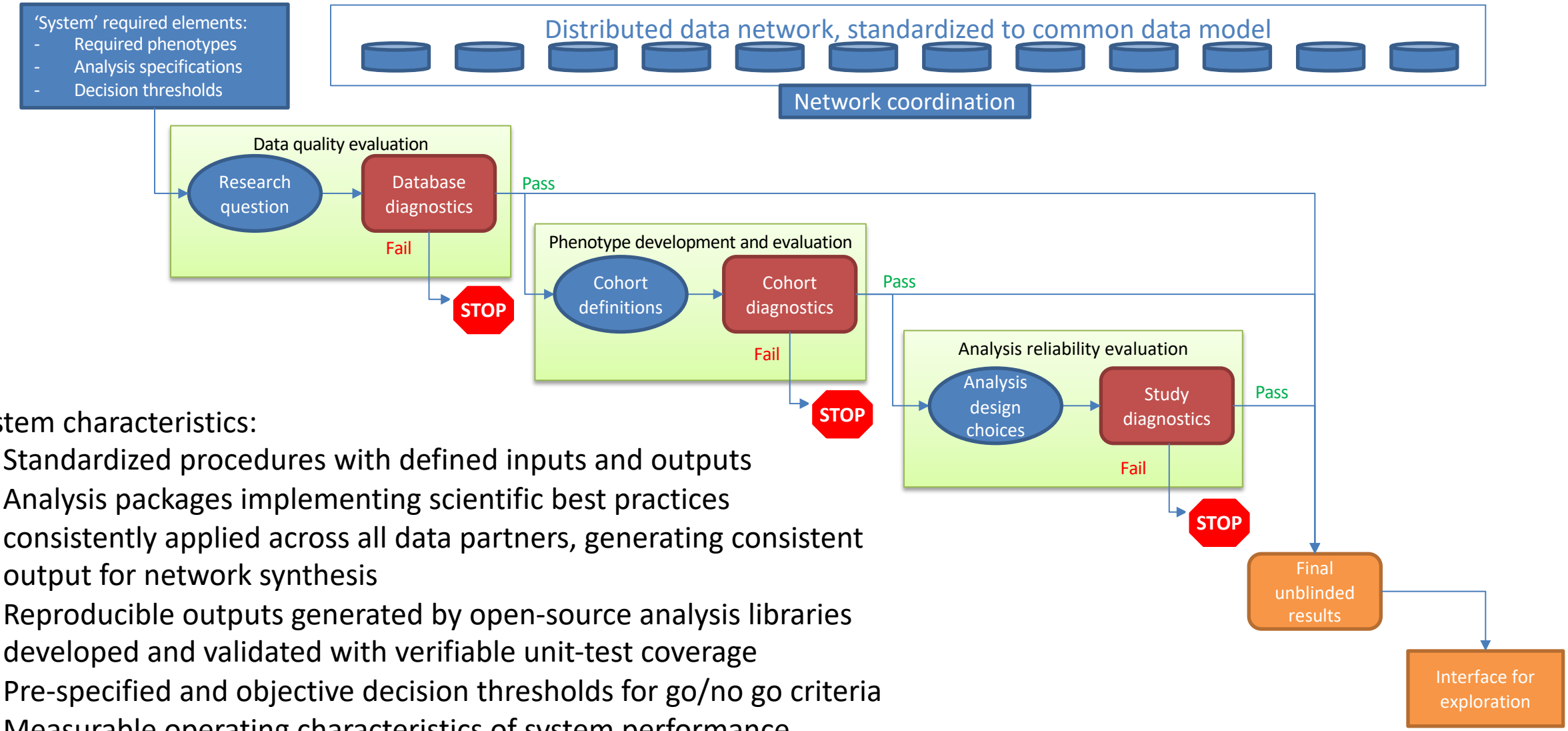
Sisyphus Challenge Week 5: Standardized analysis design

Patrick Ryan, PhD
Vice President, Observational Health Data Analytics, Johnson & Johnson
Adjunct Assistant Professor of Biomedical Informatics at Columbia University

Engineering open science systems that build trust into the real-world evidence generation and dissemination process



System characteristics:

- Standardized procedures with defined inputs and outputs
- Analysis packages implementing scientific best practices consistently applied across all data partners, generating consistent output for network synthesis
- Reproducible outputs generated by open-source analysis libraries developed and validated with verifiable unit-test coverage
- Pre-specified and objective decision thresholds for go/no go criteria
- Measurable operating characteristics of system performance

Engineering open science systems that build trust into the real-world evidence generation and dissemination process

'System' required elements:
- Required phenotypes
- Analysis specifications
- Decision thresholds

Distributed data network, standardized to common data model

Network coordination

Data quality evaluation

Research question → Database diagnostics

Database diagnostics: Pass → Phenotype development and evaluation; Fail → STOP

Phenotype development and evaluation

Cohort definitions → Cohort diagnostics

Cohort diagnostics: Pass → Analysis reliability evaluation; Fail → STOP

Analysis reliability evaluation

Analysis design choices → Study diagnostics

Study diagnostics: Pass → Final unblinded results; Fail → STOP

Final unblinded results → Interface for exploration

Week 1: Study initiation


Week 0: Research questions

VEGF→ESRD

FQ→AA




MS biologics→PML

risankizumab→CVA

Engineering open science systems that build trust into the real-world evidence generation and dissemination process

'System' required elements:
- Required phenotypes
- Analysis specifications
- Decision thresholds

Distributed data network, standardized to common data model

Network coordination

Data quality evaluation

Research question → Database diagnostics

Database diagnostics: Pass → Phenotype development and evaluation

Database diagnostics: Fail → STOP

Phenotype development and evaluation

Cohort definitions → Cohort diagnostics

Cohort diagnostics: Pass → Analysis reliability evaluation

Cohort diagnostics: Fail → STOP

Analysis reliability evaluation

Analysis design choices → Study diagnostics

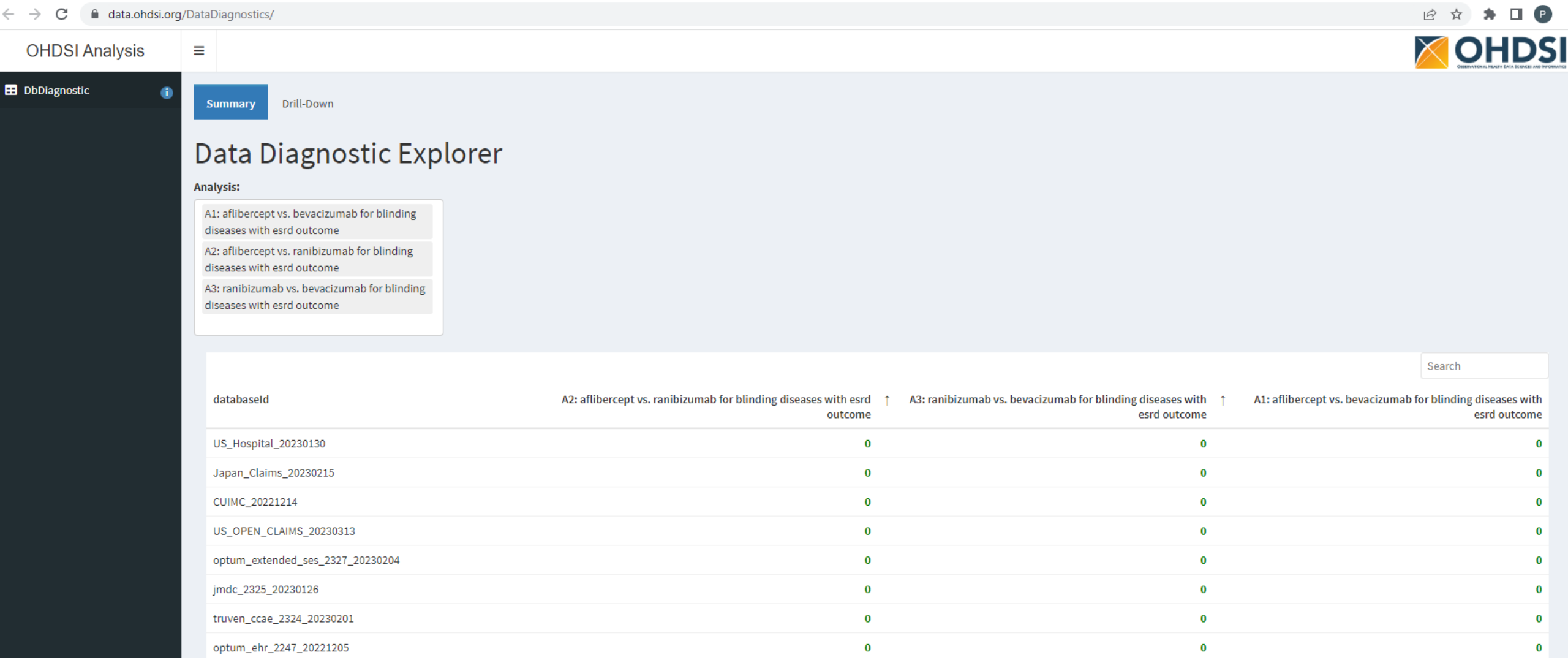
Study diagnostics: Pass → Final unblinded results

Study diagnostics: Fail → STOP

Final unblinded results → Interface for exploration

Week 2: Data diagnostics

data.ohdsi.org/DataDiagnostics



Data diagnostics:
T: antiVEGF; I: blinding disease; O: end-stage renal disease

Search

databaseId	A2: aflibercept vs. ranibizumab for blinding diseases with esrd outcome	A3: ranibizumab vs. bevacizumab for blinding diseases with esrd outcome	A1: aflibercept vs. bevacizumab for blinding diseases with esrd outcome
US_Hospital_20230130	0	0	0
Japan_Claims_20230215	0	0	0
CUIMC_20221214	0	0	0
US_OPEN_CLAIMS_20230313	0	0	0
optum_extended_ses_2327_20230204	0	0	0
jmdc_2325_20230126	0	0	0
truven_ccae_2324_20230201	0	0	0
optum_ehr_2247_20221205	0	0	0
optum_extended_dod_2323_20230201	0	0	0
truven_mdcd_2359_20230215	0	0	0
truven_mdcr_2322_20230127	0	0	0
US_PharMetrics_Plus_20230330	0	0	0
Japan_HIS_20220120	0	0	0
JHM_OMOP_20230406	1	1	0
TMUCRD_20210406	1	0	1
Klinicki_centar_Crne_Gore_20230101	1	1	
LPD_Italy_20221226	1	1	
UK_IMRD_EMIS_20230215	1	1	
UK_IMRD_THIN_20221230	1	1	
AUSOM_20220228	1	1	

1–20 of 30 rows

15 databases so far can perform are potentially feasible to conduct at least one of the antiVEGF comparisons:

- US, Japan, Taiwan
- Public + private claims, inpatient + outpatient EHR

Data diagnostics:
T: fluoroquinolone; I: UTI; O: aortic aneurysm

Search

databaseId	B6: fluoroquinolone vs. penicillin for pneumonia and risk of aortic aneurysm	B5: fluoroquinolone vs. macrolide for pneumonia and risk of aortic aneurysm	B3: fluoroquinolone vs. penicillin for urinary tract infection and risk of aortic aneurysm	B2: fluoroquinolone vs. macrolide for urinary tract infection and risk of aortic aneurysm	B1: fluoroquinolone vs. cephalosporin for urinary tract infection and risk of aortic aneurysm	B4: fluoroquinolone vs. cephalosporin for pneumonia and risk of aortic aneurysm
IQVIA_France_DA_20230201	0	0	0	0	0	0
optum_ehr_2247_20221205	0	0	0	0	0	0
UK_IMRD_EMIS_20230215	0	0	0	0	0	0
truven_mdcr_2322_20230127	0	0	0	0	0	0
Japan_HIS_20220120	0	0	0	0	0	0
IQVIA_Belgium_LPD_20221006	0	0	0	0	0	0
US_PharMetrics_Plus_20230330	0	0	0	0	0	0
LPD_Spain_20220704	0	0	0	0	0	0
Japan_Claims_20230215	0	0	0	0	0	0
France_LPD_20230118	0	0	0	0	0	0
LPD_Italy_20221226	0	0	0	0	0	0
US_OPEN_CLAIMS_20230313	0	0	0	0	0	0
optum_extended_ses_2327_20230204	0	0	0	0	0	0
IQVIA_Germany_DA_20230124	0	0	0	0		
UK_IMRD_THIN_20221230	0	0	0	0		
jmdc_2325_20230126	0	0	0	0		
truven_ccae_2324_20230201	0	0	0	0		
US_Hospital_20230130	0	0	0	0		
truven_mdcd_2359_20230215	0	0	0	0		
Australia_EMR_20230317	0	0	0	0		

20 databases so far can perform are potentially feasible to conduct at least one of the FQ analyses:

- US, UK, Belgium, Spain, France, Italy, Germany, Japan, Australia
- Public + private claims, inpatient + outpatient EHR

Data diagnostics:
T: biologics; I: multiple sclerosis; O: PML

databaseId	C2: biologics vs disease modifying treatments for multiple sclerosis and risk of PML	C1: natalizumab vs disease modifying treatments for multiple sclerosis and risk of PML
IQVIA_Germany_DA_20230124	0	0
US_OPEN_CLAIMS_20230313	0	0
truven_ccae_2324_20230201	0	0
optum_ehr_2247_20221205	0	0
optum_extended_dod_2323_20230201	0	0
truven_mdcd_2359_20230215	0	0
optum_extended_ses_2327_20230204	0	0
US_PharMetrics_Plus_20230330	0	0
jmdc_2325_20230126	0	1
CUIMC_20221214	0	1
truven_mdcr_2322_20230127	0	1
US_Hospital_20230130	1	1
LPD_Italy_20221226	1	1
Japan_Claims_20230215	1	1
UK_IMRD_EMIS_20230215	1	
JHM_OMOP_20230406	1	
RED_CDM_Tufts_20221005	1	
UK_IMRD_THIN_20221230	1	
Japan_HIS_20220120	1	
AUSOM_20220228	1	

11 databases so far can perform are potentially feasible to conduct at least one of the MS analyses:

- US, Germany, Japan
- Public + private claims, inpatient + outpatient EHR

Data diagnostics:
T: risankizumab; I: psoriasis; O: ischemic stroke

databaseId	D2: risankizumab vs. tildrakizumab for psoriasis and risk of ischemic stroke	D3: risankizumab vs. guselkumab for psoriasis and risk of ischemic stroke	D1: risankizumab vs. other biologics for psoriasis and risk of ischemic stroke
truven_ccae_2324_20230201	0	0	0
US_PharMetrics_Plus_20230330	0	0	0
US_OPEN_CLAIMS_20230313	0	0	0
optum_extended_ses_2327_20230204	0	0	0
optum_extended_dod_2323_20230201	0	0	0
optum_ehr_2247_20221205	1	0	0
truven_mdcr_2322_20230127	1	1	0
CUIMC_20221214	1	1	1
truven_mdcd_2359_20230215	1	1	1
LPD_Italy_20221226	1	1	1
JHM_OMOP_20230406	1	1	1
IQVIA_Germany_DA_20230124	1	1	1
LPD_Spain_20220704	1	1	1
Japan_Claims_20230215	1	1	1
Japan_HIS_20220120	2		
IQVIA_Belgium_LPD_20221006	2		
RED_CDM_Tufts_20221005	2		
jmdc_2325_20230126	2		
UK_IMRD_EMIS_20230215	2		
US_Hospital_20230130	2		

6 databases so far can perform are potentially feasible to conduct at least one of the PsO analyses:

- US only
- Public + private claims, inpatient + outpatient EHR

Engineering open science systems that build trust into the real-world evidence generation and dissemination process

'System' required elements:
- Required phenotypes
- Analysis specifications
- Decision thresholds

Distributed data network, standardized to common data model

Network coordination

Data quality evaluation

Research question → Database diagnostics

Database diagnostics: Pass → Phenotype development and evaluation

Database diagnostics: Fail → STOP

Phenotype development and evaluation

Cohort definitions → Cohort diagnostics

Cohort diagnostics: Fail → STOP

Cohort diagnostics: Pass → Analysis reliability evaluation

Analysis reliability evaluation

Analysis design choices → Study diagnostics

Study diagnostics: Fail → STOP

Study diagnostics: Pass → Final unblinded results

Final unblinded results → Interface for exploration

Week 3-4: Phenotype development and evaluation

Engineering open science systems that build trust into the real-world evidence generation and dissemination process

'System' required elements:
- Required phenotypes
- Analysis specifications
- Decision thresholds

Distributed data network, standardized to common data model

Network coordination

Data quality evaluation

Research question → Database diagnostics

Database diagnostics: Pass → Phenotype development and evaluation; Fail → STOP

Phenotype development and evaluation

Cohort definitions → Cohort diagnostics

Cohort diagnostics: Pass → Analysis reliability evaluation; Fail → STOP

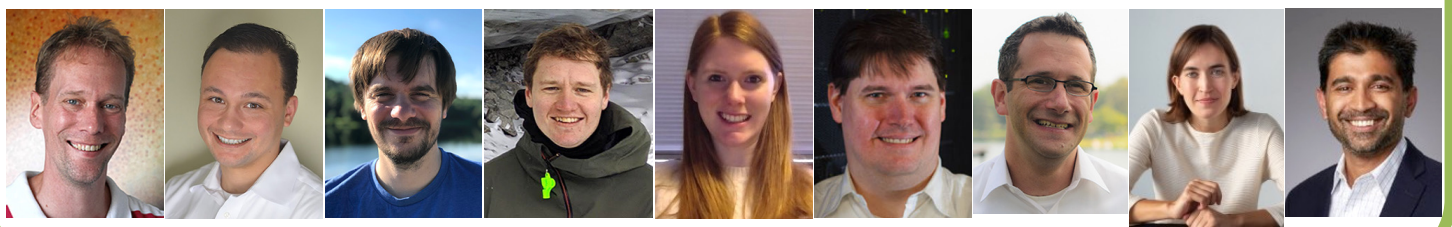
Analysis reliability evaluation

Analysis design choices → Study diagnostics

Study diagnostics: Pass → Final unblinded results; Fail → STOP

Final unblinded results → Interface for exploration

Only possible because of standardized analytics developed across our community



Week 5: Analysis design

Engineering open science systems that build trust into the real-world evidence generation and dissemination process

'System' required elements:
- Required phenotypes
- Analysis specifications
- Decision thresholds

Distributed data network, standardized to common data model

Network coordination

Data quality evaluation

Research question → Database diagnostics

Database diagnostics → Pass

Database diagnostics → Fail → STOP

Phenotype development and evaluation

Cohort definitions → Cohort diagnostics

Cohort diagnostics → Pass

Cohort diagnostics → Fail → STOP

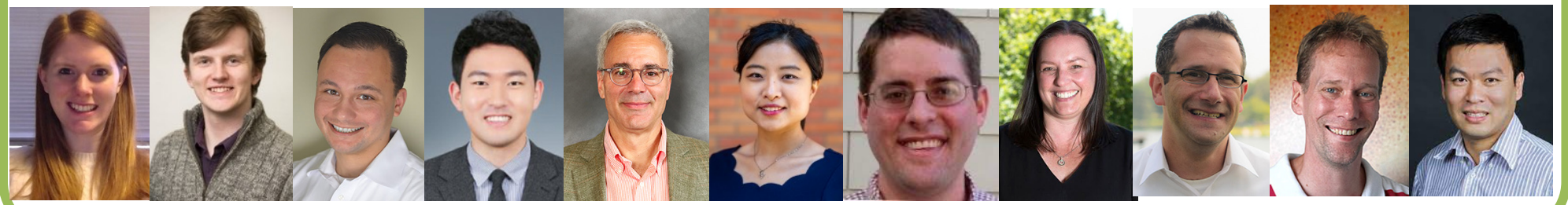
Analysis reliability evaluation

Analysis design choices → Study diagnostics

Study diagnostics → Pass

Study diagnostics → Fail → STOP

Week 6-8: Analysis execution, study diagnostics, evidence synthesis



Pass → Final unblinded results → Interface for exploration

Engineering open science systems that build trust into the real-world evidence generation and dissemination process

'System' required elements:
- Required phenotypes
- Analysis specifications
- Decision thresholds

Distributed data network, standardized to common data model

Network coordination

Data quality evaluation

Research question → Database diagnostics

Pass

Fail → STOP

Phenotype development and evaluation

Cohort definitions → Cohort diagnostics

Pass

Fail → STOP

Analysis reliability evaluation

Analysis design choices → Study diagnostics

Pass

Fail → STOP

Final unblinded results → Interface for exploration

Week 9: Results interpretation

VEGF→ESRD

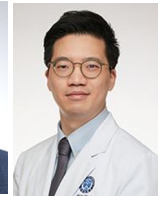
FQ→AA

MS biologics→PML

risankizumab→CVA

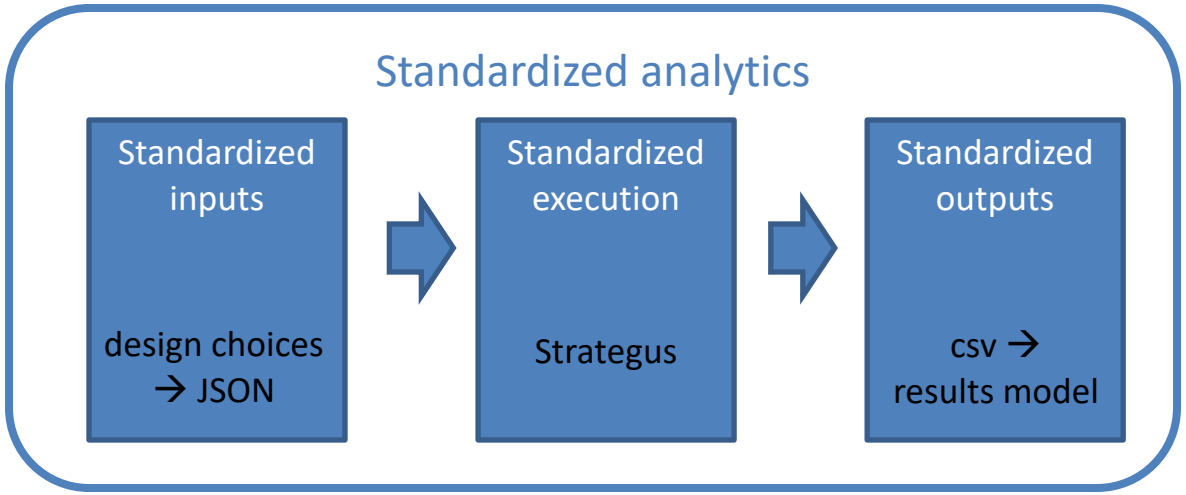






Standardized analyses currently available within Strategus pipeline

- Characterization
  - Cohort diagnostics
  - Cohort features
  - Incidence rates
  - Time-to-event
  - Dechallenge / rechallenge
- Patient-level prediction
- Population-level effect estimation
  - Comparative cohort
  - Self-controlled case-series (SCCS)



Design choices that always need to be made as input into standardized analytics

- **Target***: What exposure do we have a question about?
- **Indication(s)***: Which disease(s) is the exposure intended to treat?
- **Outcome(s)***: What event(s) would qualify as outcomes of interest?
- **Comparator(s)***: What other population(s) can be used as a proxy for counterfactual (e.g. in comparative cohort analyses)?
- **Time(s)-at-risk**: What is the span(s) of time relative to exposure start/end when the effect on the outcome is hypothesized to occur?
- **Age/sex/calendar time restrictions**
- **Negative controls:** What concepts will be used to create proxy outcomes to estimate residual systematic error and enable empirical calibration?
- **Excluded concepts:** What concepts should be excluded from propensity score modeling?

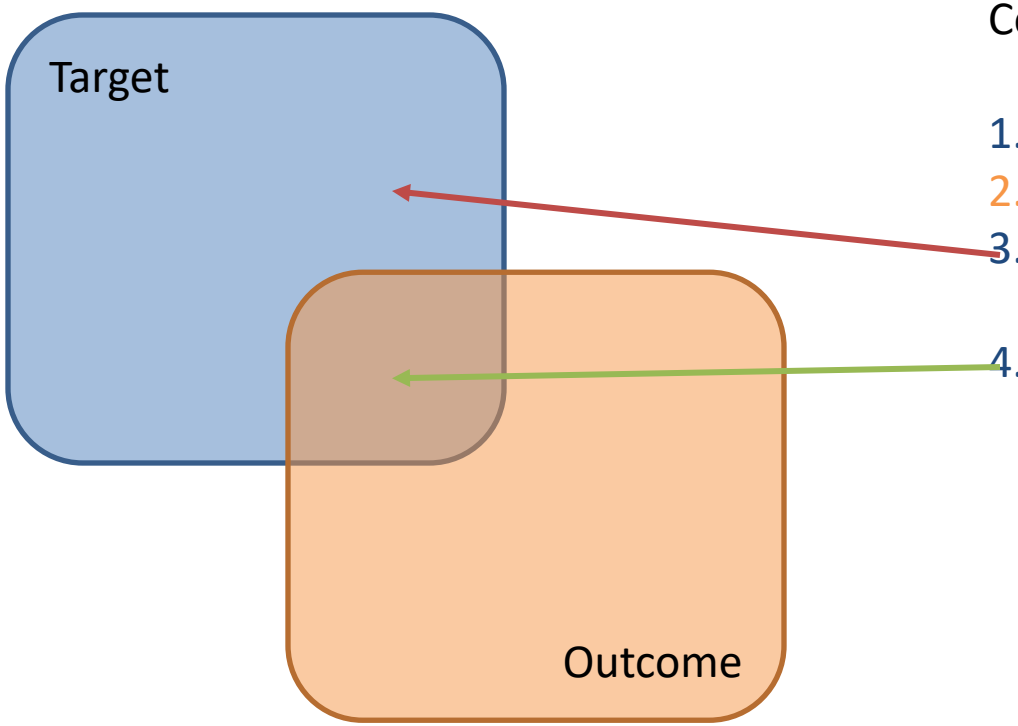
* Expressed as a **cohort**

Design choices for antiVEGF study

- **Target***:
  - T1: aflibercept exposures after new use with 3 exposures in 21-70d windows
  - T2: ranibizumab exposures after new use with 3 exposures in 21-70d windows
  - T3: bevacizumab exposures after new use with 3 exposures in 21-70d windows
- **Comparator(s)***:
  - T1 vs T2; T2 vs. T3; T1 vs. T3
- **Indication(s)***: Blinding diseases
- **Outcome(s)***: End stage renal disease
- **Time(s)-at-risk**: 'on treatment': cohort start + 1d → cohort end + 0d
- **Age/sex/calendar time restrictions:** age>=18
- **Negative controls:** candidates to review from CEM
- **Excluded concepts:** candidates to review based on comparator selector recommender

* Expressed as a **cohort**

Stratifying cohorts for characterization



Cohorts of interest:

1. Target
2. Outcome
3. Target **without** Outcome during Time-at-risk
4. Target **with** Outcome during Time-at-risk
   a. Indexed on Target
   b. Indexed on Outcome

Cohorts of interest for VEGF:

1. Aflibercept
2. End-Stage Renal Disease (ESRD)
3. Aflibercept **without** ESRD during 'on treatment' time-at-risk (start + 1d → end + 0d)
4. Aflibercept **with** ESRD during 'on treatment' time-at-risk
   a. Indexed on Aflibercept
   b. Indexed on ESRD

Characterization: CohortDiagnostics

Executed for all **target**, **comparator**, **indication** and **outcome** cohorts to evaluate measurement error in the phenotype development and evaluation process

- By default using
  - Orphan concepts - to identify potential additional concepts to include in definition
  - Visit context – to understand where care is received before/during/after cohort entry
  - Index event breakdown – to see which concepts qualify persons at cohort entry
  - Incidence rate – to characterize population-level trends in cohort by age/sex/year
  - Cohort relationship – to evaluate intersection between cohorts
  - Temporal characterization – to assess prevalence of other events before and after cohort entry

Target:

T1: aflibercept exposures after new use with 3 exposures in 21-70d windows

T2: ranibizumab exposures after new use with 3 exposures in 21-70d windows

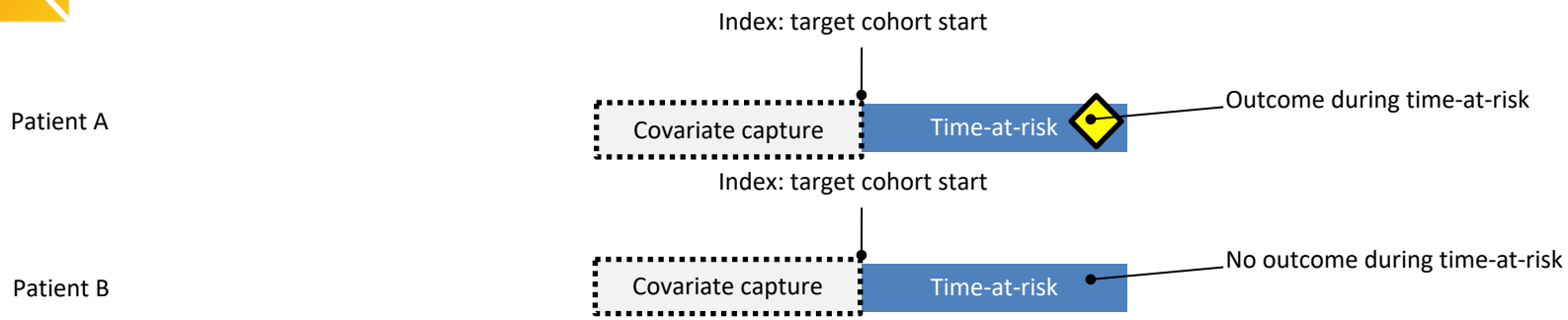
T3: bevacizumab exposures after new use with 3 exposures in 21-70d windows

Indication: Blinding diseases

Outcome: End stage renal disease

CohortDiagnostics

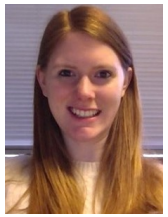
Characterization: Features of patients with and without outcome



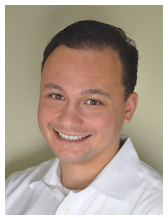
Describe patients with and without the outcome during time-at-risk

Done for the **target**, **comparator**, and **indication** cohorts, and all **outcomes** of interest

- **Target** and **comparator** are restricted:
  - To the **indication**
  - First exposure (new user)
  - Having >= 365 days of observation prior
  - Not having outcome in the **prior lookback window**
  - Applying any restriction to **age**, **sex**, or **calendar time**
- By default using
  - 365 days prior to index to capture medical history
  - FeatureExtraction's default set of features:
    - Demographics: Sex, Age group, Race, Ethnicity, Index year, Index month
    - Prior Condition group / Drug group / Procedure / Device / Measurement / Observation short term (30d) and long term (365d)
    - Risk scores: Charlson, DCSI, CHADS2VASC

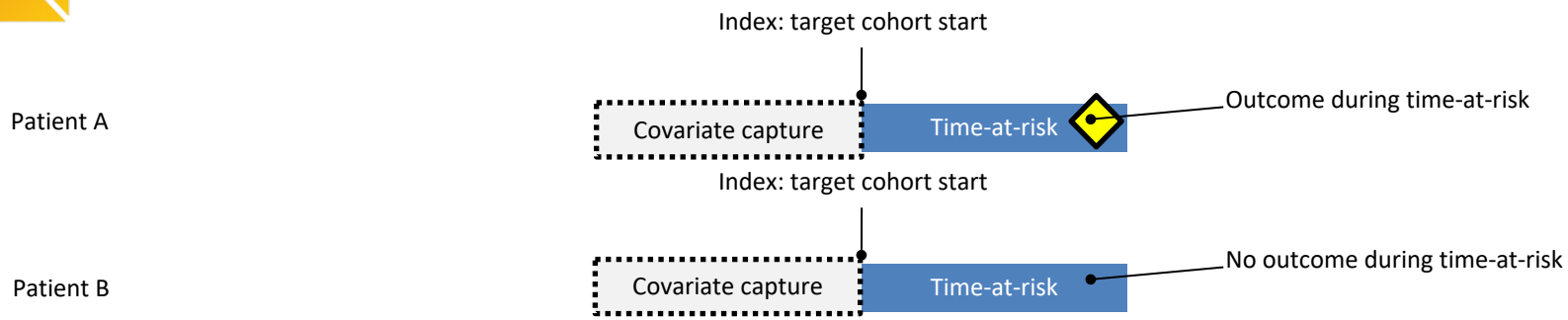


Characterization



FeatureExtraction

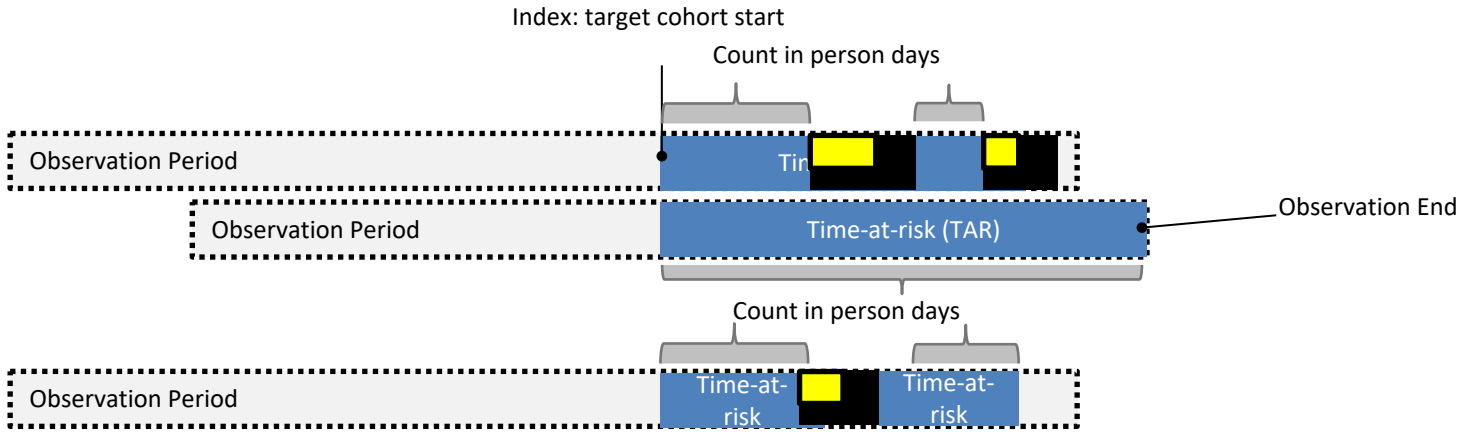
Characterization: Features of patients with and without outcome

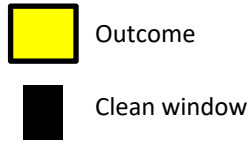


Describe patients with and without the outcome during time-at-risk

- **Target:**
  - T1: aflibercept exposures after new use with 3 exposures in 21-70d windows , with prior blinding disease, age>=18 and >365d prior observation
  - T2: ranibizumab exposures after new use with 3 exposures in 21-70d windows , with prior blinding disease , age>=18 and >365d prior observation
  - T3: bevacizumab exposures after new use with 3 exposures in 21-70d windows , with prior blinding disease , age>=18 and >365d prior observation
  - T4: Blinding disease, with , age>=18 and >365d prior observation
- **Outcome:**
  - End-stage renal disease (clean window = 9999d)
- **Time-at-risk:**
  - 'on treatment': cohort start + 1d → cohort end + 0d
- **Analysis settings:**
  - 365 days prior to index to capture medical history
  - FeatureExtraction's default set of features:
    - Demographics: Sex, Age group, Race, Ethnicity, Index year, Index month
    - Prior Condition group / Drug group / Procedure / Device / Measurement / Observation short term (30d) and long term (365d)
    - Risk scores: Charlson, DCSI, CHADS2VASC

Characterization: Incidence rates





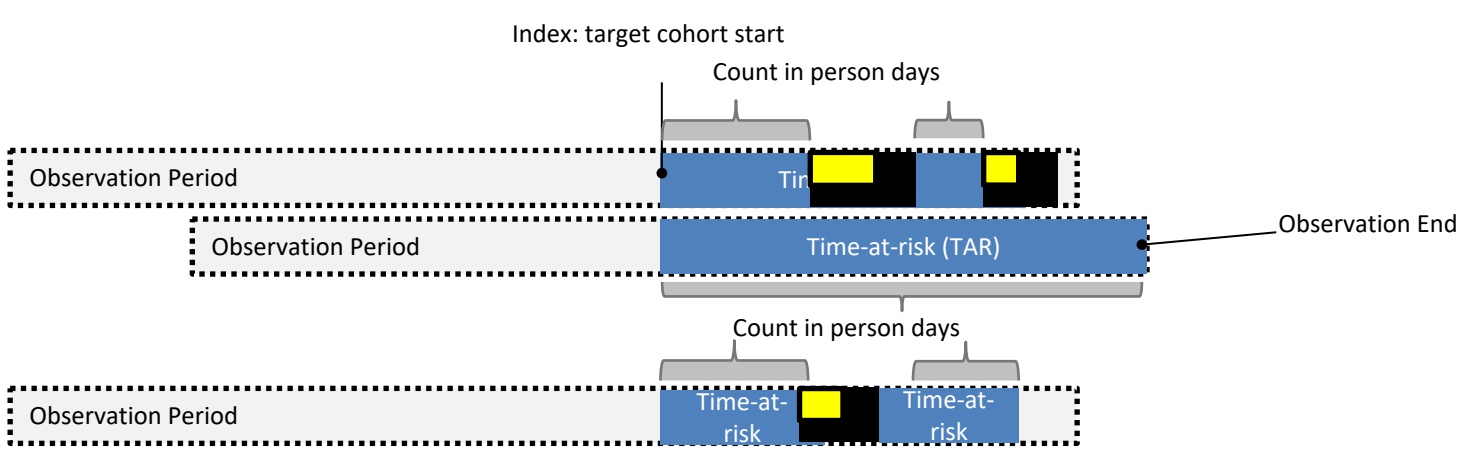
Proportion: (# people with outcome during TAR)/(# people)
Rate: (#outcomes during TAR)/(total person days)

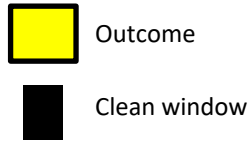
Done for the **target**, **comparator**, and **indication** cohorts, and all **outcomes** of interest

- **Target** and **comparator** are restricted:
  - To the **indication**
  - Having >= 365 days of observation prior
  - Not having outcome in the **prior lookback window**
  - Applying any restriction to **age**, **sex**, or **calendar time**
- Using **clean windows** to account for immortal time after outcome
- By default using
  - Gender/Age/Start year subgroups

CohortIncidence

Characterization: Incidence rates for VEGF

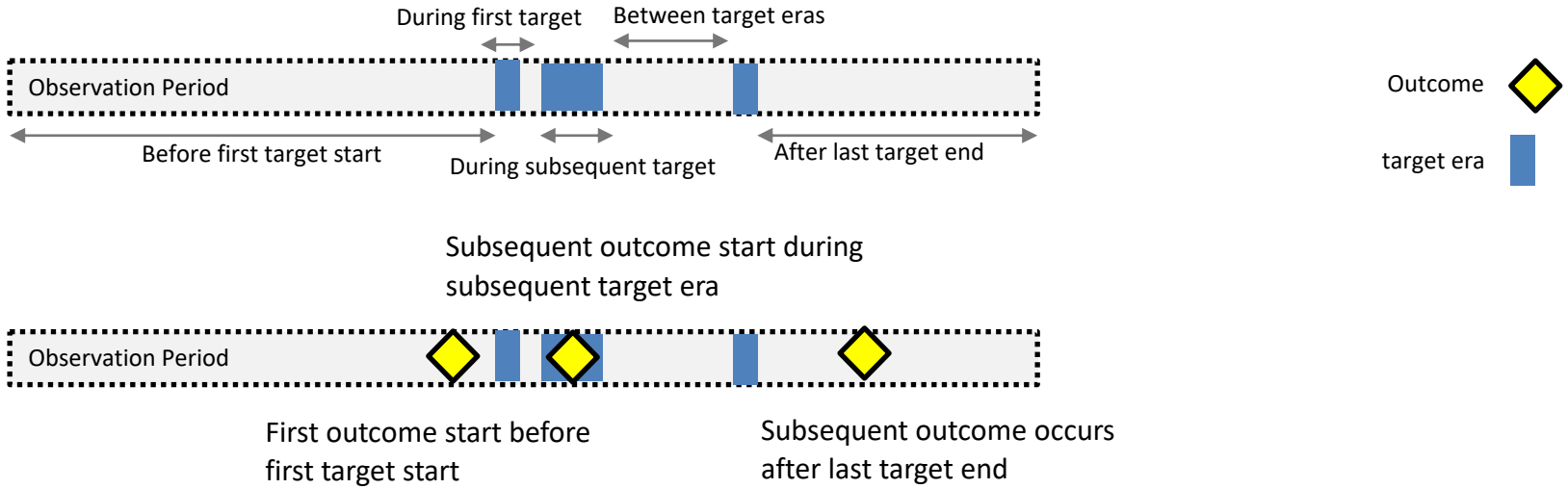




Proportion: (# people with outcome during TAR)/(# people)
Rate: (#outcomes during TAR)/(total person days)

- **Target:**
  - T1: aflibercept exposures after new use with 3 exposures in 21-70d windows , with prior blinding disease, age>=18 and >365d prior observation
  - T2: ranibizumab exposures after new use with 3 exposures in 21-70d windows , with prior blinding disease , age>=18 and >365d prior observation
  - T3: bevacizumab exposures after new use with 3 exposures in 21-70d windows , with prior blinding disease , age>=18 and >365d prior observation
  - T4: Blinding disease, with , age>=18 and >365d prior observation
- **Outcome:**
  - End-stage renal disease (clean window = 9999d)
- **Time-at-risk:**
  - 'on treatment': cohort start + 1d → cohort end + 0d
- **Strata:**
  - Gender, Age deciles, index year subgroups

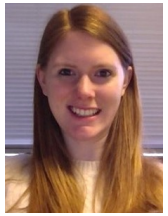
Characterization: Time-to-event



Done for the **target**, **comparator**, and **indication** cohorts, and all **outcomes** of interest

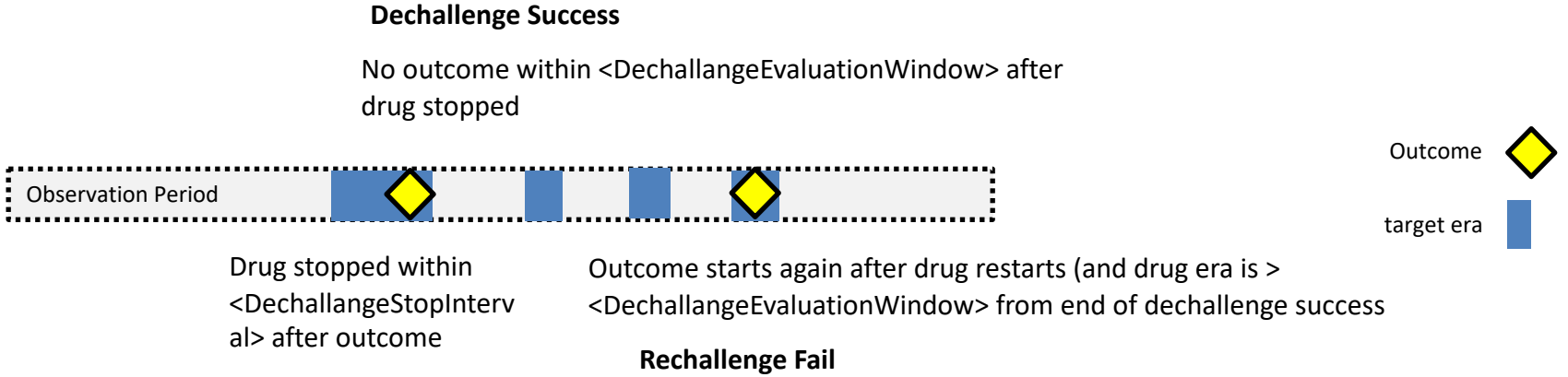
- No additional settings

- **Target:**
  - T1: aflibercept exposures after new use with 3 exposures in 21-70d windows , with prior blinding disease, age>=18 and >365d prior observation
  - T2: ranibizumab exposures after new use with 3 exposures in 21-70d windows , with prior blinding disease , age>=18 and >365d prior observation
  - T3: bevacizumab exposures after new use with 3 exposures in 21-70d windows , with prior blinding disease , age>=18 and >365d prior observation
  - T4: Blinding disease, with , age>=18 and >365d prior observation
- **Outcome:**
  - End-stage renal disease (clean window = 9999d)



Characterization

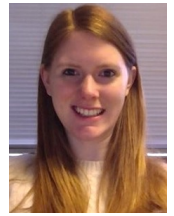
Characterization: dechallenge / rechallenge



Done for the **target** and **comparator** cohorts, and all **outcomes** of interest

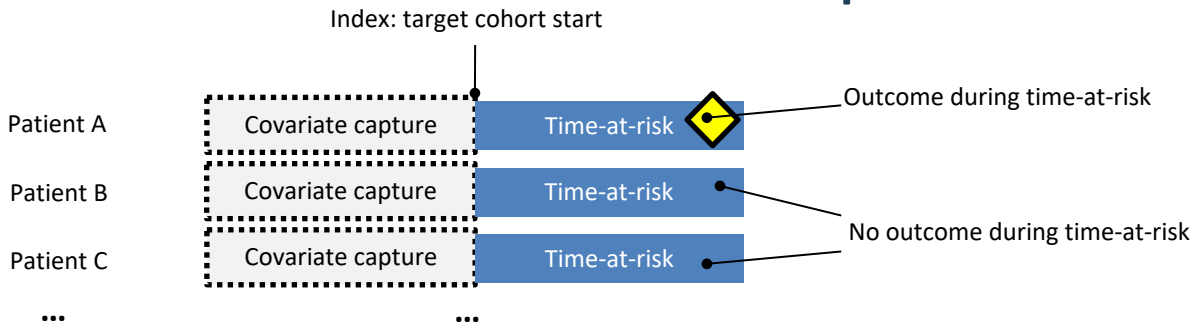
- By default using
  - DechallangeStopInterval 30 days
  - DechallangeEvaluationWindow 30 days

- **Target:**
  - T1: aflibercept exposures after new use with 3 exposures in 21-70d windows , with prior blinding disease, age>=18 and >365d prior observation
  - T2: ranibizumab exposures after new use with 3 exposures in 21-70d windows , with prior blinding disease , age>=18 and >365d prior observation
  - T3: bevacizumab exposures after new use with 3 exposures in 21-70d windows , with prior blinding disease , age>=18 and >365d prior observation
  - T4: Blinding disease, with , age>=18 and >365d prior observation
- **Outcome:**
  - End-stage renal disease (clean window = 9999d) *****RECHALLENGE not possible when event can only occur once



Characterization

Patient-level prediction

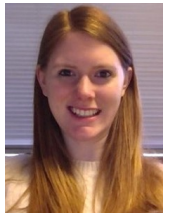


A model learns associations between covariates and the occurrence of the outcome during time-at-risk

Done for the **target** cohort, and all **outcomes** of interest

- **Target** and **comparator** are restricted:
  - To the **indication**
  - First exposure (new user)
  - Having >= 365 days of observation prior
  - Not having outcome in the **prior lookback window**
  - Applying any restriction to **age**, **sex**, or **calendar time**
- By default using
  - Features in 365 days prior, excluding index year covariates
  - Two prediction time-at-risks: 1-30 days, and 1-365 days after index
  - Model is logistic regression with LASSO regularization
  - Model developed using 75% of data and internally validated in remaining 25%
  - Model hyper-parameter selection using 3-fold cross validation
  - Do not exclude patients lost to follow-up during time-at-risk

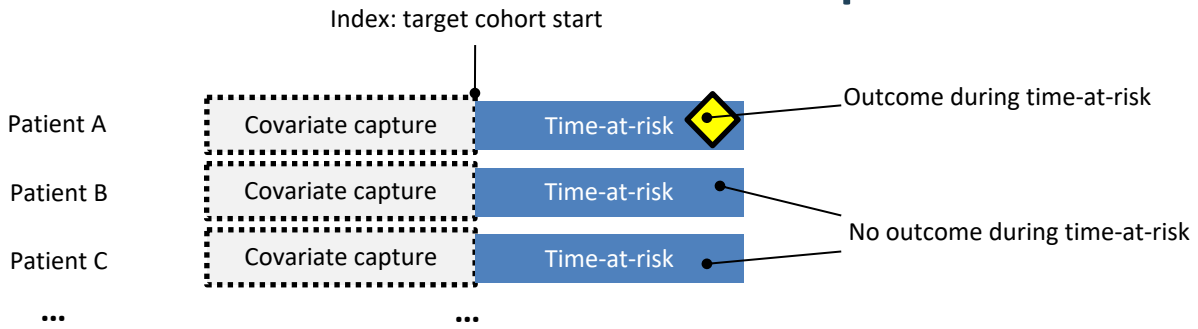
Prediction requires a sufficient number of patients with the outcome during TAR. Model development likely infeasible if <100 outcomes.



PatientLevelPrediction

Cyclops

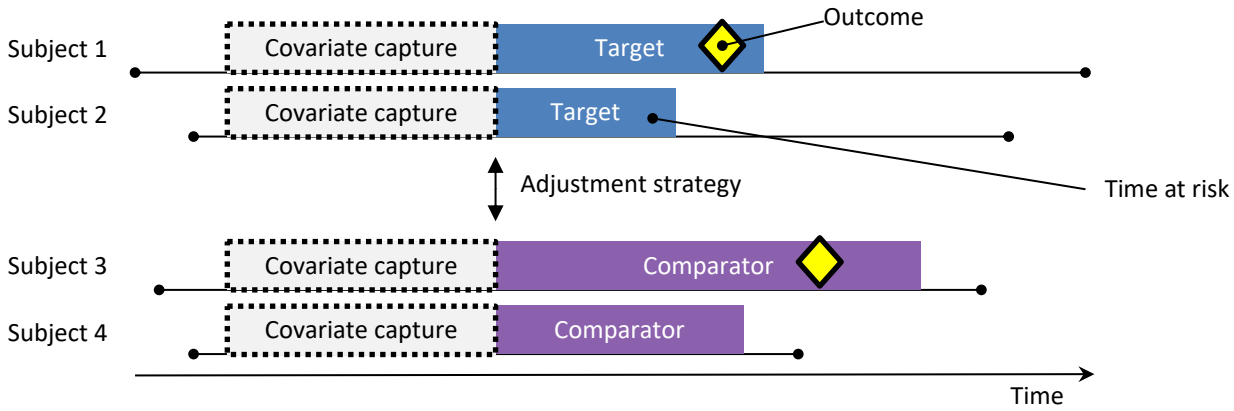
Patient-level prediction



A model learns associations between covariates and the occurrence of the outcome during time-at-risk

- **Target:**
  - T1: aflibercept exposures after new use with 3 exposures in 21-70d windows , with prior blinding disease, age>=18 and >365d prior observation
  - T2: ranibizumab exposures after new use with 3 exposures in 21-70d windows , with prior blinding disease , age>=18 and >365d prior observation
  - T3: bevacizumab exposures after new use with 3 exposures in 21-70d windows , with prior blinding disease , age>=18 and >365d prior observation
- **Outcome:**
  - End-stage renal disease (clean window = 9999d)
- **Time-at-risk:**
  - '365d fixed window': cohort start + 1d → cohort start + 365d

Causal effect estimation: comparative cohort study

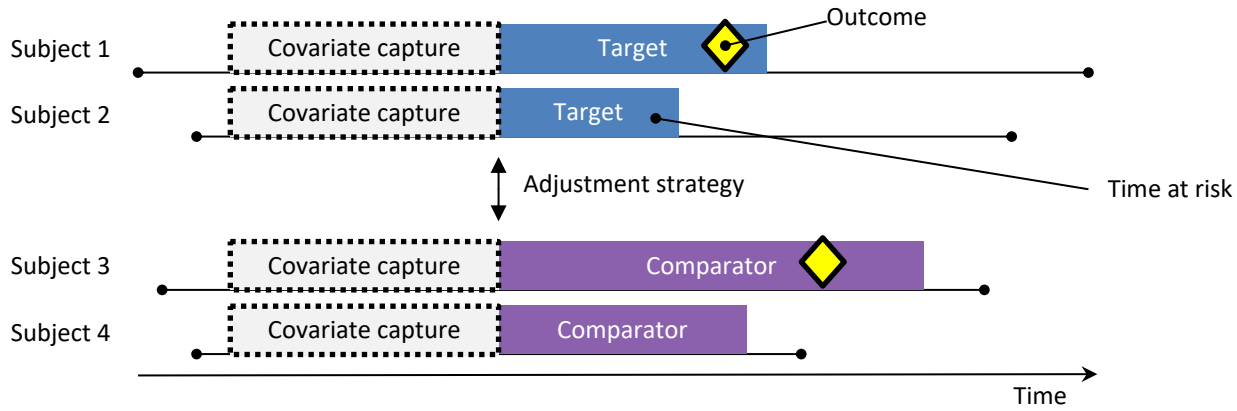


- **Target** and **comparator** are restricted:
  - To the **indication**
  - First exposure (new user)
  - Having >= 365 days of observation prior
  - Not having outcome in the **prior lookback window**
  - Applying any restriction to **age**, **sex**, or **calendar time**
- By default using
  - Large-scale propensity scores (PS)
  - 1:1 PS matching
  - Cox proportional hazards model
  - A large set of negative control outcomes

CohortMethod

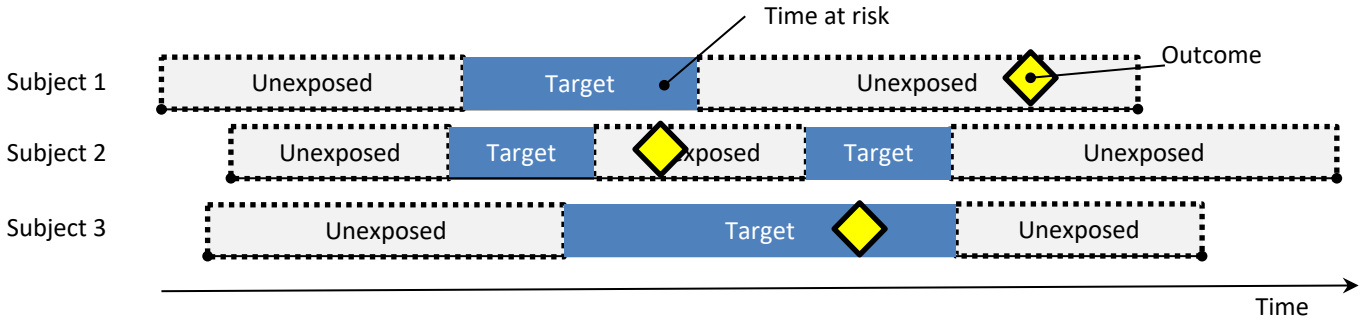
Cyclops

Causal effect estimation: comparative cohort study



- **Target / Comparators:**
  - T1: aflibercept exposures after new use with 3 exposures in 21-70d windows , with prior blinding disease, age>=18 and >365d prior observation
  - T2: ranibizumab exposures after new use with 3 exposures in 21-70d windows , with prior blinding disease , age>=18 and >365d prior observation
  - T3: bevacizumab exposures after new use with 3 exposures in 21-70d windows , with prior blinding disease , age>=18 and >365d prior observation
  - T1 vs. T2; T1 vs. T3; T2 vs. T3
- **Outcome:**
  - End-stage renal disease (clean window = 9999d)
- **Time-at-risk:**
  - 'on treatment': cohort start + 1d → cohort end + 0d
- **Analysis settings:**
  - Large-scale propensity scores (PS)
  - 1:1 PS matching
  - Cox proportional hazards model
  - Negative control outcomes, as recommended by CEM **** to be reviewed

Causal effect estimation: Self-controlled case-series



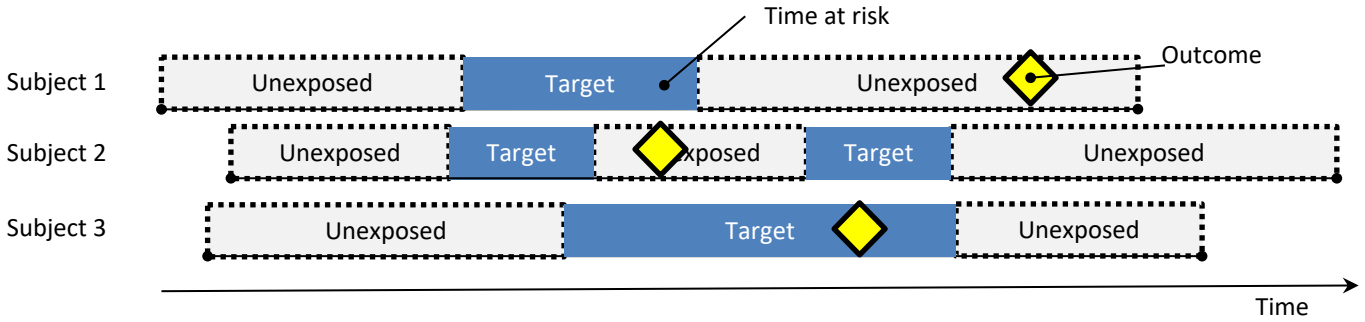
- Patient time is restricted to
  - Time when having the **indication**
  - Excluding first 365 days after observation period start (to ensure first observed outcome is first in patient's history)
  - Applying any restriction to **age**, **sex**, or **calendar time**
- By default using
  - Pre-exposure window of 30 days (account for (contra) indication)
  - Spline for calendar time
  - First outcome only (to avoid dependency between outcome occurrences)
  - A large set of negative control outcomes

SCCS can be appropriate for any exposure and outcome, as long as certain assumptions are met (which we check via our diagnostics)

SelfControlledCaseSeries

Cyclops

Causal effect estimation: Self-controlled case-series



- **Targets:**
  - T1: aflibercept exposures after new use with 3 exposures in 21-70d windows
  - T2: ranibizumab exposures after new use with 3 exposures in 21-70d windows
  - T3: bevacizumab exposures after new use with 3 exposures in 21-70d windows
- Indications:
  - Blinding disease
- Restrictions:
  - Age >= 18
- Analysis settings:
  - Excluding first 365 days after observation period start
  - Pre-exposure window of 30 days
  - Spline for calendar time
  - First outcome only
  - Negative control outcomes, as recommended by CEM **** to be reviewed

Demo Strategus specifications

Homework for VEGF team

- Review negative control conceptset
- Revise protocol to reflect the analyses to perform
- Draft Methods section in manuscript