Walkthrough of implementing a cohort study using OHDSI tools

Cardiovascular, Bleeding, and Mortality Risks in Elderly Medicare Patients Treated With Dabigatran or Warfarin for Nonvalvular Atrial Fibrillation

David J. Graham, MD, MPH; Marsha E. Reichman, PhD; Michael Wernecke, BA; Rongmei Zhang, PhD; Mary Ross Southworth, PharmD; Mark Levenson, PhD; Ting-Chang Sheu, MPH; Katrina Mott, MHS; Margie R. Goulding, PhD; Monika Houstoun, PharmD, MPH; Thomas E. MaCurdy, PhD; Chris Worrall, BS; Jeffrey A. Kelman, MD, MMSc

Background—The comparative safety of dabigatran versus warfarin for treatment of nonvalvular atrial fibrillation in general practice settings has not been established.

Methods and Results—We formed new-user cohorts of propensity score–matched elderly patients enrolled in Medicare who initiated dabigatran or warfarin for treatment of nonvalvular atrial fibrillation between October 2010 and December 2012. Among 134414 patients with 37587 person-years of follow-up, there were 2715 primary outcome events. The hazard ratios (95% confidence intervals) comparing dabigatran with warfarin (reference) were as follows: ischemic stroke, 0.80 (0.67–0.96); intracranial hemorrhage, 0.34 (0.26–0.46); major gastrointestinal bleeding, 1.28 (1.14–1.44); acute myocardial infarction, 0.92 (0.78–1.08); and death, 0.86 (0.77–0.96). In the subgroup treated with dabigatran 75 mg twice daily, there was no difference in risk compared with warfarin for any outcome except intracranial hemorrhage, in which case dabigatran risk was reduced. Most patients treated with dabigatran 75 mg twice daily appeared not to have severe renal impairment, the intended population for this dose. In the dabigatran 150-mg twice daily subgroup, the magnitude of effect for each outcome was greater than in the combined-dose analysis.

Conclusions—In general practice settings, dabigatran was associated with reduced risk of ischemic stroke, intracranial hemorrhage, and death and increased risk of major gastrointestinal hemorrhage compared with warfarin in elderly patients with nonvalvular atrial fibrillation. These associations were most pronounced in patients treated with dabigatran 150 mg twice daily, whereas the association of 75 mg twice daily with study outcomes was indistinguishable from warfarin except for a lower risk of intracranial hemorrhage with dabigatran. (***Circulation***. **2015;131:157-164. DOI: 10.1161/CIRCULATIONAHA.114.012061.**)

Key Words: anticoagulant ■ pharmacoepidemiology ■ safety ■ thrombin inhibitor ■ warfarin

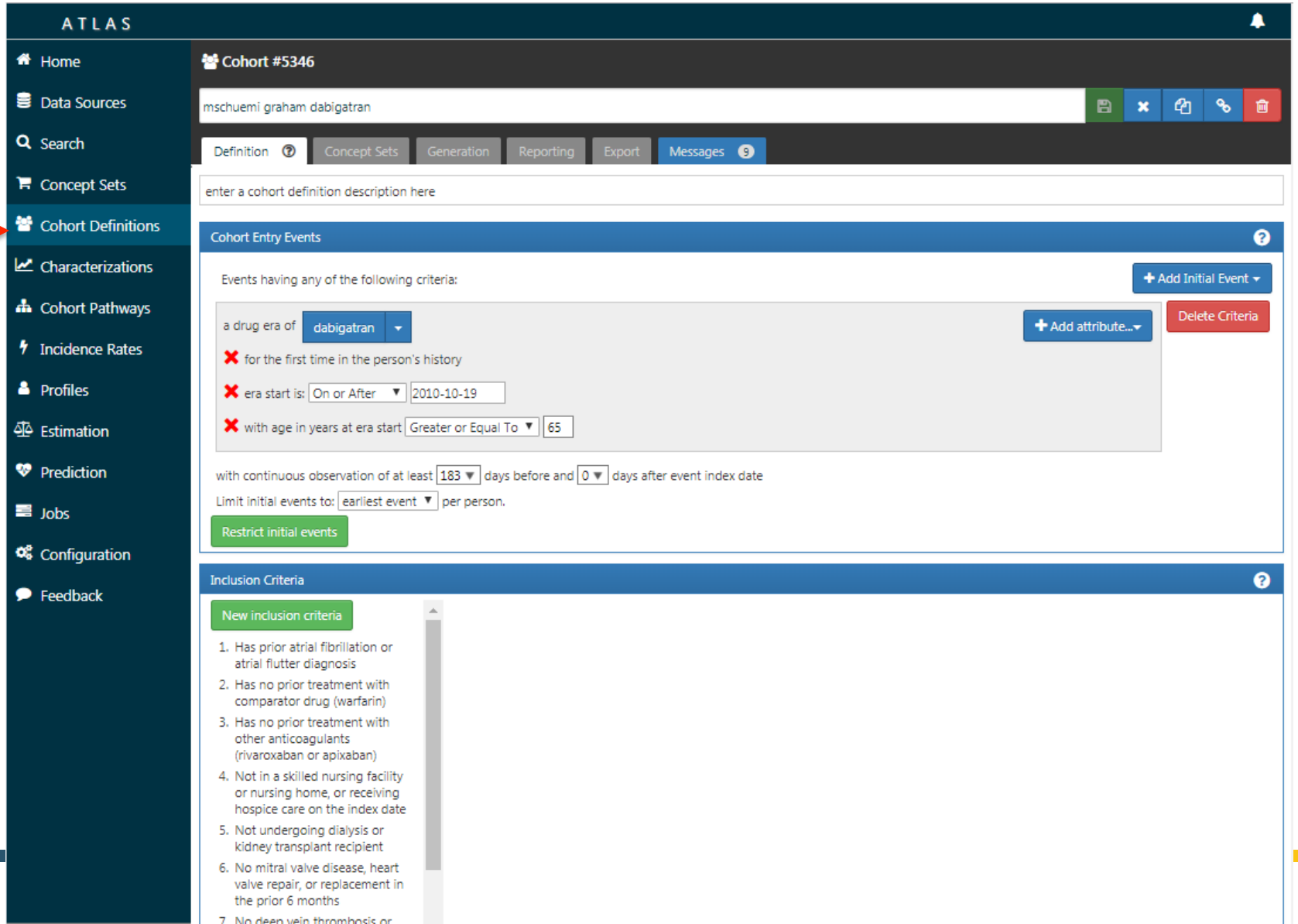
What is the design used by Graham et al?

Input parameter	Design choice
Target cohort (T)	dabigatran new users with prior atrial fibrillation
Comparator cohort (C)	warfarin new users with prior atrial fibrillation
Outcome cohort (O)	Ischemic stroke
Time-at-risk	1 day after cohort start → cohort end
Model specification	1:1 propensity score-matched univariable conditional Cox proportional hazards

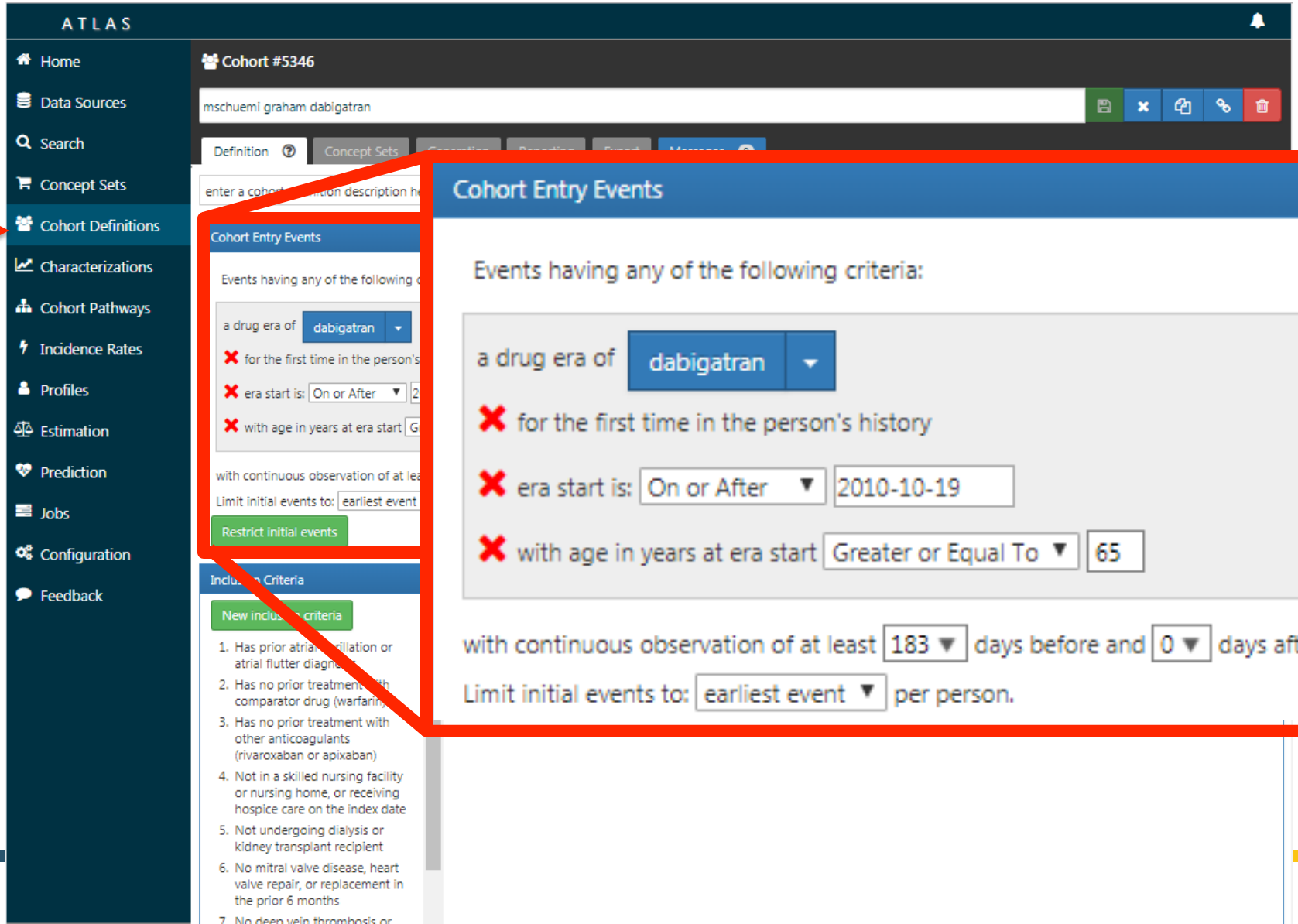
Graham et al. description of the cohort(s)

A new-user retrospective cohort design was used to compare patients initiating dabigatran or warfarin for the treatment of nonvalvular AF.[10] We identified all patients with any inpatient or outpatient diagnoses of AF or atrial flutter based on *International Classification of Diseases, Ninth Revision* coding who also filled at least 1 prescription for either drug from October 19, 2010 (US dabigatran approval date) through December 31, 2012, the study end date. Patients were excluded if they had <6 months of enrollment in Medicare before their index dispensing, were aged <65 years, received prior treatment with a study medication or rivaroxaban or apixaban (anticoagulants approved during the study), were in a skilled nursing facility or nursing home, or were receiving hospice care on the date of their cohort-qualifying prescription. Patients were also excluded if they had a hospitalization that extended beyond the index dispensing date. Patients discharged from the hospital on the same day as their index dispensing were included. Patients undergoing dialysis and kidney transplant recipients were also excluded. Additionally, because warfarin is approved for indications other than AF, we excluded patients with diagnoses indicating the presence of mitral valve disease, heart valve repair or replacement, deep vein thrombosis, pulmonary embolism, or joint replacement surgery in the preceding 6 months.

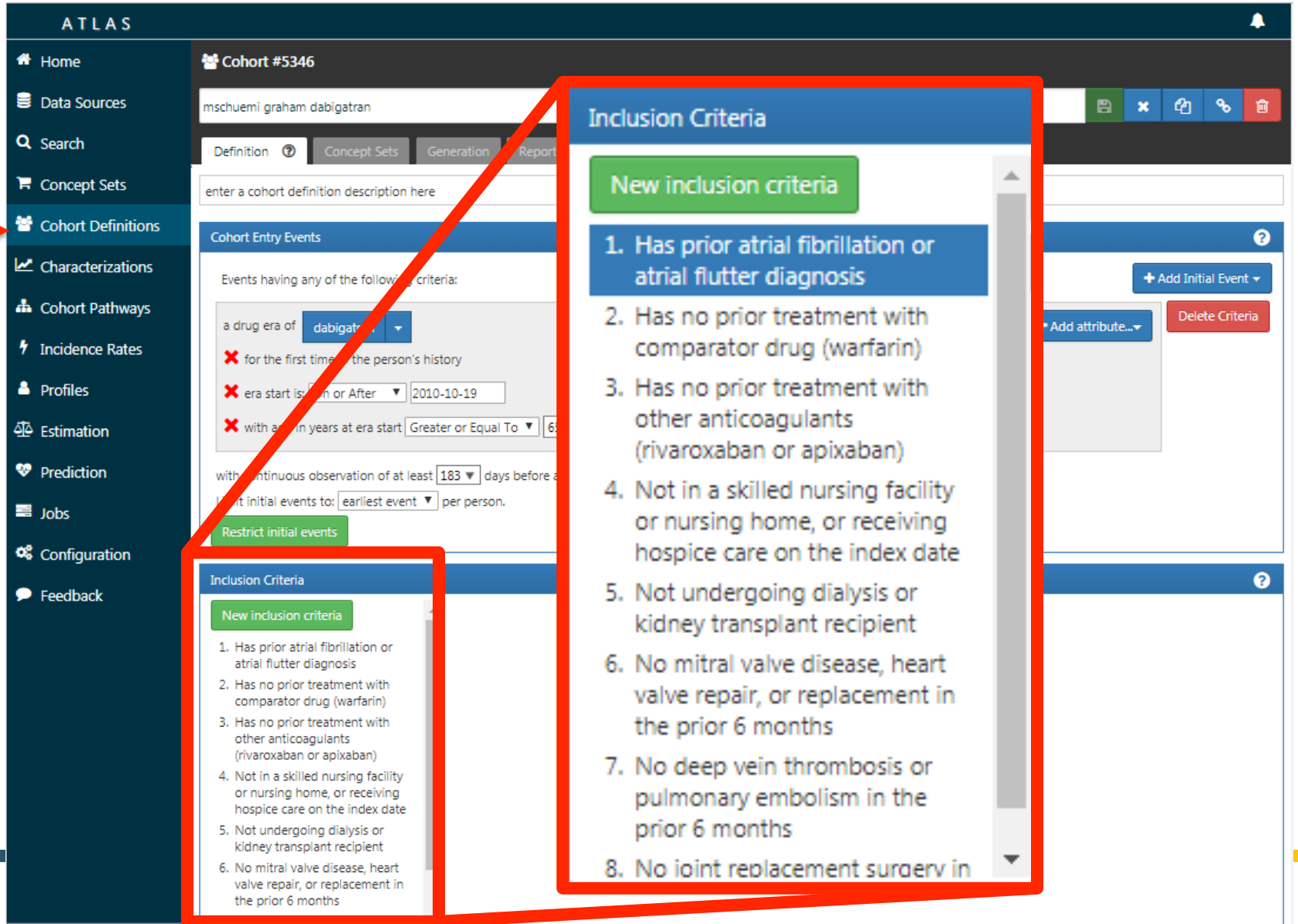
Graham et al. replication: Designing the target cohort in ATLAS



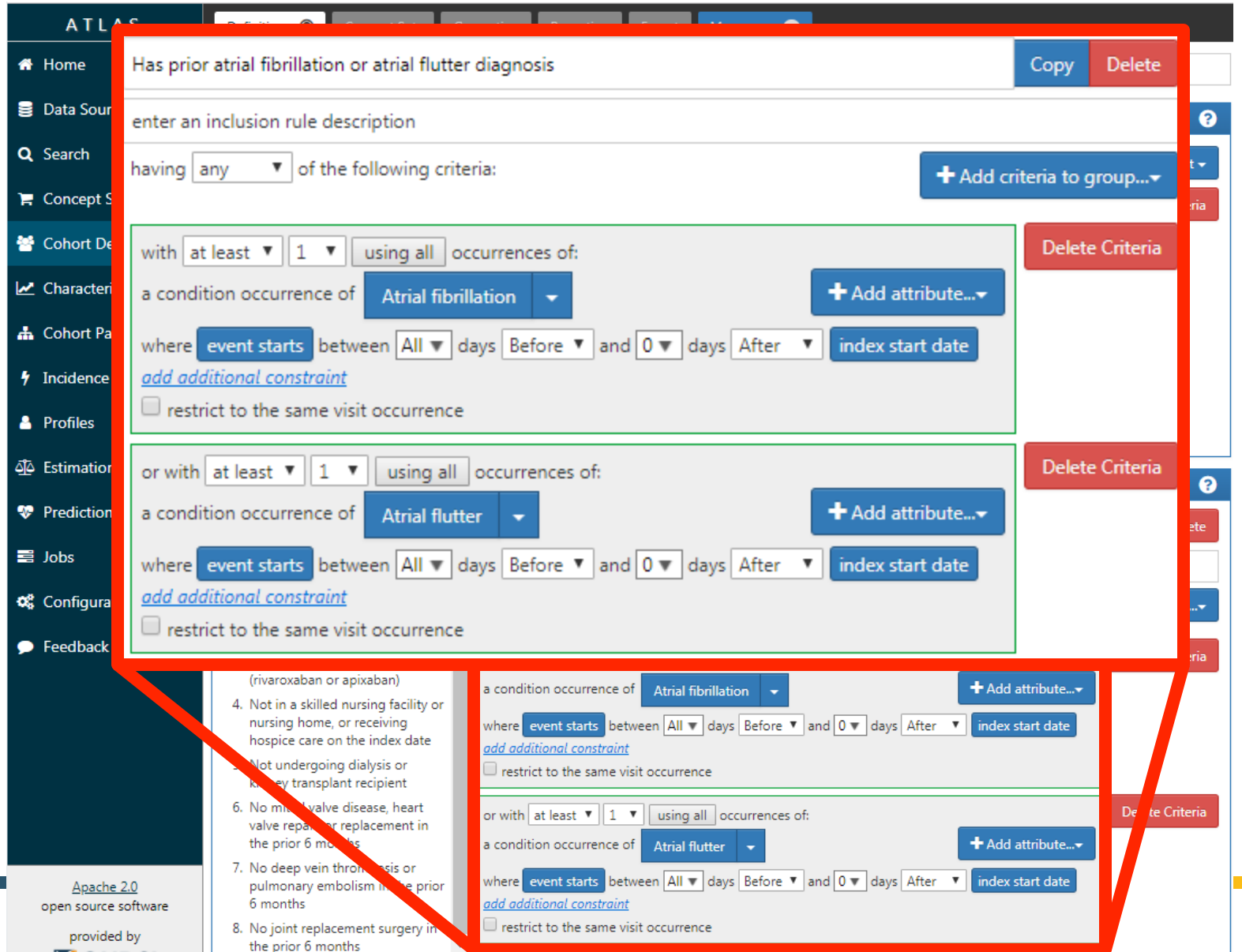
Graham et al. replication: Designing the target cohort in ATLAS



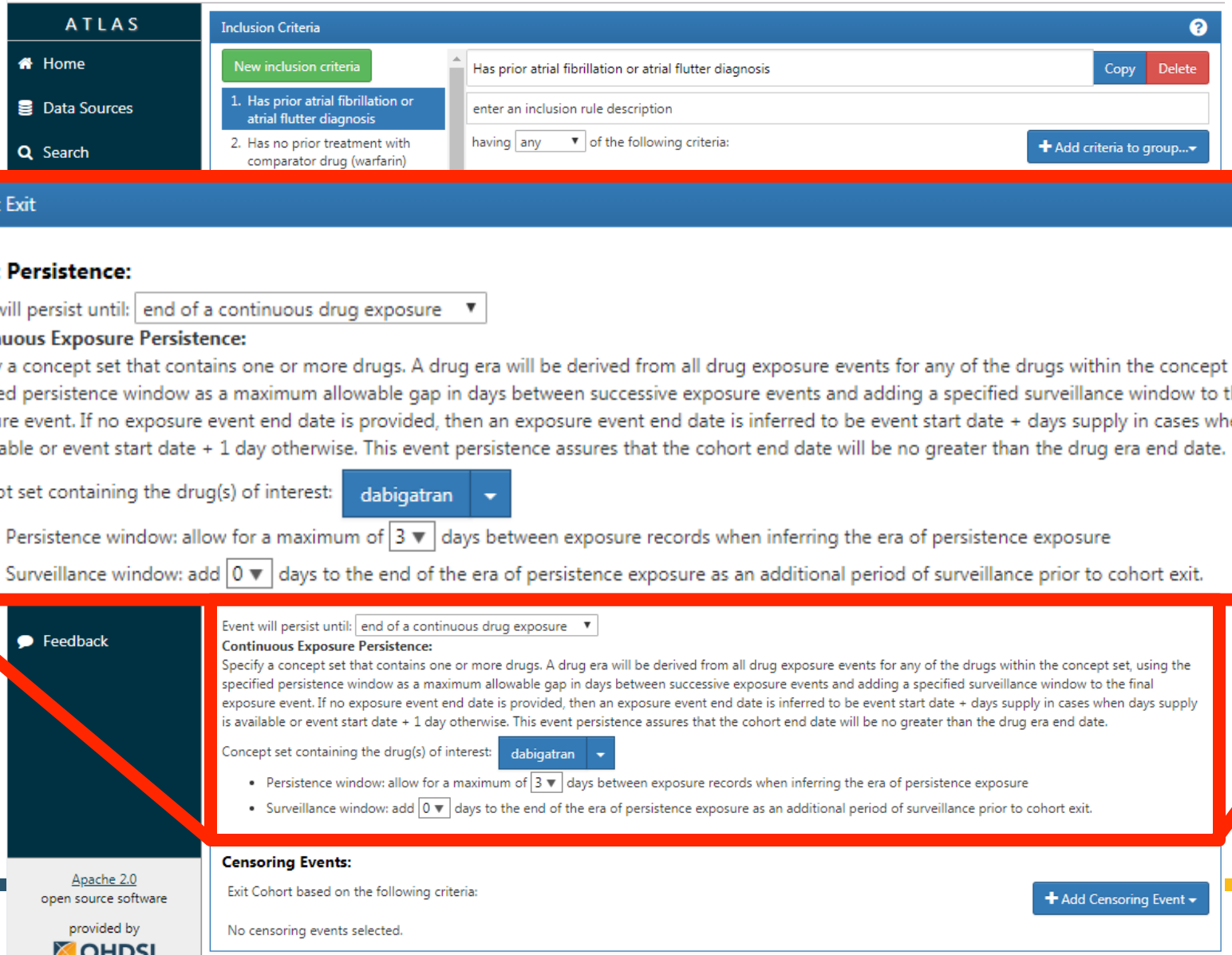
Graham et al. replication: Designing the target cohort in ATLAS



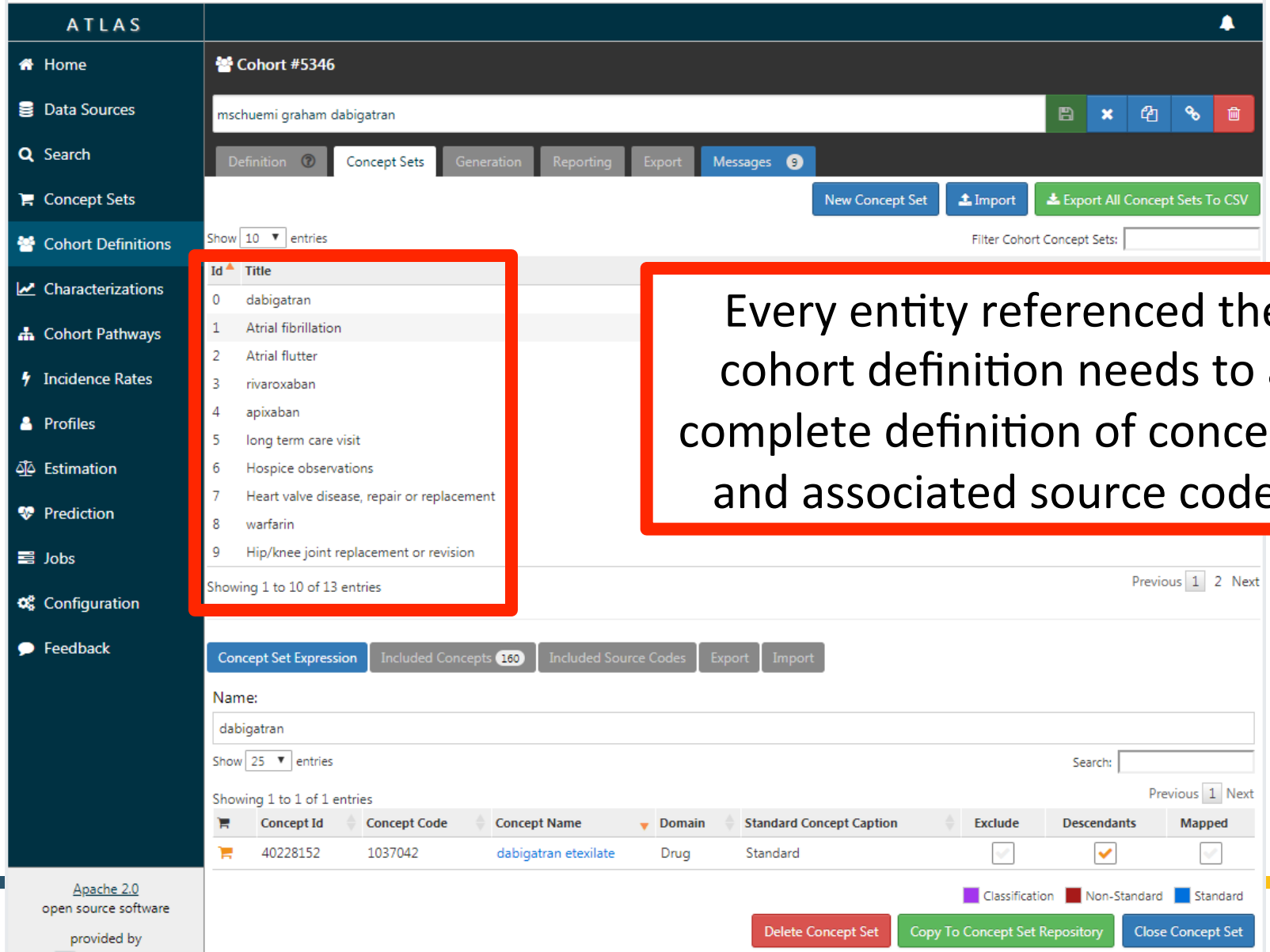
Graham et al. replication: Designing the target cohort in ATLAS



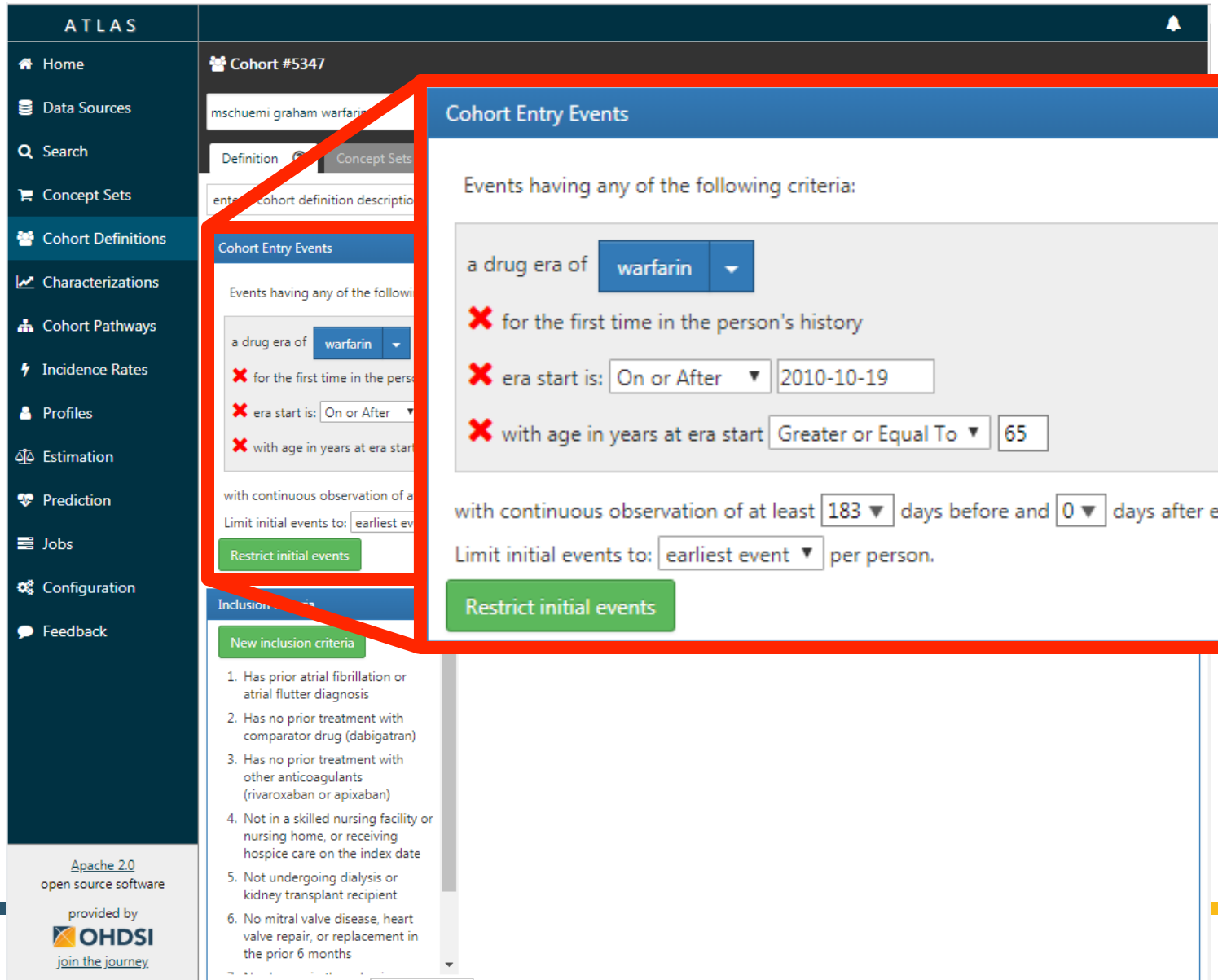
Graham et al. replication: Designing the target cohort in ATLAS



Graham et al. replication: Designing the target cohort in ATLAS



Graham et al. replication: Designing the target cohort in ATLAS



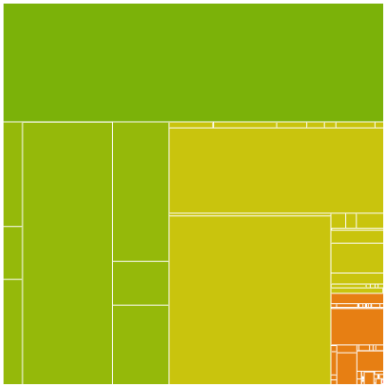
Graham et al. replication: Evaluating the impact of inclusion criteria on the comparator cohort in ATLAS

Inclusion Report for Truven MDCR (v779)

		Match Rate	Matches	Total
	Summary Statistics:	31.13%	56,648	182,001

	Inclusion Rule	N	% Remain	% Diff
1.	Has prior atrial fibrillation or atrial flutter diagnosis	86,995	47.80%	52.20%
2.	Has no prior treatment with comparator drug (dabigatran)	83,161	45.69%	2.11%
3.	Has no prior treatment with other anticoagulants (rivaroxaban or apixaban)	78,222	42.98%	2.71%
4.	Not in a skilled nursing facility or nursing home, or receiving hospice care on the index date	78,161	42.95%	0.03%
5.	Not undergoing dialysis or kidney transplant recipient	76,110	41.82%	1.13%
6.	No mitral valve disease, heart valve repair, or replacement in the prior 6 months	69,645	38.27%	3.55%
7.	No deep vein thrombosis or pulmonary embolism in the prior 6 months	59,195	32.52%	5.74%
8.	No joint replacement surgery in the prior 6 months	56,648	31.13%	1.40%

Population Visualization Switch to attrition view



Switch to intersect view

Attrition Visualization

Graham et al. description of the outcomes

Study Outcomes

The primary outcomes were ischemic stroke, major bleeding with specific focus on intracranial and gastrointestinal bleeding, and AMI. Secondary outcomes were all hospitalized bleeding events and mortality. The *International Classification of Diseases, Ninth Revision, Clinical Modification* codes used to define these outcomes are listed in Table II in the online-only Data Supplement. The codes defining ischemic stroke have a positive predictive value (PPV) of 88% to 95%.[18–20] Major bleeding was defined as

Table 2. International Classification of Disease, 9th edition, Clinical Modification (ICD 9-CM) codes used to define study outcomes.

Outcome	ICD-9 Codes	Position	Setting
AMI	410 (all)	1st or 2nd	IP only
Ischemic stroke	433.x1, 434.x (except subcode: x0), 436	1st	IP only

Graham et al. description of the outcomes

Study Outcomes

The primary outcomes were ischemic stroke, major bleeding with specific focus on intracranial and gastrointestinal bleeding, and AMI. Secondary outcomes were all hospitalized bleeding events and mortality. The *International Classification of Diseases, Ninth Revision, Clinical Modification* codes used to define these outcomes are listed in Table II in the online-only Data Supplement. The codes defining ischemic stroke have a positive predictive value (PPV) of 88% to 95%[18–20] Major bleeding was defined as

…and look, ischemic stroke is a 'validated' outcome!

Table 2. International Classification of Disease, 9th edition, Clinical Modification (ICD 9-CM) codes used to define study outcomes.

Outcome	ICD-9 Codes	Position	Setting
AMI	410 (all)	1st or 2nd	IP only
Ischemic stroke	433.x1, 434.x (except subcode: x0), 436	1st	IP only

Sidebar: just how good was that validation?

Ref 18 from Graham et al.:

Validating Administrative Data in Stroke Research

David L. Tirschwell, MD; W.T. Longstreth, Jr, MD

Background and Purpose—Research based on administrative data has advantages, including large numbers, consistent data, and low cost. This study was designed to compare different methods of stroke classification using administrative data.

Methods—Administrative hospital discharge data and medical record review of 206 patients were used to evaluate 3 algorithms for classifying stroke patients. These algorithms were based on all (algorithm 1), the first 2 (algorithm 2), or the primary (algorithm 3) administrative discharge diagnosis code(s). The diagnoses after review of medical record data were considered the gold standard. Then, using a large administrative data set, we compared patients with a primary discharge diagnosis of stroke with patients with their stroke discharge diagnosis code in a nonprimary position.

Results—Compared with the gold standard, algorithm 1 had the highest κ for classifying ischemic stroke, with a sensitivity of 86%, specificity of 95%, positive predictive value of 90%, and κ=0.82. Algorithm 3 had the highest κ values for intracerebral hemorrhage and subarachnoid hemorrhage. For intracerebral hemorrhage, the sensitivity was 85%, specificity was 96%, positive predictive value was 89%, and κ=0.82. For subarachnoid hemorrhage, those values were 90%, 97%, 94%, and 0.88, respectively. Nonprimary position ischemic stroke patients had significantly greater comorbidity and 30-day mortality (odds ratio, 3.2) than primary position ischemic stroke patients.

Conclusions—Stroke classification in these administrative data were optimal using all discharge diagnoses for ischemic stroke and primary discharge diagnosis only for intracerebral and subarachnoid hemorrhage. Selecting ischemic stroke patients on the basis of primary discharge diagnosis may bias administrative samples toward more benign, unrepresentative outcomes and should be avoided. (***Stroke.*** **2002;33:2465-2470.)**

http://www.ncbi.nlm.nih.gov/pubmed/12364739

Sidebar: just how good was that validation?

Ref 18 from Graham et al.:

Validating Administrative Data in Stroke Research

David L. Tirschwell, MD; W.T. Longstreth, Jr, MD

Background and Purpose—Research based on administrative data has advantages, including large numbers, consistent data, and low cost. This study was designed to compare different methods of stroke classification using administrative data.

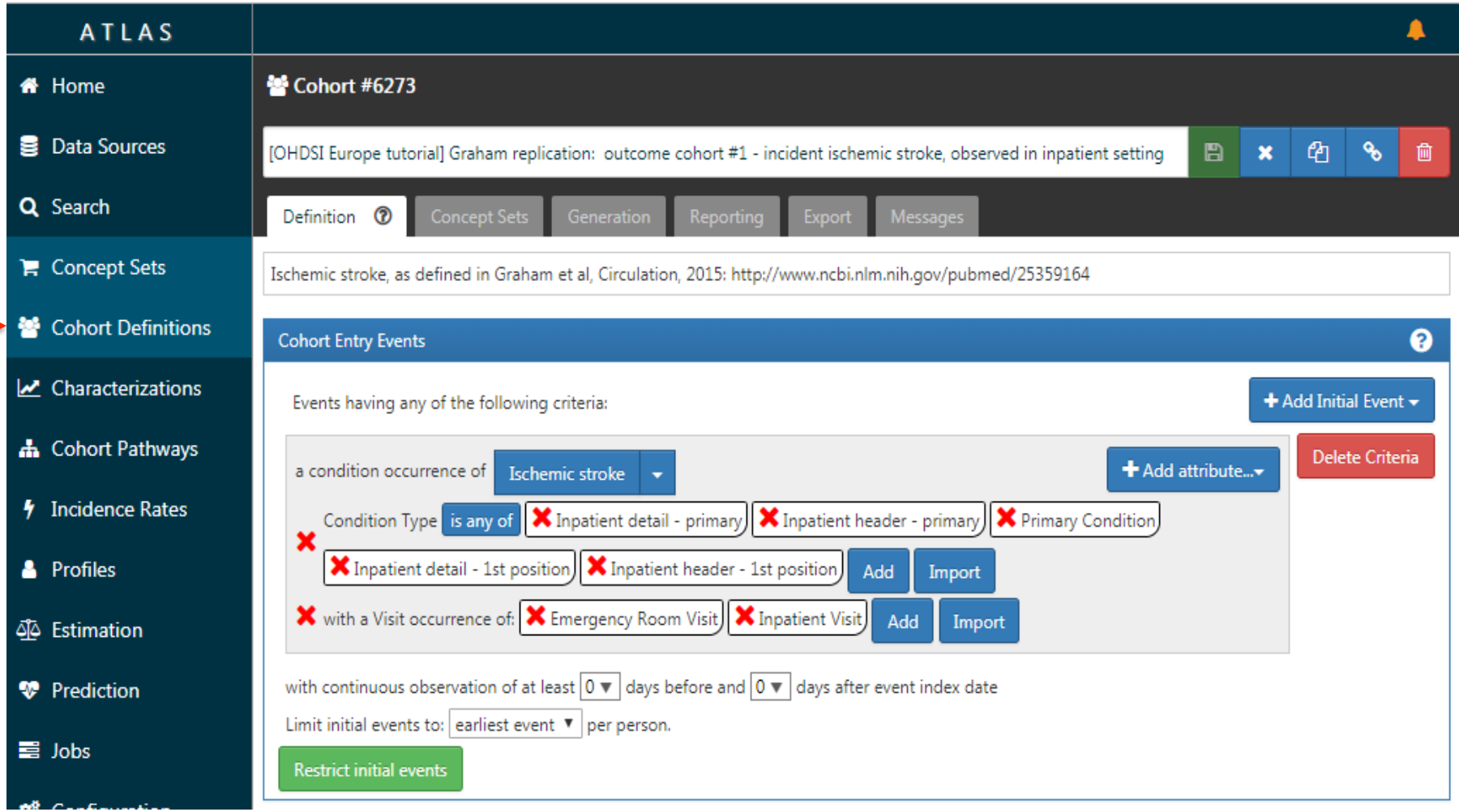
Methods—Administrative hospital discharge data and medical record review of 206 patients were used to evaluate 3 algorithms for classifying stroke patients. These algorithms were based on all (algorithm 1), the first 2 (algorithm 2), or the primary (algorithm 3) administrative discharge diagnosis code(s). The diagnoses after review of medical record data were considered the gold standard. Then, using a large administrative data set, we compared patients with a primary discharge diagnosis of stroke with patients with their stroke discharge diagnosis code in a nonprimary position.

Results—Compared with the gold standard, algorithm 1 had the highest κ for classifying ischemic stroke, with a sensitivity of 86%, specificity of 95%, positive predictive value of 90%, and κ=0.82. Algorithm 3 had the highest κ values for intracerebral hemorrhage and subarachnoid hemorrhage. For intracerebral hemorrhage, the sensitivity was 85%, specificity was 9
90%, 97%, 94%
comorbidity and
Conclusions—Stro
stroke and prima
patients on the
unrepresentative

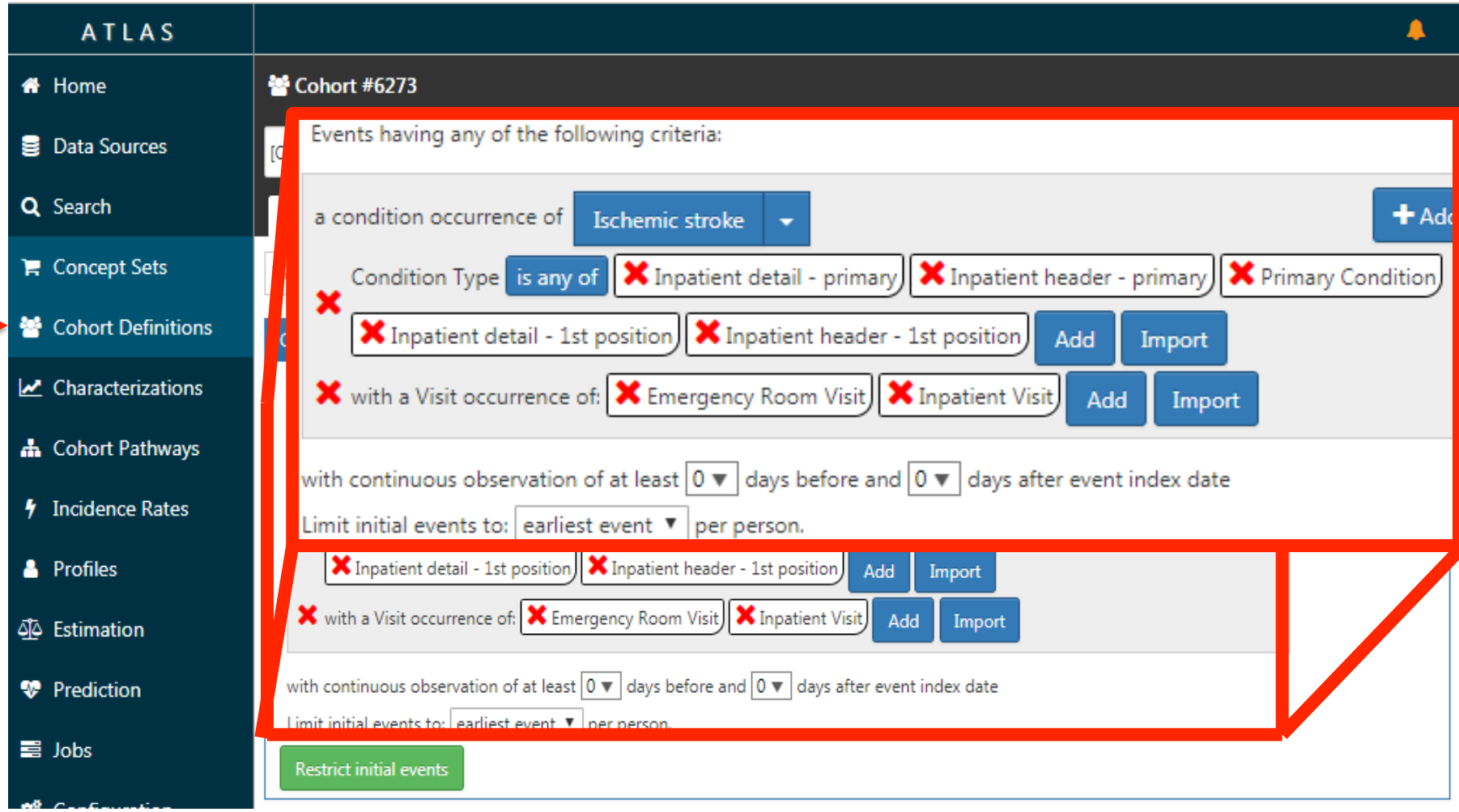
Conclusions- Stroke classification in these administrative data were **optimal using all discharge diagnoses** for ischemic stroke….Selecting ischemic stroke patients on the basis of **primary discharge diagnosis may bias** administrative samples toward more benign, unrepresentative outcomes and **should be avoided**.

http://www.ncbi.nlm.nih.gov/pubmed/12364739

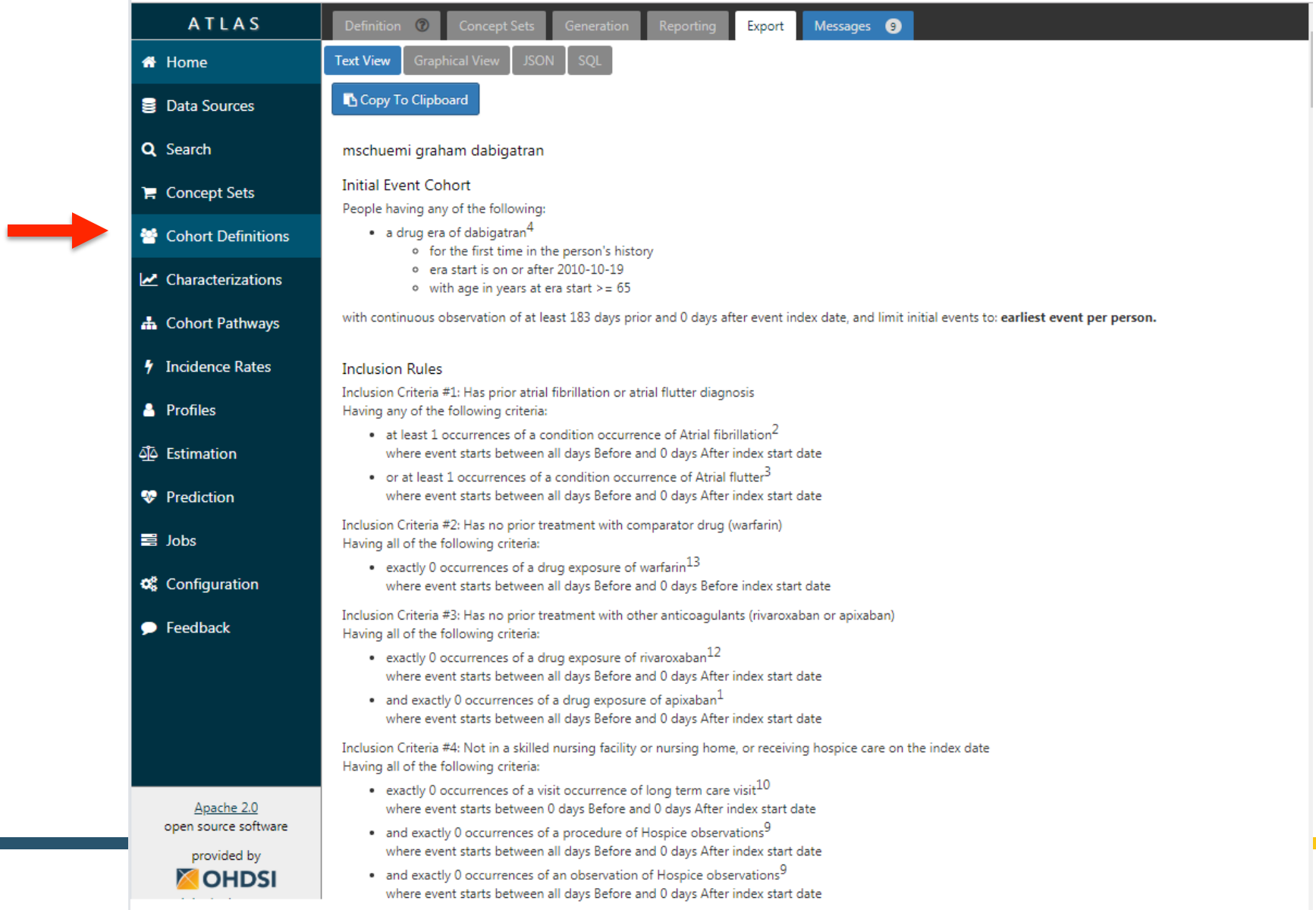
Graham et al. replication: Designing the outcome cohort in ATLAS



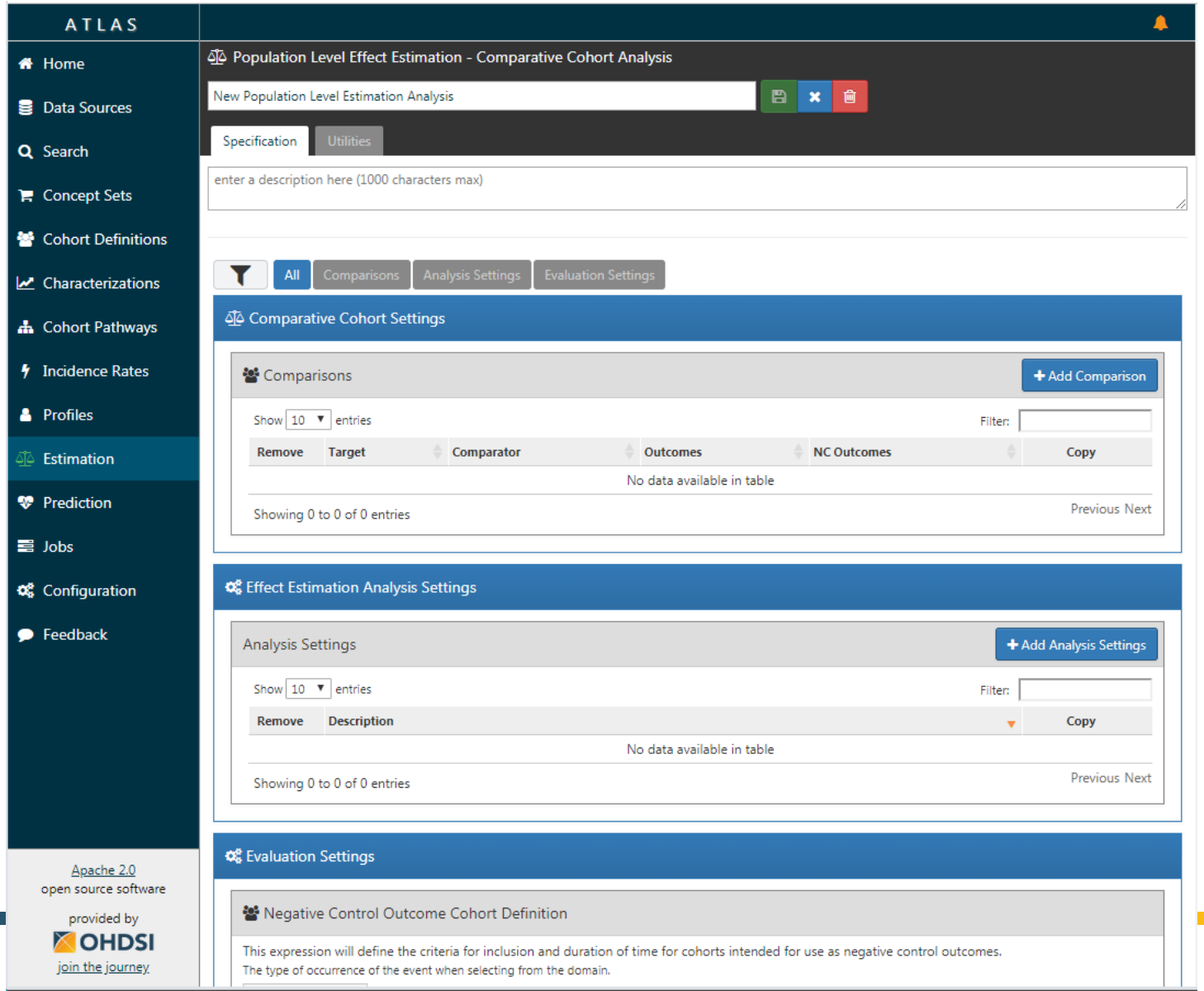
Graham et al. replication: Designing the outcome cohort in ATLAS



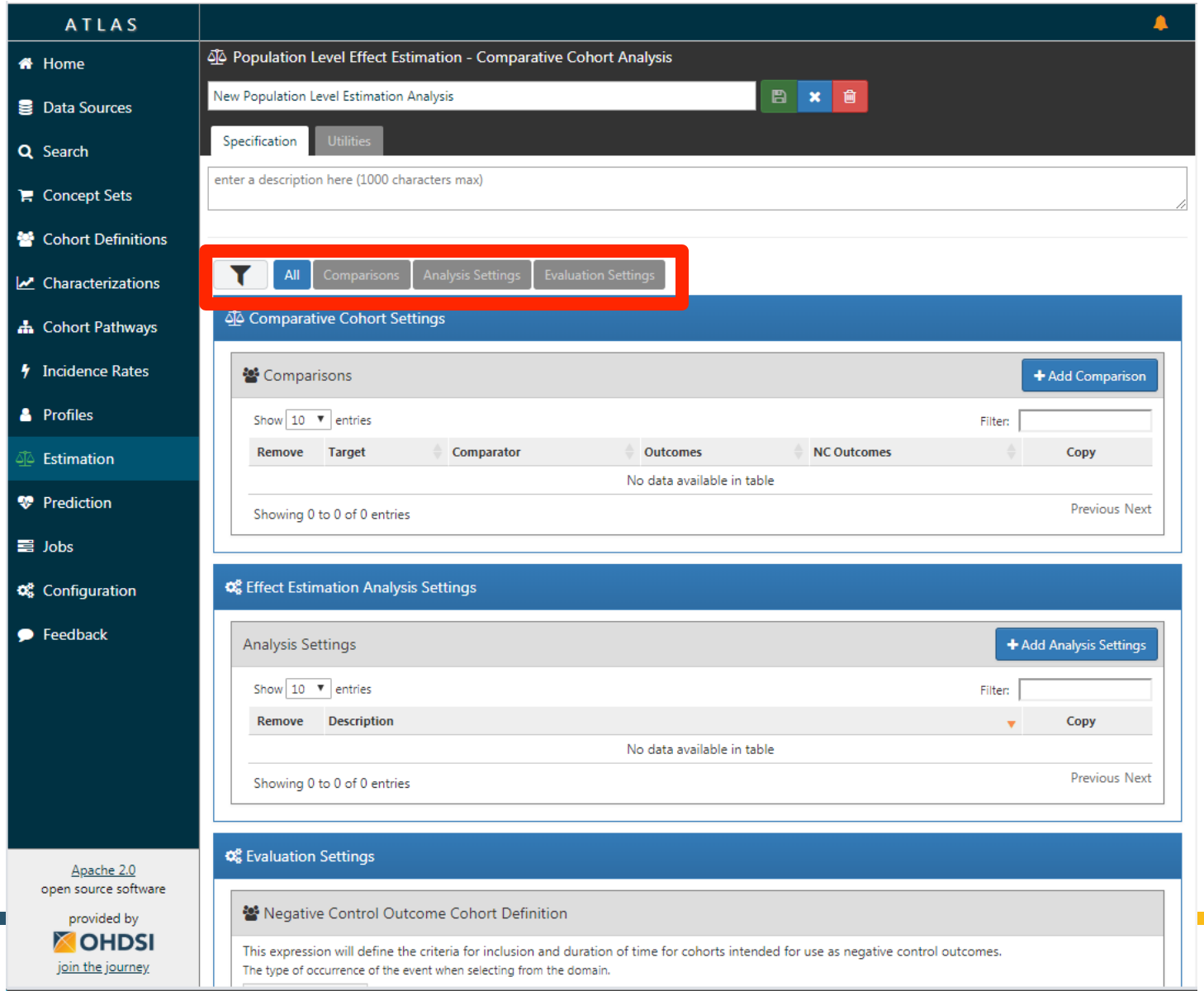
Graham et al. replication: Cohort exports



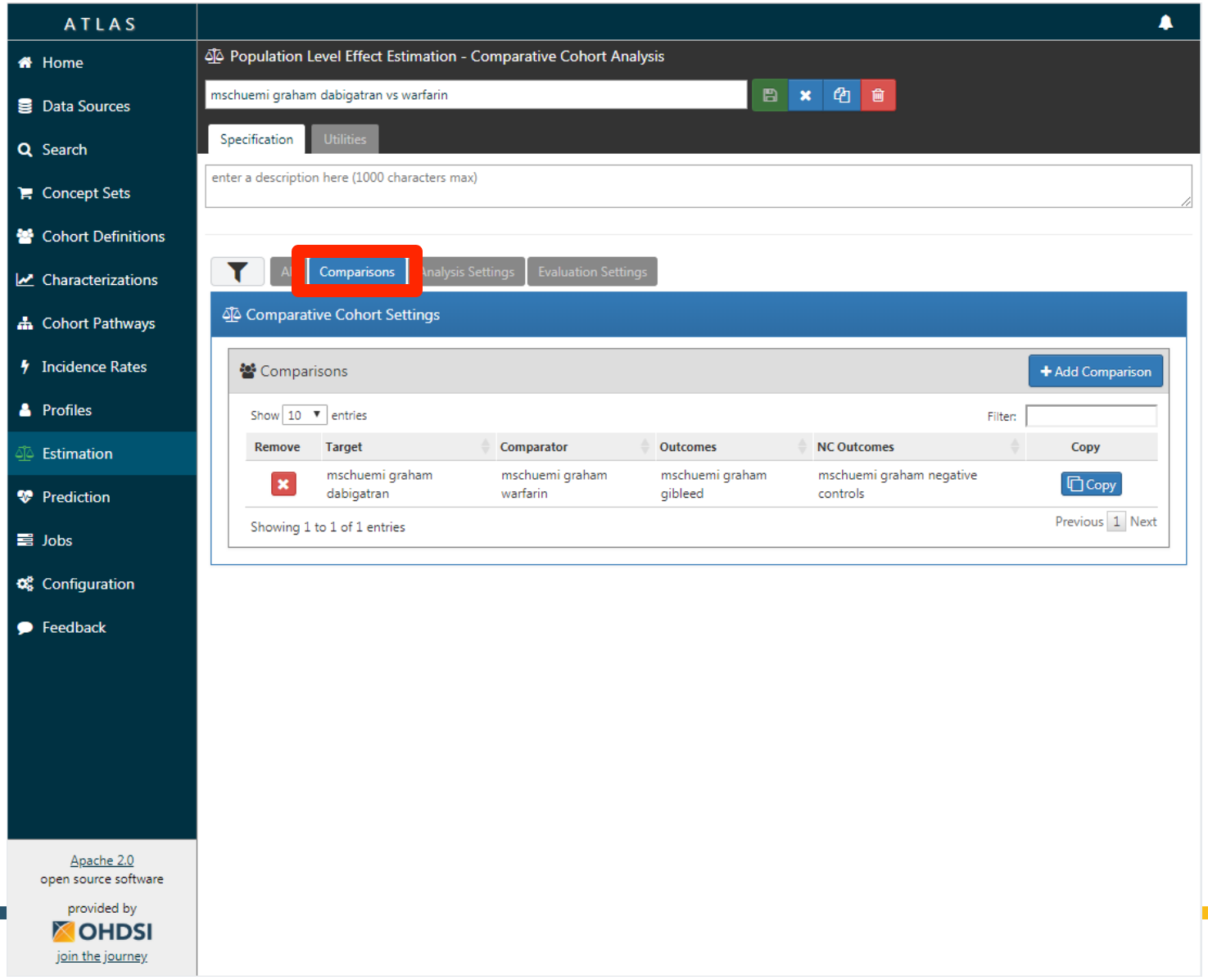
Graham et al. replication: Designing the study in ATLAS



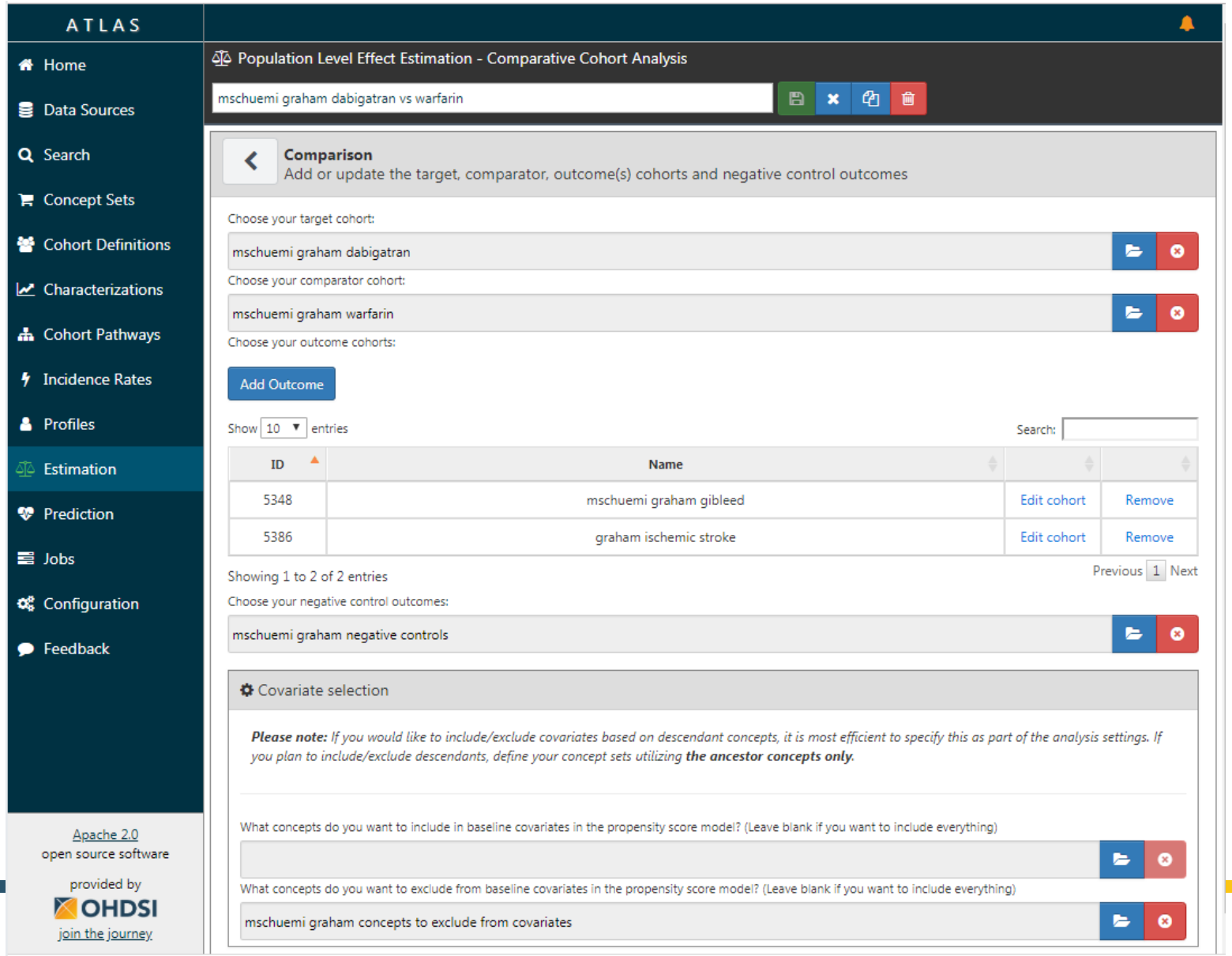
Graham et al. replication: Designing the full study in ATLAS



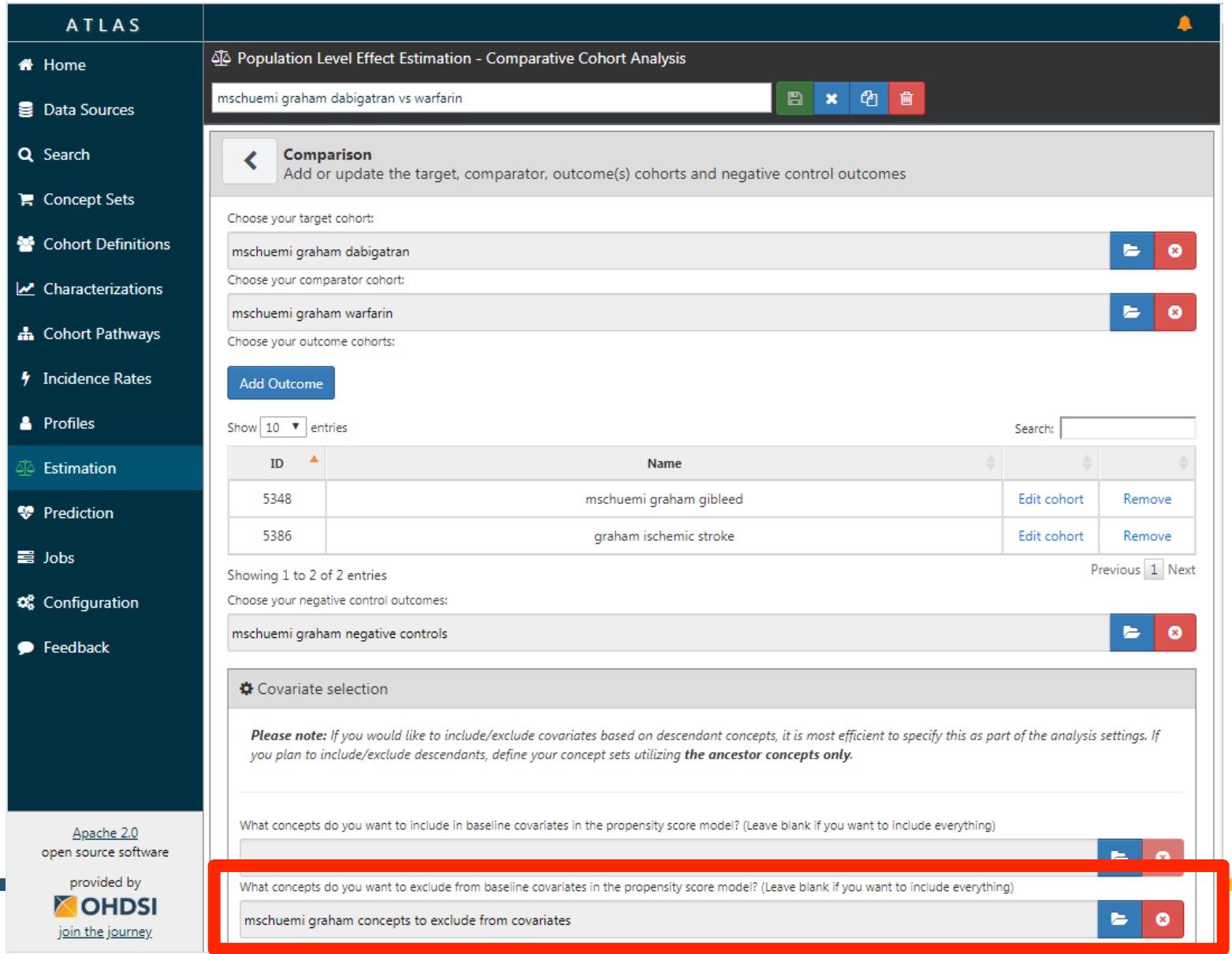
Graham et al. replication: Specifying the comparison



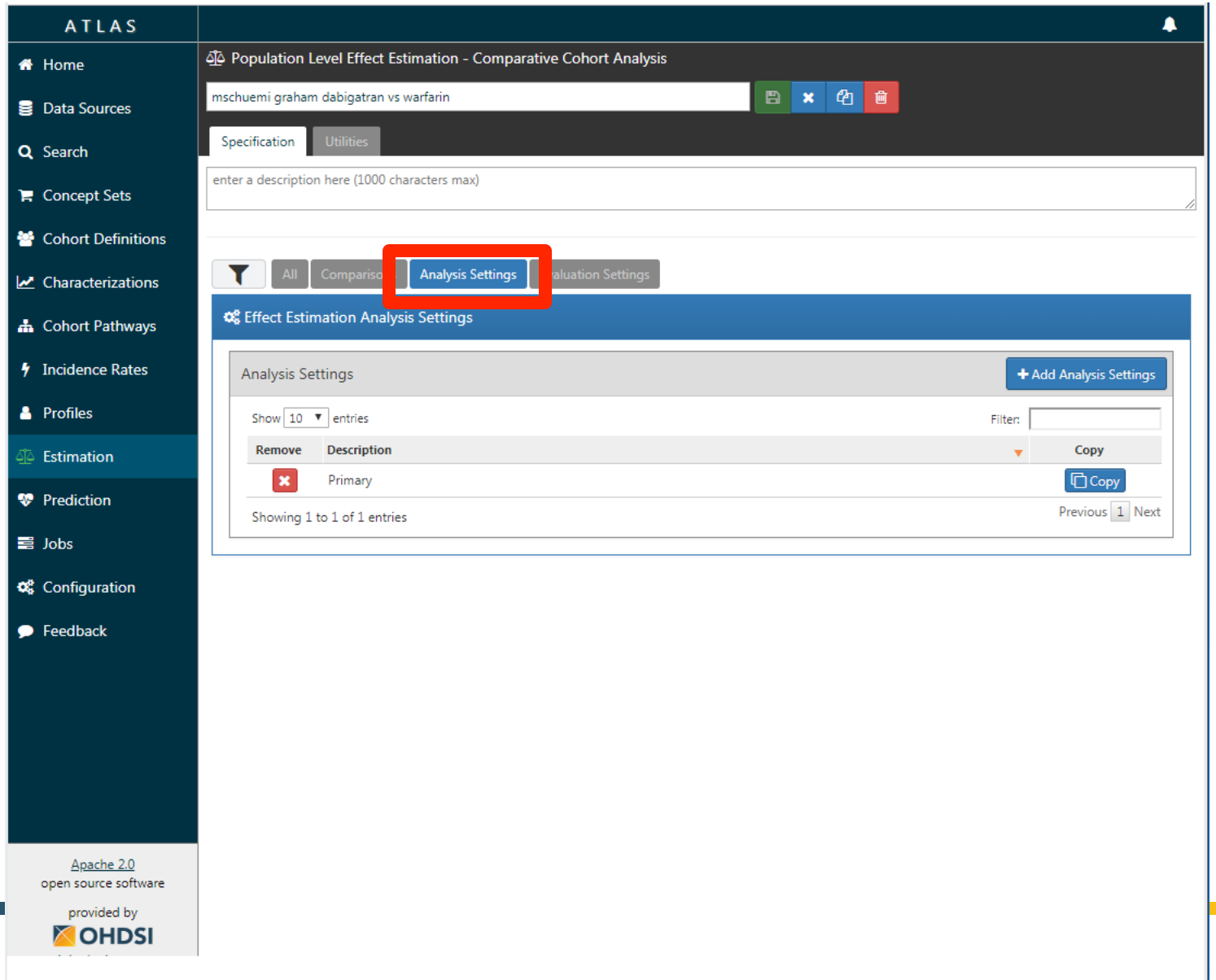
Graham et al. replication: Specifying the comparison



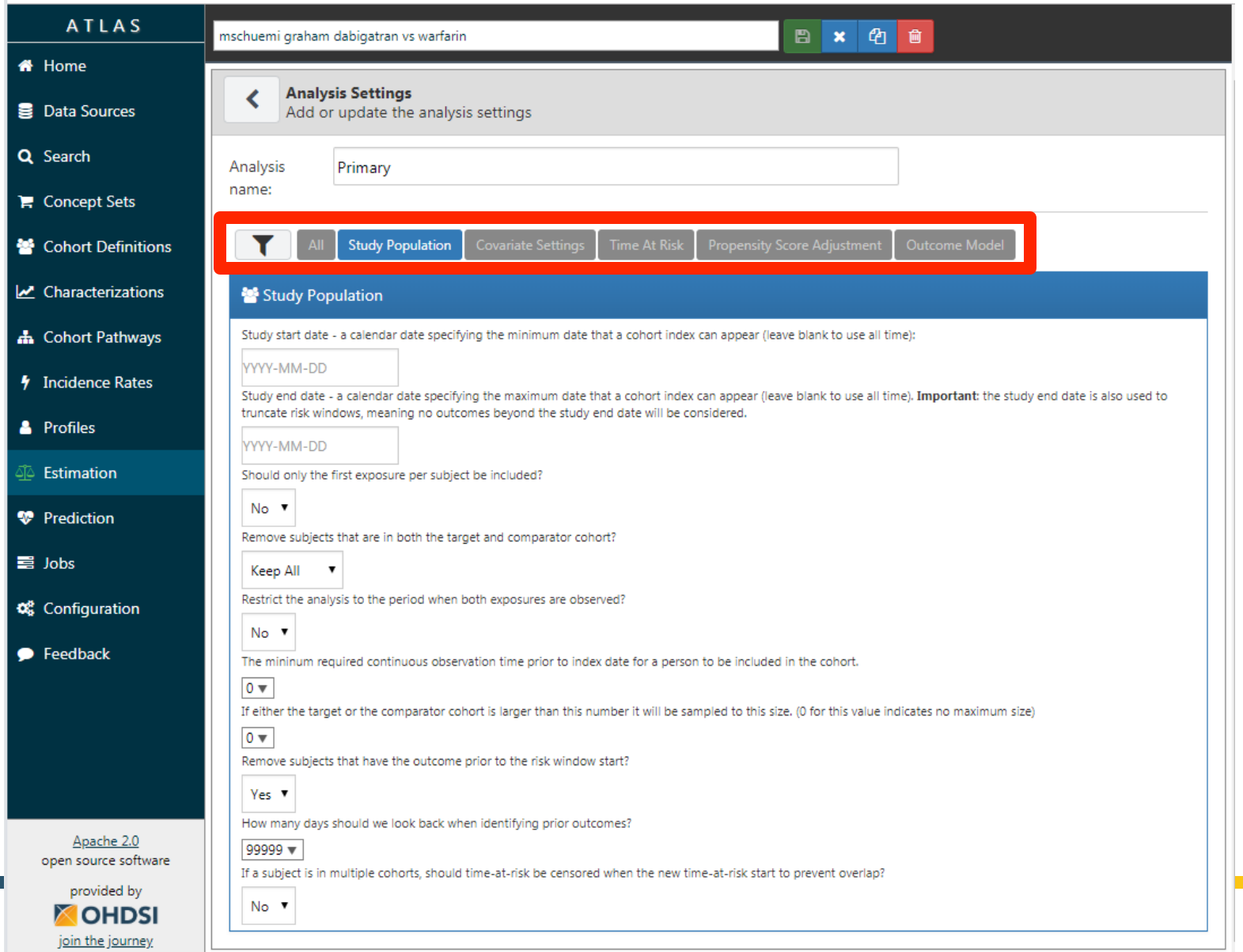
Graham et al. replication: Specifying the comparison



Graham et al. replication: Specifying the analysis settings



Graham et al. replication: Specifying the analysis settings

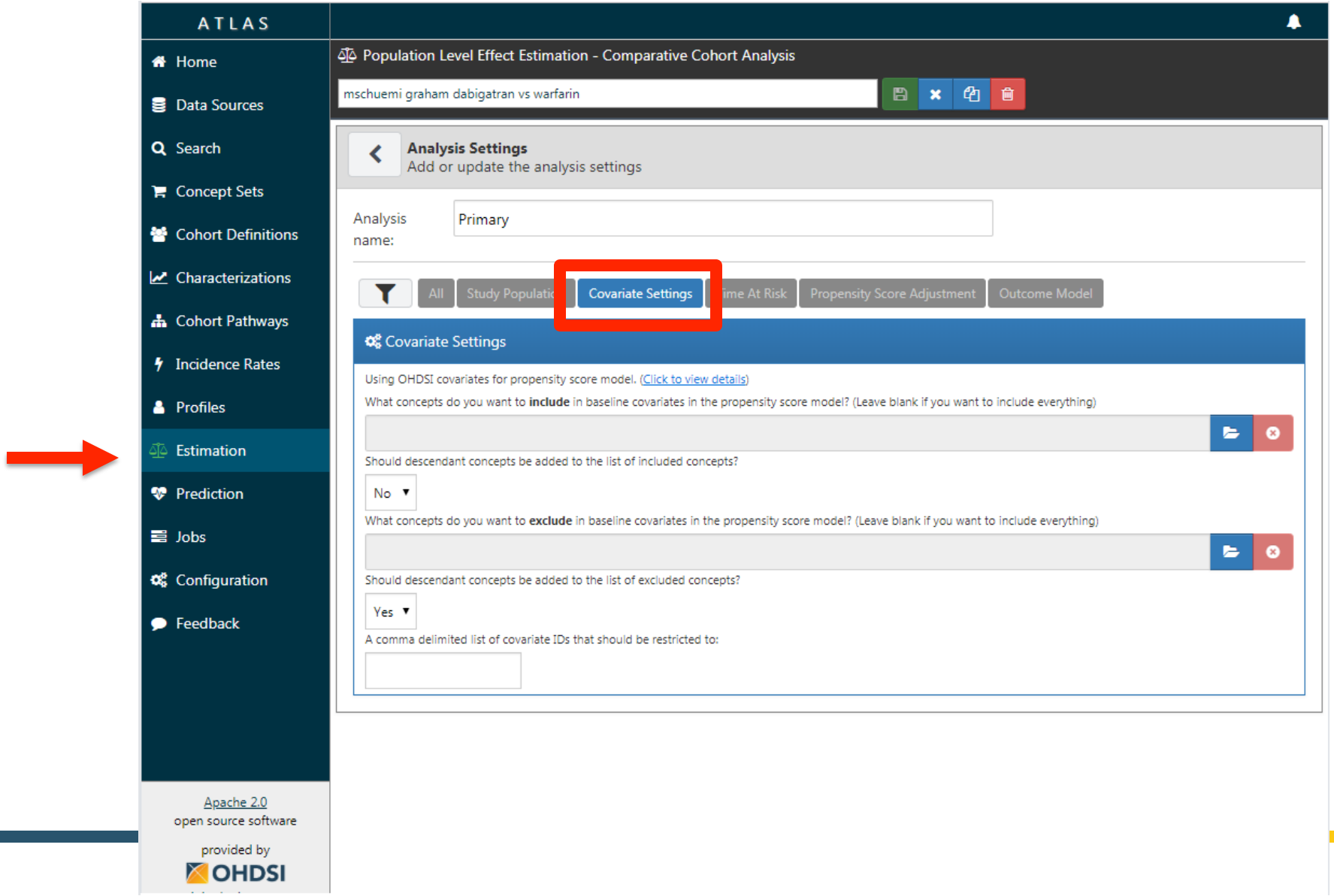


Graham et al. replication: covariates for confounding adjustment

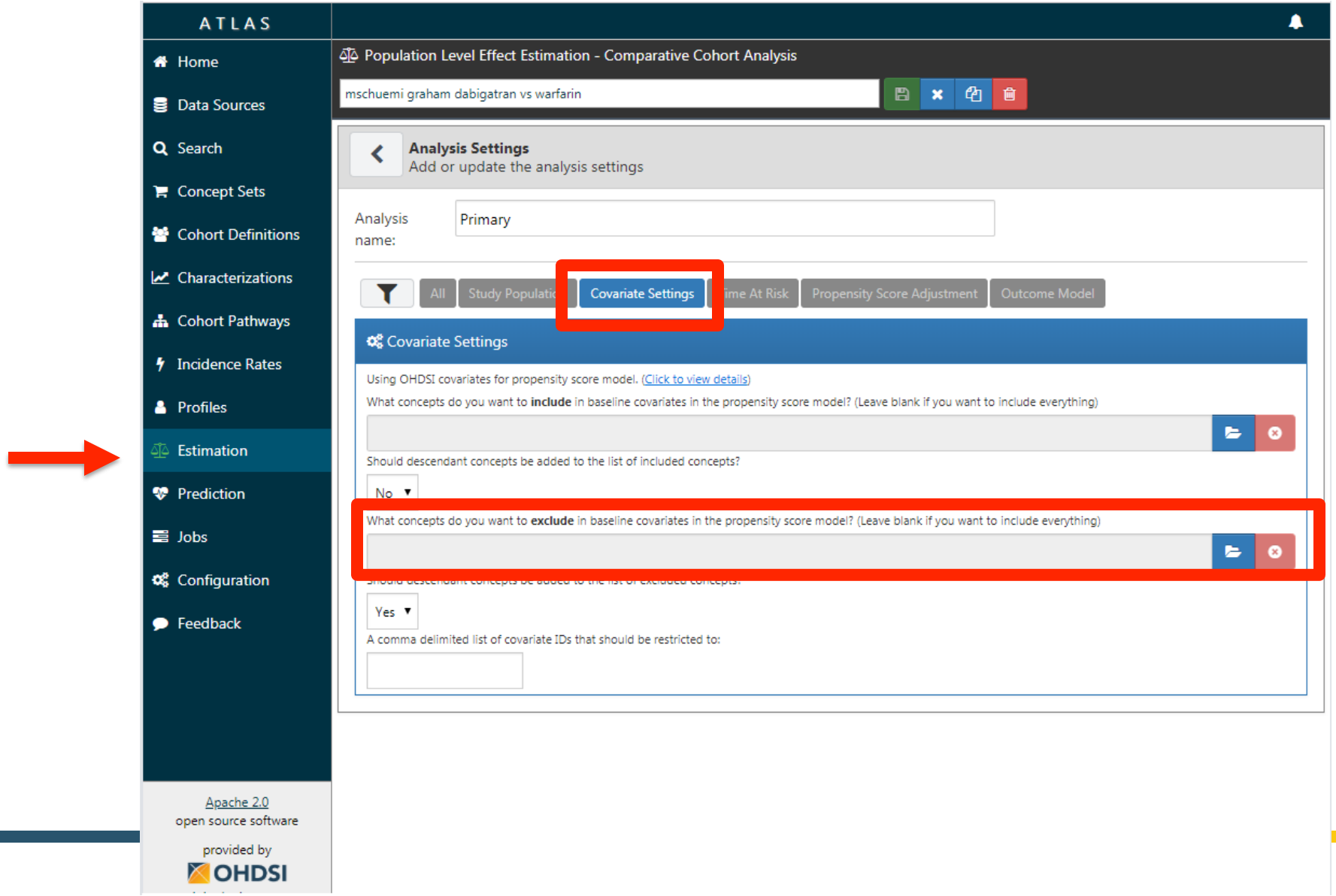
Claims data on chronic medical conditions, cardiovascular risk factors, risk factors for bleeding events, and healthcare utilization were collected for each patient during the 6 months preceding their cohort-qualifying prescription fill. We also collected data on prescriptions for medications used for treatment of cardiovascular disease and other chronic medical conditions, as well as potentially interacting medications that might alter warfarin or dabigatran pharmacokinetics. Finally, to the extent possible using claims data, we calculated the $CHADS_2$ score,[11] which predicts the risk of stroke in patients with AF, and the HAS-BLED score,[12,13] which predicts the risk of bleeding in patients with AF treated with warfarin.

To reduce confounding due to imbalance in study covariates, propensity score matching was used.[14–16] Unconditional logistic regression was used to estimate the predicted probability of patients initiating dabigatran therapy given their sociodemographic characteristics, baseline medical comorbidities, medications used during the preceding 6 months, prescriber characteristics, and other potentially relevant variables (Table 1 and

Graham et al. replication: covariates for confounding adjustment



Graham et al. replication: covariates for confounding adjustment



Graham et al. replication: covariates for confounding adjustment

- Leverages FeatureExtraction package
- Default settings create covariates for all drug and condition group concepts, and procedure, measurement, observation, and device exposure concepts during 2 lookback window
- Defaults also include demographics and 4 risk indexes

Select Covariates...

	Gender	Age	Age Groups	Race	Ethnicity	Index Year	Index Month	Prior Observation Time	Post Observation Time	Time In Cohort	Index Year & Month
Demographics	☑	☐	☑	☑	☑	☑	☑	☐	☐	☐	☐

Time bound covariates

Set the time windows for the time bound covariates in days relative to the cohort index

	Any Time Prior	Long Term	Medium Term	Short Term	End Days
Time Windows	All Time	-365	-180	-30	0

Set the time bound era covariates

Domain	Any Time Prior	Long Term (-365 days)	Medium Term (-180 days)	Short Term (-30 days)	Overlapping	Era Start: Long Term (-365 days)	Era Start: Medium Term (-180 days)	Era Start: Short Term (-30 days)
Condition	☐	☐	☐	☐	☐	☐	☐	☐
Condition Group	☐	☑	☐	☑	☐	☐	☐	☐
Drug	☐	☐	☐	☐	☐	☐	☐	☐
Drug Group	☐	☑	☐	☑	☑	☐	☐	☐

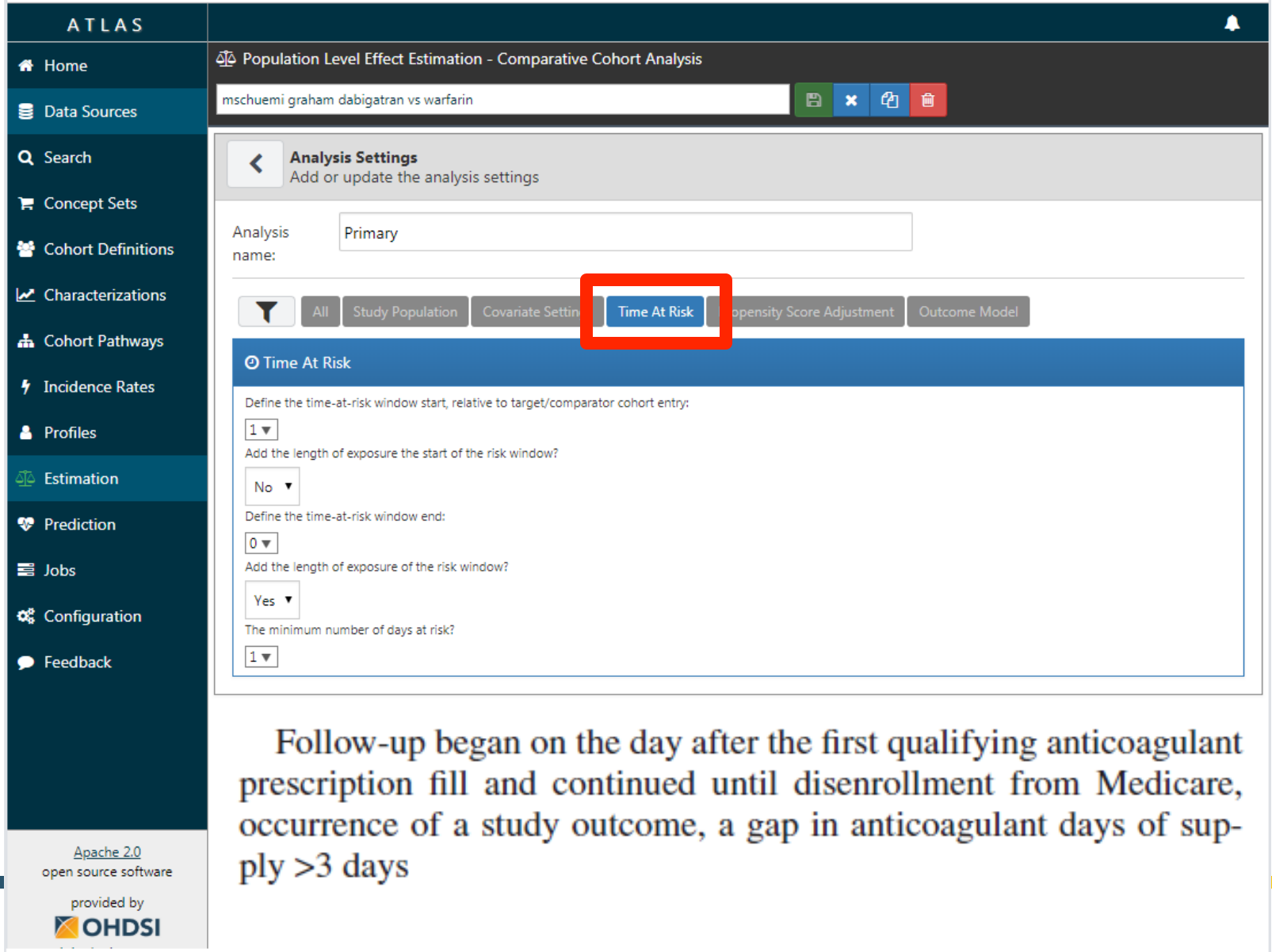
Set the time bound covariates

Domain	Any Time Prior	Long Term (-365 days)	Medium Term (-180 days)	Short Term (-30 days)	Distinct Count: Long Term (-365 days)	Distinct Count: Medium Term (-180 days)	Distinct Count: Short Term (-30 days)
Condition	☐	☐	☐	☐	☐	☐	☐
Condition - Primary Inpatient	☐	☐	☐	☐			
Drug	☐	☐	☐	☐	☐	☐	☐
Procedure	☐	☑	☐	☑	☐	☐	☐
Measurement	☐	☑	☐	☑	☐	☐	☐
Measurement - Value	☐	☐	☐	☐			
Measurement - Range Group	☐	☑	☐	☐			
Observation	☐	☑	☐	☑	☐	☐	☐
Device	☐	☑	☐	☑			
Visit - Count		☐	☐	☐			
Visit - Concept Count		☐	☐	☐			

Set the index score covariates

Index Score Type	
$CHADS_2$	☑
CHA_2DS_2VASc	☑
DCSI	☑
Charlson	☑

Graham et al. replication: specifying time-at-risk

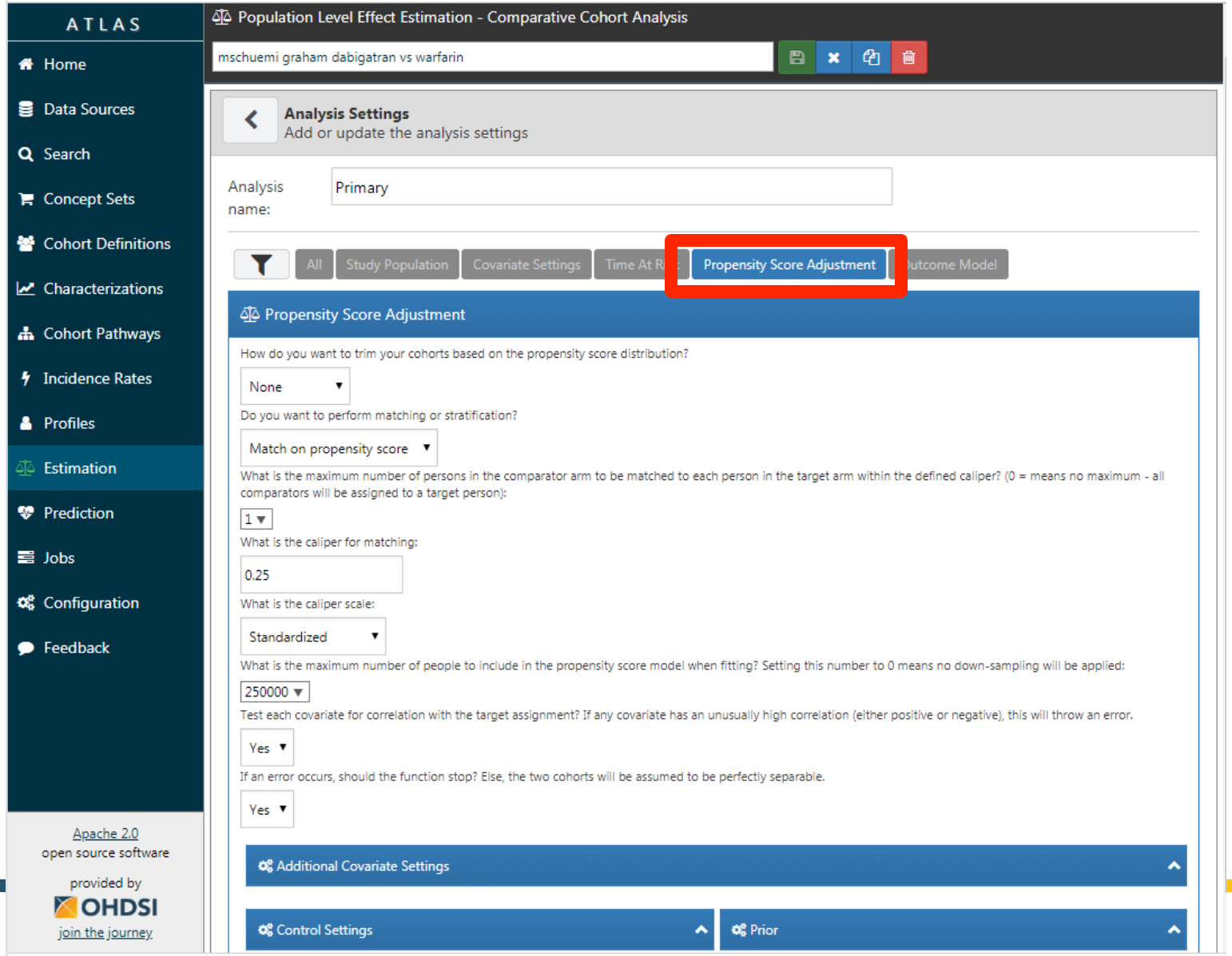


Follow-up began on the day after the first qualifying anticoagulant prescription fill and continued until disenrollment from Medicare, occurrence of a study outcome, a gap in anticoagulant days of supply >3 days

Graham et al. replication: specifying the propensity score model

Dabigatran users were propensity score matched to warfarin users in a 1:1 ratio with the use of a greedy matching algorithm. The balance of measured covariates between the matched cohorts was assessed with the standardized mean difference, a measure not influenced by sample size and thus useful for comparing cohorts in large observational studies.[17] A standardized mean difference of ≤0.1 indicates a negligible difference in the measured variables between groups.[17]

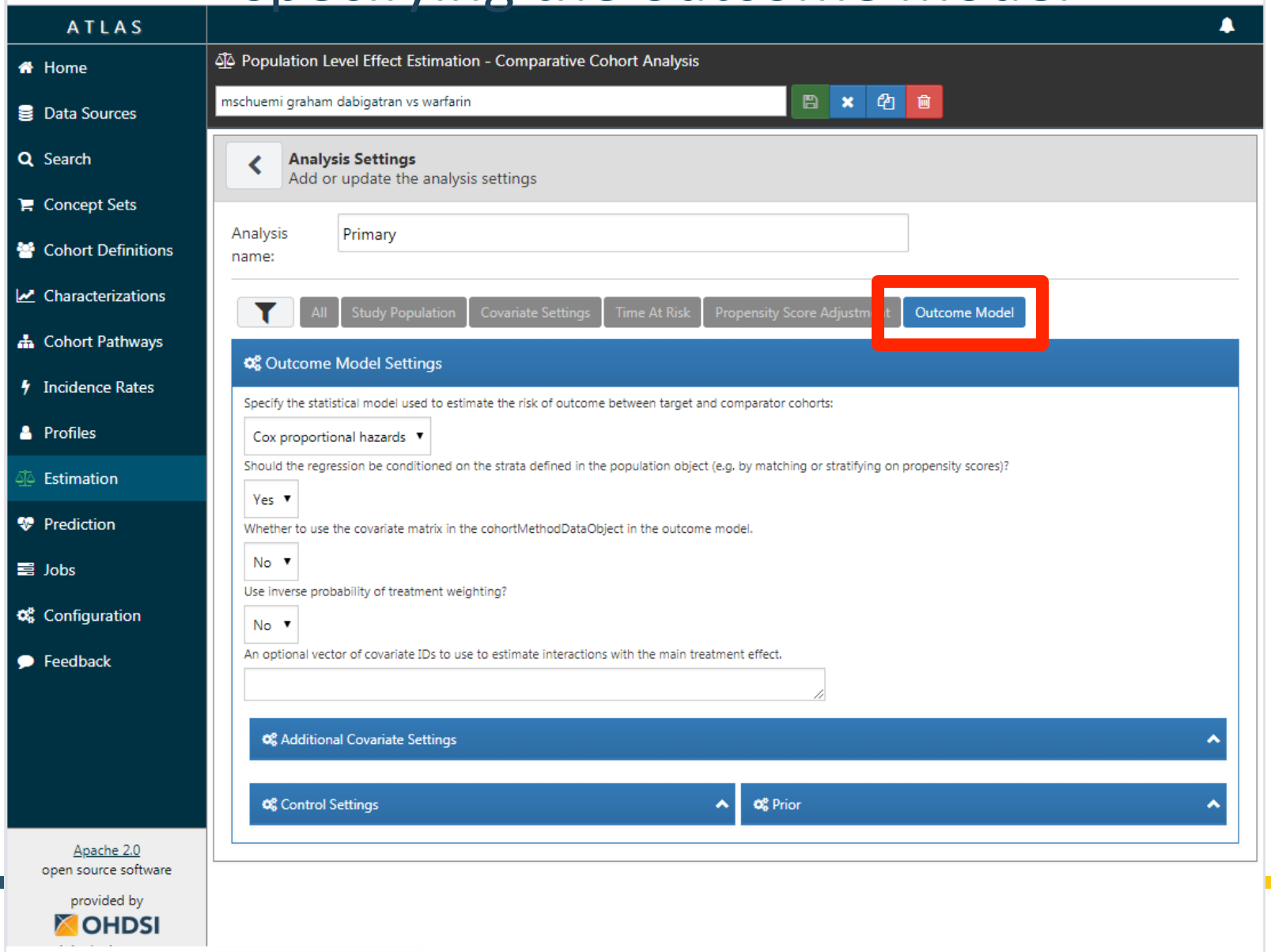
Graham et al. replication: specifying the propensity score model



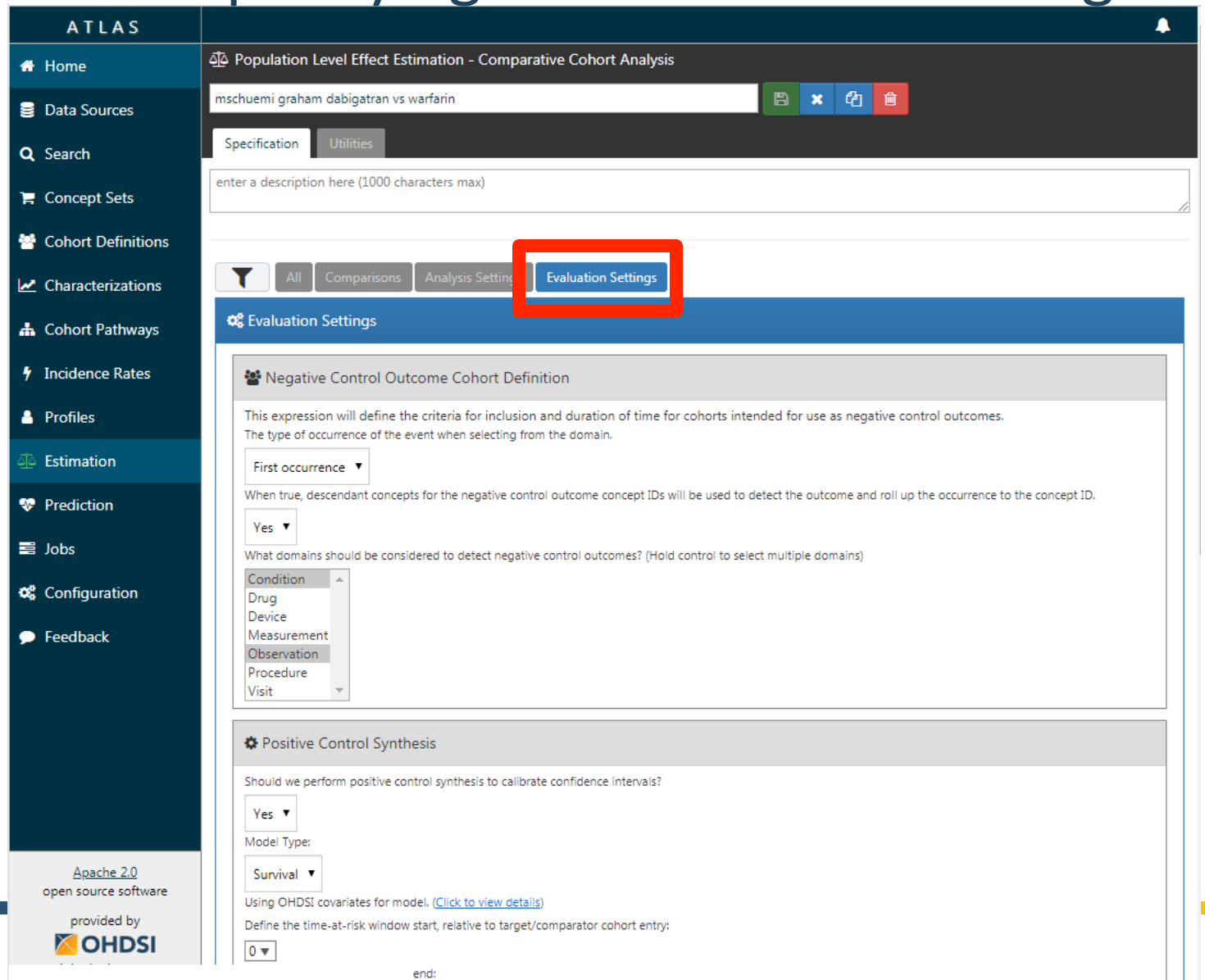
Graham et al. replication: specifying the outcome model

Analyses were performed on the propensity score–matched cohorts, thereby accounting for the potential confounding factors shown in Table 1 and in the online-only Data Supplement. Incidence rates were estimated with the use of event counts and exposure follow-up time. Kaplan–Meier plots were generated to characterize the contour of risk over time for each outcome. Cox proportional hazards regression was used to compare time to event in dabigatran compared with warfarin (reference) cohorts.

Graham et al. replication: specifying the outcome model



Graham et al. replication: specifying the evaluation settings



Graham et al. replication: exporting the study package for execution