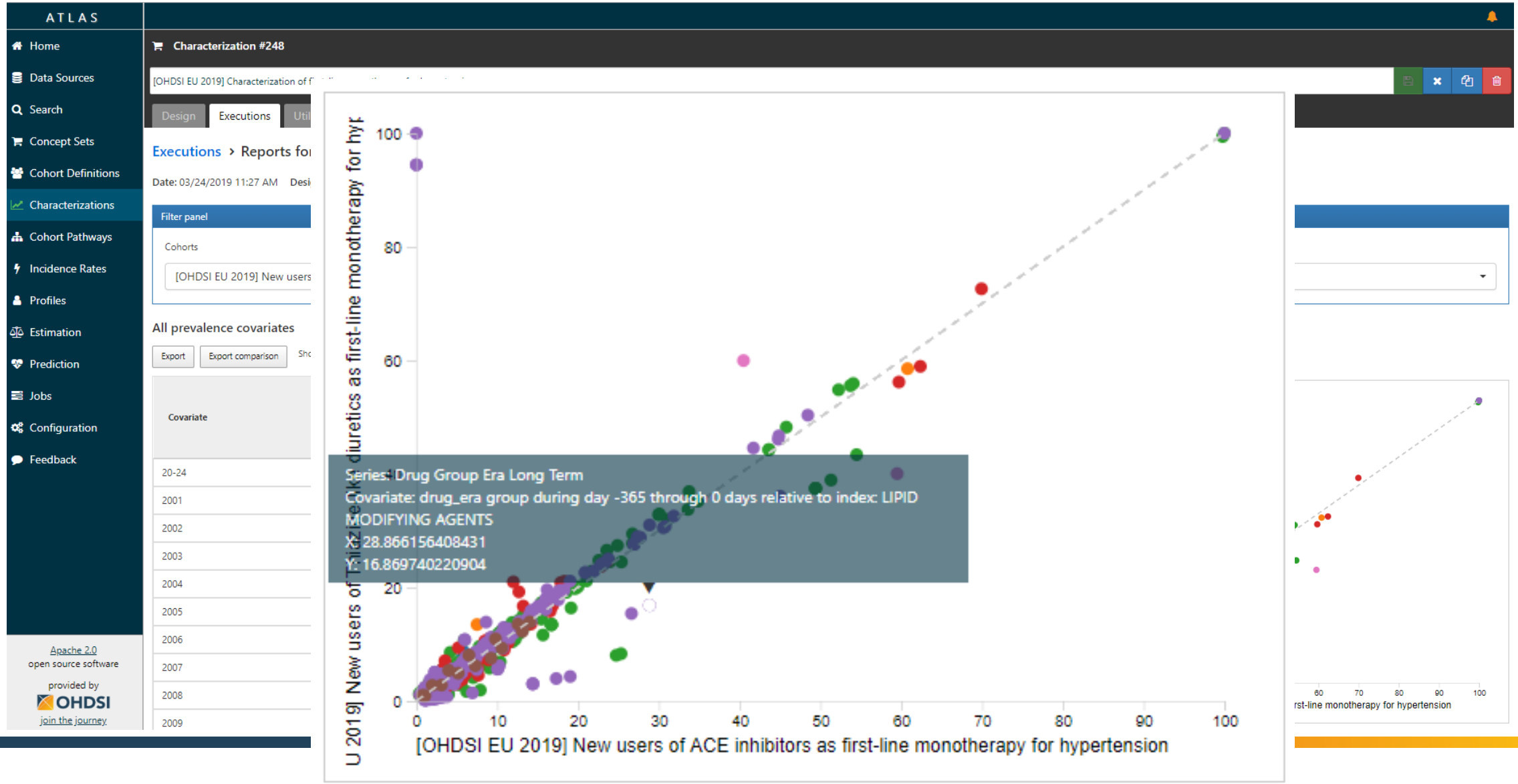
Cohort definition - results



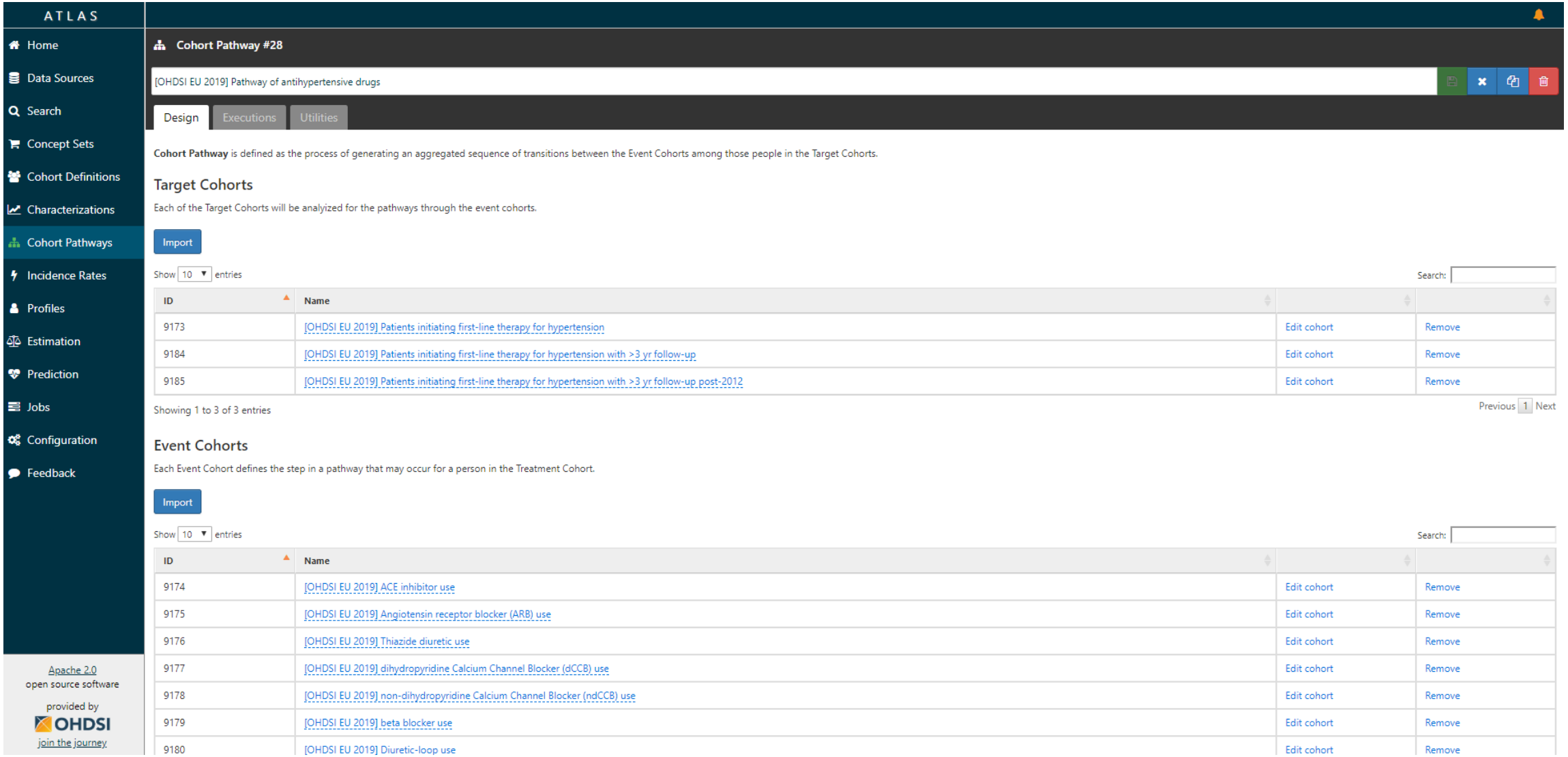
Cohort characterization - design



Cohort characterization - results



Cohort pathways - design



Cohort pathways - results

Pathways Analysis

Filter panel

Cohorts

3 items selected

Pathways Analysis for [OHDSI EU 2019] Pathway of antihypertensive drugs

Legend

Target Cohort

[OHDSI EU 2019] Patien
with >3 yr follow-up po

- Target cohort coun
- Persons with pathw
- Persons with pathw

Event Cohorts

- [OHDSI EU 2019]
- [OHDSI EU 2019]
- [OHDSI EU 2019]
- [OHDSI EU 2019]
- [OHDSI EU 2019]
- [OHDSI EU 2019]
- [OHDSI EU 2019]
- [OHDSI EU 2019] (dCCB) use
- [OHDSI EU 2019] (ndCCB) use

% Diff

54.3%

Legend

Target Cohort

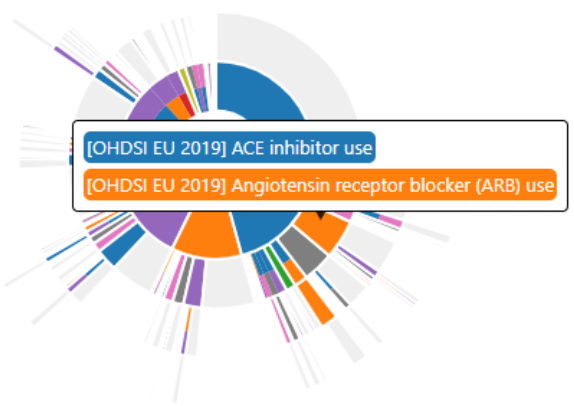
[OHDSI EU 2019] Patients initiating first-line therapy for hypertension with >3 yr follow-up post-2012

- Target cohort count: 157,234
- Persons with pathways count: 157,234
- Persons with pathways portion: 100.0%

Event Cohorts

- [OHDSI EU 2019] ACE inhibitor use
- [OHDSI EU 2019] Angiotensin receptor blocker (ARB) use
- [OHDSI EU 2019] Diuretic-loop use
- [OHDSI EU 2019] Diuretic-potassium sparing use
- [OHDSI EU 2019] Thiazide diuretic use
- [OHDSI EU 2019] alpha-1 blocker use
- [OHDSI EU 2019] beta blocker use
- [OHDSI EU 2019] dihydropyridine Calcium Channel Blocker (dCCB) use
- [OHDSI EU 2019] non-dihydropyridine Calcium Channel Blocker (ndCCB) use

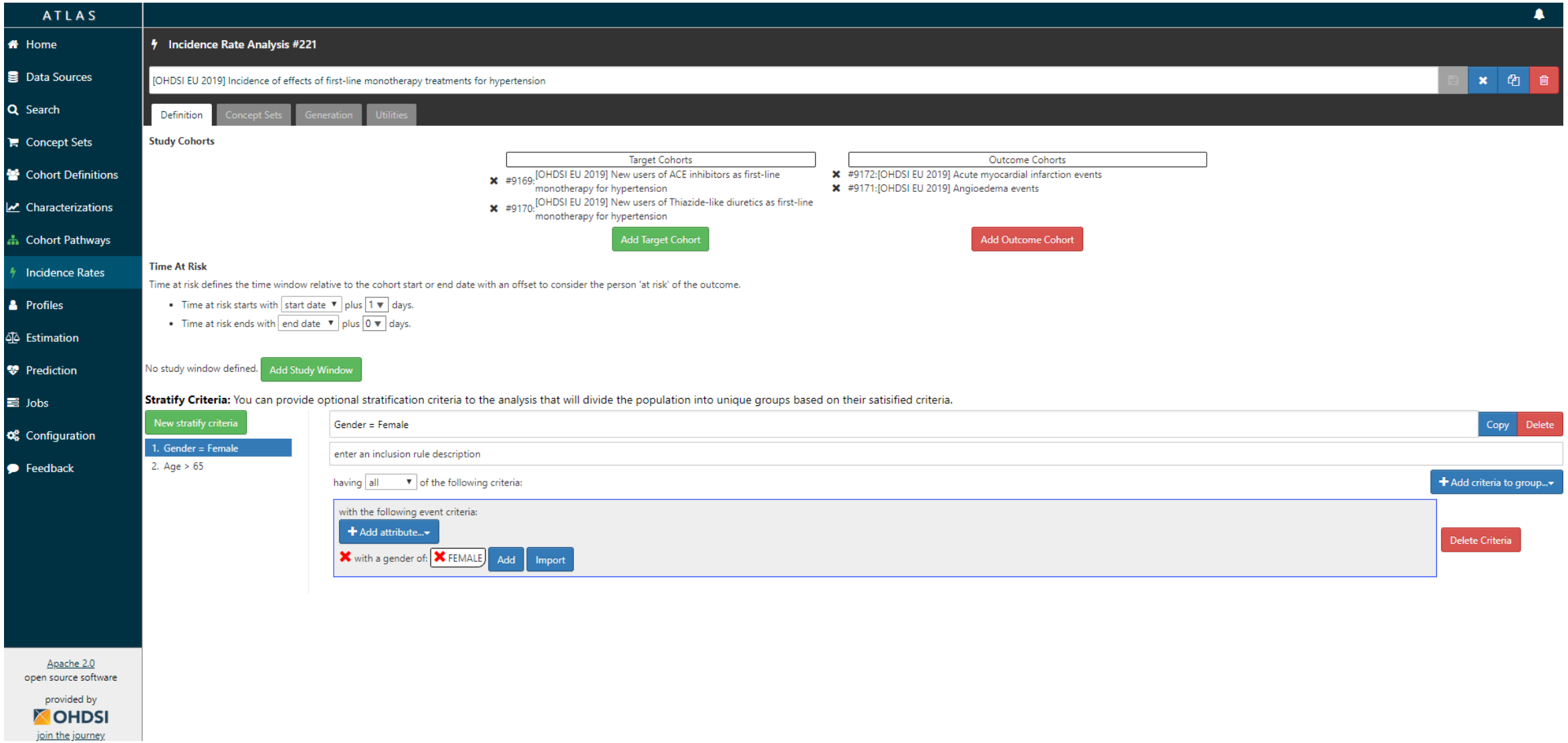
Sunburst plot



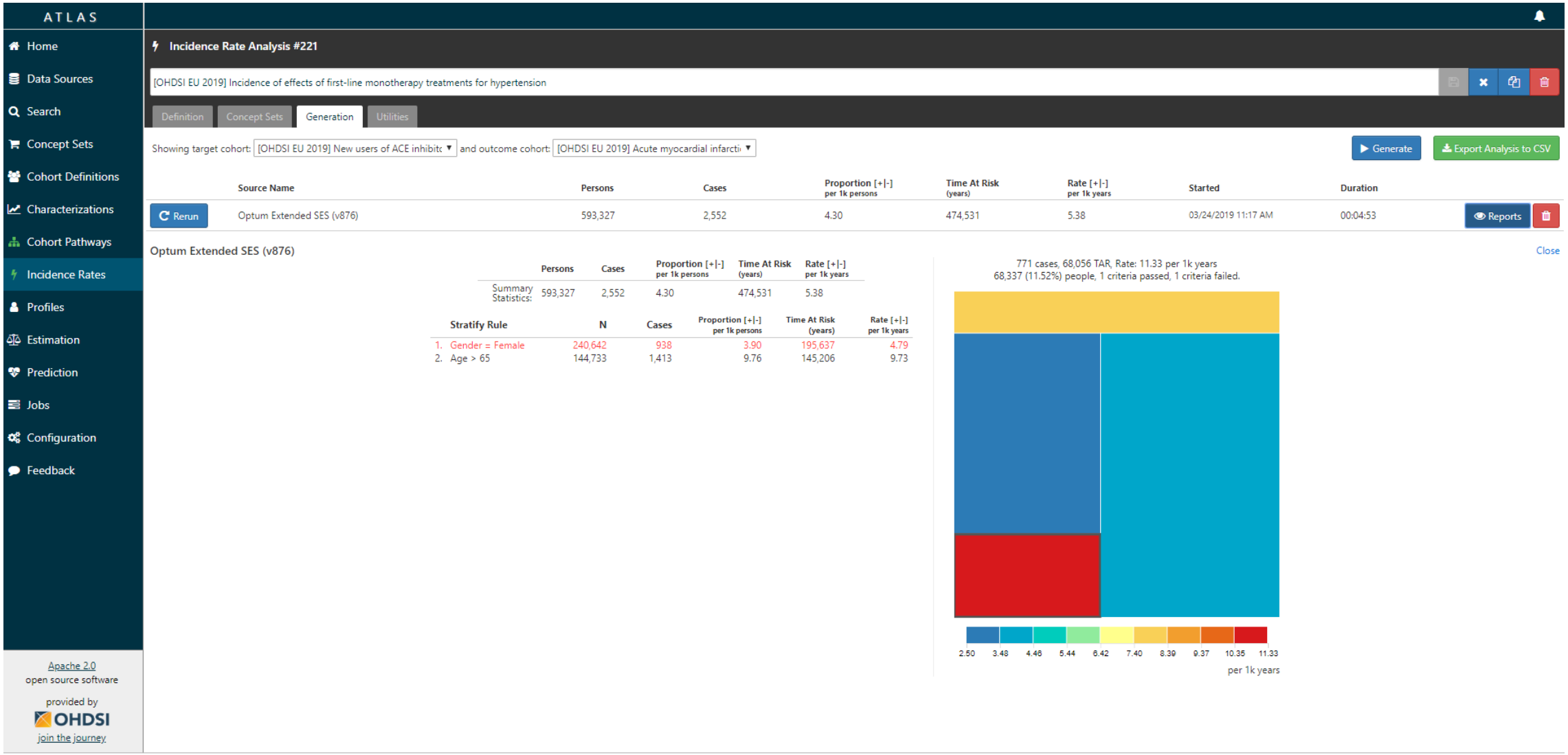
Path details

Name	% Remain	% Diff
[OHDSI EU 2019] ACE inhibitor use	45.7%	54.3%
[OHDSI EU 2019] Angiotensin receptor blocker (ARB) use	4.3%	41.5%

Incidence rate - design



Incidence rate - results



The journey through OHDSI's standardized analytics: Patient-level Prediction

Jenna Reps, PhD

Janssen Research and Development

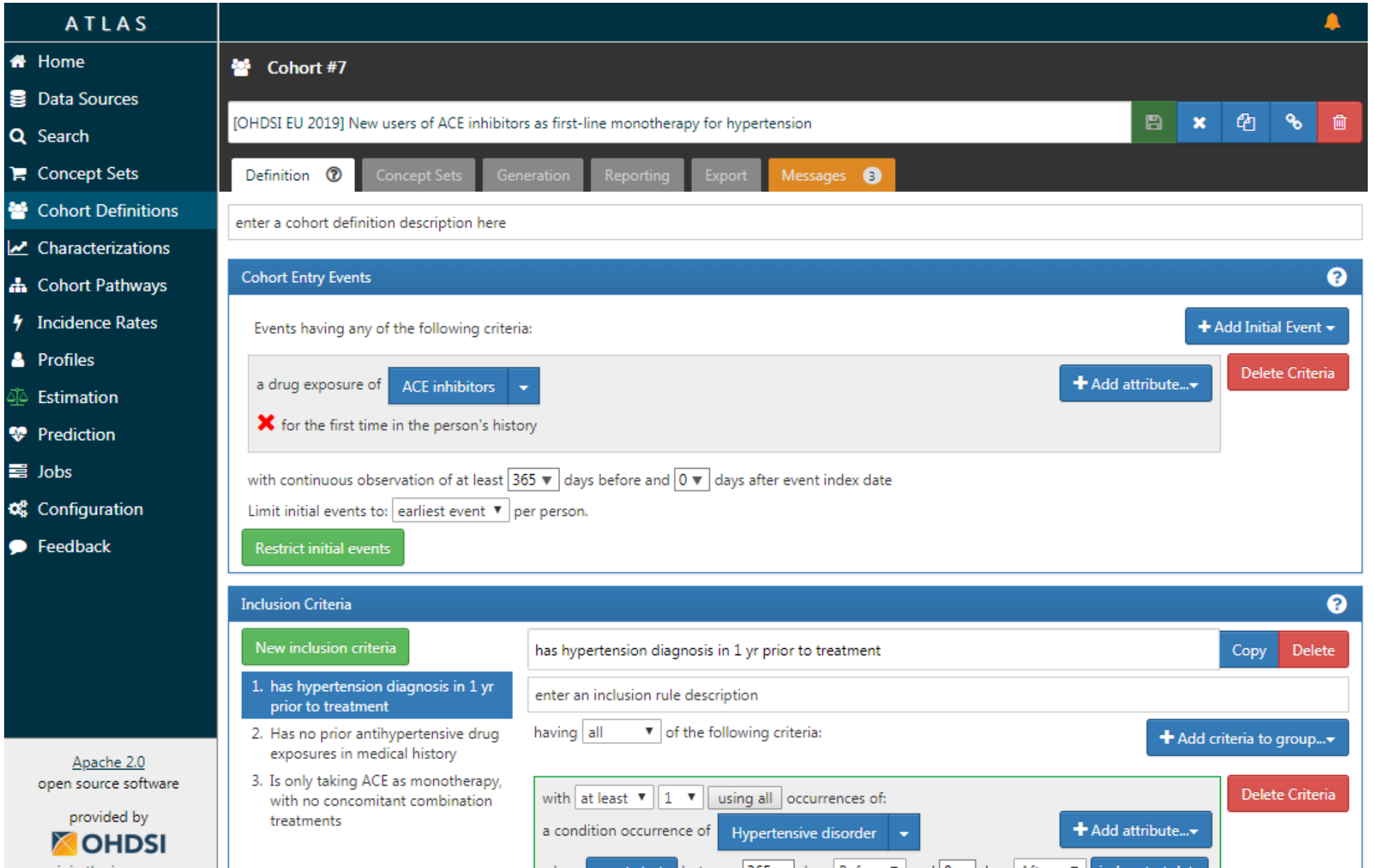
Journey through Patient-Level Prediction

- At the point in time that I am first given ACE inhibitors for hypertension, what is my risk of experiencing angioedema or acute myocardial infarction within a year?

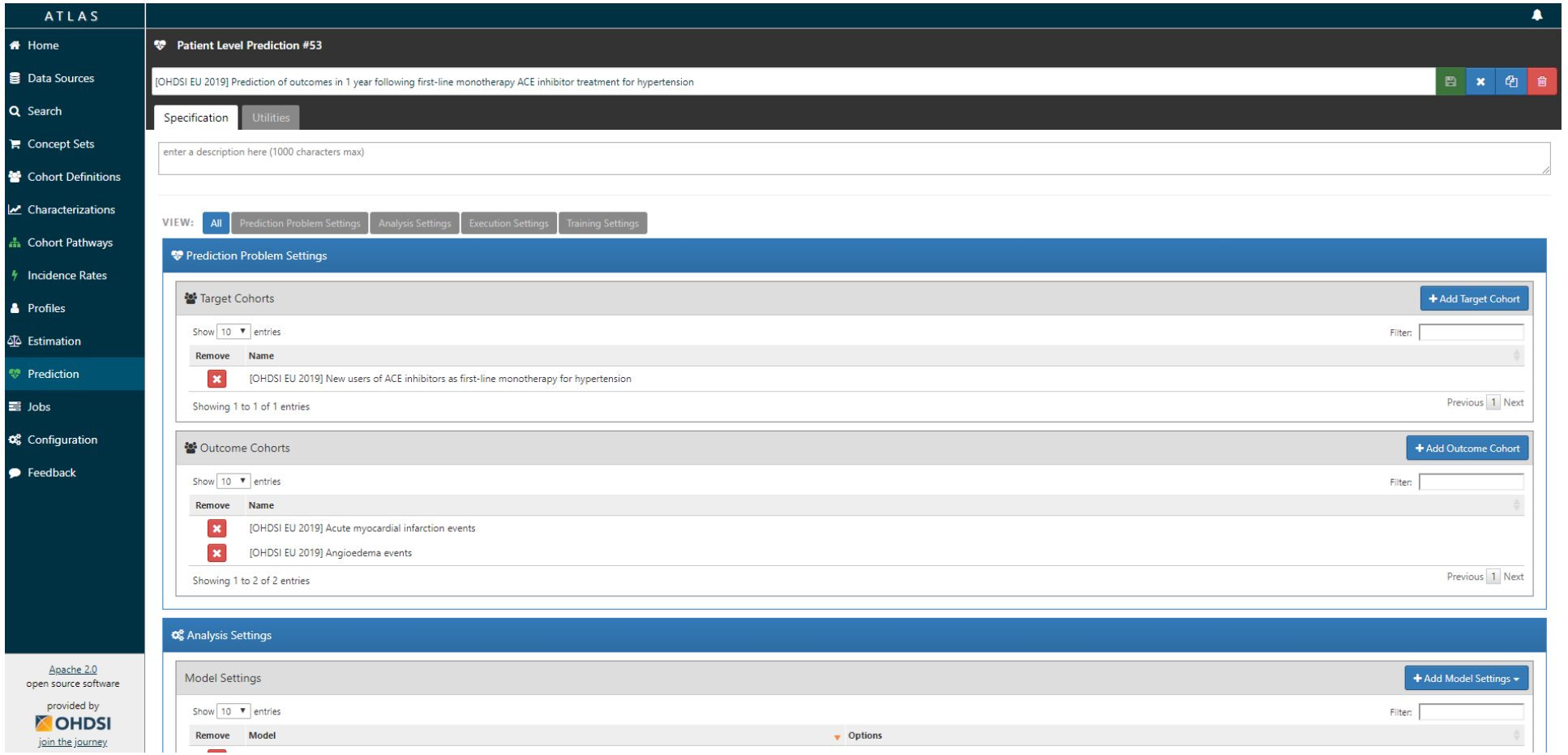
Characterization can show **Population risk**

Patient-Level Prediction can show **my risk**

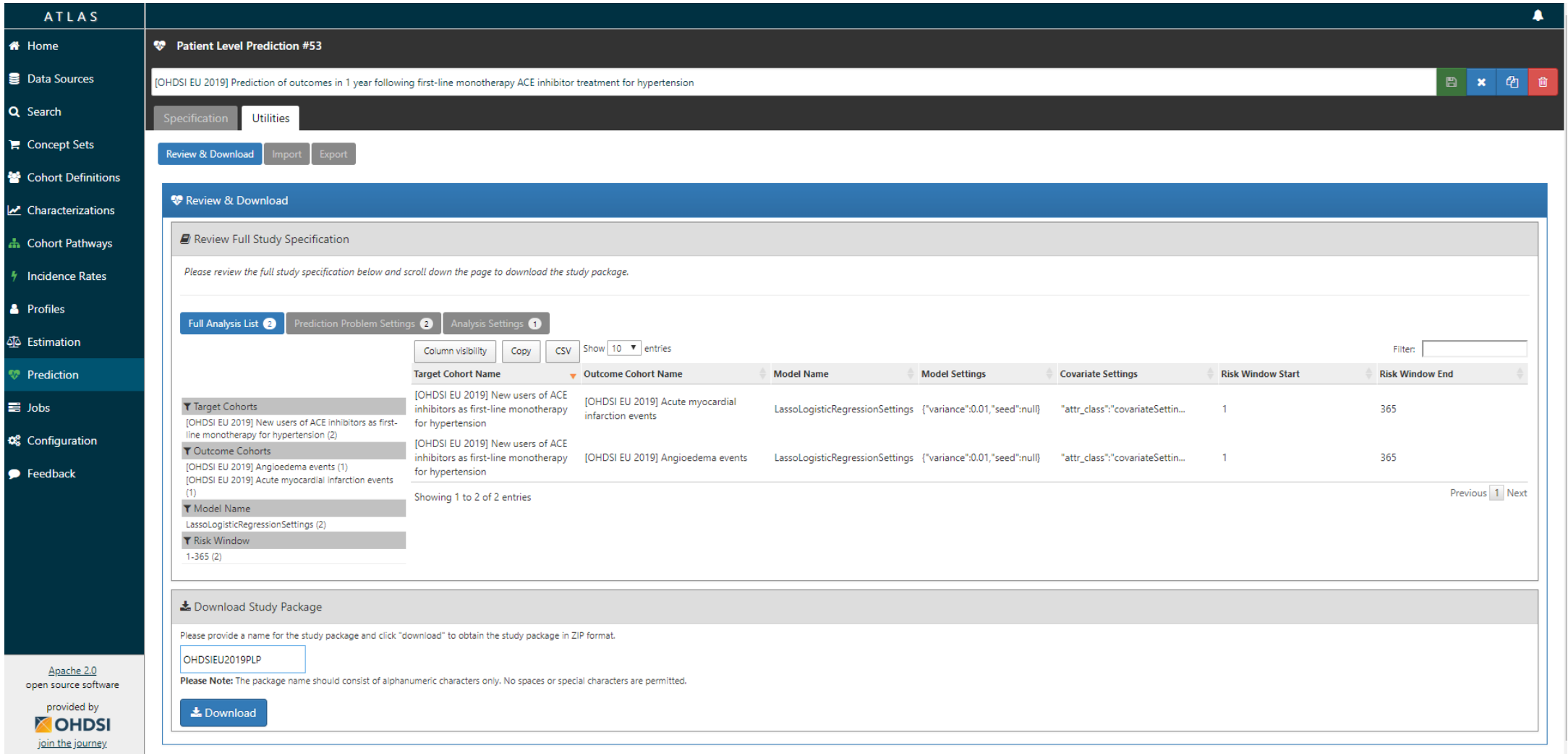
Define cohorts



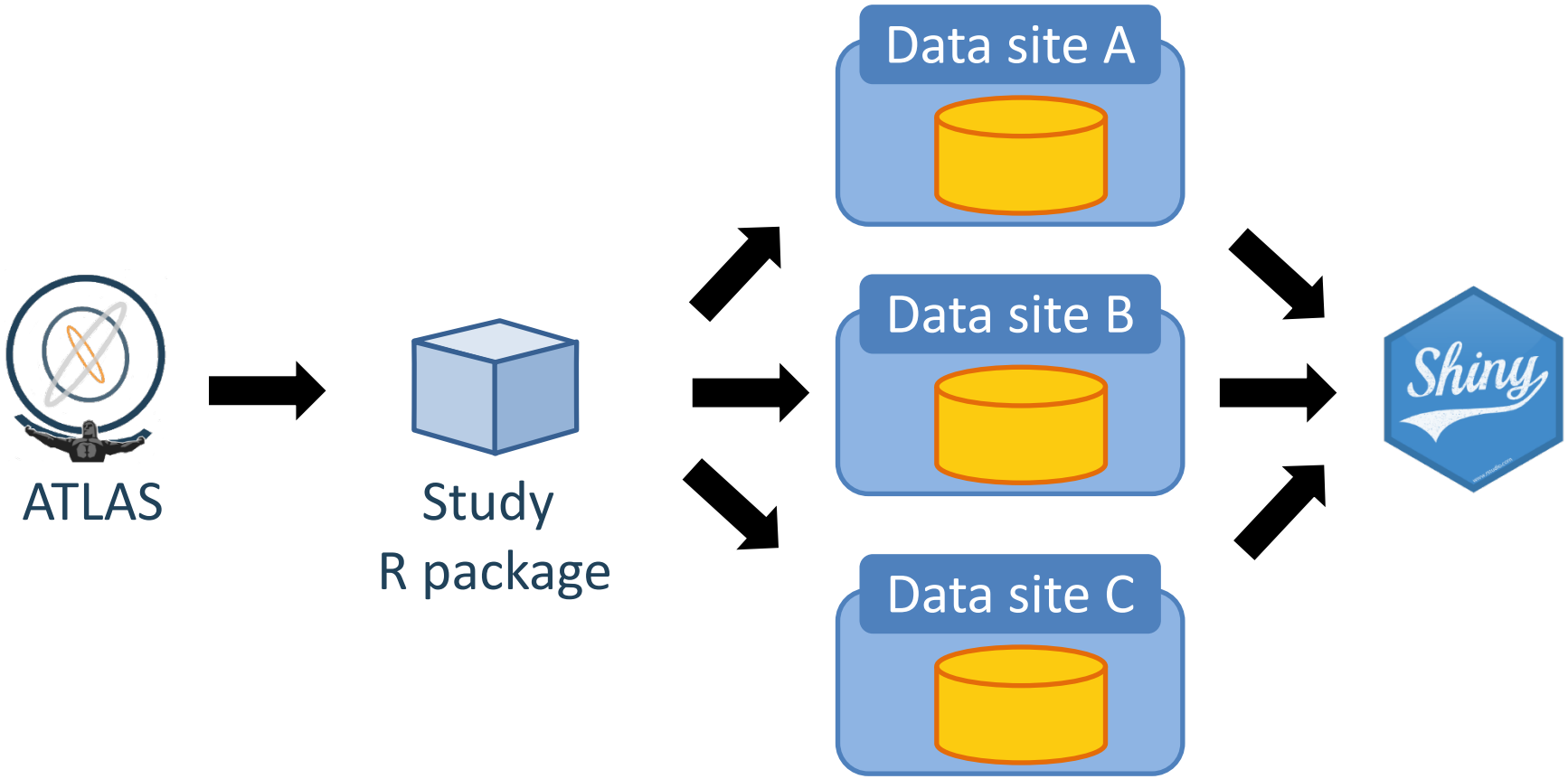
Patient-level prediction- design



Patient-level prediction- generation



Study design to results



Results: Overview

Results | Model Settings | Population Settings | Covariate Settings

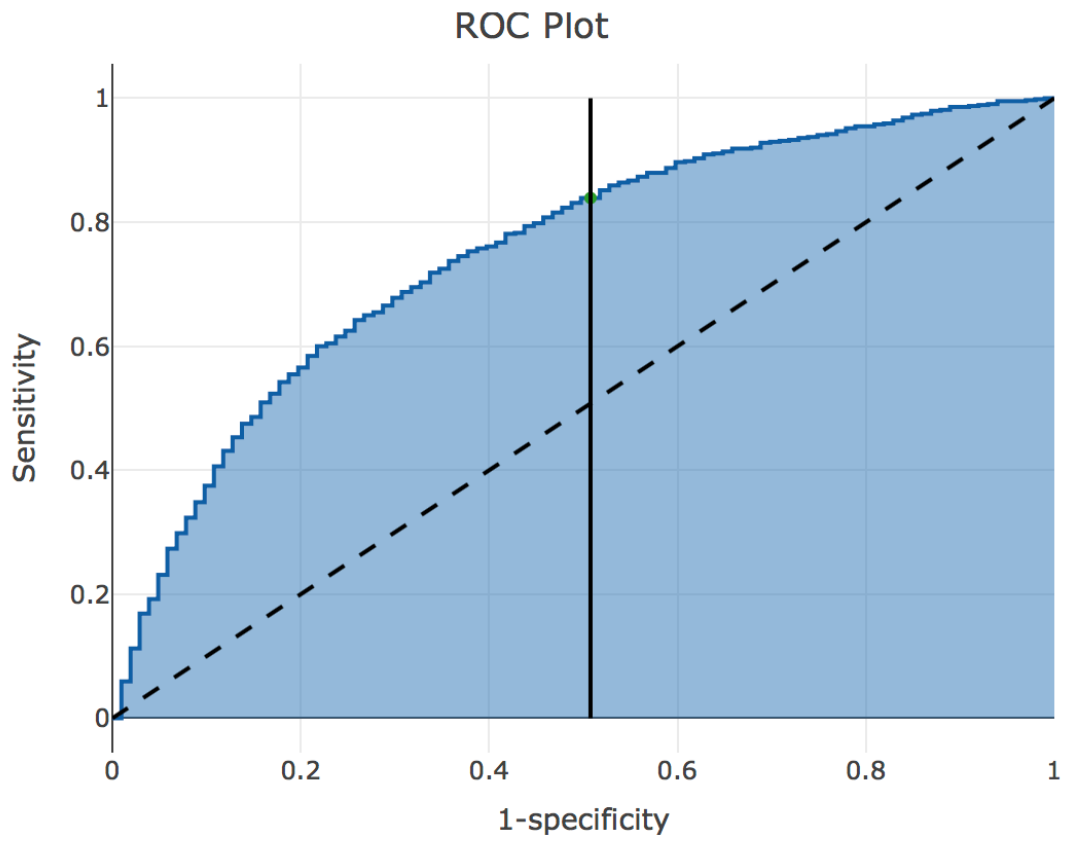
Show 10 entries Search:

Analysis	Dev	Val	T	O	Model	TAR start	TAR end	AUC	AUPRC	T Size	O Count	O Incidence (%)
Analysis_1	friendly database name	friendly database name	[OHDSI EU 2019] New users of ACE inhibitors as first-line monotherapy for hypertension	[OHDSI EU 2019] Acute myocardial infarction events	Lasso Logistic Regression	1	365	0.75642	0.02763	87772	640	0.72916
Analysis_2	friendly database name	friendly database name	[OHDSI EU 2019] New users of ACE inhibitors as first-line monotherapy for hypertension	[OHDSI EU 2019] Angioedema events	Lasso Logistic Regression	1	365	0.65238	0.00355	87637	142	0.16203

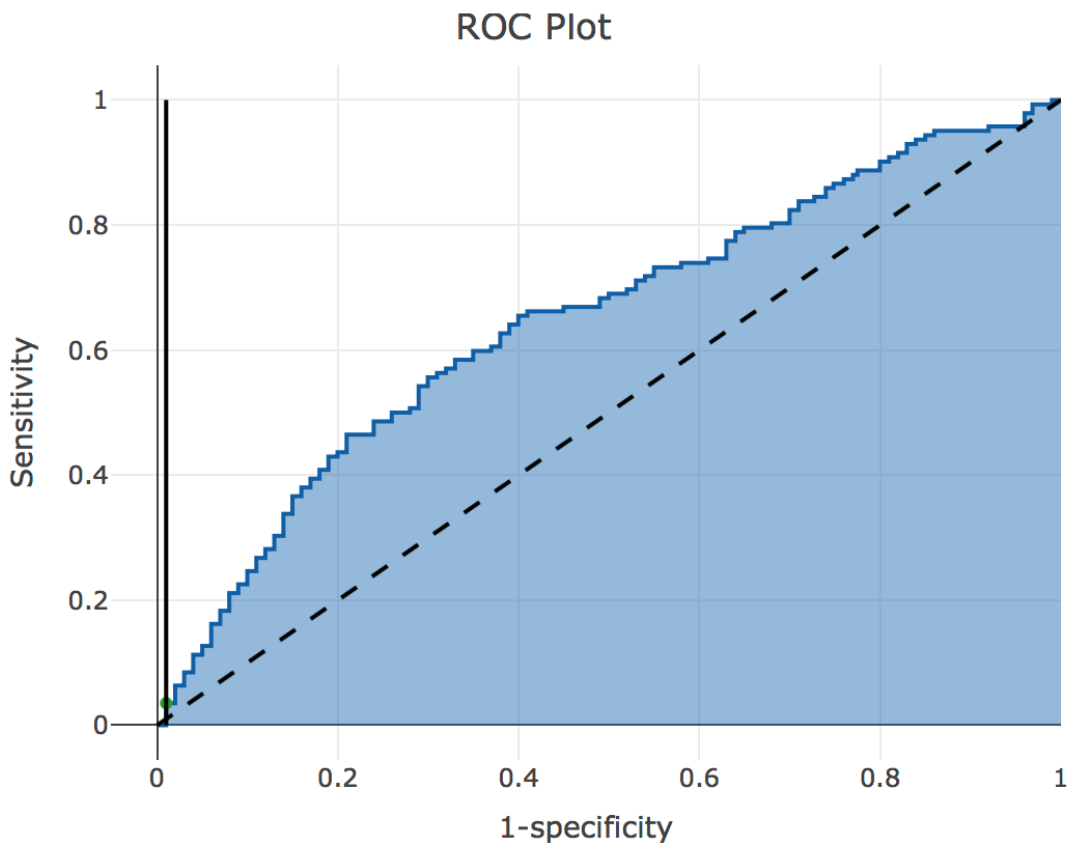
Showing 1 to 2 of 2 entries Previous 1 Next

Results: ROC Plots

ROC Plot

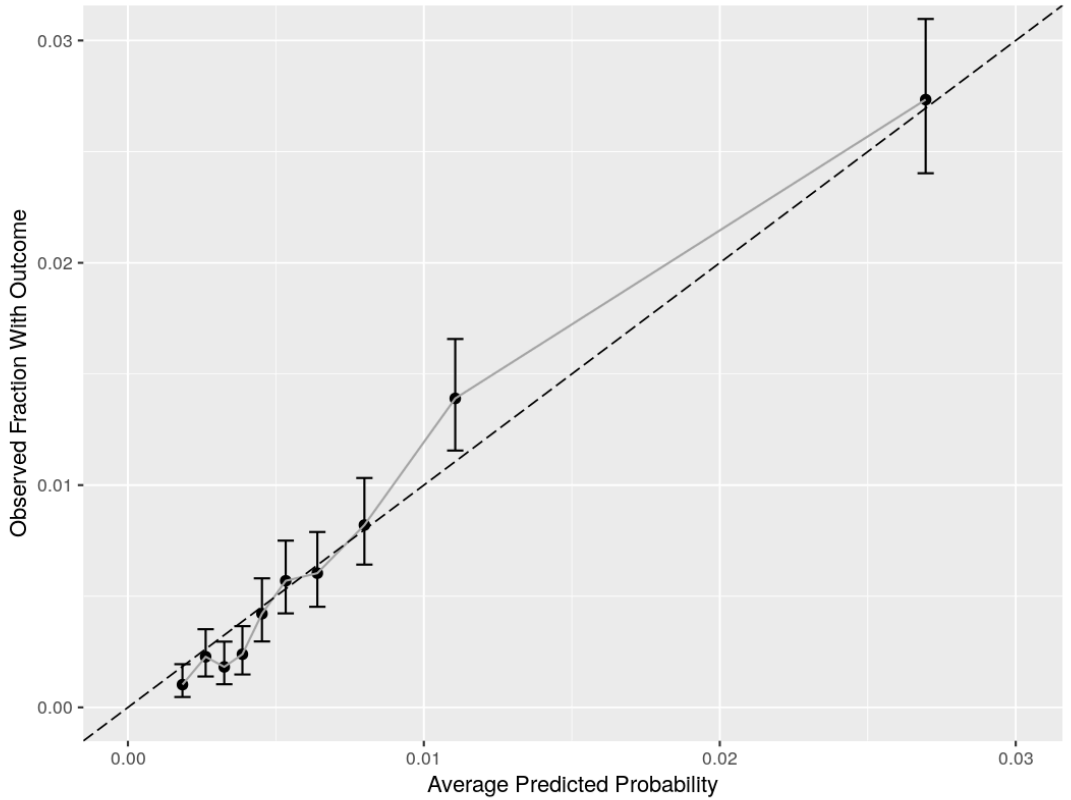


ROC Plot

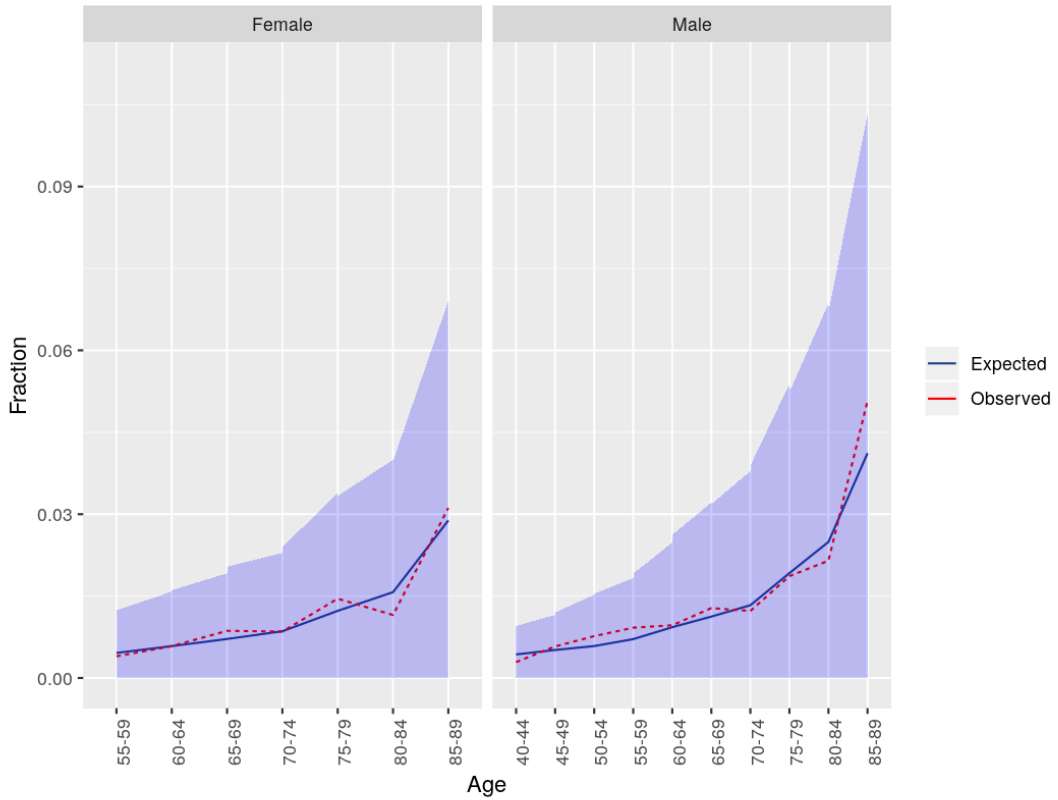


Results: Calibration Plots

Calibration Plot

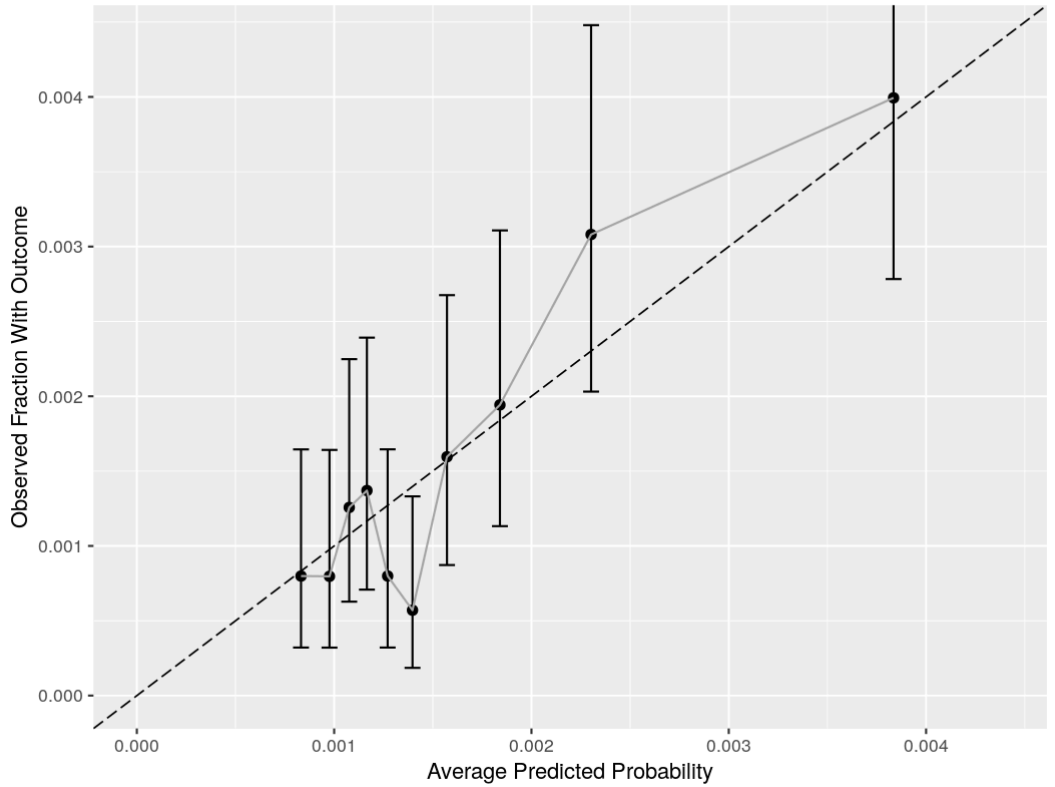


Demographic Plot

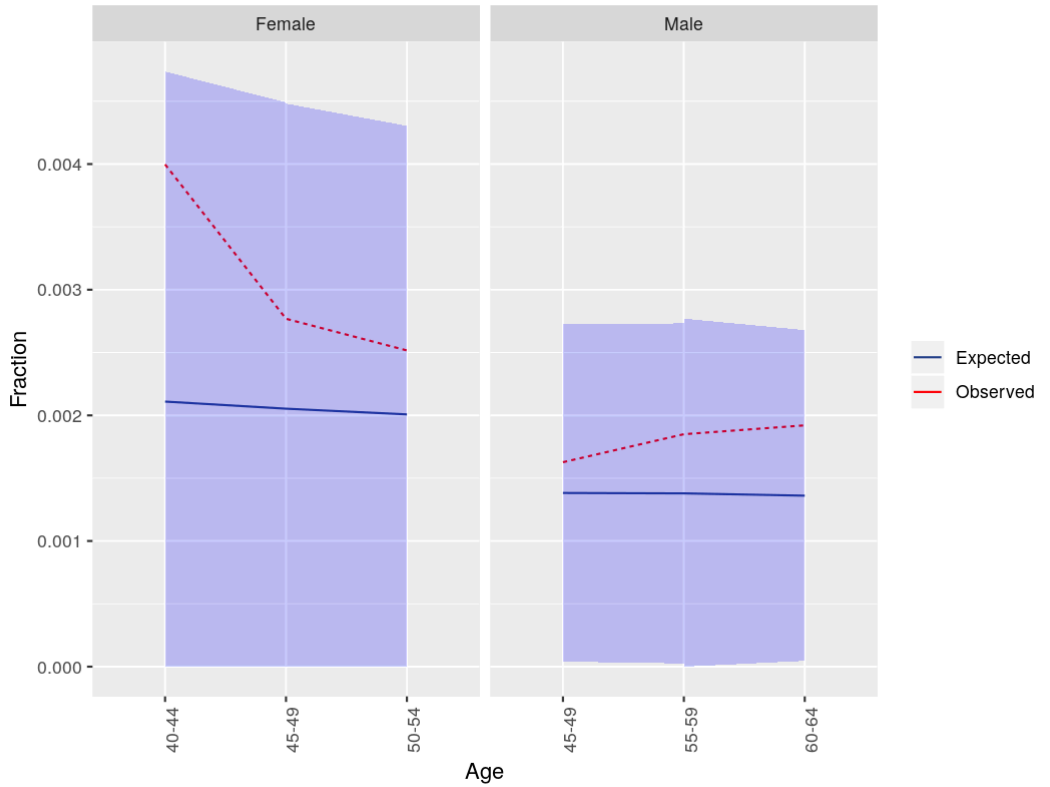


Results: Calibration Plots

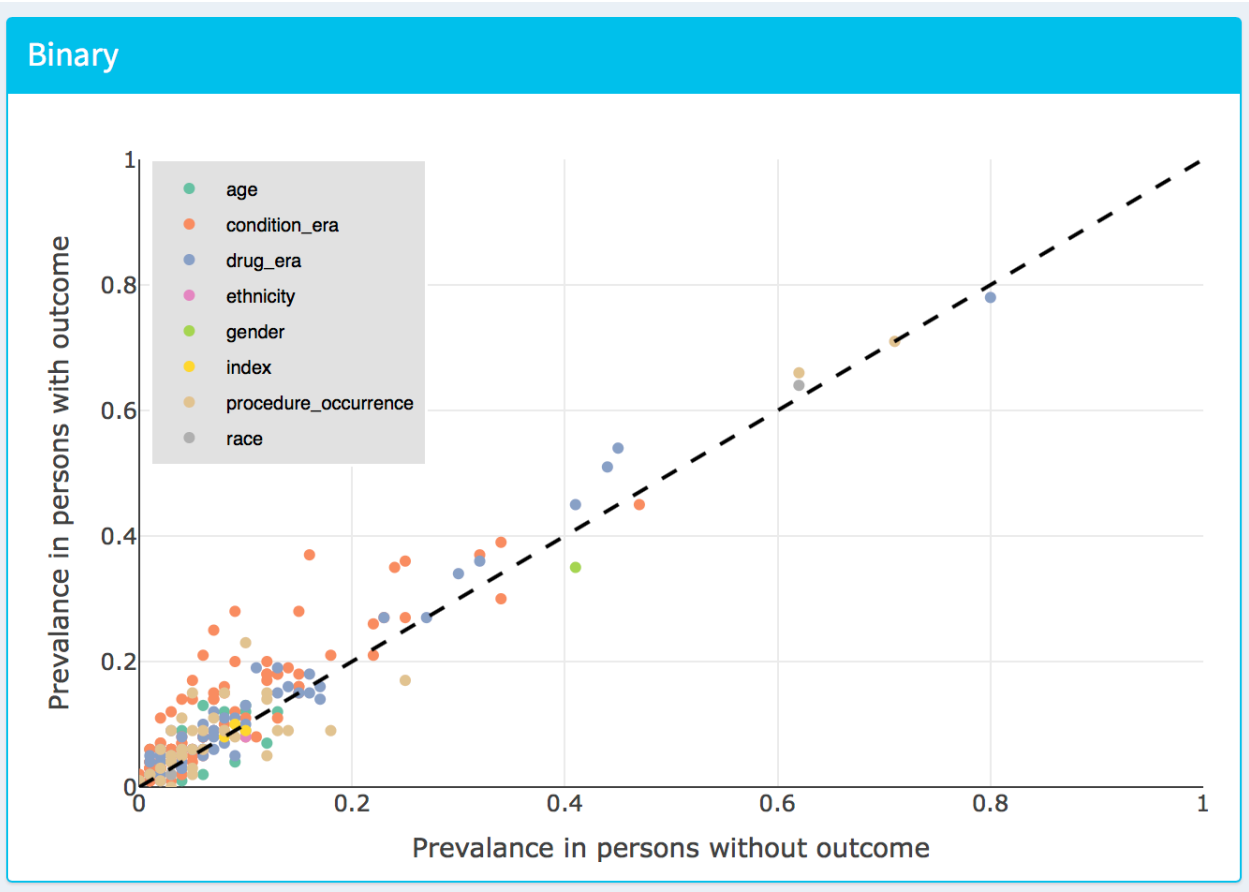
Calibration Plot

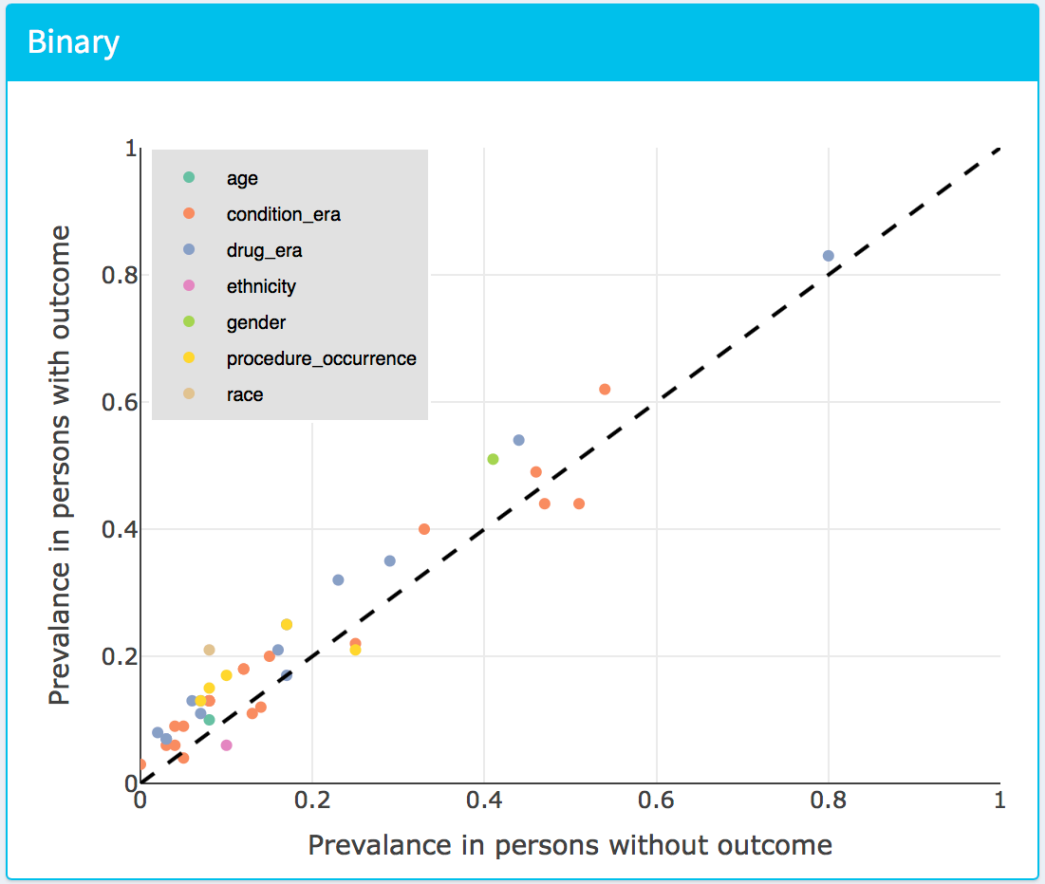


Demographic Plot



Results: Models





Results: Models

Model Table

Download Model

Show 10 entries

Search:

	Covariate Name	Value	Outcome Mean	Non-outcome Mean
32	age group: 85-89	1.61	0.06	0.01
31	age group: 80-84	1.12	0.09	0.04
30	age group: 75-79	0.93	0.13	0.06
29	age group: 70-74	0.66	0.12	0.08
28	age group: 65-69	0.53	0.13	0.1
26	age group: 60-64	0.4	0.12	0.1
20543	condition_era group during day -365 through 0 days relative to index: Aneurysm of artery of trunk	0.28	0.03	0.01
688	condition_era group during day -365 through 0 days relative to index: Chest pain	0.25	0.28	0.15
23527	condition_era group during day -365 through 0 days relative to index: Nicotine dependence	0.25	0.14	0.07
4	Charlson index - Romano adaptation	0.24	2.49	1.36

Showing 1 to 10 of 29,137 entries

Previous 1 2 3 4 5 … 2914 Next

Results: Models

Model Table

Download Model

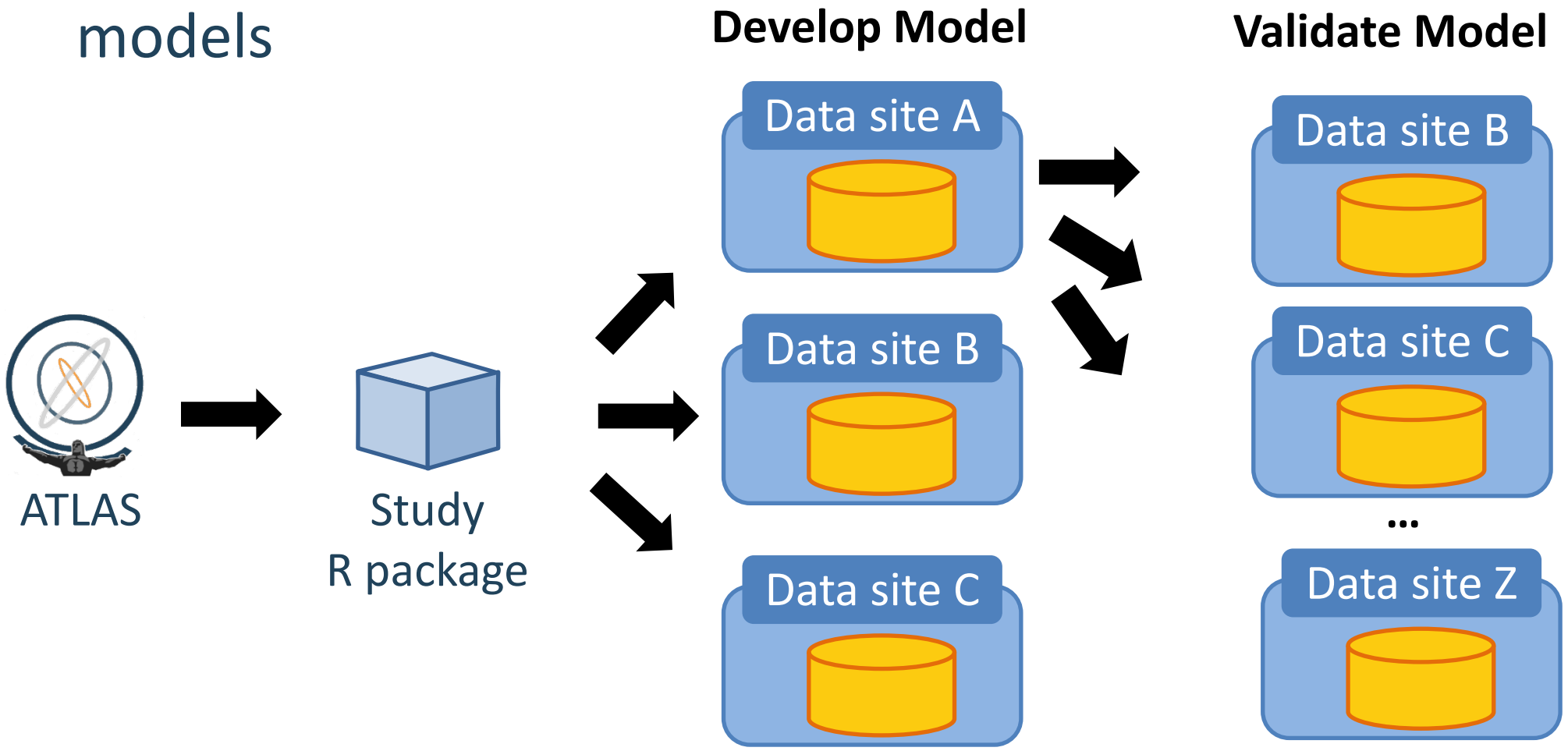
Show 10 entries

Search:

	Covariate Name	Value	Outcome Mean	Non-outcome Mean
35	race = Black or African American	0.84	0.21	0.08
4263	condition_era group during day -365 through 0 days relative to index: Angioedema	0.64	0.03	0
26059	drug_era group during day -365 through 0 days relative to index: ESTROGENS	0.38	0.13	0.06
4457	condition_era group during day -365 through 0 days relative to index: Edema	0.34	0.09	0.04
8262	drug_era group during day -365 through 0 days relative to index: Prednisone	0.27	0.13	0.07
14895	procedure_occurrence during day -365 through 0 days relative to index: Radiologic examination, chest, 2 views, frontal and lateral	0.26	0.25	0.17
25784	drug_era group during day -365 through 0 days relative to index: ANTIEMETICS AND ANTINAUSEANTS	0.22	0.07	0.03
16014	procedure_occurrence during day -365 through 0 days relative to index: Emergency department visit for the evaluation and management of a patient, which requires these 3 key components: An expanded problem focused history; An expanded problem focused examination; and Medical decision making of moderate complexity. Counseling	0.18	0.13	0.07
23443	condition_era group during day -365 through 0 days relative to index: Mass of trunk	0.17	0.2	0.15
27949	condition_era group during day -365 through 0 days relative to index: Allergic condition	0.15	0.18	0.12

Final thoughts

- Standardization can enable efficient external validation of models



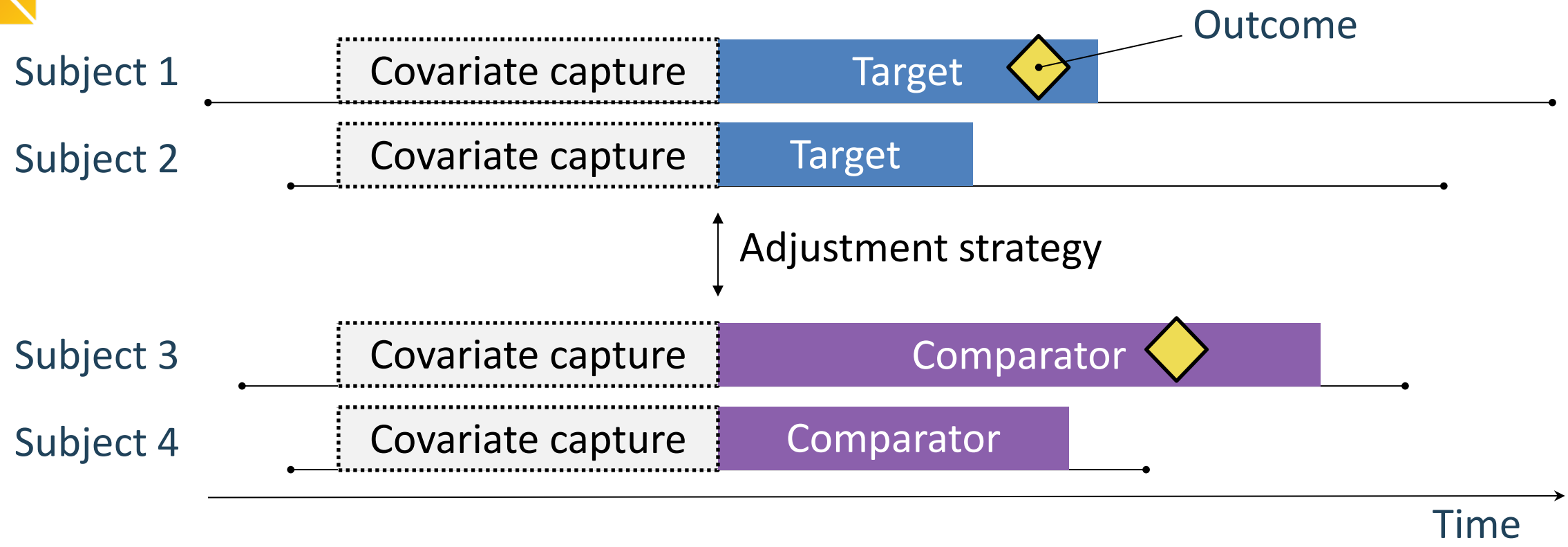
The journey through OHDSI's standardized analytics: Population-level Effect Estimation

Martijn Schuemie, PhD

Janssen Research & Development

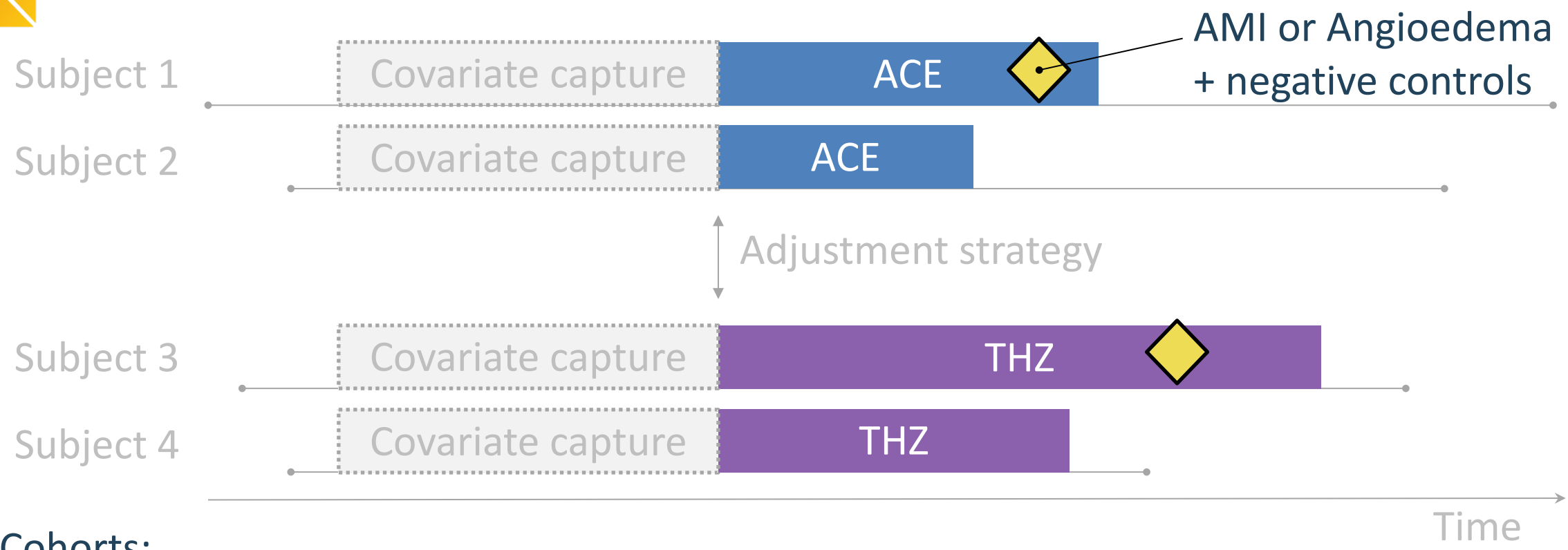
UCLA

New-user cohort method



Compare the hazard of the **Outcome** in **Target** to **Comparator**

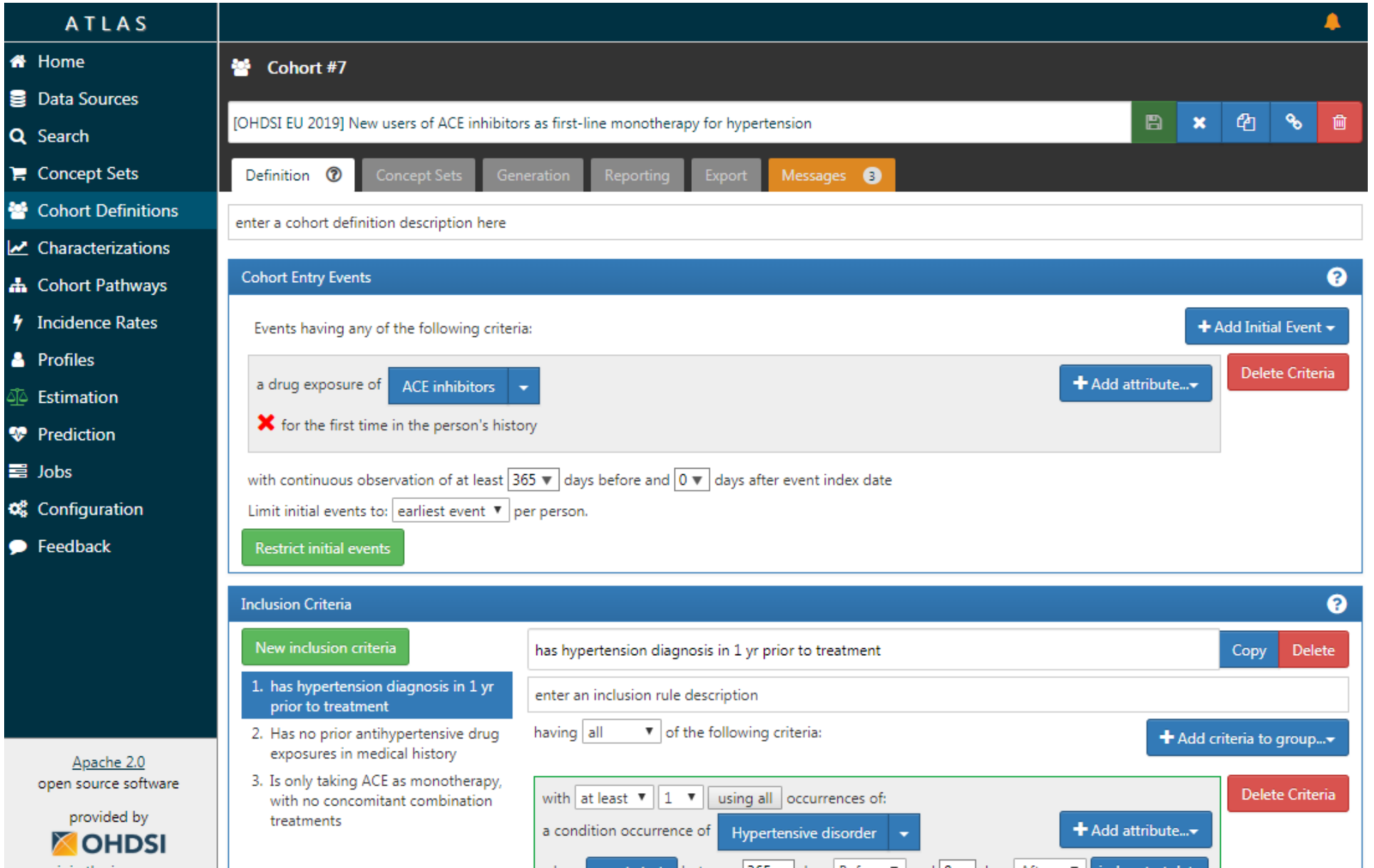
Hypertension



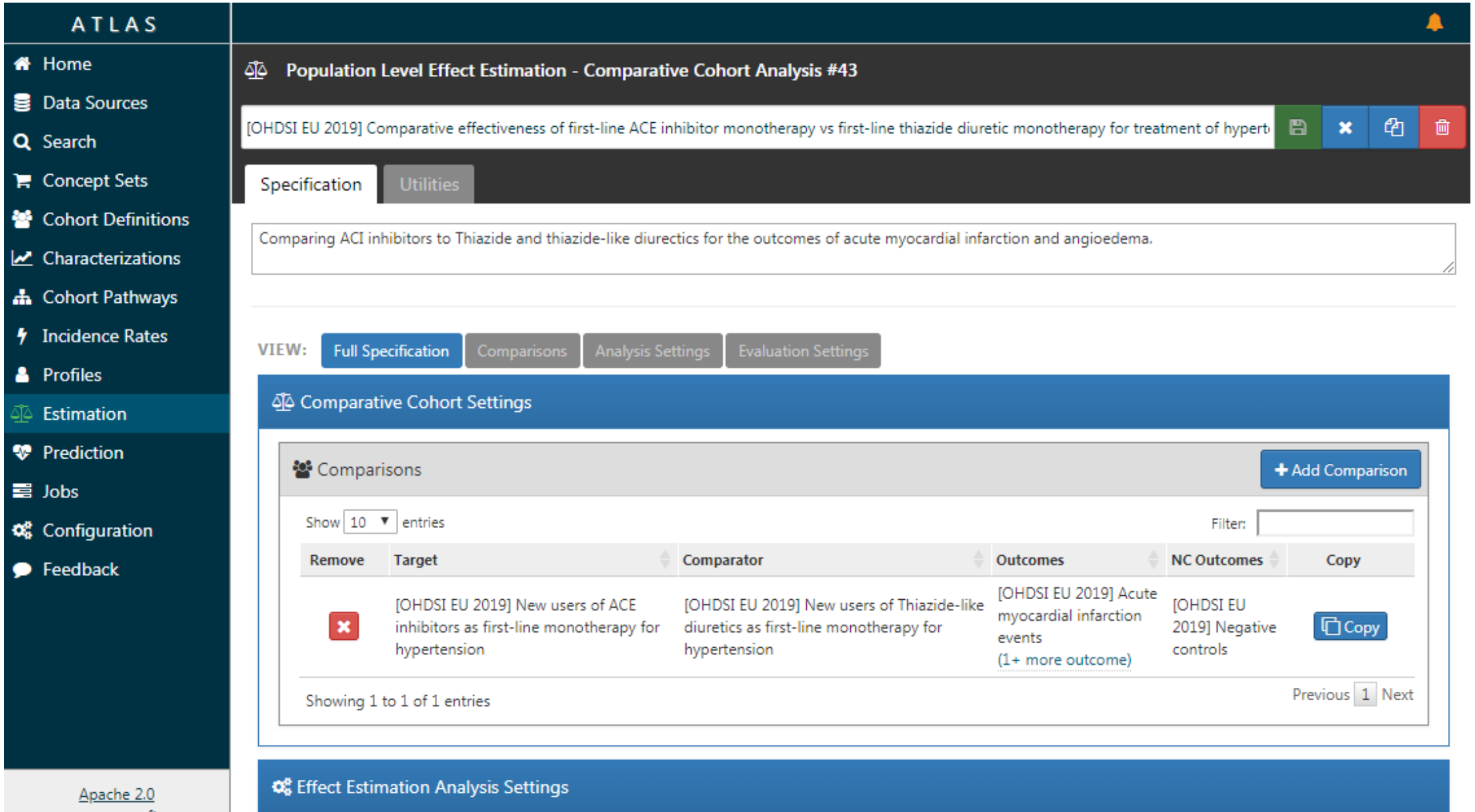
Cohorts:

- New users of ACE
- New users of THZ
- AMI
- Angioedema

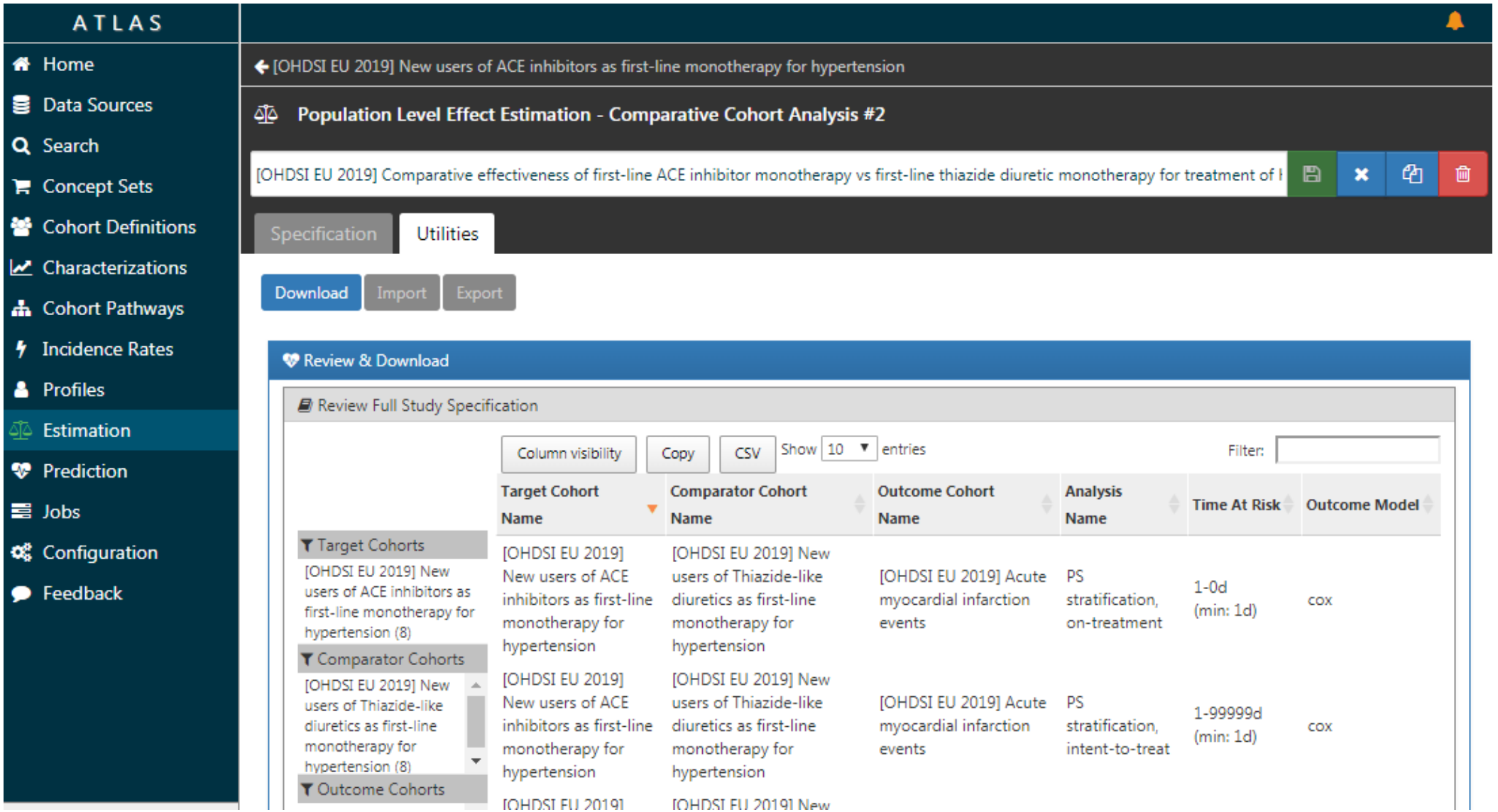
Define cohorts



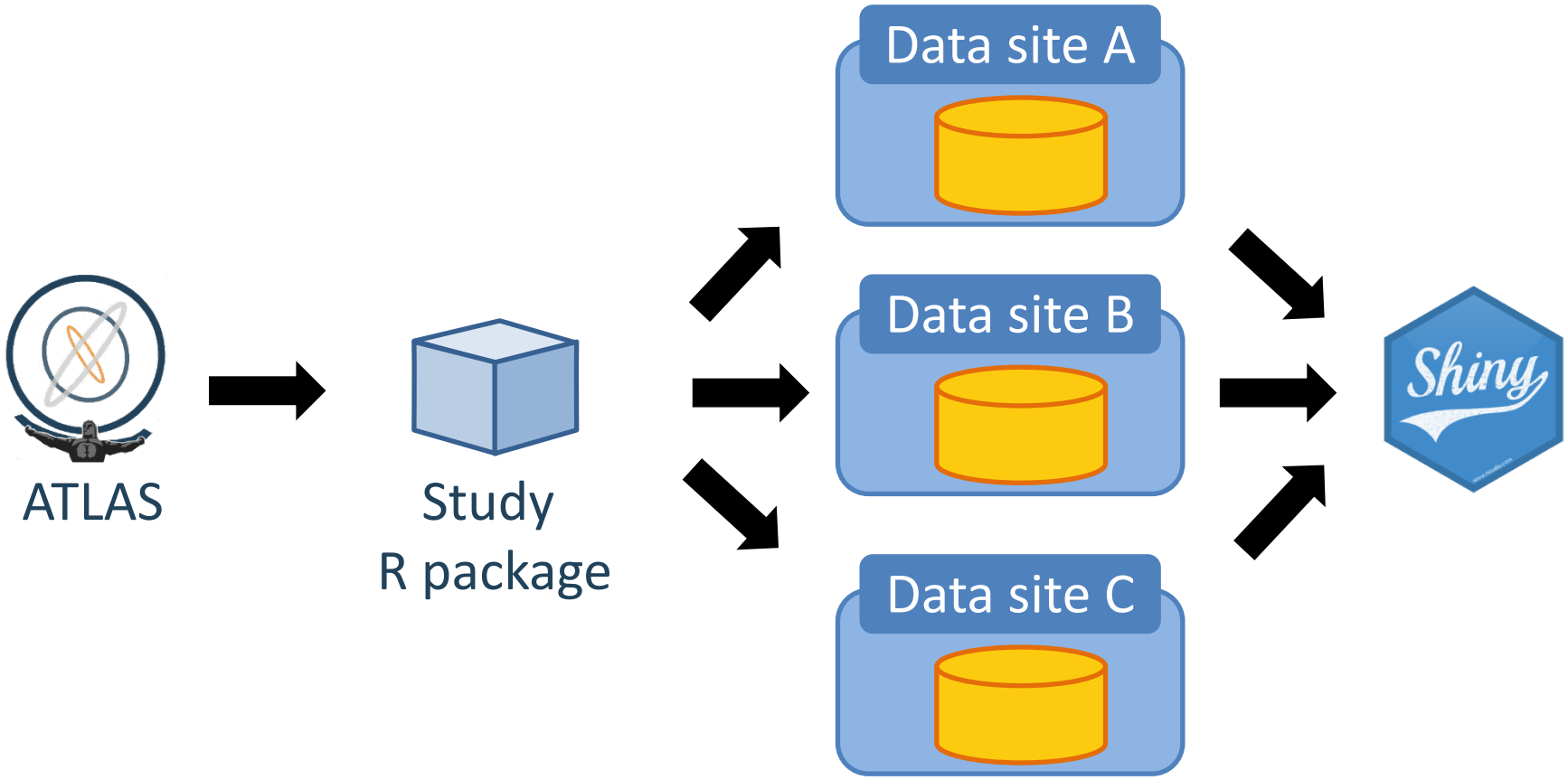
Design study



Study package generation



Study design to results



Results: Overview

LEGEND basic viewer

About | Specific research questions

Indication: Hypertension

Exposure group: Drug class

☐ Include combination exposures

Target: ACE inhibitors

Comparator: Thiazide or thiazide-like diuretics

Outcome: Acute myocardial infarction

Data source
- ☑ CCAE
- ☑ CUMC
- ☑ IMSG
- ☑ JMDC
- ☑ MDCD
- ☑ MDCR
- ☑ NHIS_NSC
- ☑ Optum
- ☑ Panther
- ☑ Meta-analysis

Analysis
- ☑ PS stratification, on-treatment
- ☐ PS stratification, intent-to-treat
- ☐ PS matching, on-treatment
- ☐ PS matching, intent-to-treat

Show 15 entries

Analysis	Data source	HR	LB	UB	P	Cal.HR	Cal.LB	Cal.UB	Cal.P
PS stratification, on-treatment	CCAE	1.12	0.99	1.26	0.07	1.17	0.97	1.46	0.10
PS stratification, on-treatment	CUMC	0.94	0.59	1.51	0.79	1.03	0.65	1.76	0.84
PS stratification, on-treatment	MDCD	1.43	1.02	2.06	0.04	1.45	1.03	2.09	0.03
PS stratification, on-treatment	MDCR	1.07	0.91	1.27	0.42	1.10	0.91	1.36	0.29
PS stratification, on-treatment	Meta-analysis	1.16	1.08	1.23	0.00	1.19	1.04	1.40	0.01
PS stratification, on-treatment	NHIS_NSC	0.79	0.30	2.13	0.64	0.75	0.27	2.06	0.54
PS stratification, on-treatment	Optum	1.19	1.08	1.32	0.00	1.23	1.05	1.49	0.01
PS stratification, on-treatment	Panther	1.22	1.02	1.46	0.03	1.20	0.97	1.59	0.03

Showing 1 to 8 of 8 entries Previous 1 Next

Power | Propensity scores | Covariate balance | Systematic error | Forest plot

Source	HR (95% CI)	Calibrated HR (95% CI)
CCAE	1.12 (0.99-1.26)	1.17 (0.97-1.46)
CUMC	0.94 (0.59-1.51)	1.03 (0.65-1.76)
MDCD	1.43 (1.02-2.06)	1.45 (1.03-2.09)
MDCR	1.07 (0.91-1.27)	1.10 (0.91-1.36)
NHIS_NSC	0.79 (0.30-2.13)	0.75 (0.27-2.06)
Optum	1.19 (1.08-1.32)	1.23 (1.05-1.49)
Panther	1.22 (1.02-1.46)	1.20 (0.97-1.59)
Summary (I^2 = 0.00)	1.16 (1.08-1.23)	1.19 (1.04-1.40)

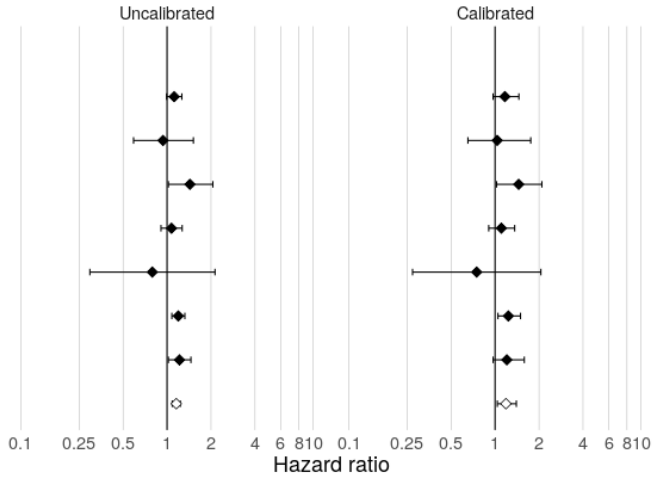


Figure 6. Forest plot showing the per-database and summary hazard ratios (and 95 percent confidence intervals) comparing ACE inhibitors to Thiazide or thiazide-like diuretics for the outcome of Acute myocardial infarction, using stratification. Estimates are shown both before and after empirical calibration. The I2 is computed on the uncalibrated estimates.

Results: Power

Power | Attrition | Population characteristics | Propensity scores | Covariate balance | Systematic error | Kaplan-Meier

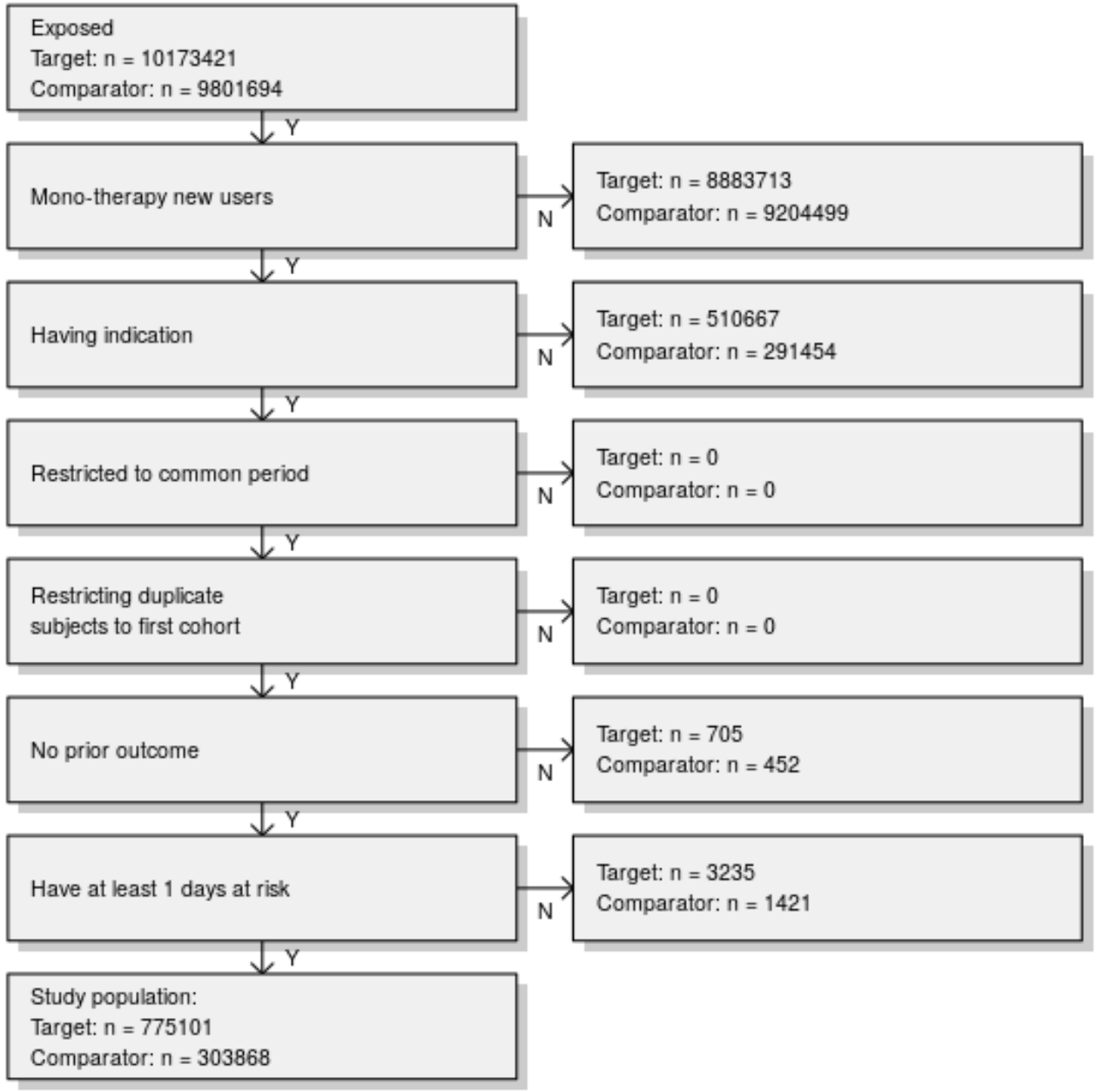
Table 1a. Number of subjects, follow-up time (in years), number of outcome events, and event incidence rate (IR) per 1,000 patient years (PY) in the target (*ACE inhibitors*) and comparator (*Thiazide or thiazide-like diuretics*) group after stratification, as well as the minimum detectable relative risk (MDRR). Note that the IR does not account for any stratification.

Target subjects	Comparator subjects	Target years	Comparator years	Target events	Comparator events	Target IR (per 1,000 PY)	Comparator IR (per 1,000 PY)	MDRR
775,101	303,868	588,000	213,511	784	169	1.33	0.79	1.22

Table 1b. Time (days) at risk distribution expressed as minimum (min), 25th percentile (P25), median, 75th percentile (P75), and maximum (max) in the target (*ACE inhibitors*) and comparator (*Thiazide or thiazide-like diuretics*) cohort after stratification.

Cohort	Min	P10	P25	Median	P75	P90	Max
Target	1	29	38	116	530	721	6,218
Comparator	1	29	29	95	486	676	6,101

Results: Attrition



Results: Patient characteristics

Power | Attrition | Population characteristics | Propensity scores | Covariate balance | Systematic error | Kaplan-Meier

Table 2. Select characteristics before and after stratification, showing the (weighted) percentage of subjects with the characteristics in the target (*ACE inhibitors*) and comparator (*Thiazide or thiazide-like diuretics*) group, as well as the standardized difference of the means.

Characteristic	Before stratification			After stratification		
	Target	Comparator		Target	Comparator	
	%	%	Std. diff	%	%	Std. diff
Age group						
10-14	0.2	0.1	0.02	0.1	0.1	0.00
15-19	0.7	0.6	0.02	0.7	0.7	0.00
20-24	1.4	1.6	-0.02	1.4	1.5	0.00
25-29	2.6	3.5	-0.06	2.8	2.7	0.00
40-44	12.1	13.4	-0.04	12.4	12.3	0.00
45-49	16.1	16.3	-0.01	16.2	15.9	0.01
50-54	18.7	17.7	0.03	18.5	18.4	0.00
55-59	18.3	16.2	0.06	17.8	18.0	0.00
60-64	15.5	13.2	0.06	15.0	15.3	-0.01
65-69	1.3	1.1	0.02	1.3	1.3	0.00
Gender: female	38.4	60.7	-0.46	44.7	45.2	-0.01
Medical history: General						
Acute respiratory disease	24.5	26.1	-0.04	25.0	25.4	-0.01
Attention deficit hyperactivity disorder	1.0	1.1	-0.01	1.1	1.0	0.00

Results: Propensity score

Power | Attrition | Population characteristics | Propensity scores | Covariate balance | Systematic error | Kaplan-Meier

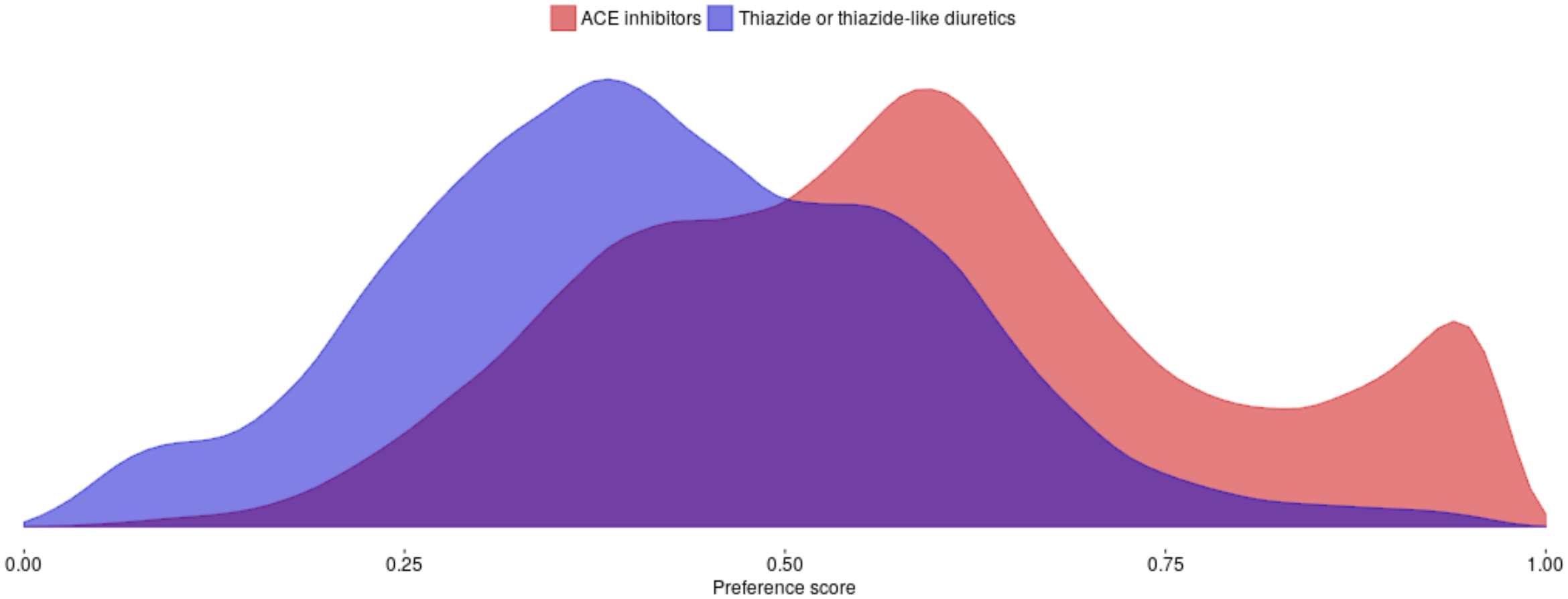


Figure 2. Preference score distribution. The preference score is a transformation of the propensity score that adjusts for differences in the sizes of the two treatment groups. A higher overlap indicates subjects in the two groups were more similar in terms of their predicted probability of receiving one treatment over the other.

Results: Covariate balance

Power | Attrition | Population characteristics | Propensity scores | Covariate balance | Systematic error | Kaplan-Meier

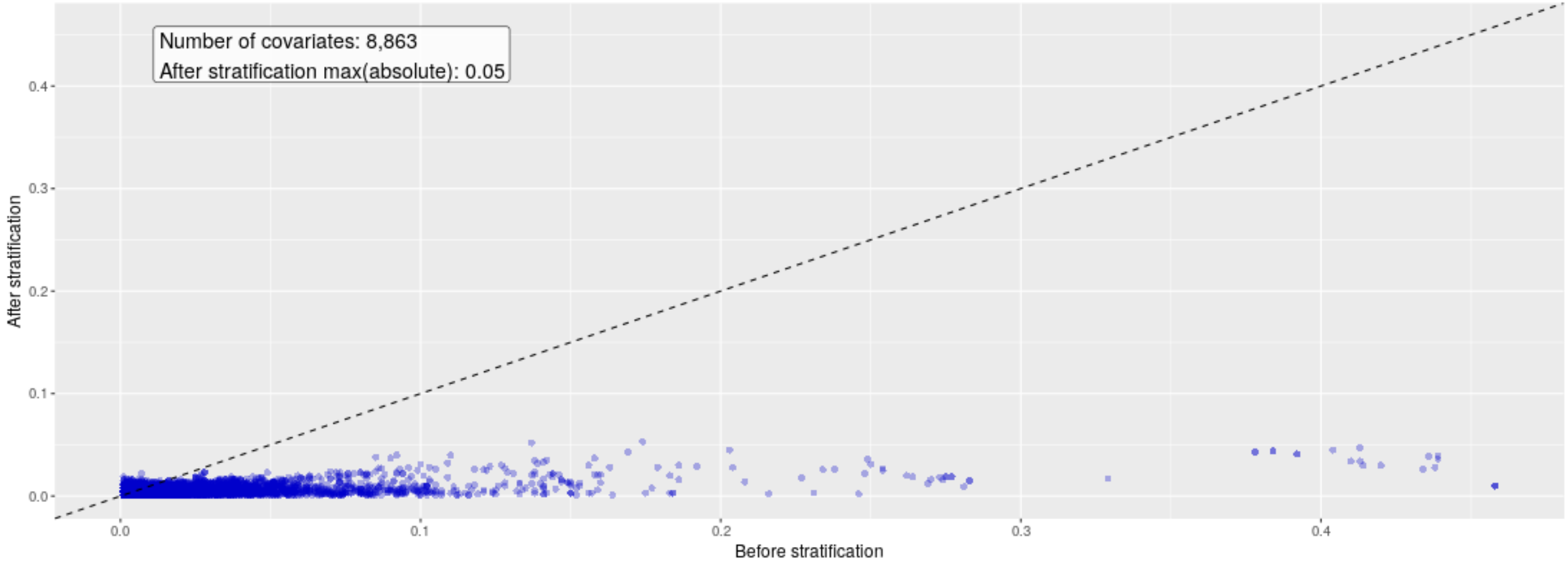


Figure 3. Covariate balance before and after stratification. Each dot represents the standardized difference of means for a single covariate before and after stratification on the propensity score. Move the mouse arrow over a dot for more details.

Results: Estimates for negative and positive controls

Power | Attrition | Population characteristics | Propensity scores | Covariate balance | Systematic error | Kaplan-Meier

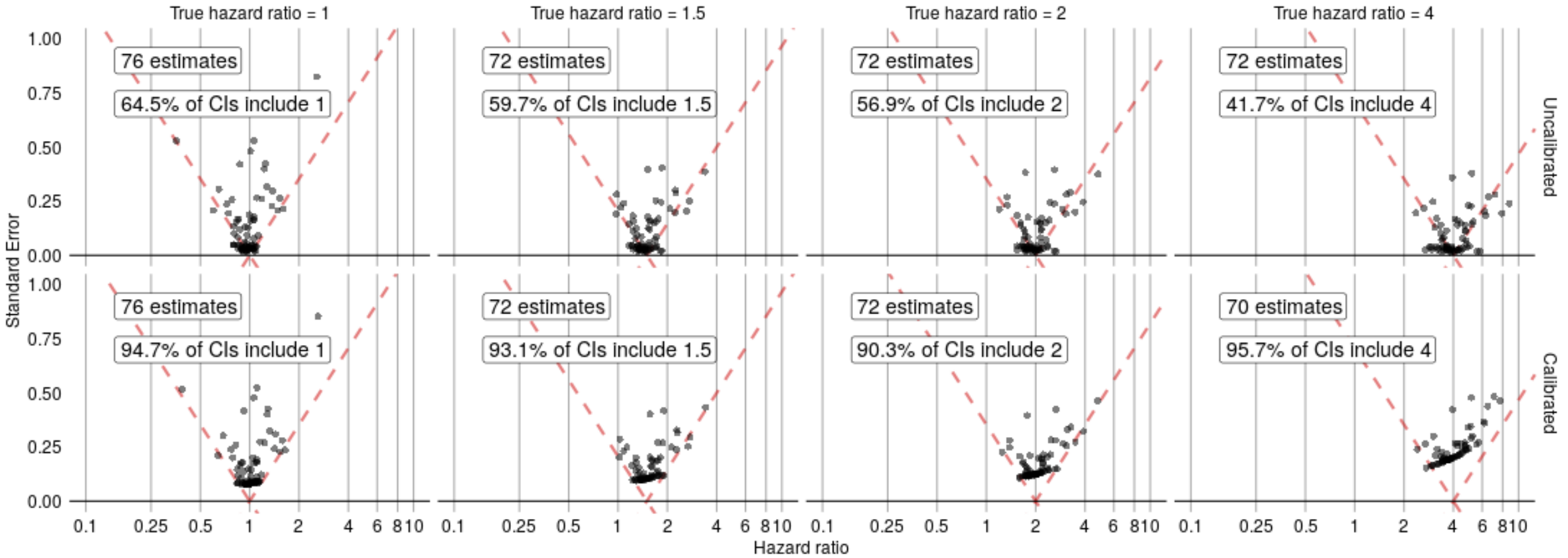
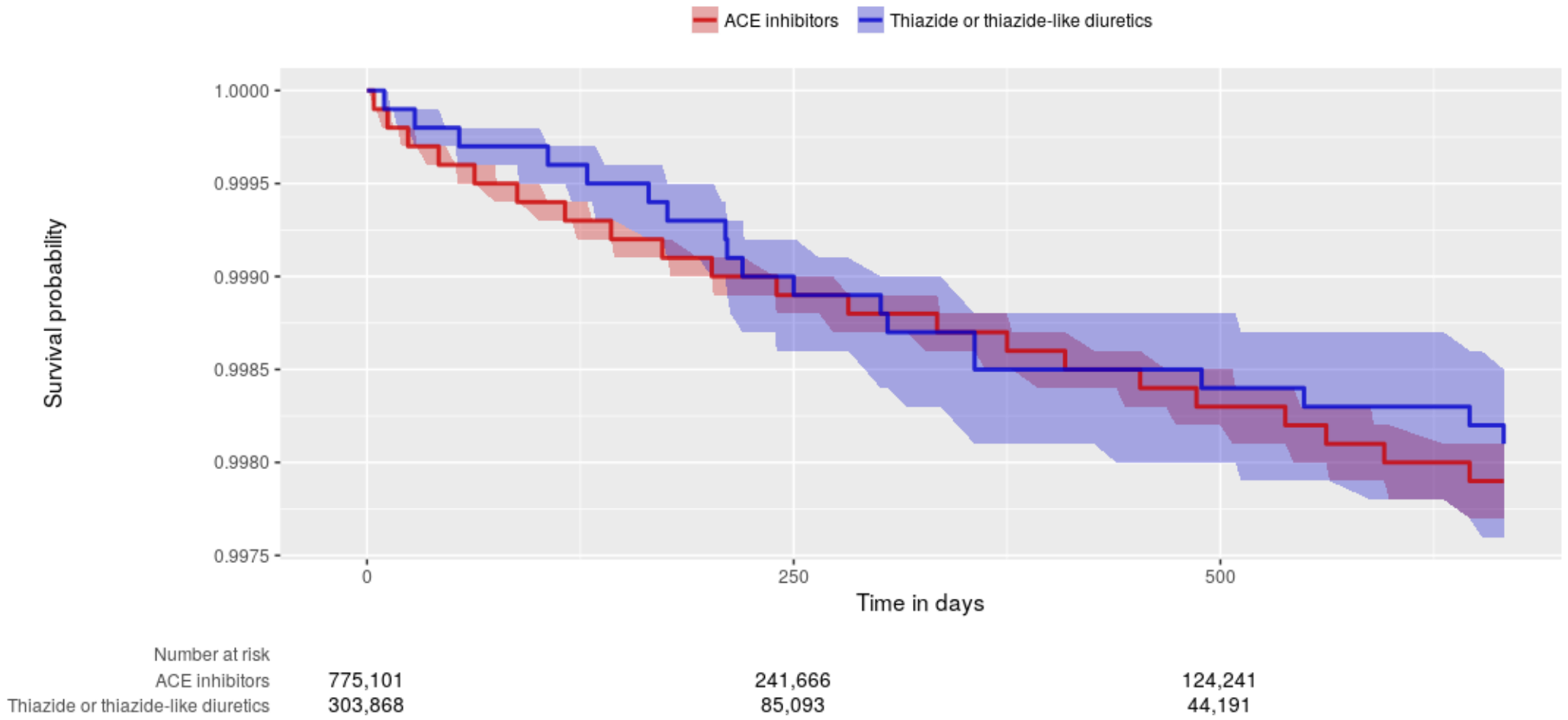


Figure 4. Systematic error. Effect size estimates for the negative controls (true hazard ratio = 1) and positive controls (true hazard ratio > 1), before and after calibration. Estimates below the diagonal dashed lines are statistically significant (alpha = 0.05) different from the true effect size. A well-calibrated estimator should have the true effect size within the 95 percent confidence interval 95 percent of times.

Result: Kaplan-Meier plot

Power | Attrition | Population characteristics | Propensity scores | Covariate balance | Systematic error | Kaplan-Meier



Number at risk			
ACE inhibitors	775,101	241,666	124,241
Thiazide or thiazide-like diuretics	303,868	85,093	44,191

Figure 5. Kaplan Meier plot, showing survival as a function of time. This plot is adjusted for the propensity score stratification: The target curve (*ACE inhibitors*) shows the actual observed survival. The comparator curve (*Thiazide or thiazide-like diuretics*) applies reweighting to approximate the counterfactual of what the target survival would look like had the target cohort been exposed to the comparator instead. The shaded area denotes the 95 percent confidence interval.

Result: Forest plot

Power | Propensity scores | Covariate balance | Systematic error | Forest plot

Source	HR (95% CI)	Calibrated HR (95% CI)
CCAE	1.96 (1.65-2.34)	2.01 (1.55-2.80)
CUMC	1.08 (0.52-2.29)	1.17 (0.60-2.49)
MDCD	4.57 (3.00-7.19)	4.52 (2.85- NA)
MDCR	2.44 (1.52-4.13)	2.46 (1.49-4.17)
Optum	1.99 (1.62-2.47)	2.03 (1.56-2.80)
Panther	2.70 (1.97-3.78)	2.31 (1.62-4.03)
Summary (I^2 = 0.73)	2.32 (1.79-3.00)	2.26 (1.68-3.30)

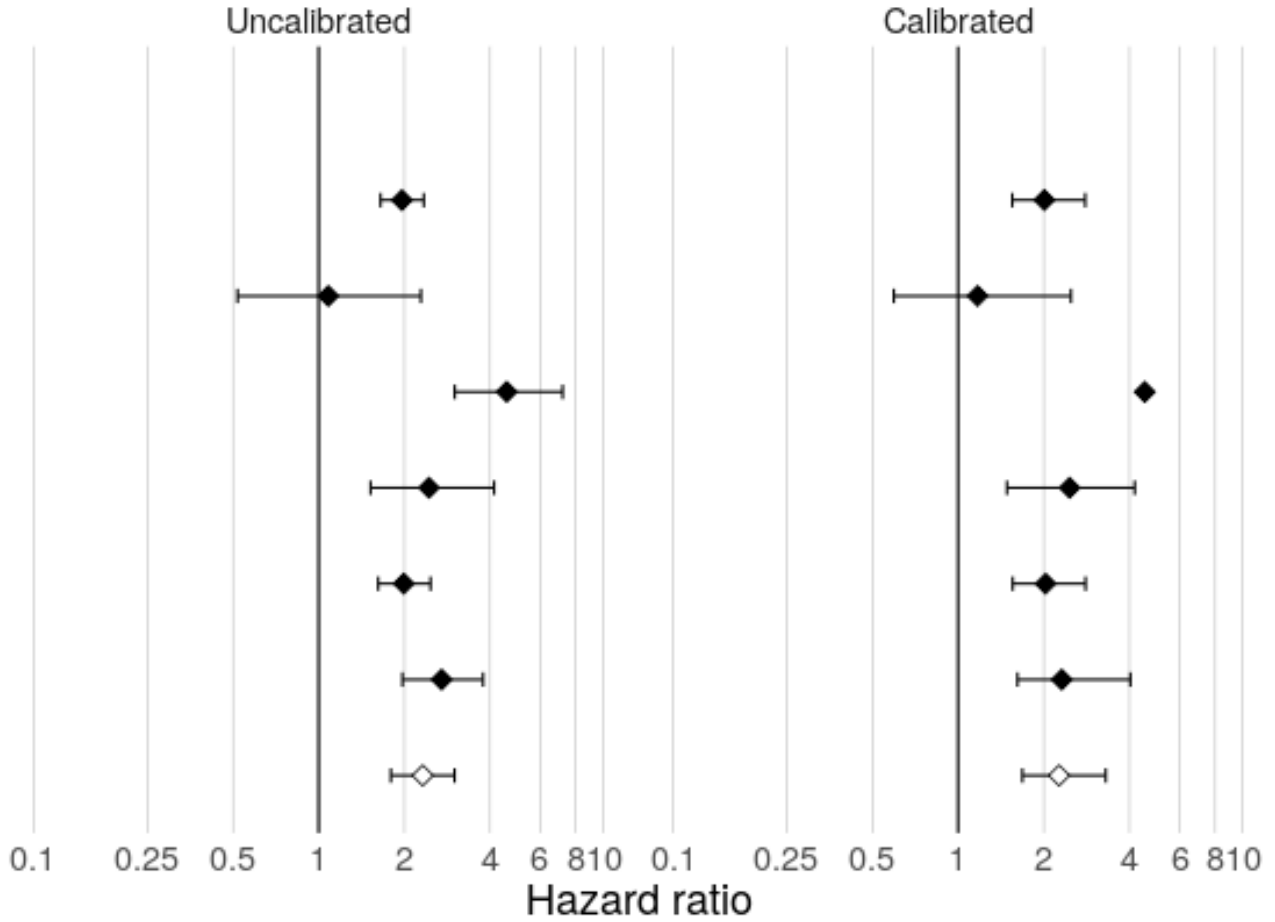


Figure 6. Forest plot showing the per-database and summary hazard ratios (and 95 percent confidence intervals) comparing ACE inhibitors to Thiazide or thiazide-like diuretics for the outcome of Angioedema, using stratification. Estimates are shown both before and after empirical calibration. The I2 is computed on the uncalibrated estimates.

Final thoughts

- Standardization does not mean one-size-fits-all
- Exposes design choices
- Enables:
  - Consistency
  - Transparency
  - Reproducibility
  - Efficiency