Embrace Real-world Data and Enhance Your Analytics Approaches with New SAS' Open Platform

phuse·share
connect·share·advance

Around the world
throughout 2019

Single Day Events

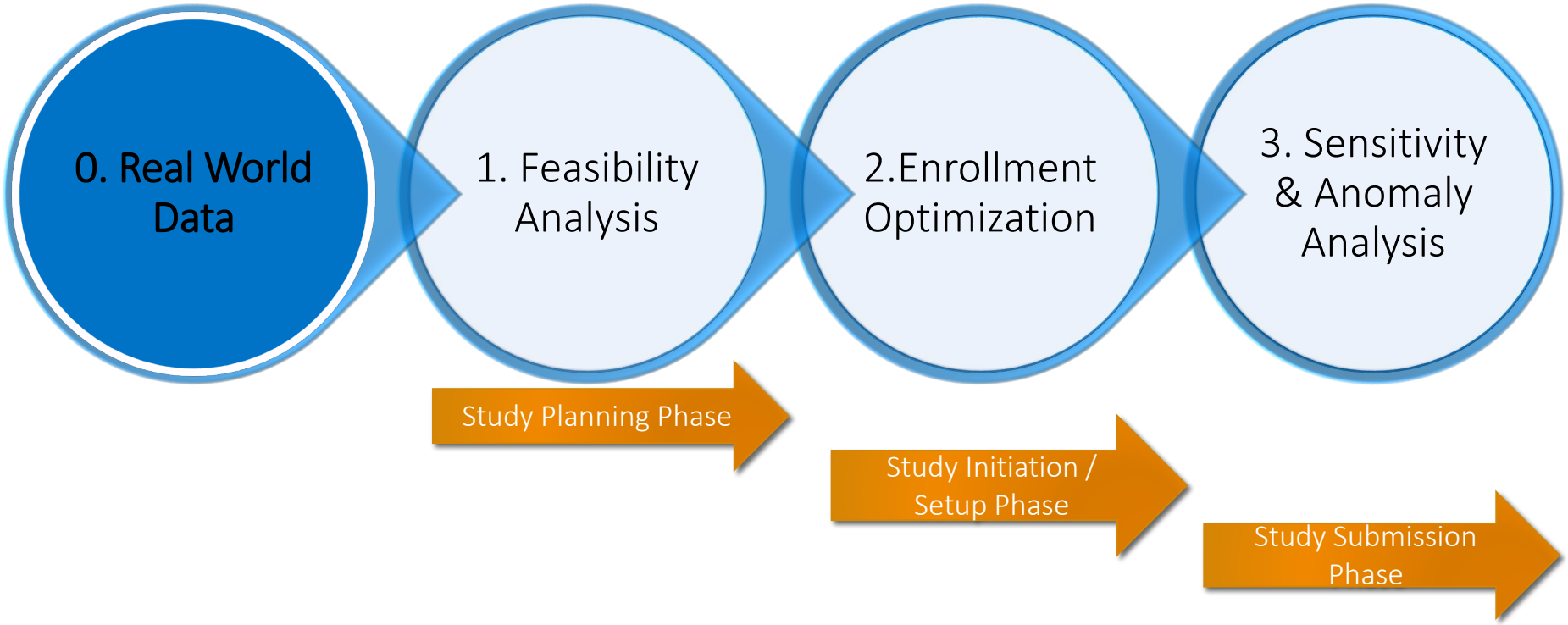
Utrecht, the Netherlands

10th October 2019

Stijn Rogiers – Principal Industry Consultant, Global Health and Life Sciences Practice
Mark Lambrecht – Director, Global Health and Life Sciences Practice

sas

Scope



RWE is a hot topic today …

Maximize product value: Beyond efficacy and safety

Value Message

Society
- Equal access to healthcare
- Return to work
- Caregiver burden
- Productivity

Payer/HTA
- Comparative effectiveness
- Cost-utility
- Budget Impact
- Long-term Outcomes

Patient
- Morbidity, Mortality
- Symptoms
- Functionality
- Preferences
- Quality of Life

Economic
- Direct/Indirect Cost
- Cost:Outcomes
- Cost Consequences
- Resource Utilization
- QALY

Provider
- Effectiveness
- Tolerability
- Safety
- Compliance
- Adherence
- Convenience

Clinical
- Efficacy
- Safety
- Morbidity
- Mortality

Case Study: Maximize Product Value with Real World Evidence, Maria Kubin, VP Bayer Healthcare –Eye for Pharma

Recent Industry Activities

US FDA

Framework for FDA's Real World Evidence Program (Dec 2018)

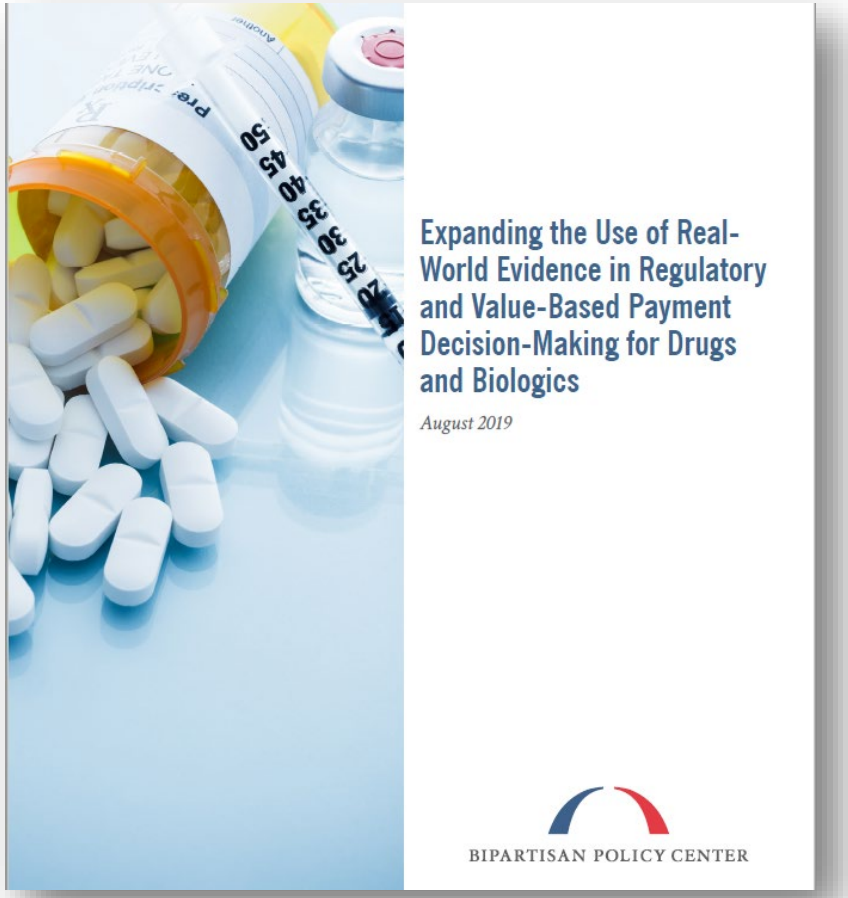
National Academies of Science (US)

1. Real-World Evidence Generation and Evaluation of Therapeutics (Workshop 2017)
2. Examining the Impact of Real-World Evidence on Medical Product Development I. Incentives (Feb 2018)
3. Examining the Impact of Real-World Evidence on Medical Product Development II. Practical Approaches (Jul 2018)

The Academy of Medical Sciences (UK)

1. Real-World Evidence Generation Scoping Roundtable (Jan 2018)
2. Next steps for using real world evidence (May 2018)

Recent Industry Activities



Conclusion

Real-world evidence generated from the application of appropriate research methods and analysis of data derived from clinical, claims, and patient-facing software and systems can play a significant role in improving and modernizing both the drug evaluation and approval process and new value-based payment arrangements in the United States. Progress has already been made by the FDA in establishing a framework for the use of real-world evidence for regulatory decision-making. Significant progress has also been made by ONC and CMS in advancing interoperability and use of common standards in systems, particularly through the development of recently proposed rules. Recent efforts by HHS to promote regulatory reforms can also play a key role in modernizing evidence development.

Regulatory and payment decision-making will be informed and improved by improving regulatory clarity; increasing support for real-world evidence activities; enabling access to and analysis of data; improving its reliability and relevance; expanding pilot and demonstration projects; clearing regulatory barriers; building capacity within federal agencies; and promoting collaboration among payers, regulators, and other key stakeholders around the generation and use of real-world evidence. These steps will ultimately help to accelerate the availability of safe, effective, and affordable medicines for patients in need.

Reference: https://bipartisanpolicy.org/report/expanding-the-use-of-realworld-evidence-in-regulatory-and-value-based-payment-decision-making-for-drugs-and-biologics/

HMA
Heads of Medicines Agencies

EUROPEAN MEDICINES AGENCY
SCIENCE MEDICINES HEALTH

Big Data – Challenges and Opportunities

Moving forward with recommendations from the HMA-EMA Joint Big Data taskforce

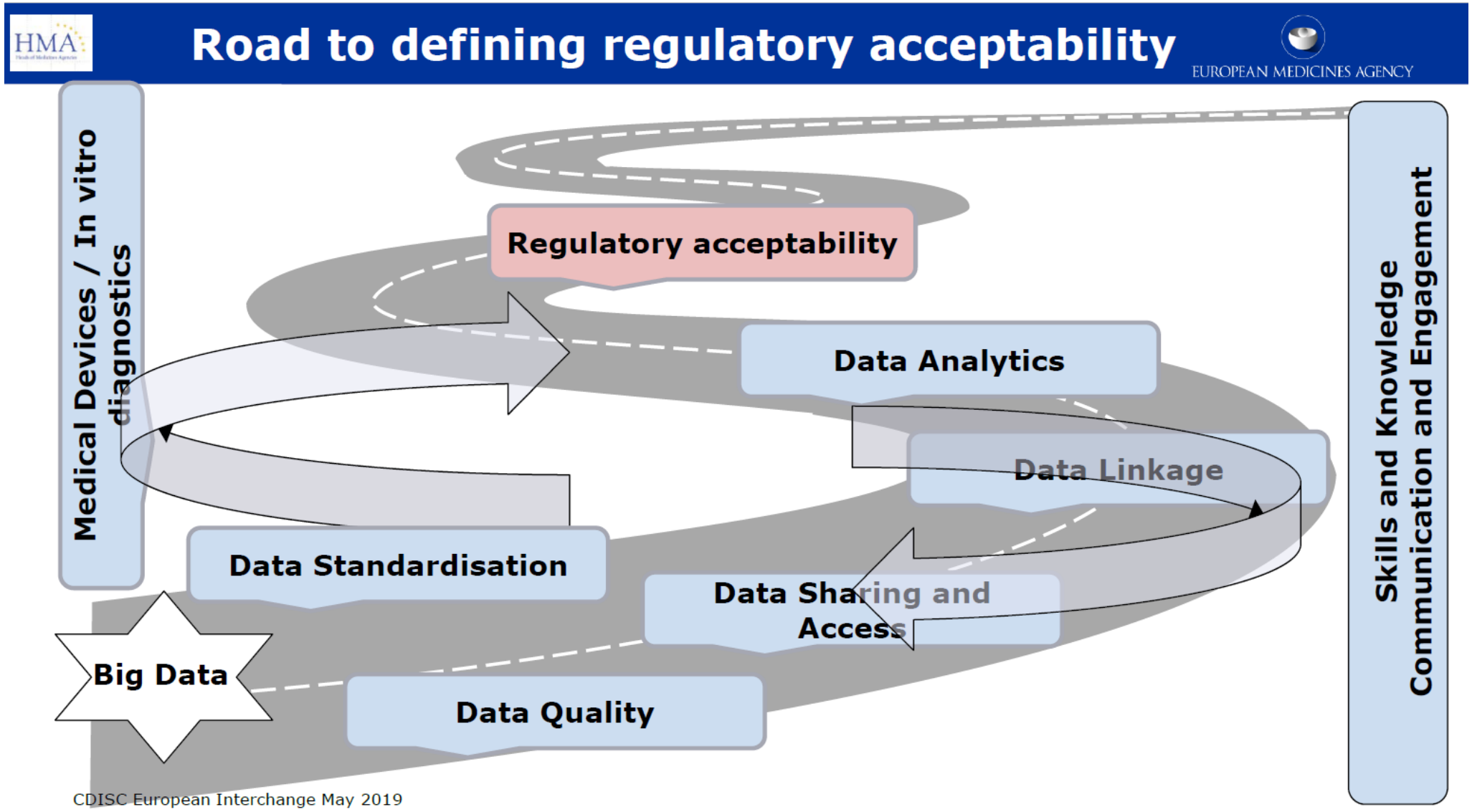
Dr Alison Cave
Principal Scientific Administrator
Pharmacovigilance and Epidemiology Department

CDISC European Exchange May 2019

An agency of the European Union

sas

HMA-EMA



Recent Examples (Jan, 2018)

J&J performed a pragmatic real-world trial, a randomized trial that used measures that are collected in clinical practice

"INVEGA SUSTENNA® (paliperidone palmitate), a once-monthly schizophrenia treatment, is the first and only antipsychotic to have the U.S. Food and Drug Administration (FDA) ***approve the inclusion of real-world data in its product labeling****"*

https://www.janssen.com/landmark-schizophrenia-data-bring-hope-breaking-cycle-hospitalization-and-incarceration-receive-fda

What's scary, and appealing, about real-world evidence (StatNews 2019)

Reference: "*Real World Evidence in Clinical development, life cycle management and/or repurposing*"
(Julijana Dukanovic, Andrew Leary – Dr. Regenold GmbH, SAS HLS Exec Forum Basel, 25 June 2019)

Recent Examples (Apr, 2019)

Pfizer Uses EHR Data to Support Expanded Indication for Breast Cancer Drug

*"**The approval is based on RWD** from electronic health records and post-marketing reports of Ibrance™ in male patients sourced from three databases: IQVIA Insurance database, Flatiron Health Breast Cancer database and the Pfizer global safety database"*

https://www.raps.org/news-and-articles/news-articles/2019/4/pfizer-uses-ehr-data-to-support-expanded-indicatio

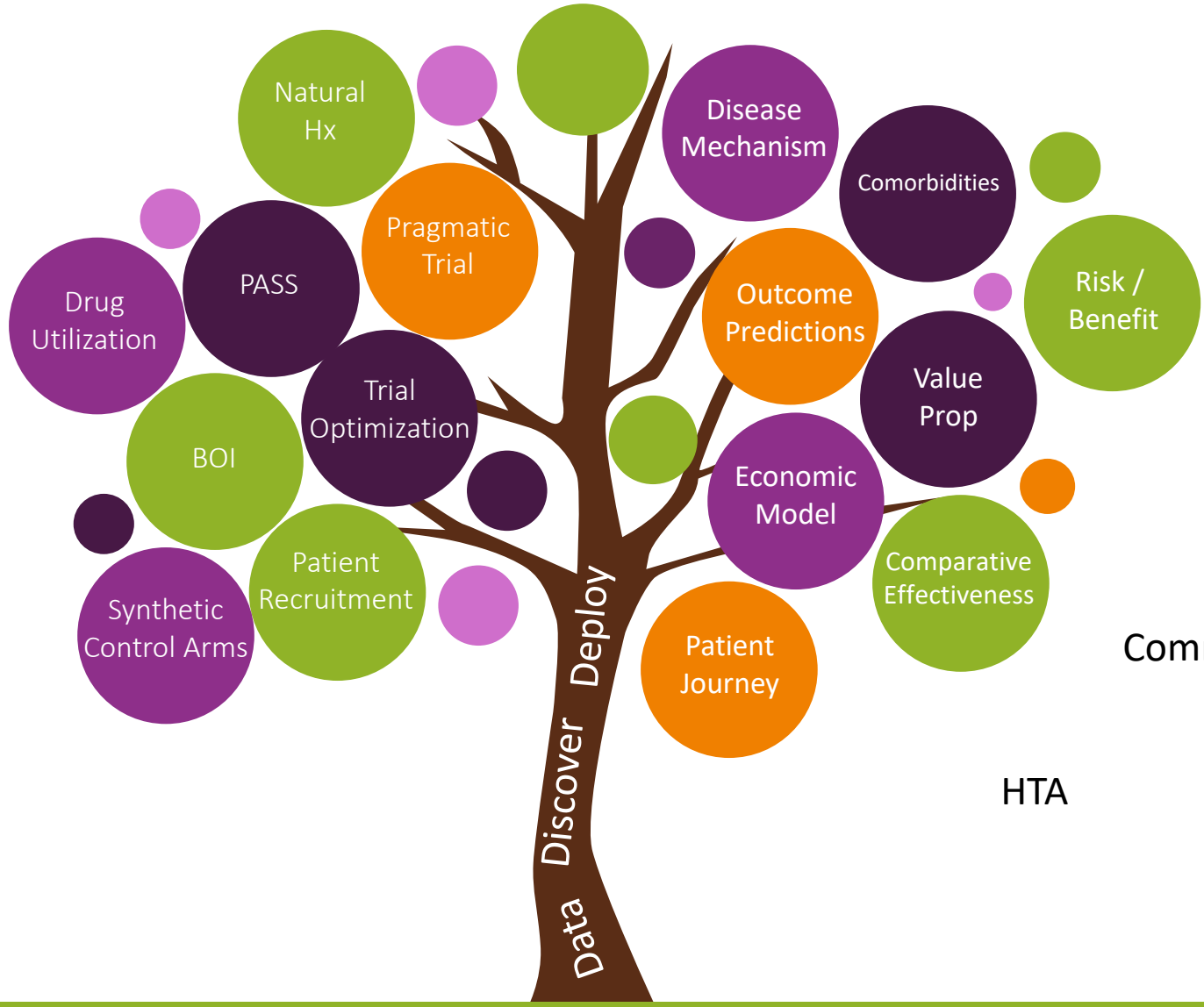
Reference: "*Real World Evidence in Clinical development, life cycle management and/or repurposing*"
(Julijana Dukanovic, Andrew Leary – Dr. Regenold GmbH, SAS HLS Exec Forum Basel, 25 June 2019)

Epidemiology

Pharmacovigilance

Regulatory Affairs

R&D



Medical Affairs

HEOR

Market Access

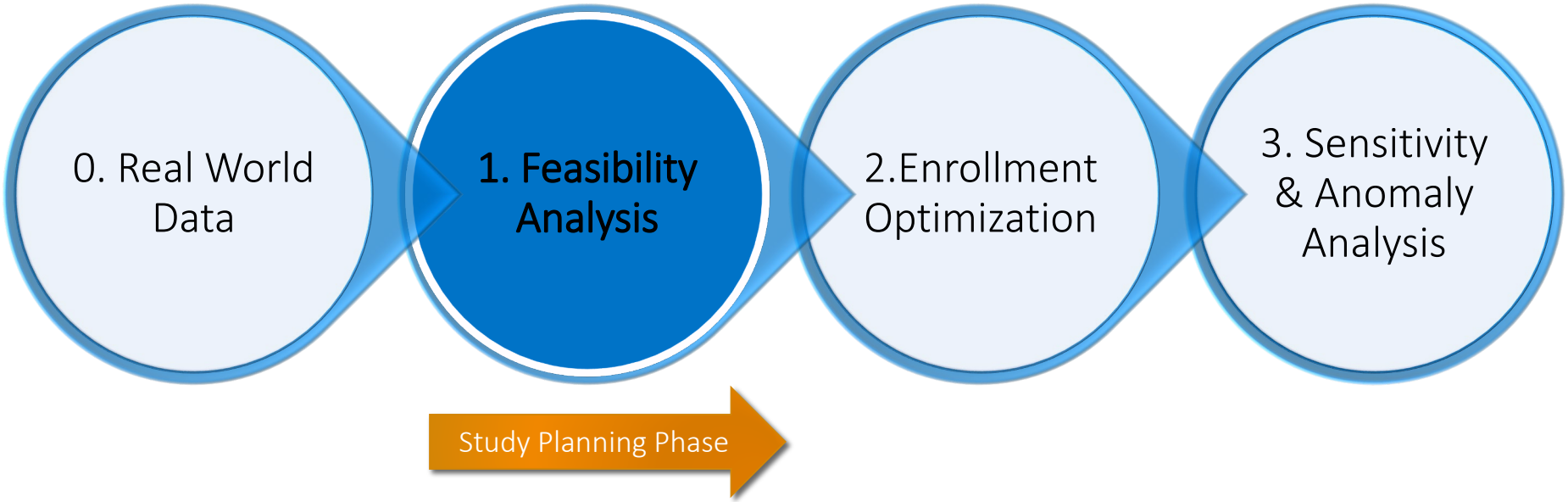
Commercial

HTA

PRO CLAIMS RX COA MORTALITY mHEALTH CONSUMER

EMR LABS REGISTRY SAFETY DIGITAL SOCIAL MEDIA

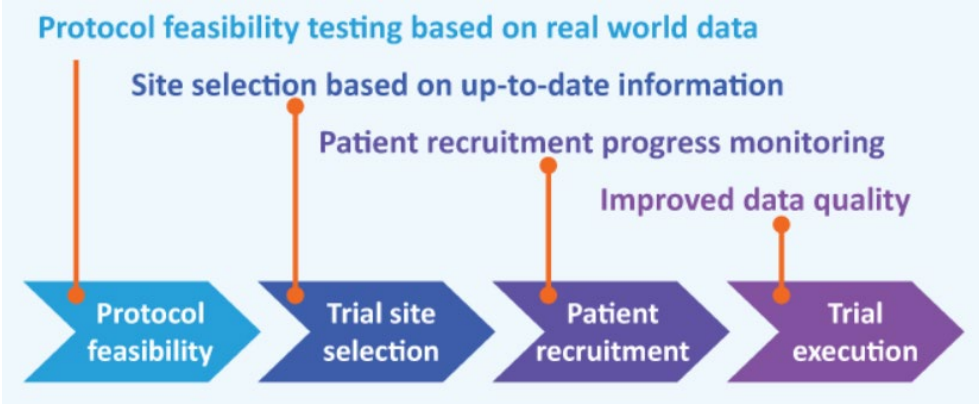
Scope



Feasibility Analysis

How can sponsors accurately **measure the feasibility of a protocol** and plan accordingly to achieve **timely and cost-effective enrollment and the quality data** to advance their pipelines?

- True clinical trial feasibility includes **strategic, scientific, operational, and patient recruitment** considerations
- Focus is on the **enrollability** of a clinical trial
- **Traditional approach** in our industry
  - ✓ Historical Data / Past Experiences, Previous Studies (CRO, Pharma)
  - ✓ Discussions with potential sites and Patient Organizations
  - ✓ Educated guesses



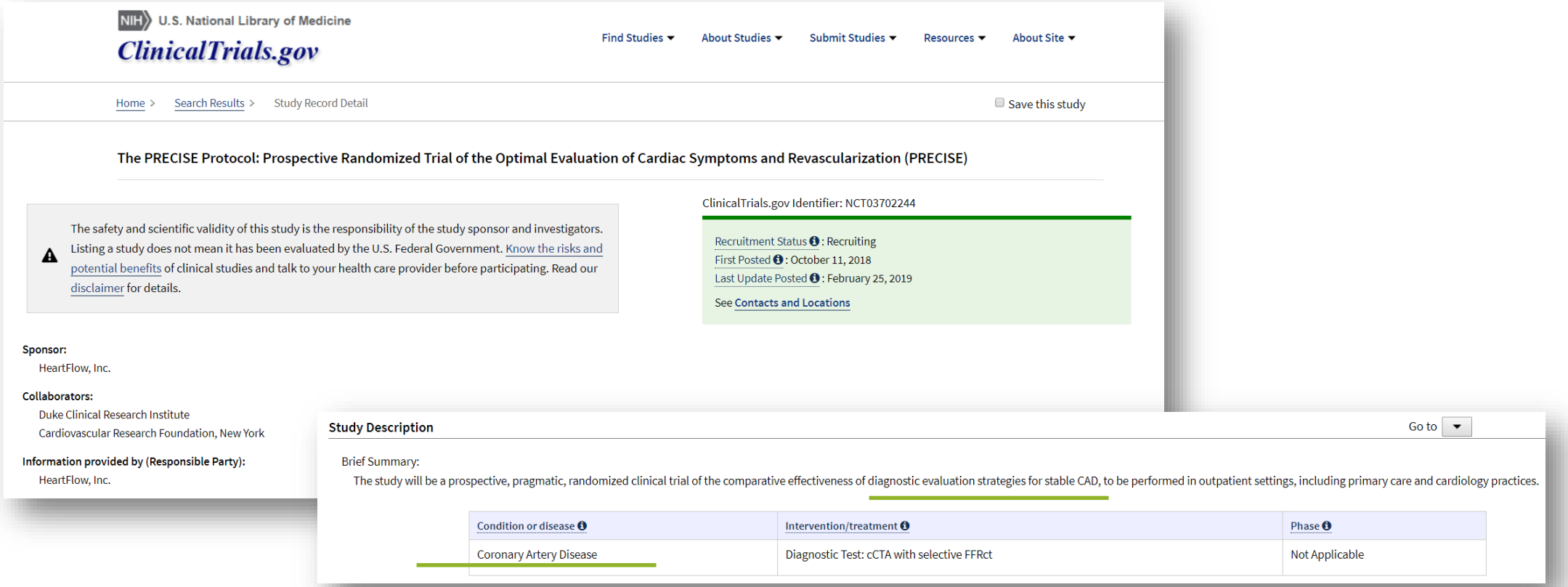
0. Real World Data

1. Feasibility Analysis

Study Planning Phase

Hypothesis Generation

(example study / protocol)



NIH U.S. National Library of Medicine

ClinicalTrials.gov

Find Studies | About Studies | Submit Studies | Resources | About Site

Home > Search Results > Study Record Detail

Save this study

The PRECISE Protocol: Prospective Randomized Trial of the Optimal Evaluation of Cardiac Symptoms and Revascularization (PRECISE)

The safety and scientific validity of this study is the responsibility of the study sponsor and investigators. Listing a study does not mean it has been evaluated by the U.S. Federal Government. Know the risks and potential benefits of clinical studies and talk to your health care provider before participating. Read our disclaimer for details.

ClinicalTrials.gov Identifier: NCT03702244

Recruitment Status : Recruiting
First Posted : October 11, 2018
Last Update Posted : February 25, 2019

See Contacts and Locations

Sponsor:
HeartFlow, Inc.

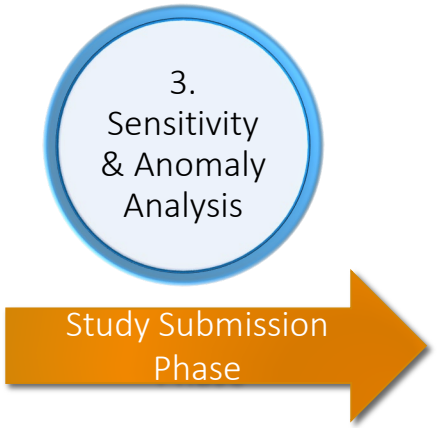
Collaborators:
Duke Clinical Research Institute
Cardiovascular Research Foundation, New York

Information provided by (Responsible Party):
HeartFlow, Inc.

Study Description Go to

Brief Summary:
The study will be a prospective, pragmatic, randomized clinical trial of the comparative effectiveness of diagnostic evaluation strategies for stable CAD, to be performed in outpatient settings, including primary care and cardiology practices.

Condition or disease	Intervention/treatment	Phase
Coronary Artery Disease	Diagnostic Test: cCTA with selective FFRct	Not Applicable



Hypothesis Generation (example study protocol)

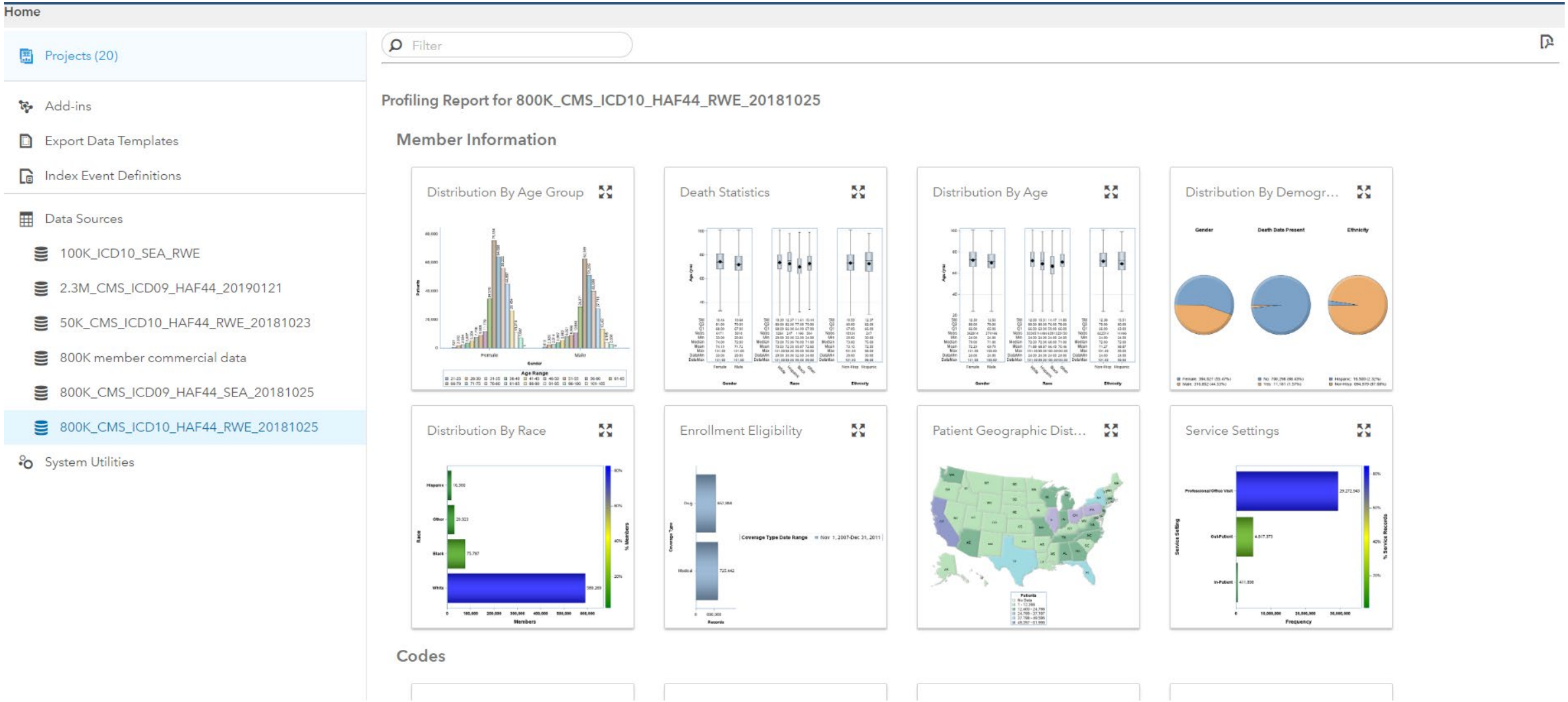
[Framingham Heart Study Wikipedia]

Framingham Heart Study

From Wikipedia, the free encyclopedia

The **Framingham Heart Study** is a long-term, ongoing cardiovascular cohort study on residents of the city of Framingham, Massachusetts. The study began in 1948 with 5,209 adult subjects from Framingham, and is now on its third generation of participants.[1] Prior to it almost nothing was known about the "epidemiology of hypertensive or arteriosclerotic cardiovascular disease".[2] Much of the now-common knowledge concerning heart disease, such as the effects of diet, exercise, and common medications such as aspirin, is based on this longitudinal study. It is a project of the National Heart, Lung, and Blood Institute, in collaboration with (since 1971) Boston University.[1] Various health professionals from the hospitals and universities of Greater Boston staff the project.

Profiling the data



Creation ‘Index Event’ Definition

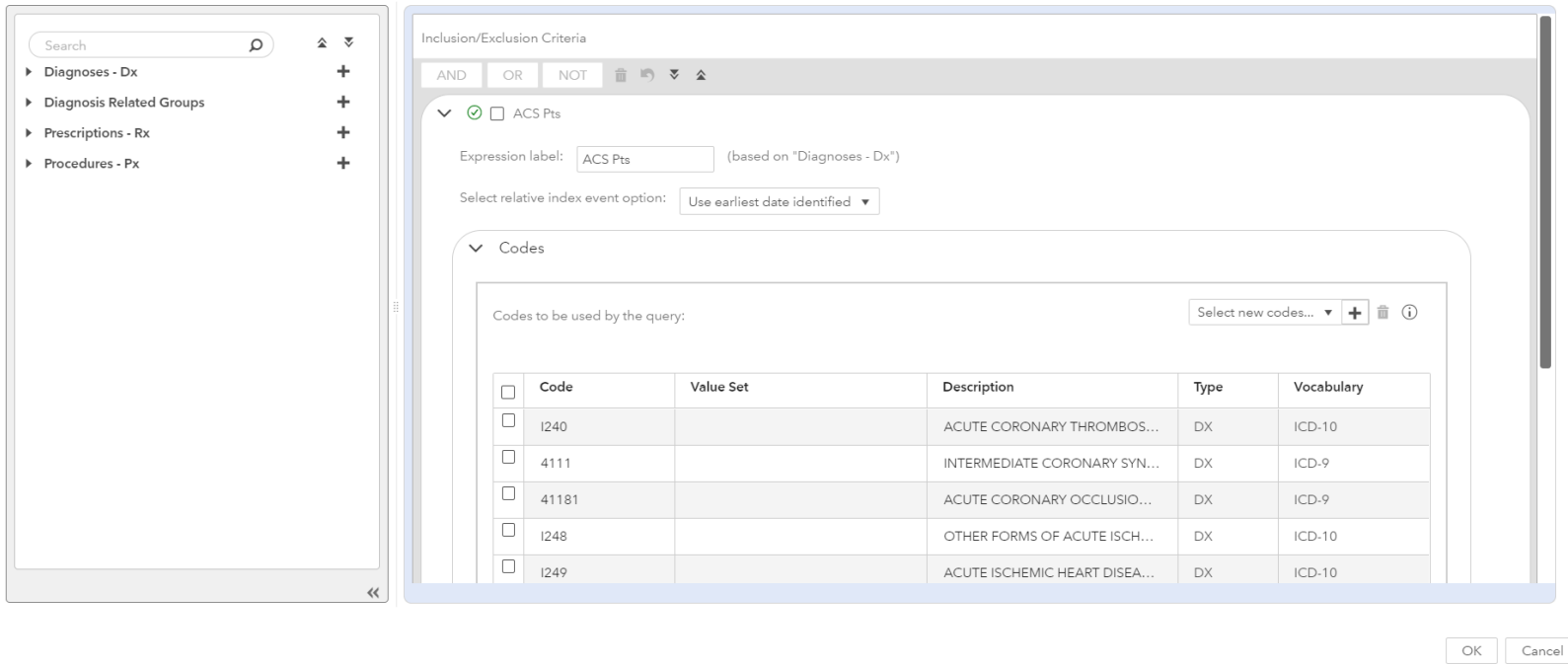
Acute coronary syndrome (ACS) Patients

Criteria

Inclusion Criteria:

- Age ≥18 years
- Stable typical or atypical symptoms suggesting possible coronary artery disease (CAD) with further non-emergent testing or elective catheterization recommended to evaluate the presence of suspected CAD
- Safe performance of cCTA:
  1. Creatinine clearance ≥45 ml/min
  2. For a female participant of childbearing potential, a pregnancy test must be performed with negative results known within 7 days prior to randomization
- Willingness to comply with all aspects of the protocol, including adherence to the assigned strategy and follow-up visits
- Ability to provide written informed consent

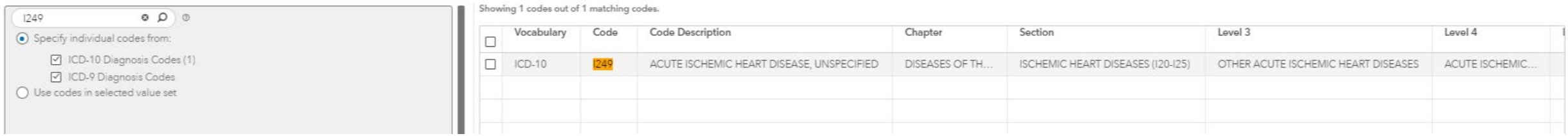
Index Event Definition: ACS Pts

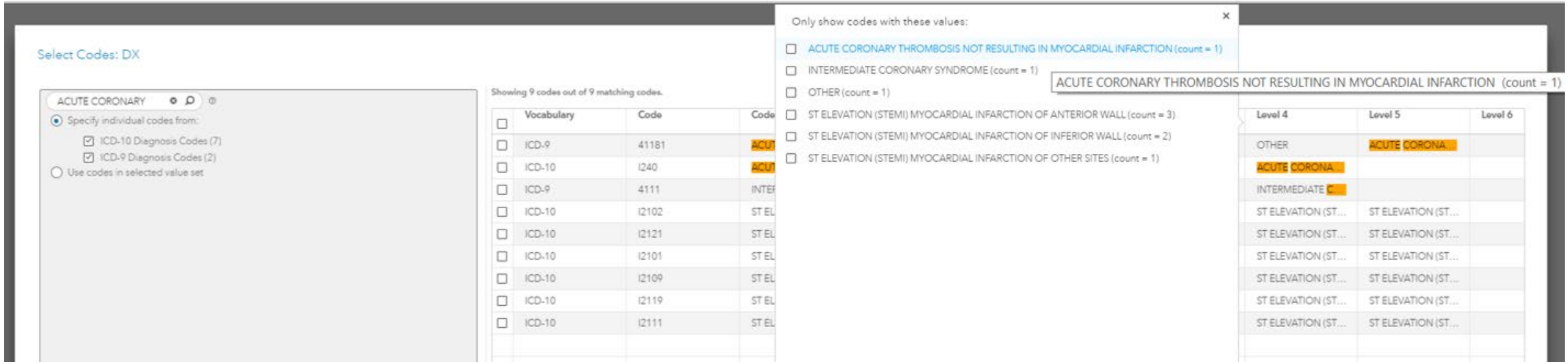


Creation ‘Index Event’ Definition

Acute coronary syndrome (ACS) Patients – Cont’d

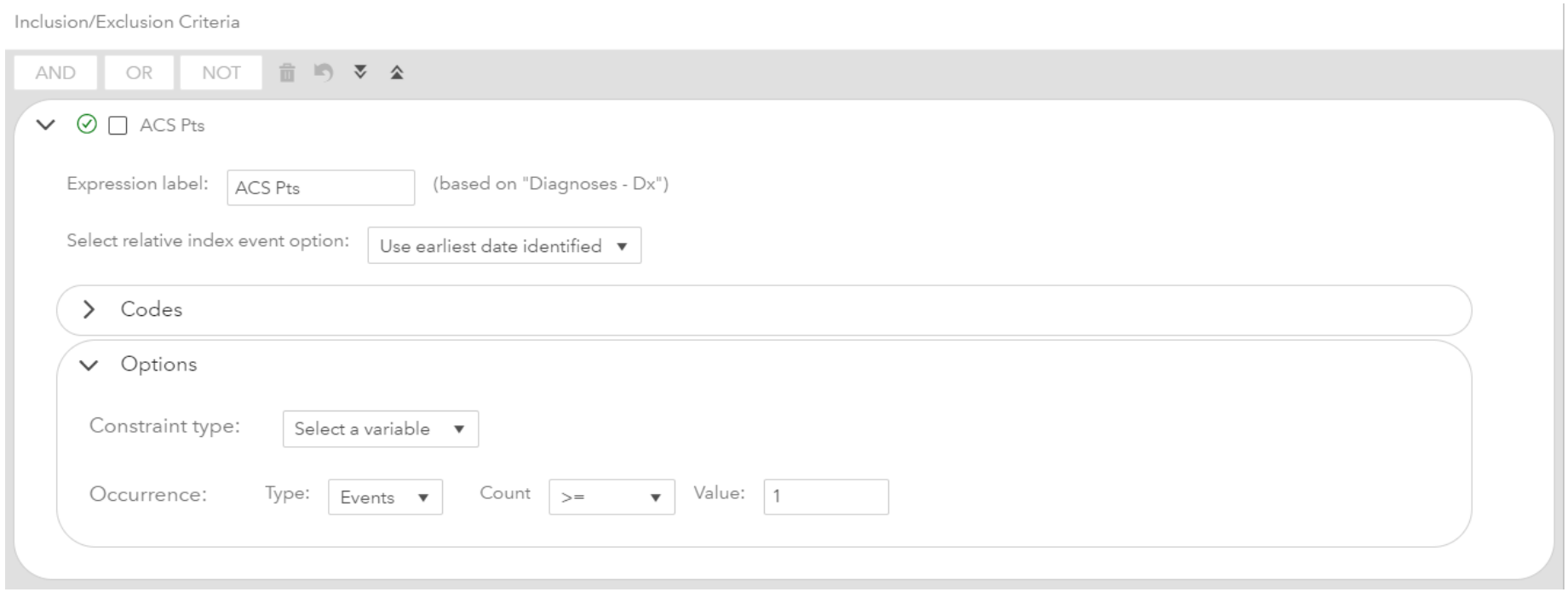
Select Codes: DX





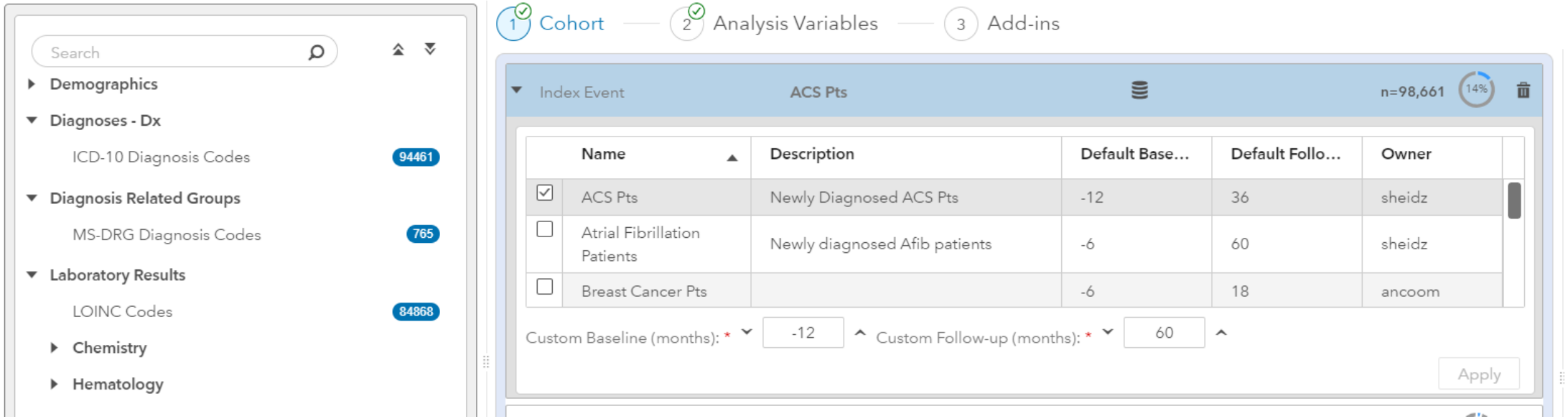
Creation ‘Index Event’ Definition

Acute coronary syndrome (ACS) Patients – Cont’d



Cohort Generation

Selection ‘Index Event’ Definition



Cohort Generation

Inclusion Criteria: Age at Study Start (>=18)

Criteria

Inclusion Criteria:

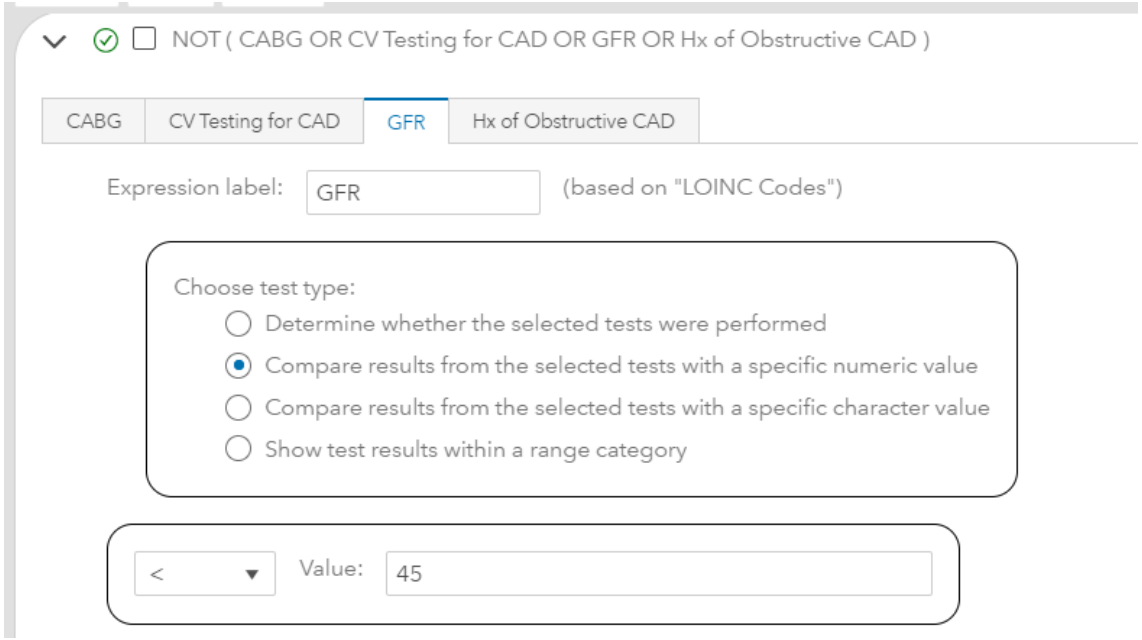
- Age ≥18 years
- Stable typical or atypical symptoms suggesting possible coronary artery disease (CAD) with further non-emergent testing or elective catheterization recommended to evaluate the presence of suspected CAD
- Safe performance of cCTA:
    1. Creatinine clearance ≥45 ml/min
    2. For a female participant of childbearing potential, a pregnancy test must be performed with negative results known within 7 days prior to randomization
- Willingness to comply with all aspects of the protocol, including adherence to the assigned strategy and follow-up visits
- Ability to provide written informed consent

Cohort Generation

Exclusion Criteria: GFR (<45)

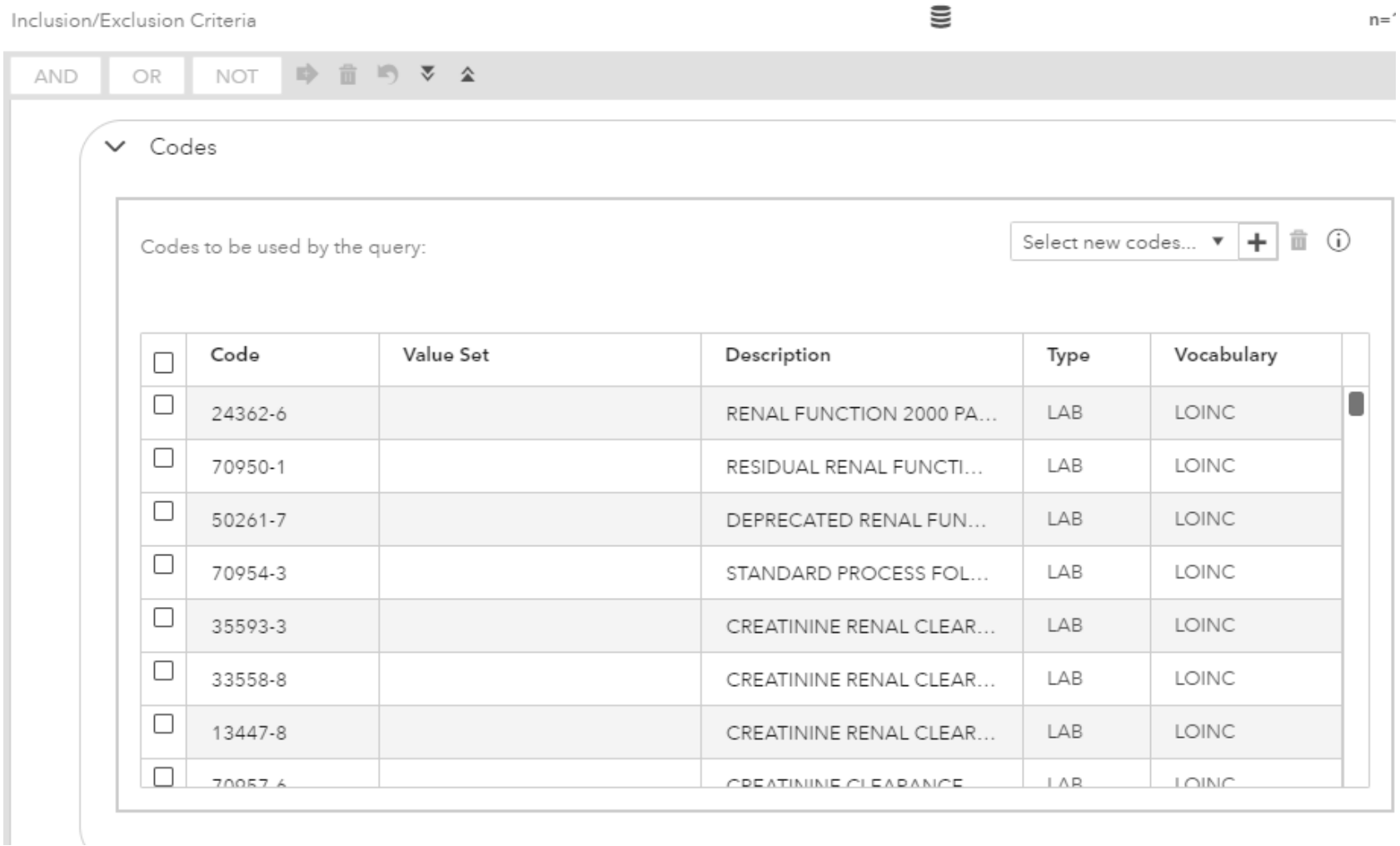
Exclusion Criteria:

- Acute chest pain
- Unstable clinical status
- Noninvasive or invasive CV testing for CAD within 1 year
- Lifetime history of any obstructive CAD (no prior CABG or PCI, stenosis ≥50%), or known EF ≤40% or moderate to severe valvular or congenital cardiac disease
- Contraindications to cCTA including but not limited to estimated creatinine clearance (GFR) <45 ml/min
- Any condition leading to possible inability to comply with the protocol
- Exceeds the weight or size limit for cCTA or cardiac catheterization at the site
- Life expectancy less than 2 years due to non-cardiovascular comorbidities
- Enrolled in an investigational trial that involves a non-approved cardiac drug or device which has not reached its primary endpoint
- Any condition that might interfere with the study procedures or follow-up



Cohort Generation

Exclusion Criteria: GFR (<45) – Cont'd



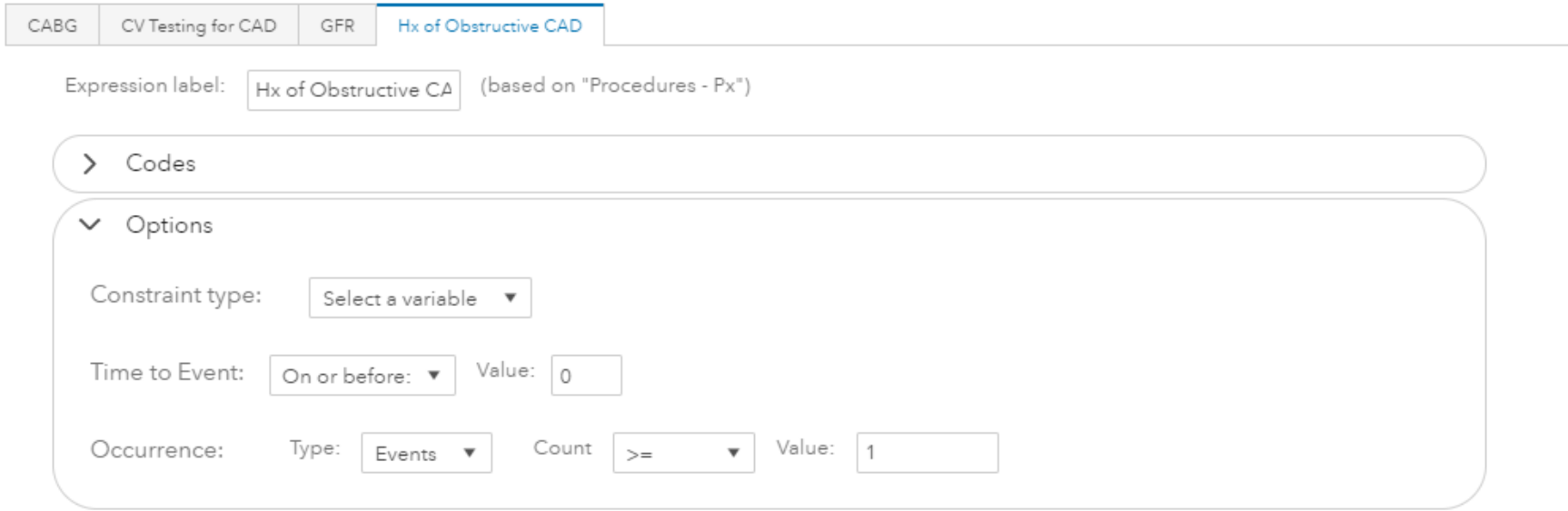
Code	Value Set	Description	Type	Vocabulary
24362-6		RENAL FUNCTION 2000 PA...	LAB	LOINC
70950-1		RESIDUAL RENAL FUNCTI...	LAB	LOINC
50261-7		DEPRECATED RENAL FUN...	LAB	LOINC
70954-3		STANDARD PROCESS FOL...	LAB	LOINC
35593-3		CREATININE RENAL CLEAR...	LAB	LOINC
33558-8		CREATININE RENAL CLEAR...	LAB	LOINC
13447-8		CREATININE RENAL CLEAR...	LAB	LOINC
70957-6		CREATININE CLEARANCE	LAB	LOINC

Cohort Generation

Exclusion Criteria: Lifetime history of any obstructive CAD

Exclusion Criteria:

- Acute chest pain
- Unstable clinical status
- Noninvasive or invasive CV testing for CAD within 1 year
- Lifetime history of any obstructive CAD (no prior CABG or PCI, stenosis ≥50%), or known EF ≤40% or moderate to severe valvular or congenital cardiac disease
- Contraindications to cCTA including but not limited to estimated creatinine clearance (GFR) <45 ml/min
- Any condition leading to possible inability to comply with the protocol
- Exceeds the weight or size limit for cCTA or cardiac catheterization at the site
- Life expectancy less than 2 years due to non-cardiovascular comorbidities
- Enrolled in an investigational trial that involves a non-approved cardiac drug or device which has not reached its primary endpoint
- Any condition that might interfere with the study procedures or follow-up

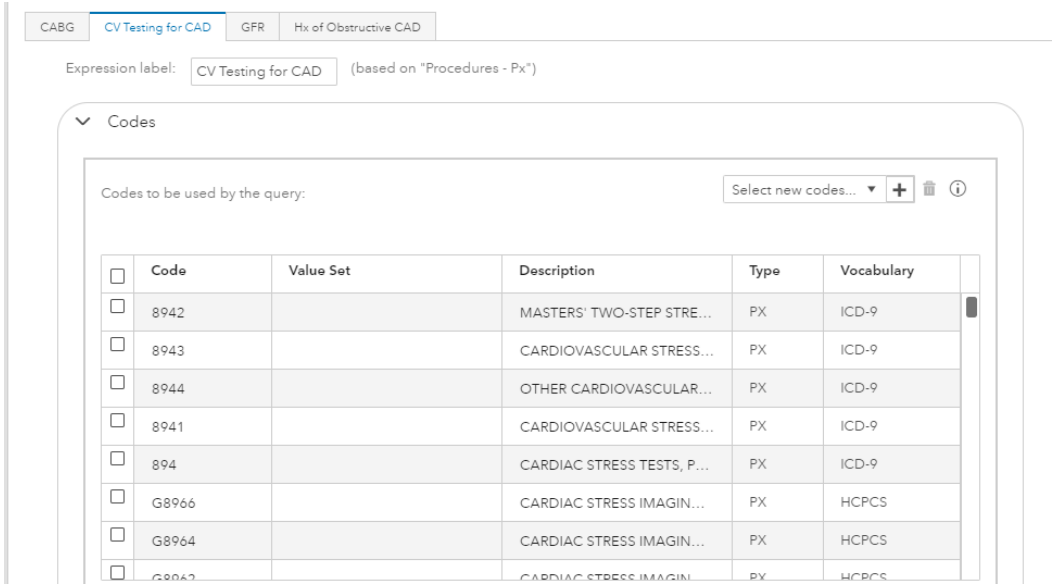


Cohort Generation

Exclusion Criteria: CV testing for CAD within 1 year

Exclusion Criteria:

- Acute chest pain
- Unstable clinical status
- Noninvasive or invasive CV testing for CAD within 1 year
- Lifetime history of any obstructive CAD (no prior CABG or PCI, stenosis ≥50%), or known EF ≤40% or moderate to severe valvular or congenital cardiac disease
- Contraindications to cCTA including but not limited to estimated creatinine clearance (GFR) <45 ml/min
- Any condition leading to possible inability to comply with the protocol
- Exceeds the weight or size limit for cCTA or cardiac catheterization at the site
- Life expectancy less than 2 years due to non-cardiovascular comorbidities
- Enrolled in an investigational trial that involves a non-approved cardiac drug or device which has not reached its primary endpoint
- Any condition that might interfere with the study procedures or follow-up

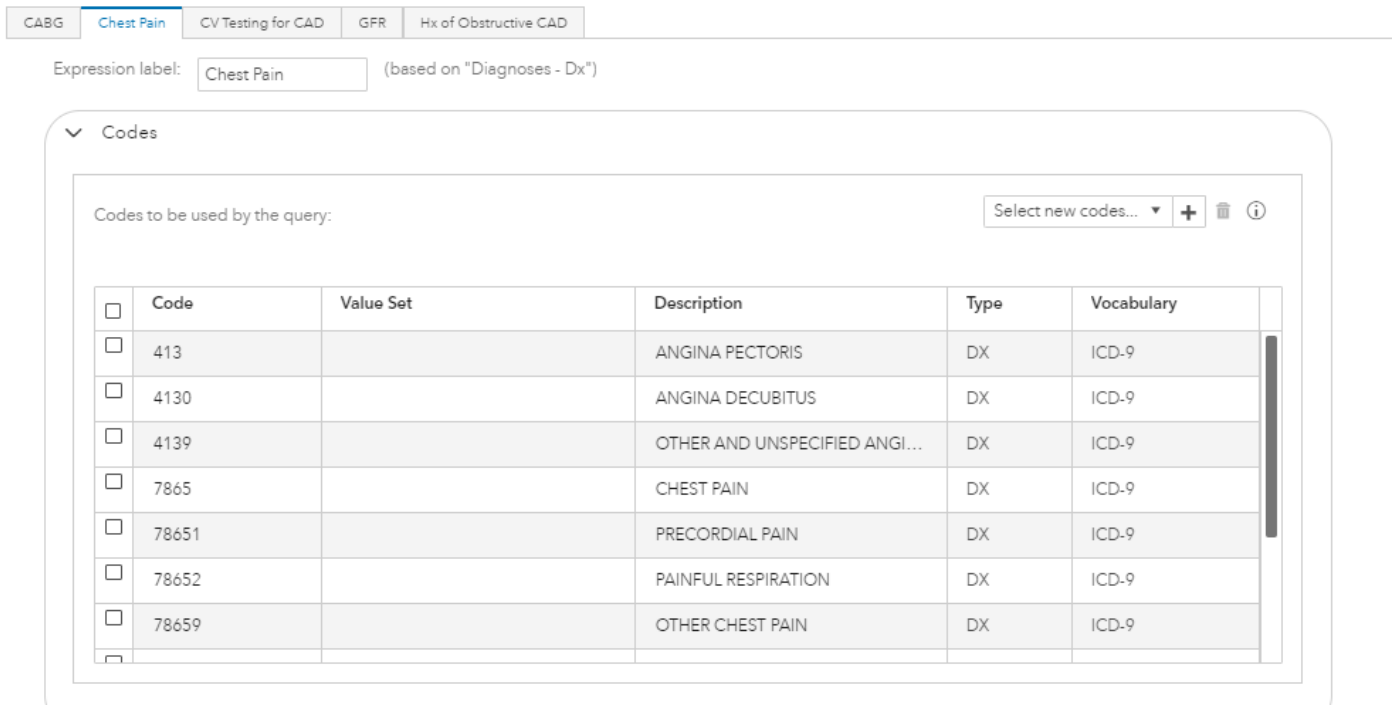


Cohort Generation

Exclusion Criteria: Acute Chest Pain

Exclusion Criteria:

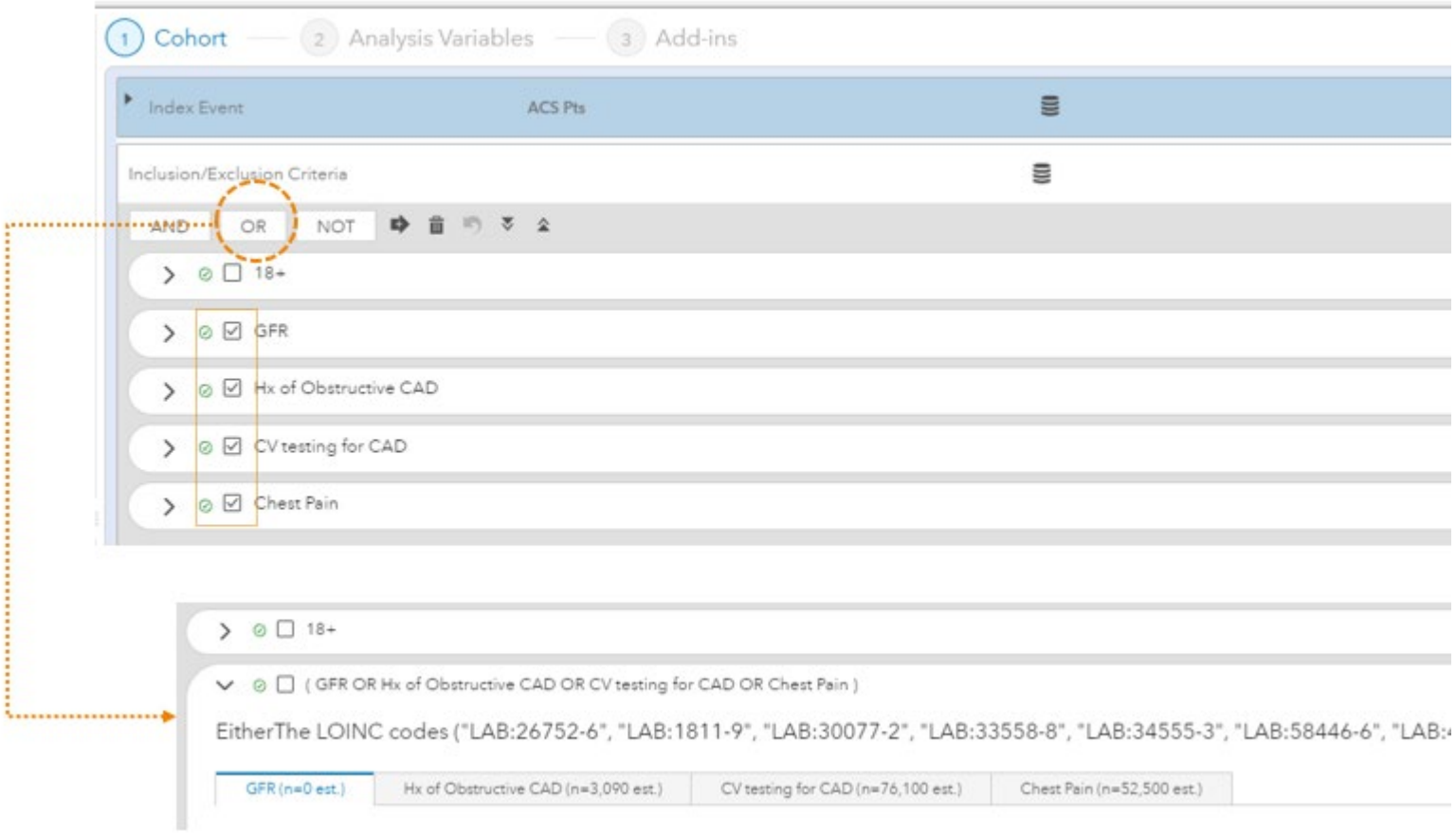
- Acute chest pain
- Unstable clinical status
- Noninvasive or invasive CV testing for CAD within 1 year
- Lifetime history of any obstructive CAD (no prior CABG or PCI, stenosis ≥50%), or known EF ≤40% or moderate to severe valvular or congenital cardiac disease
- Contraindications to cCTA including but not limited to estimated creatinine clearance (GFR) <45 ml/min



Code	Value Set	Description	Type	Vocabulary
413		ANGINA PECTORIS	DX	ICD-9
4130		ANGINA DECUBITUS	DX	ICD-9
4139		OTHER AND UNSPECIFIED ANGI...	DX	ICD-9
7865		CHEST PAIN	DX	ICD-9
78651		PRECORDIAL PAIN	DX	ICD-9
78652		PAINFUL RESPIRATION	DX	ICD-9
78659		OTHER CHEST PAIN	DX	ICD-9

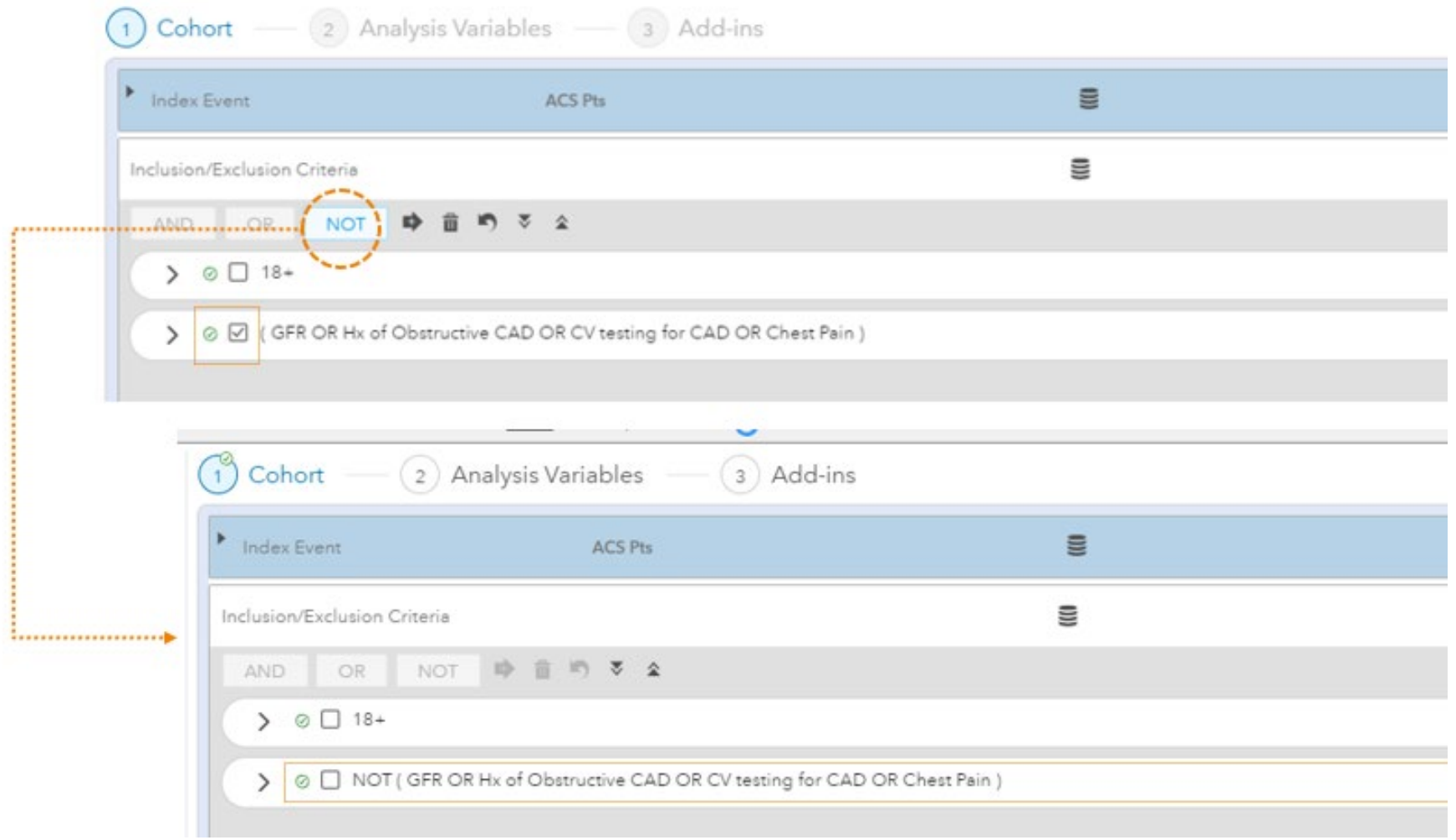
Cohort Generation

Exclusion Criteria (OR)



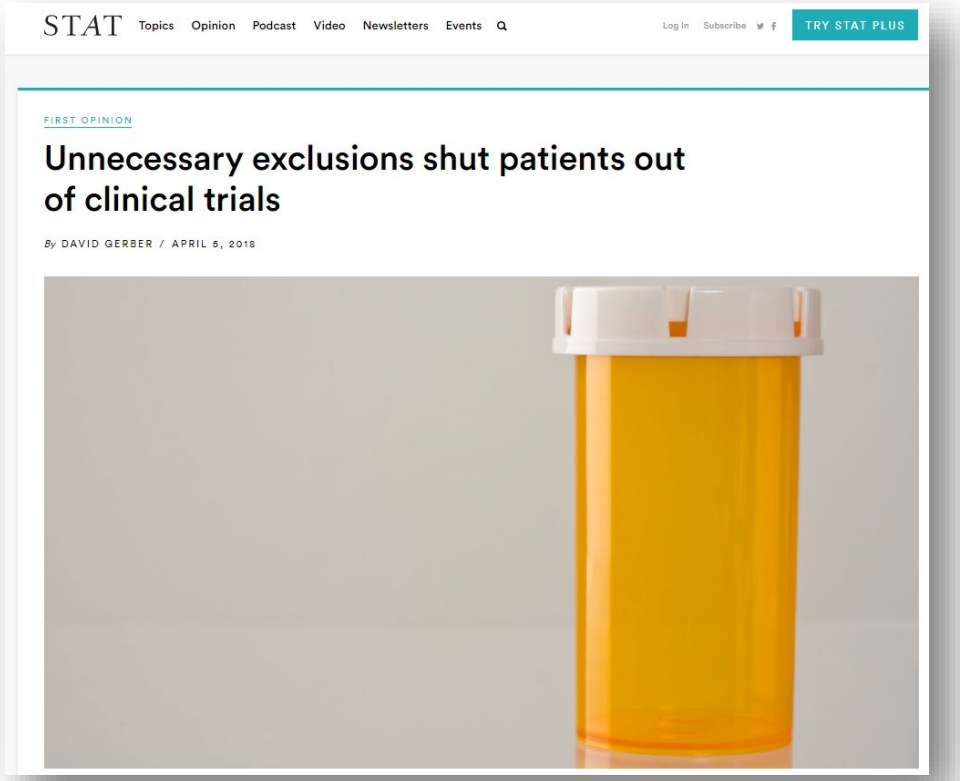
Cohort Generation

Exclusion Criteria (NOT)



Cohort Generation

What if scenario's… What if we loosen exclusion criteria ?



https://www.statnews.com/2018/04/05/clinical-trials-excluding-patients/

Not long ago, I examined an otherwise healthy and active man in his 50s with newly diagnosed metastatic lung cancer. He had no heart disease, no kidney disease, no liver disease. He came to my practice at UT Southwestern Medical Center specifically for clinical trial options.

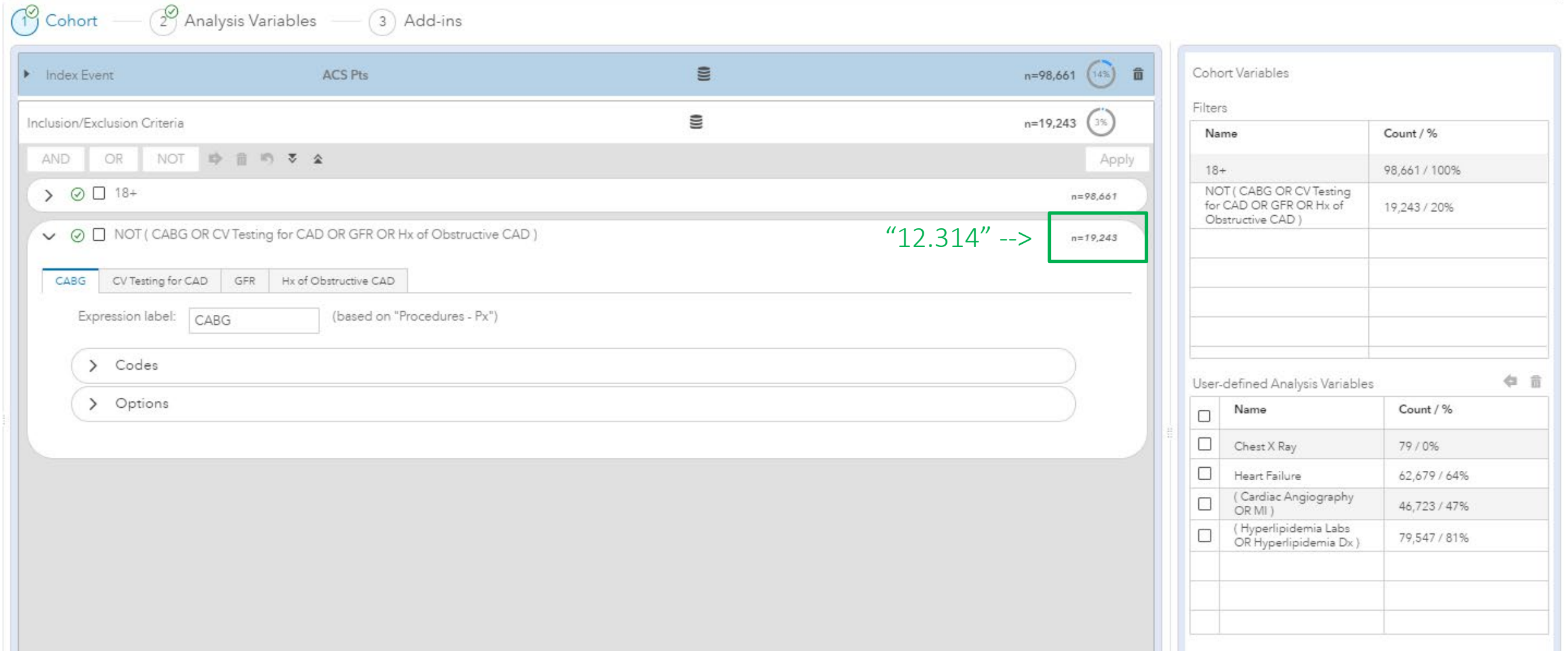
Disappointingly, one detail in his medical history excluded him from all of our available trials. Four years earlier, he had been treated for stage 1 prostate cancer, a disease so prevalent and nonaggressive that the U.S. government no longer recommends routine screening for it. I couldn't think of a single way in which the patient's prior experience with cancer would interfere with treatments or assessments on a lung cancer trial. And yet I couldn't enroll him.

Unfortunately, many sick patients are being denied the opportunity to test new drugs that might save their lives. It's time for this to change.

Running clinical trials effectively and efficiently is critical to medical progress, not just in cancer care but across disciplines. Allowing more patients to be enrolled in trials will speed the medical innovation process, allow more sick people to access potentially beneficial therapies, and produce more generalized results. Spurred by this month's FDA meeting, lead researchers, sponsors, and regulators must rethink their approach to clinical trial eligibility (5 Apr 2018).

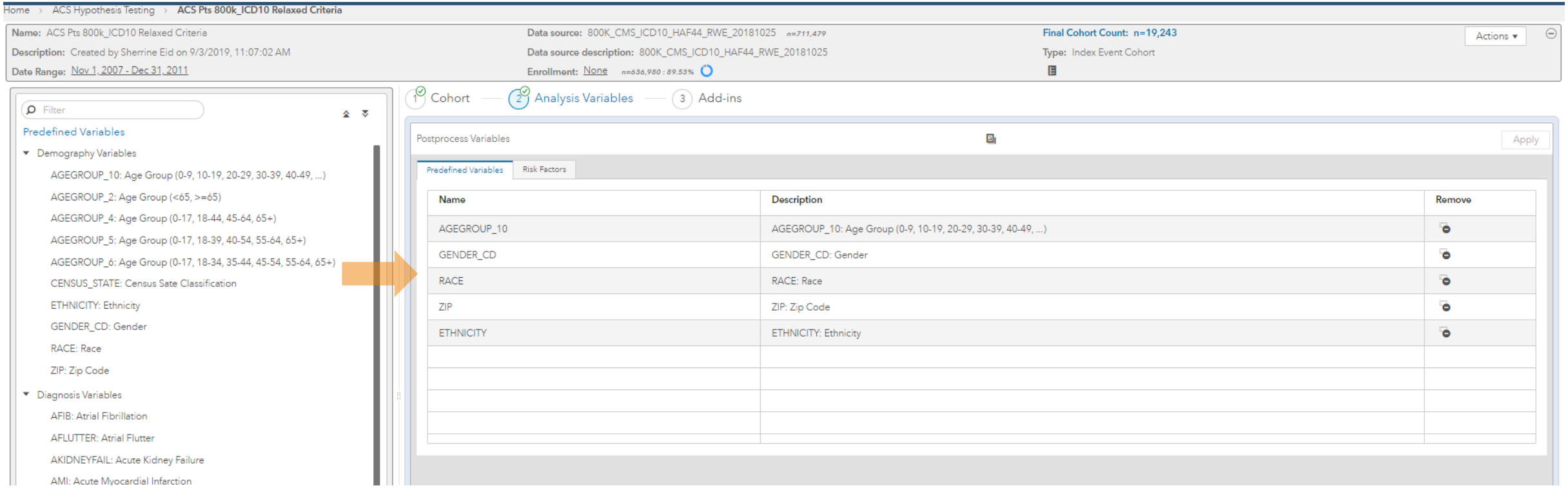
Cohort Generation

What if scenario's… What if we loosen exclusion criteria ?



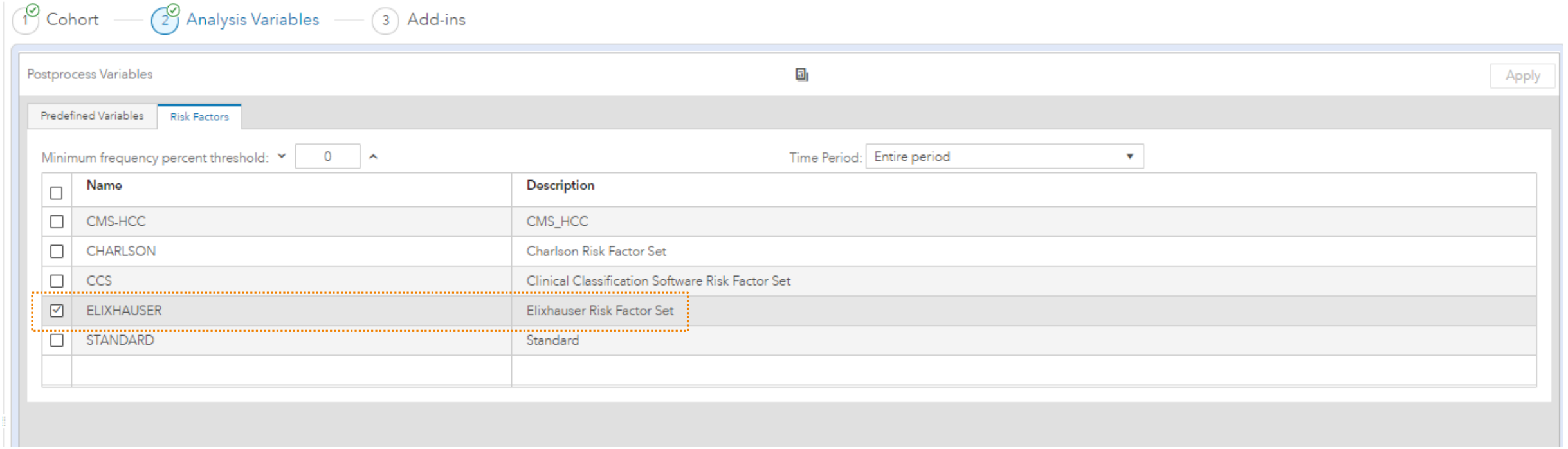
Analysis Variables

Select Pre-defined Analysis Variables



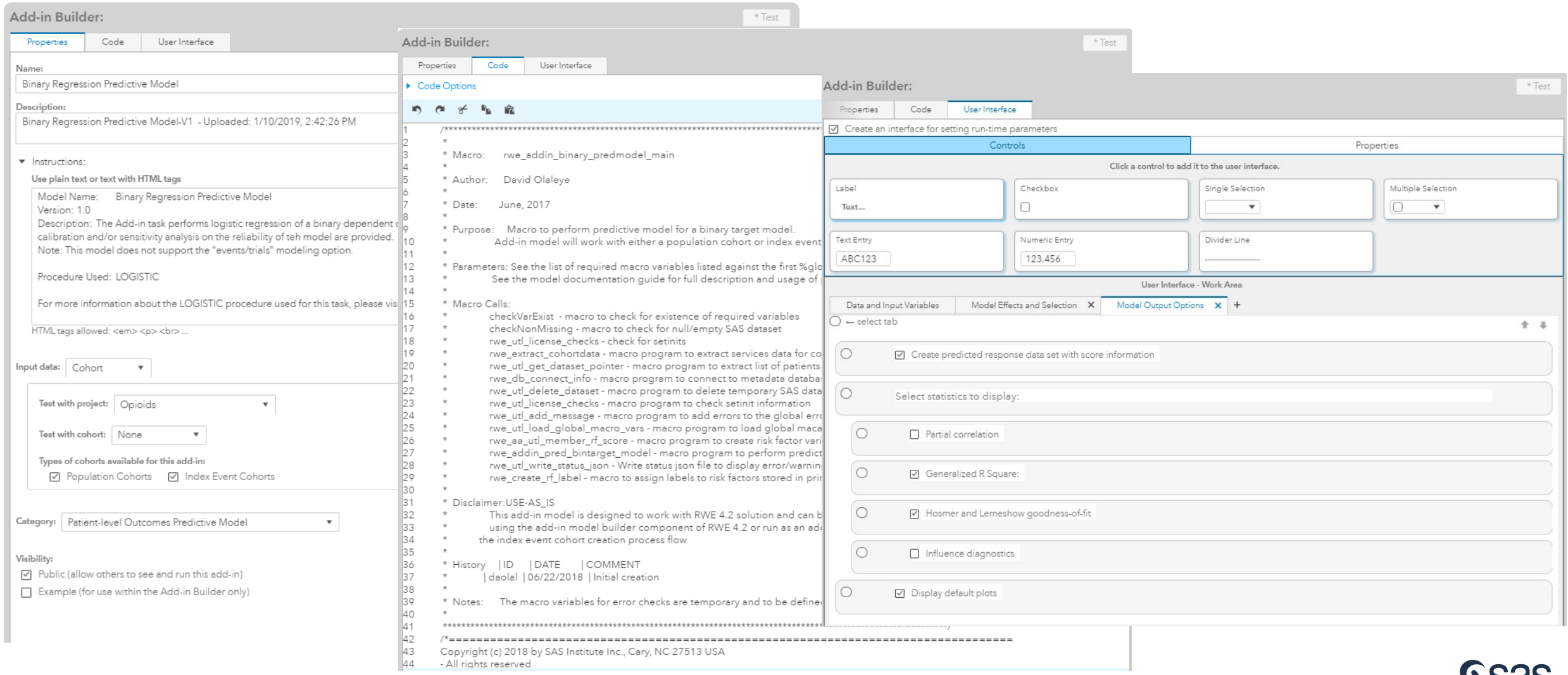
Analysis Variables

Risk Factors



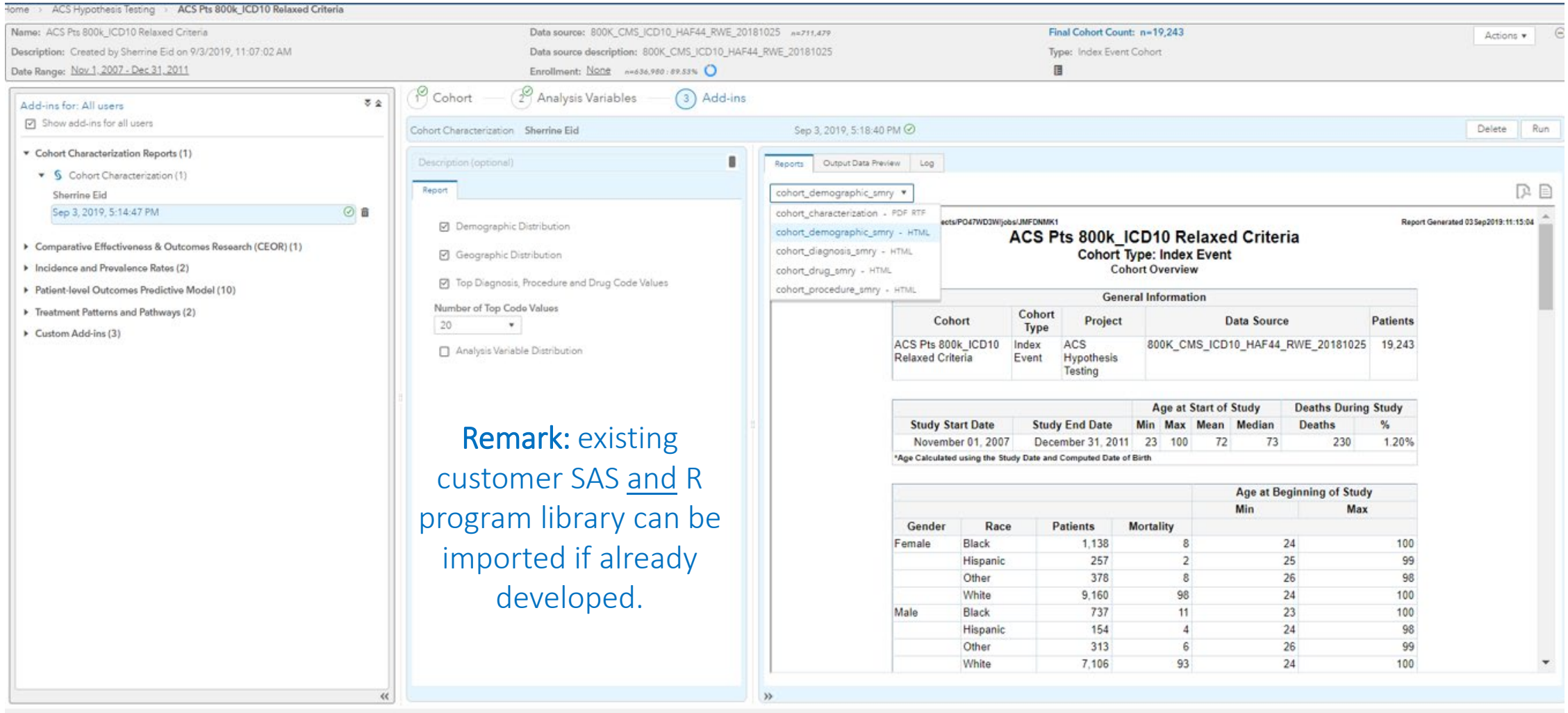
Model & UI development

Use SAS (industry) and company specific models



Model & UI development

Cohort Characterization Report

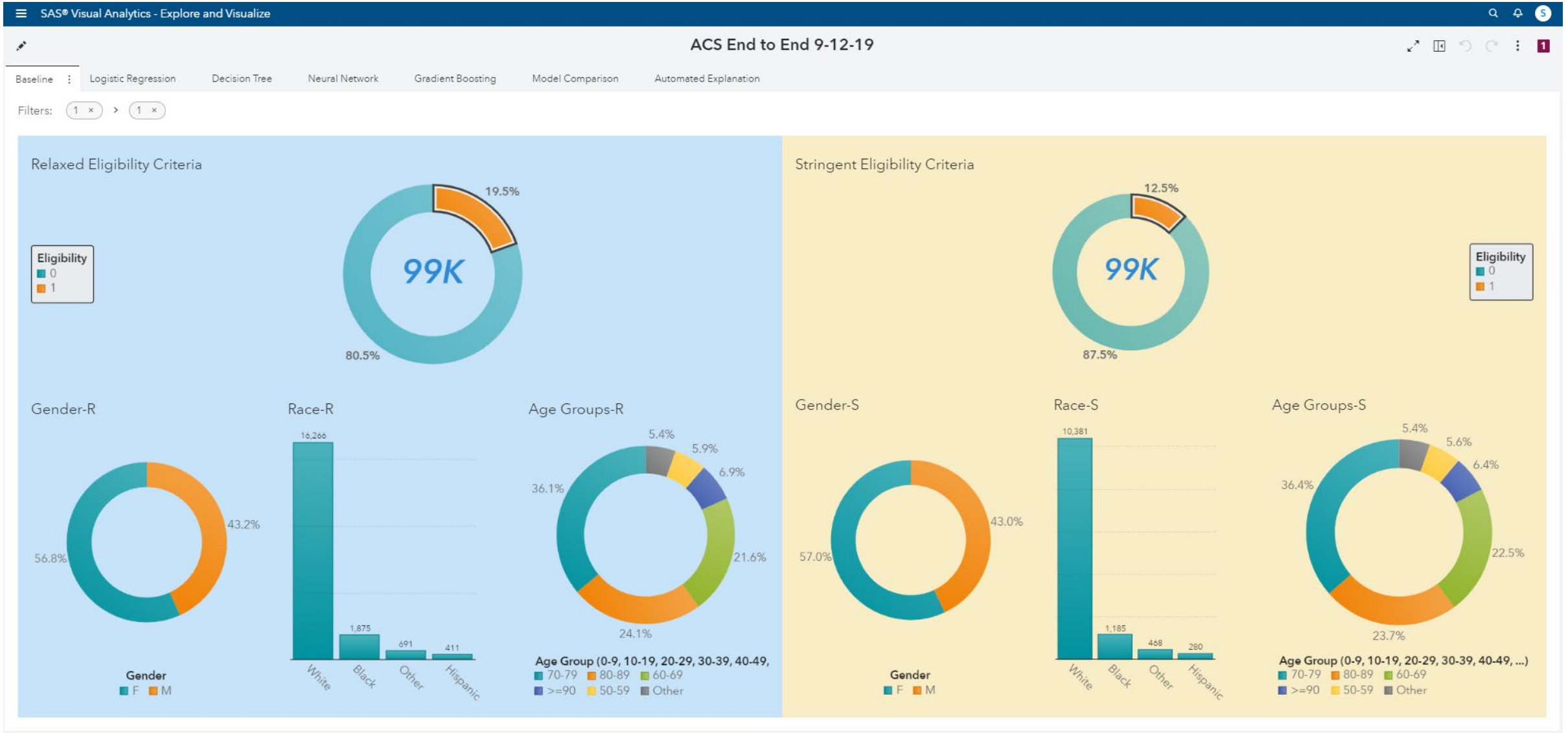


Remark: existing customer SAS and R program library can be imported if already developed.

Further Visualization / Exploration, Data Mining & Machine Learning

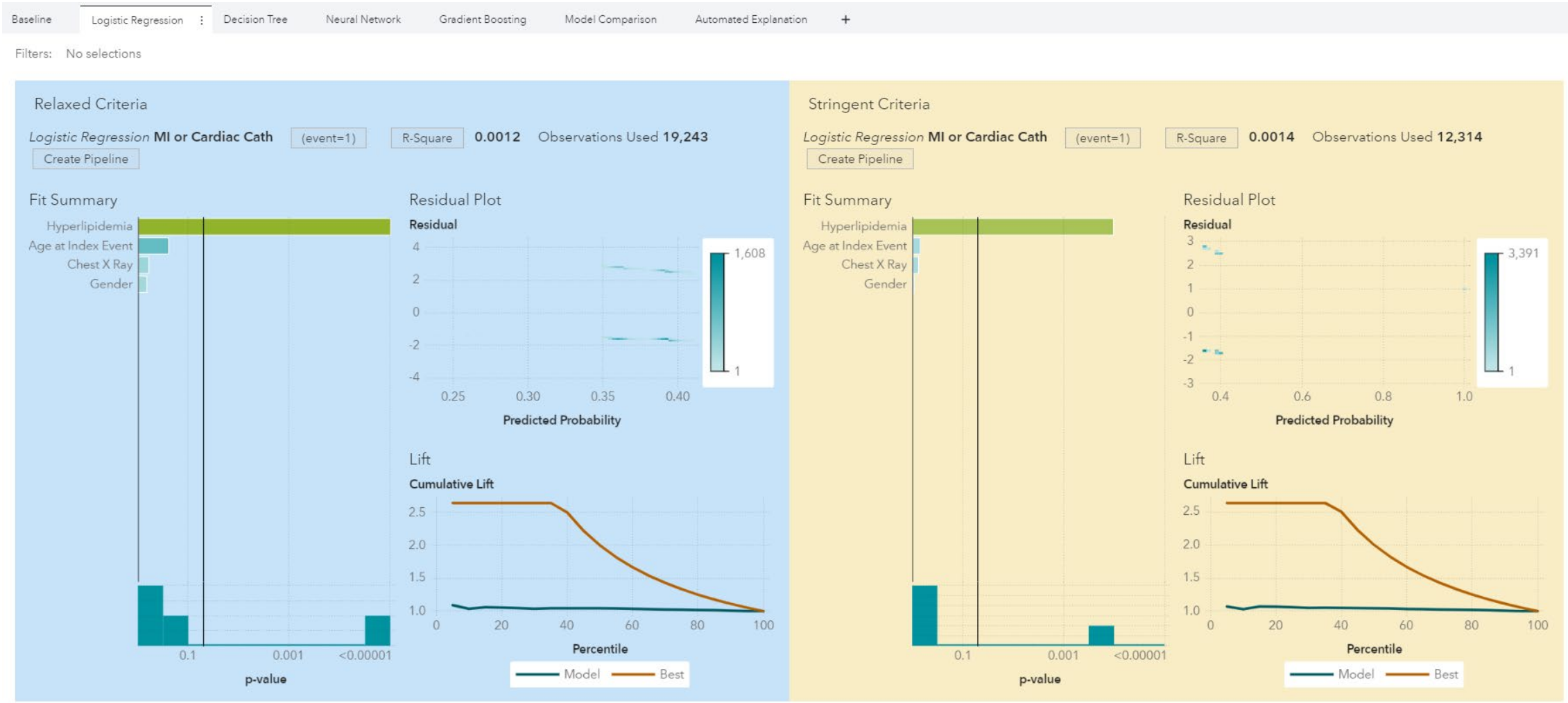
… Collaboration …

Visualization / Exploration



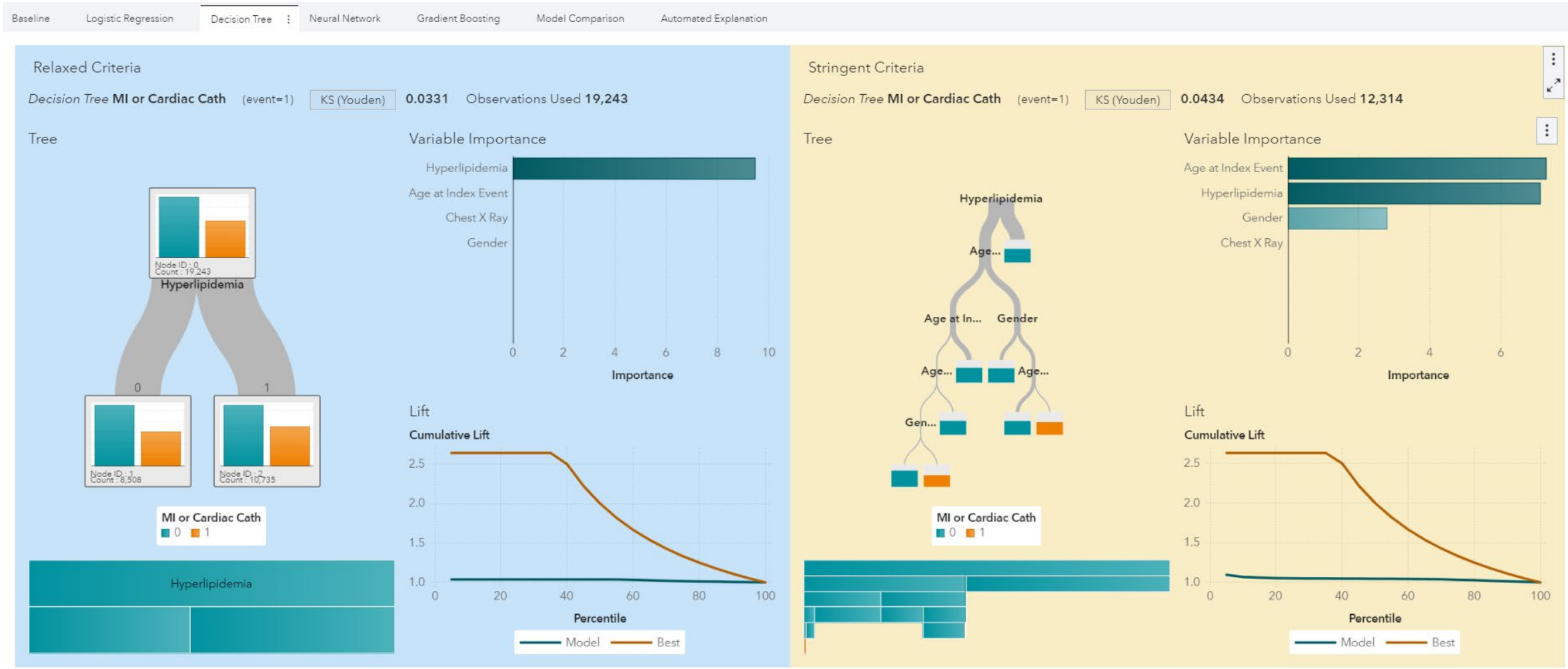
Data Mining & Machine Learning

Logistic Regression



Data Mining & Machine Learning

Decision Tree



Data Mining & Machine Learning

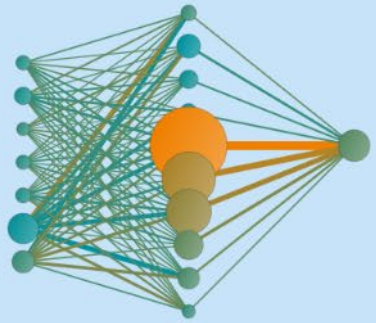
Neural Network

Baseline | Logistic Regression | Decision Tree | Neural Network | Gradient Boosting | Model Comparison | Automated Explanation

Relaxed Criteria

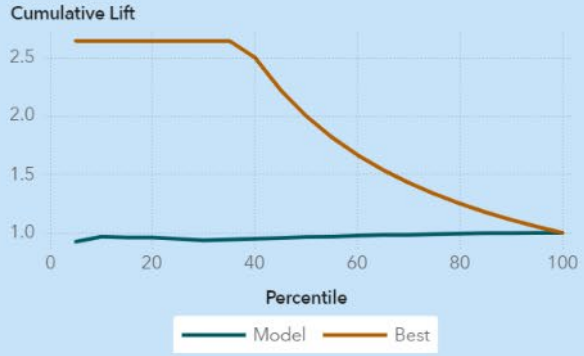
Neural Network **MI or Cardiac Cath** (event=1) KS (Youden) **0.0000** Observations Used **19,243**

Network



Iteration Plot

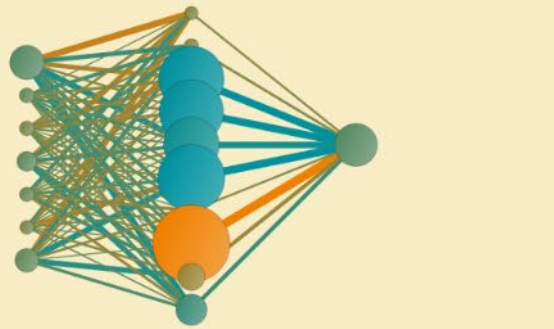
Lift



Stringent Criteria

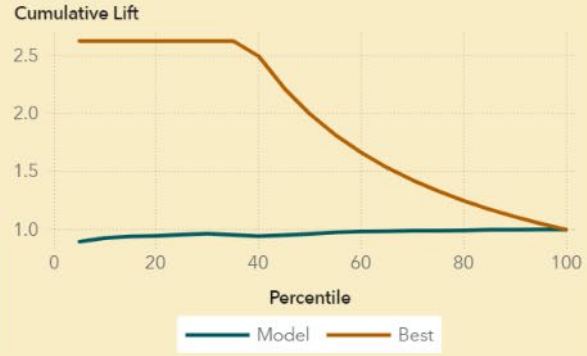
Neural Network **MI or Cardiac Cath** (event=1) KS (Youden) **0.0000** Observations Used **12,314**

Network



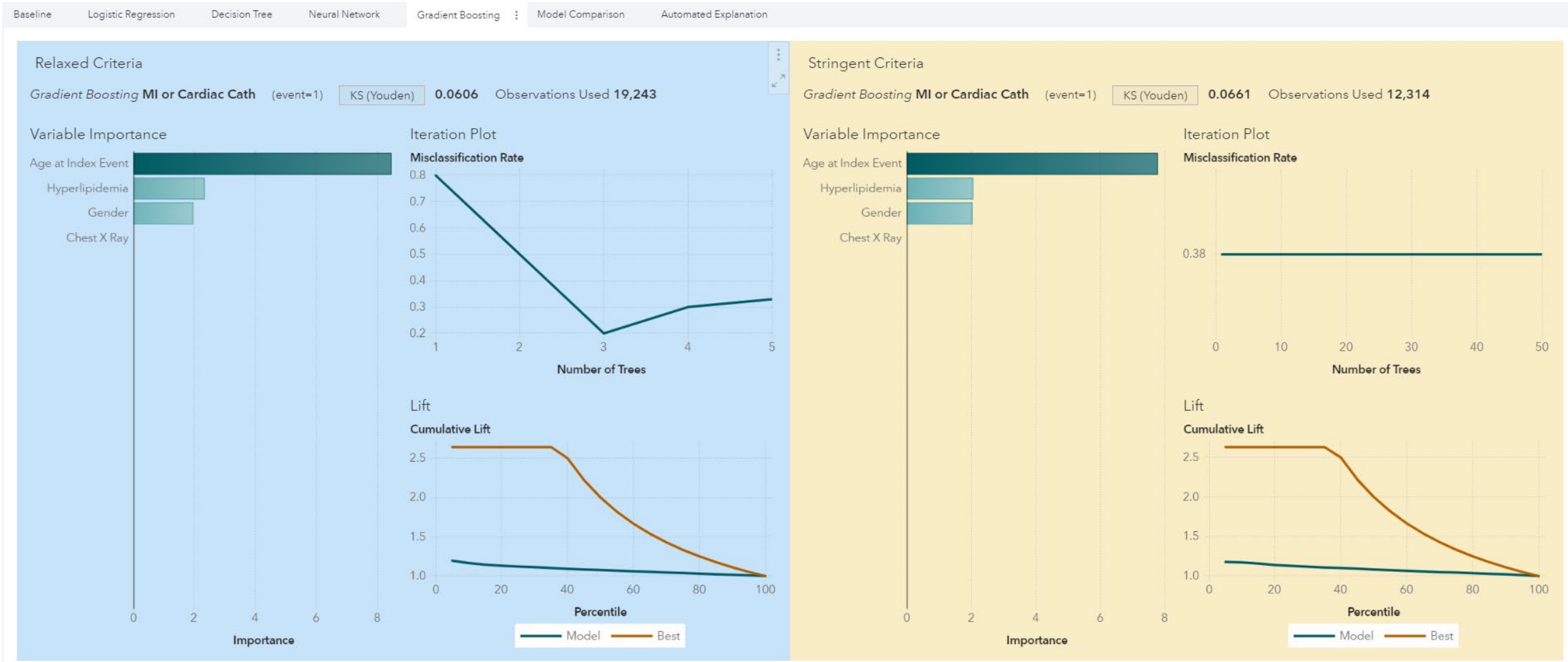
Iteration Plot

Lift



Data Mining & Machine Learning

Gradient Boosting



Data Mining & Machine Learning

Model (automatic) Comparison

Data Mining & Machine Learning (NLG)

Automated Explanation

Baseline | Logistic Regression | Decision Tree | Neural Network | Gradient Boosting | Model Comparison | Automated Explanation

Relaxed Criteria

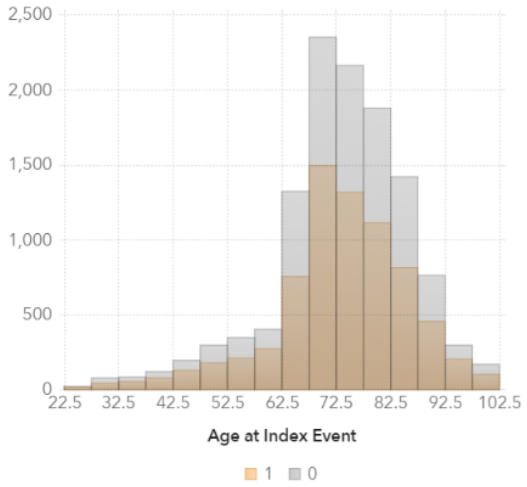
What are the characteristics of MI or Cardiac Cath?

0 | 1

MI or Cardiac Cath has a 37.86% chance (7.3K of 19K) of being 1. It's the least common MI or Cardiac Cath value.

What factors are most related to MI or Cardiac Cath?

What is the relationship between MI or Cardiac Cath and Age at Index Event?



The average Age at Index Event when MI or Cardiac Cath is 1 is 73, with a minimum of 24 and a maximum of 101. The average Age at Index Event when MI or Cardiac Cath is 0 is 73, with a minimum of 24 and a maximum of 101. Average Age at Index Event is 73, and it ranges from 24 to 101.

What are the groups based on Age at Index Event by the chance of MI or Cardiac Cath being 1?

High | Low

56.25% If Age at Index Event is between 81 and 83, Race is Hispanic or Other, then MI or Cardiac Cath has a 56.25% chance (36 out of 64 cases) of being 1.

51.67% If Race is Hispanic or White, Age at Index Event is between 61 and 62, then MI or Cardiac Cath has a 51.67% chance (62 out of 120 cases) of being 1.

45.79% If Race is Hispanic or White, Age at

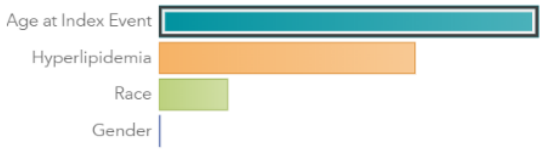
Stringent Criteria

What are the characteristics of MI or Cardiac Cath?

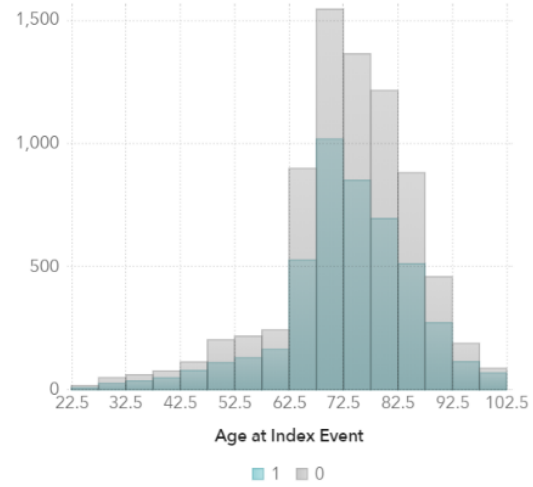
0 | 1

MI or Cardiac Cath has a 37.99% chance (4.7K of 12K) of being 1. It's the least common MI or Cardiac Cath value.

What factors are most related to MI or Cardiac Cath?



What is the relationship between MI or Cardiac Cath and Age at Index Event?



The average Age at Index Event when MI or Cardiac Cath is 1 is 73, with a minimum of 25 and a maximum of 101. The average Age at Index Event when MI or Cardiac Cath is 0 is 73, with a minimum of 24 and a maximum of 101. Average Age at Index Event is 73, and it ranges from 24 to 101.

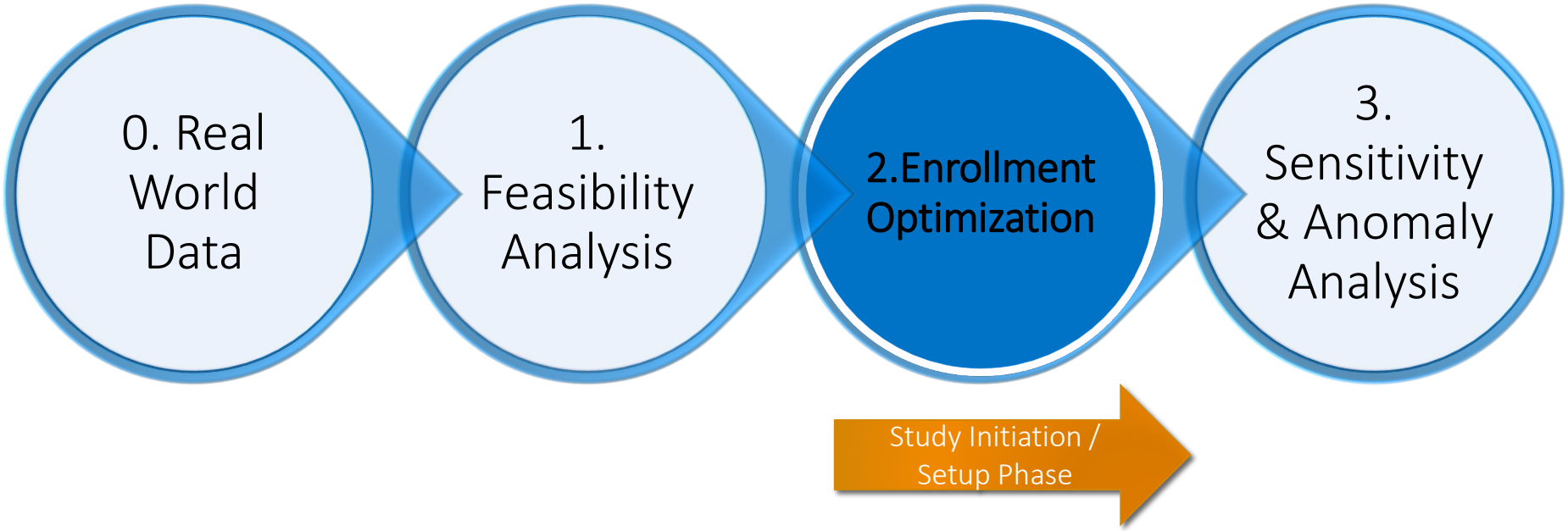
What are the groups based on Age at Index Event by the chance of MI or Cardiac Cath being 1?

High | Low

38.98% If Age at Index Event is less than 74, then MI or Cardiac Cath has a 38.98% chance (2.6K out of 6.5K cases) of being 1.

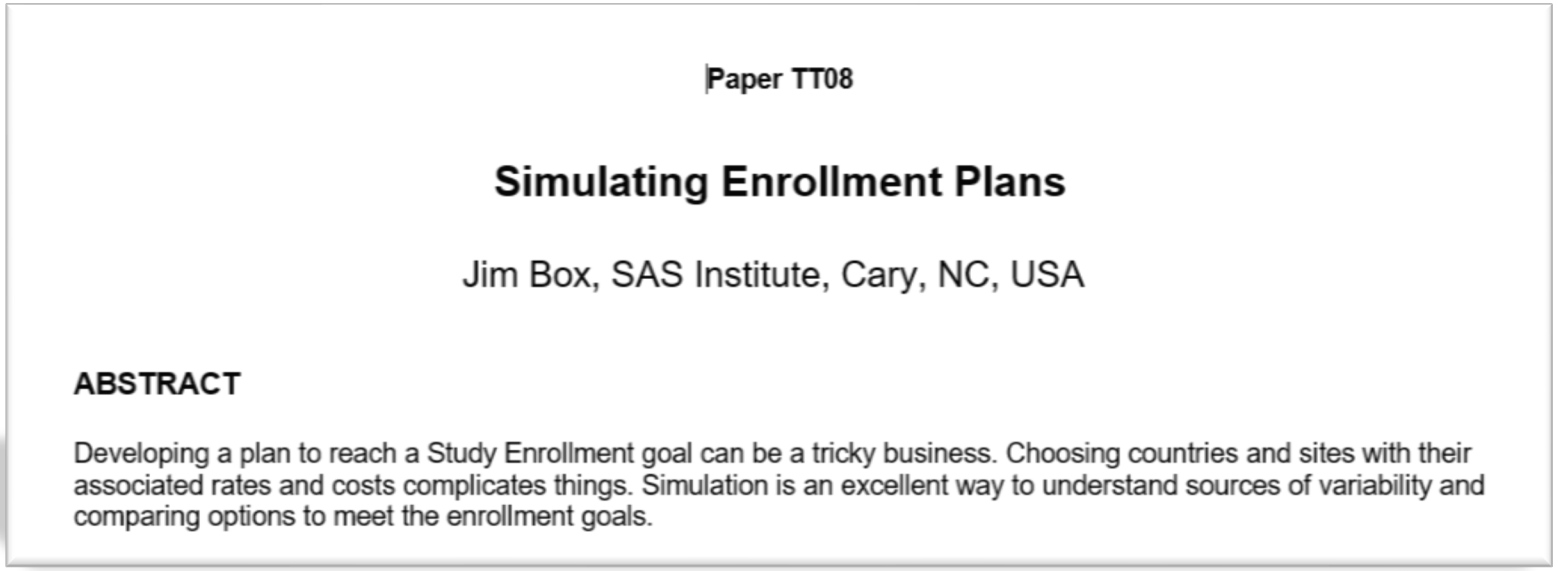
36.86% If Age at Index Event is greater than or equal to 74, then MI or Cardiac Cath has a 36.86% chance (2.1K out of 5.8K cases) of being 1.

Scope



Enrollment Optimization

Phuse EU connect 2018



Paper TT08

Simulating Enrollment Plans

Jim Box, SAS Institute, Cary, NC, USA

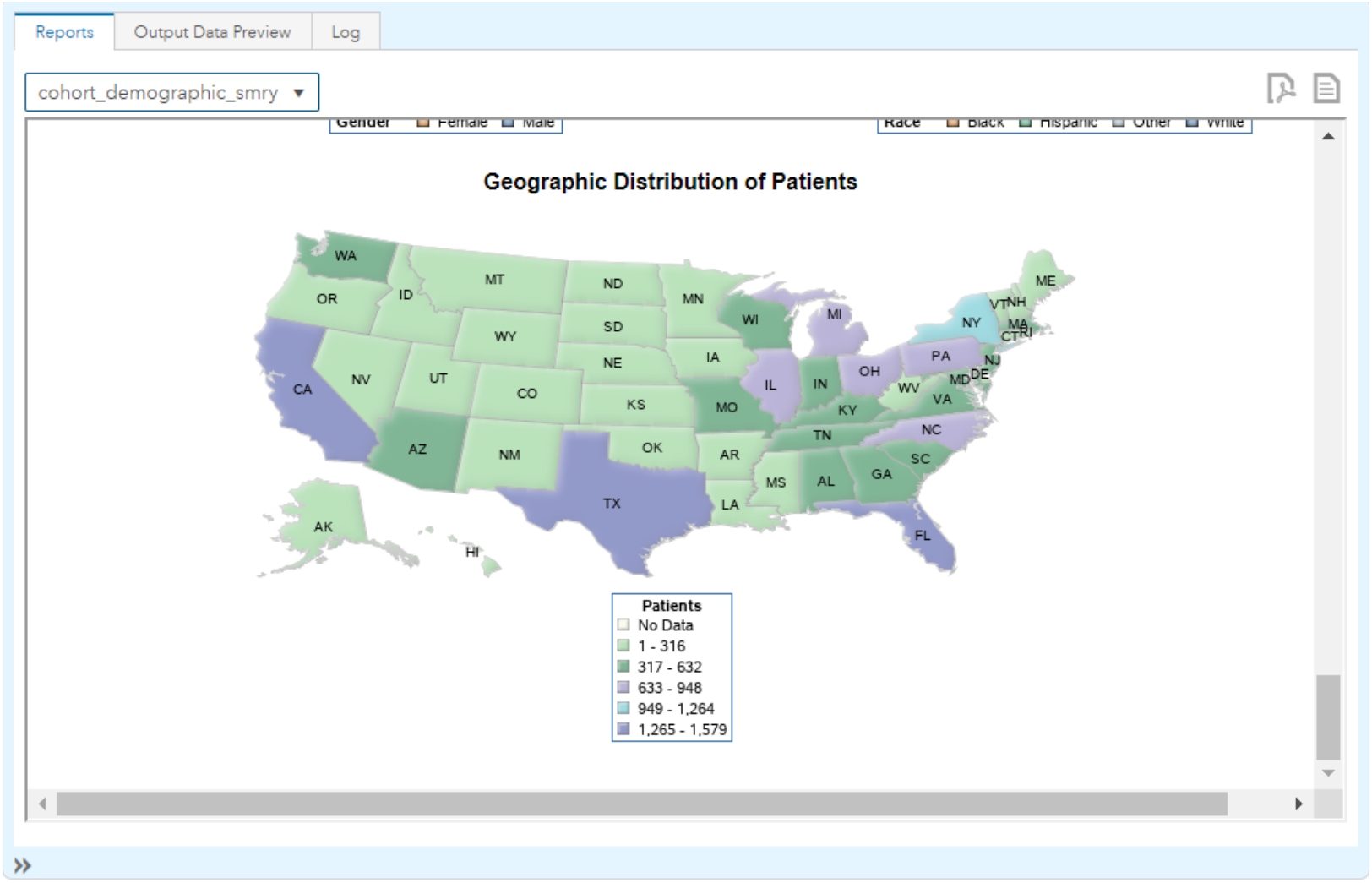
ABSTRACT

Developing a plan to reach a Study Enrollment goal can be a tricky business. Choosing countries and sites with their associated rates and costs complicates things. Simulation is an excellent way to understand sources of variability and comparing options to meet the enrollment goals.

https://www.lexjansen.com/phuse/2018/tt/TT08.pdf

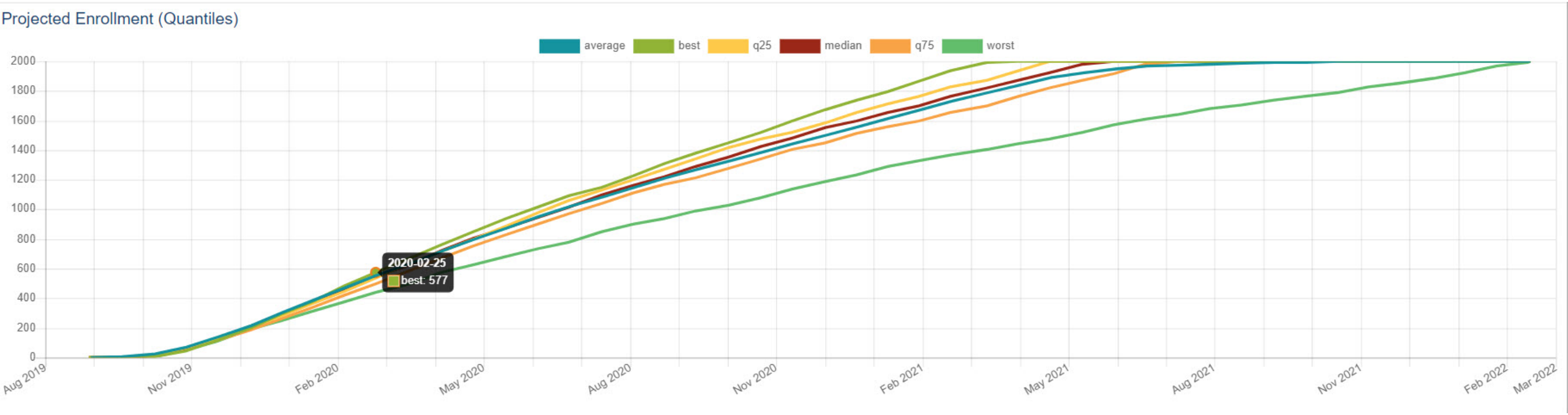
Enrollment Optimization

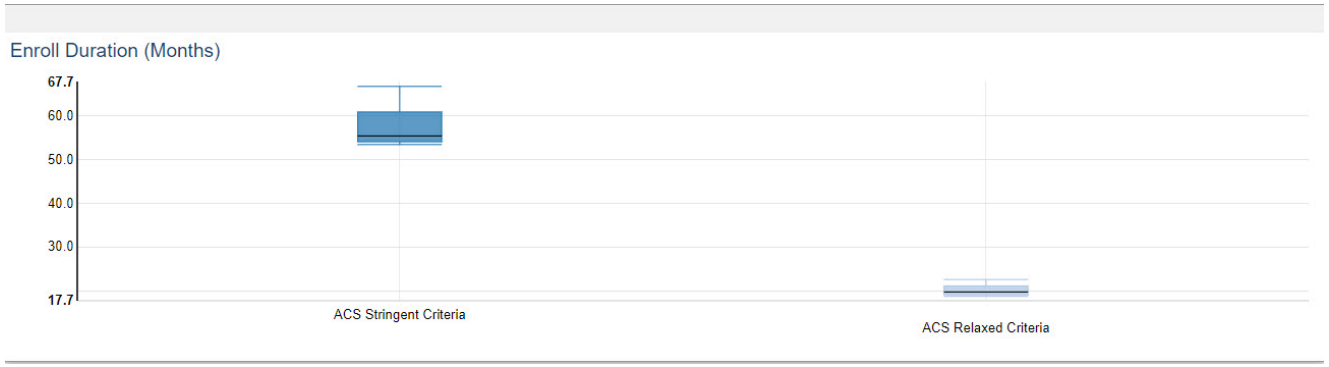
State/Site Selection



Enrollment Optimization

Relaxed Criteria





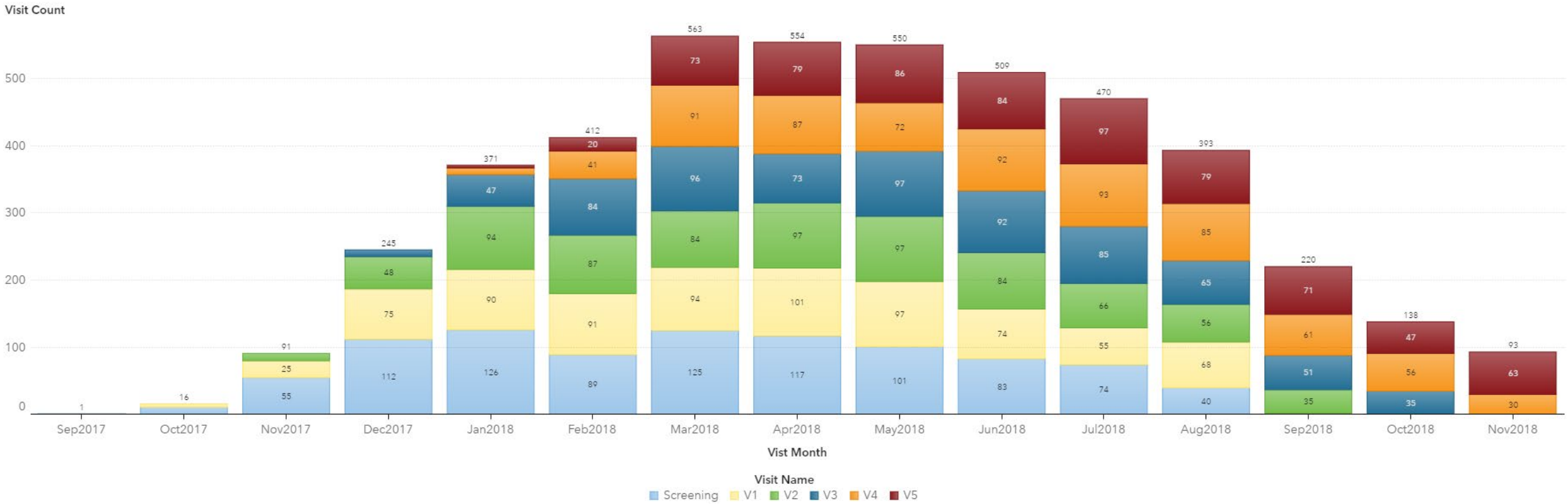
Enrollment Optimization

Monthly Visits Projection

Monthly Visit Count | Milestone Chart | Key Dates | Duration Distribution | Individual Enrollment Graphs | Site Initiation | Monthly Site Cost Forecasting | Page 8

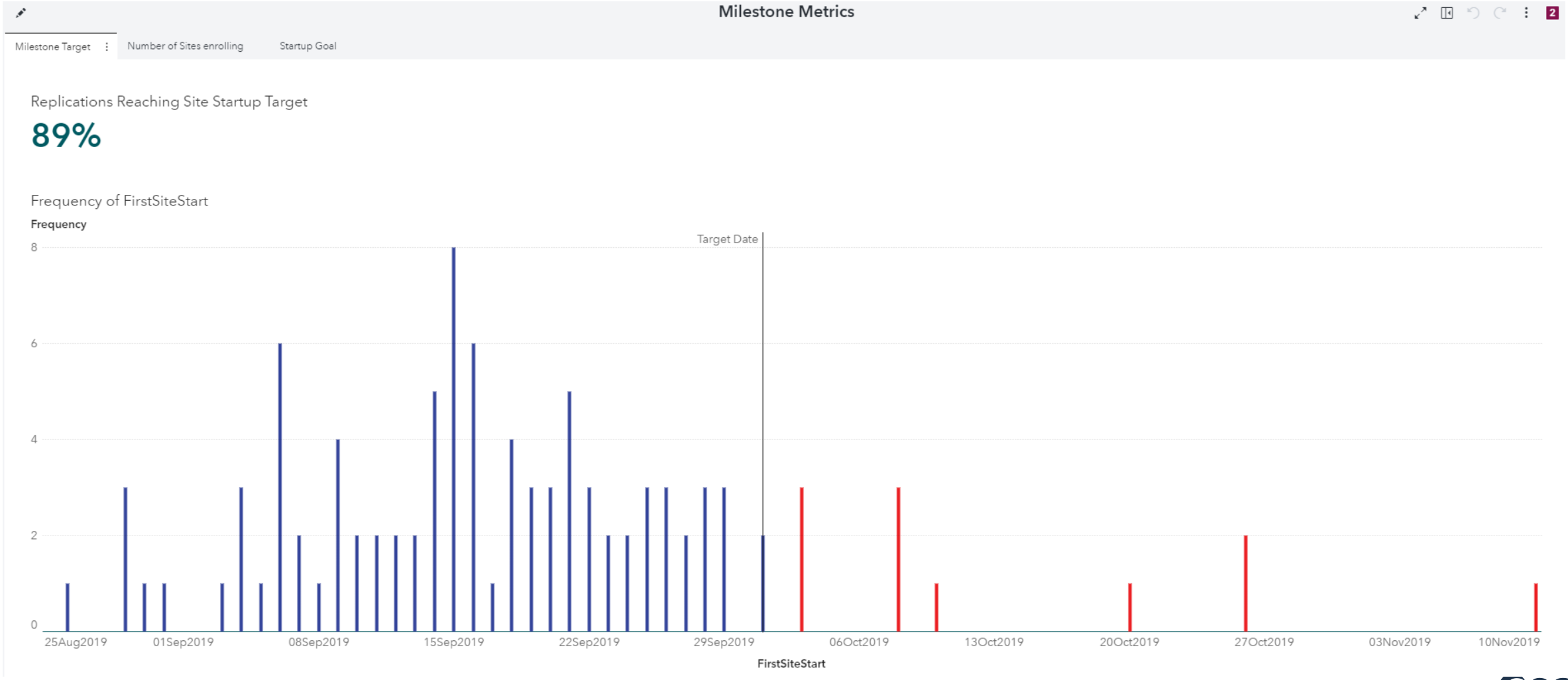
Derry Medical Center	Heisenberg Medical Center	Lassen General Hospital	Princeton - Plainsboro Teaching Hospital	Sacred Heart Hospital	Shasta Regional Hospital	St. Eligus Hospital	Tower Medical Group	Twin Pines Medical Center	Western Regional Hospital

Screening	V1	V2	V3	V4	V5



Enrollment Optimization

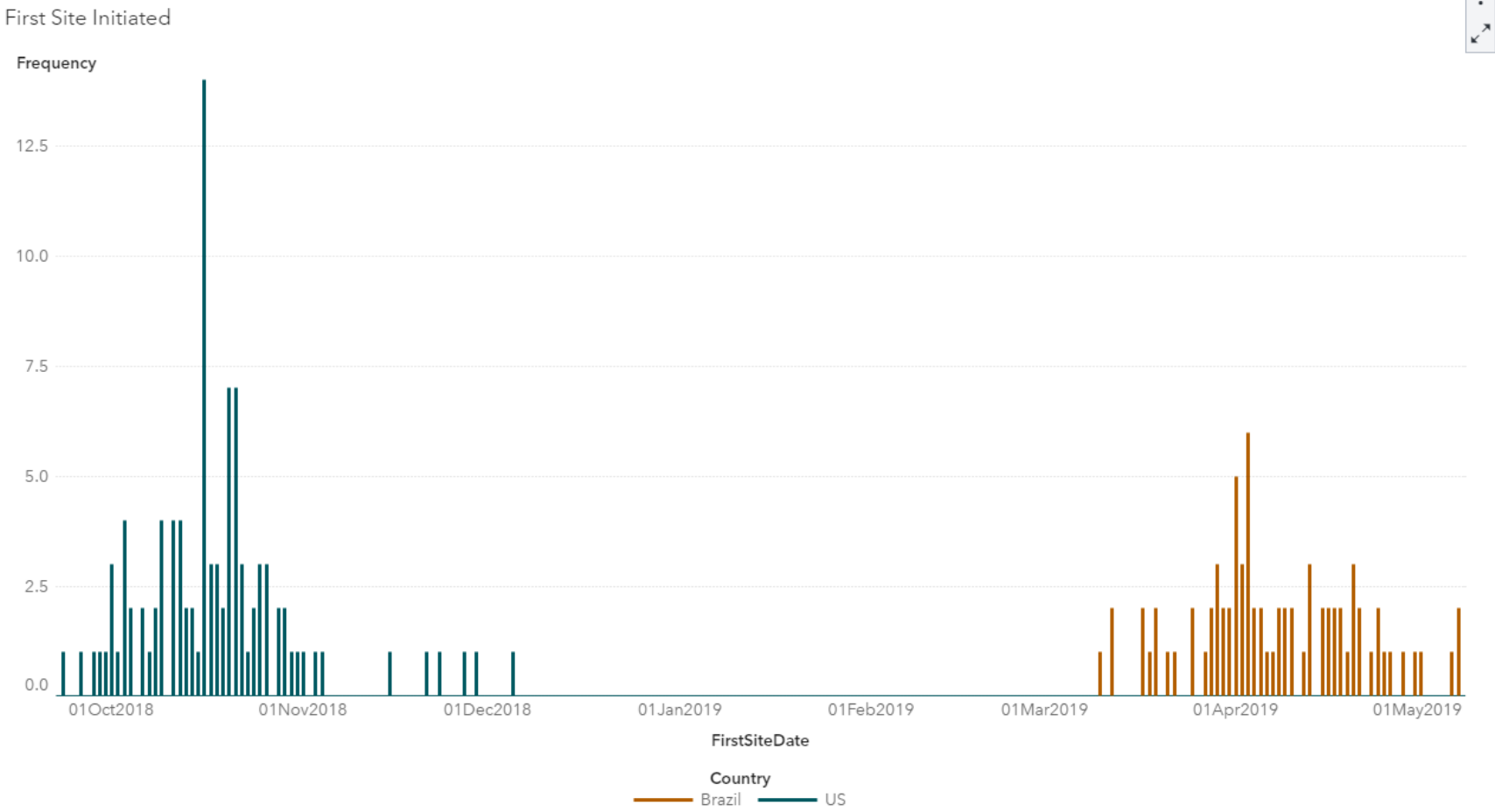
Milestone Metrics

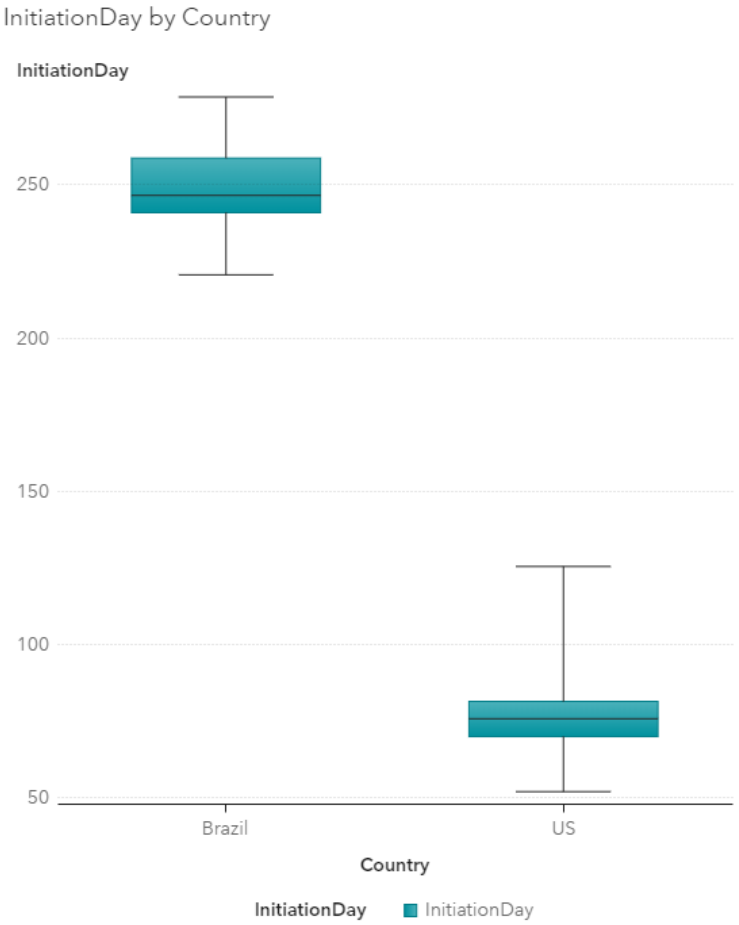


Enrollment Optimization

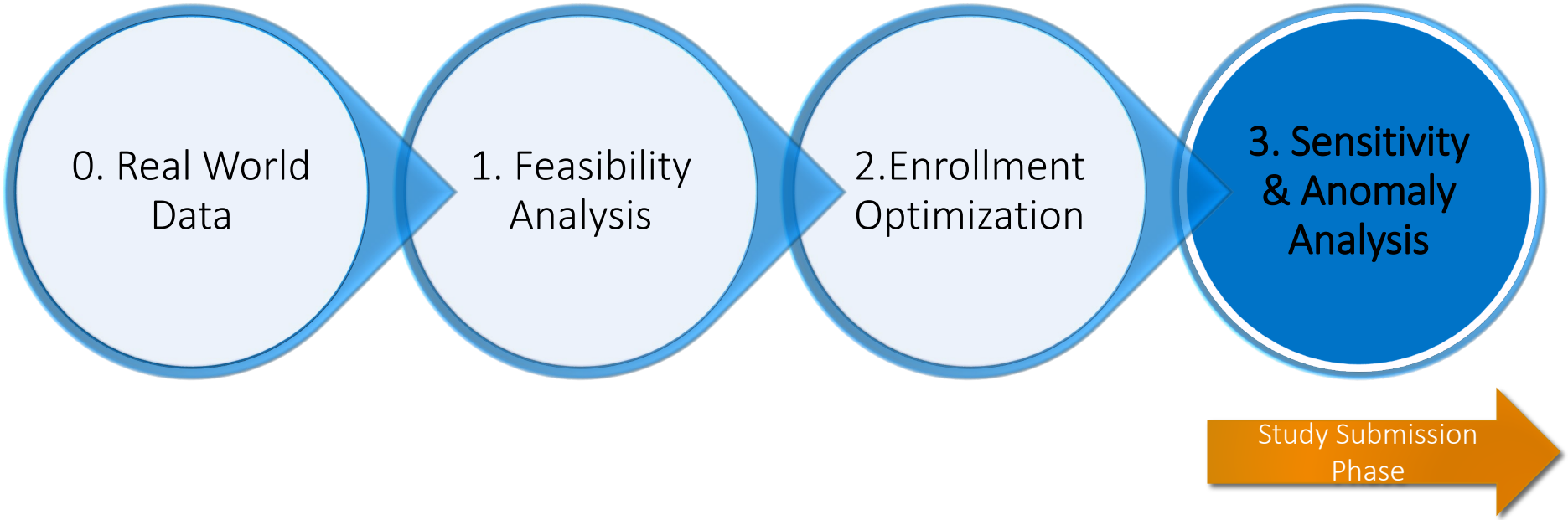
What if scenario's …

Monthly Visit Count | Milestone Chart | Key Dates | Duration Distribution | Individual Enrollment Graphs | Site Initiation | Monthly Site Cost Forecasting | Page 8



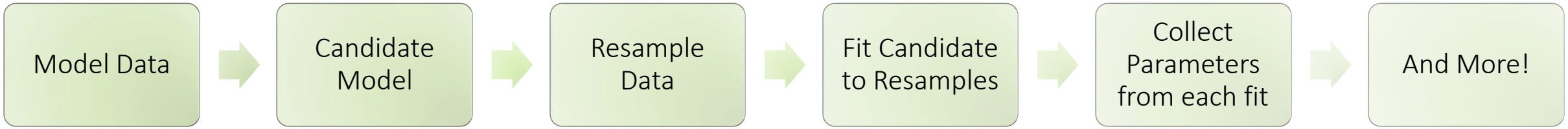


Scope



SAS Viya: Using CAS for faster inference

Review a workflow

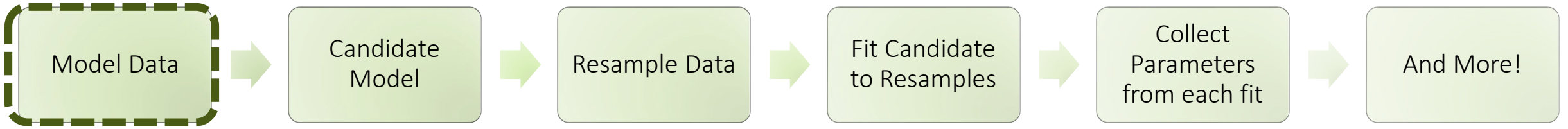


CAS = Cloud Analytics Services

SAS Viya: Using CAS for faster inference

- Framingham Heart Study – Original Cohort
  - 1948 Study commissioned by Congress (USA)
  - 5209 women and men
  - Age at entry from 30 up to 62
  - No history of heart attack or stroke
  - Some participants were doctors and nurses. They signed up to set an example for patients to participate
  - Variables Considered: 1 row per patient
    - Status = Dead, Alive
    - Measurements: Sex, AgeAtStart, Height, Weight, Diastolic, Systolic, MRW, Smoking, AgeAtDeath, Cholesterol
    - Graded variables: Chol_Status, BP_Status, Weight_Status, Smoking_Status

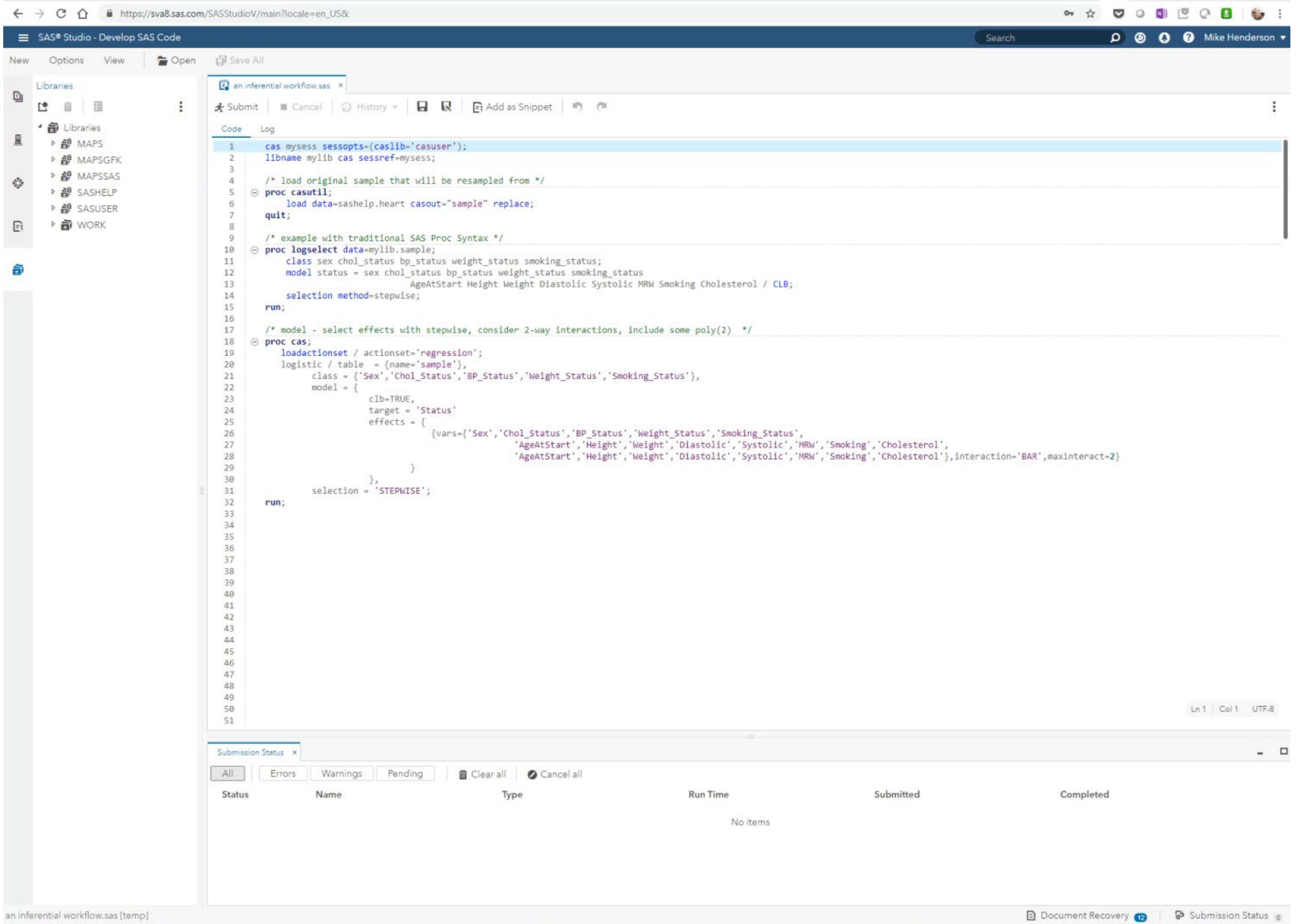
SAS Viya: Using CAS for faster inference



- SAS PROC syntax compared to CASL Syntax
  - SAS Proc
    - From SAS Interface
    - Familiar to SAS Users
  - Viya CASL syntax
    - From SAS Interface with PROC CAS;
    - From other languages with SWAT library: Python, Java, LUA, R, REST
    - More flexible model specification!

```
/* example with traditional SAS Proc Syntax */
proc logselect data=mylib.sample;
    class sex chol_status bp_status weight_status smoking_status;
    model status = sex chol_status bp_status weight_status smoking_status
                            AgeAtStart Height Weight Diastolic Systolic MRW Smoking Cholesterol / CLB;
    selection method=stepwise;
run;
```

```
/* model - select effects with stepwise, consider 2-way interactions, include some poly(2)  */
proc cas;
    loadactionset / actionset='regression';
    logistic / table  = {name='sample'},
          class = {'Sex','Chol_Status','BP_Status','Weight_Status','Smoking_Status'},
          model = {
                    clb=TRUE,
                    target = 'Status'
                    effects = {
                              {vars={'Sex','Chol_Status','BP_Status','Weight_Status','Smoking_Status',
                                      'AgeAtStart','Height','Weight','Diastolic','Systolic','MRW','Smoking','Cholesterol',
                                      'AgeAtStart','Height','Weight','Diastolic','Systolic','MRW','Smoking','Cholesterol'},interaction='BAR',maxinteract=2}
                    }
                    },
          selection = 'STEPWISE';
run;
```


https://sva8.sas.com/SASStudioV/main?locale=en_US&
SAS® Studio - Develop SAS Code
Search
Mike Henderson
New
Options
View
Open
Save All
Libraries
Libraries
MAPS
MAPSGFK
MAPSSAS
SASHELP
SASUSER
WORK
an inferential workflow.sas
Submit
Cancel
History
Add as Snippet
Code
Log
1 cas mysess sessopts=(caslib='casuser');
2 libname mylib cas sessref=mysess;
3
4 /* load original sample that will be resampled from */
5 proc casutil;
6 load data=sashelp.heart casout="sample" replace;
7 quit;
8
9 /* example with traditional SAS Proc Syntax */
10 proc logselect data=mylib.sample;
11 class sex chol_status bp_status weight_status smoking_status;
12 model status = sex chol_status bp_status weight_status smoking_status
13 AgeAtStart Height Weight Diastolic Systolic MRW Smoking Cholesterol / CLB;
14 selection method=stepwise;
15 run;
16
17 /* model - select effects with stepwise, consider 2-way interactions, include some poly(2) */
18 proc cas;
19 loadactionset / actionset='regression';
20 logistic / table = {name='sample'},
21 class = {'Sex','Chol_Status','BP_Status','Weight_Status','Smoking_Status'},
22 model = {
23 clb=TRUE,
24 target = 'Status'
25 effects = {
26 {vars={'Sex','Chol_Status','BP_Status','Weight_Status','Smoking_Status',
27 'AgeAtStart','Height','Weight','Diastolic','Systolic','MRW','Smoking','Cholesterol',
28 'AgeAtStart','Height','Weight','Diastolic','Systolic','MRW','Smoking','Cholesterol'},interaction='BAR',maxinteract=2}
29 }
30 },
31 selection = 'STEPWISE';
32 run;
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
Ln 1
Col 1
UTF-8
Submission Status
All
Errors
Warnings
Pending
Clear all
Cancel all
Status
Name
Type
Run Time
Submitted
Completed
No items
an inferential workflow.sas [temp]
Document Recovery
12
Submission Status
0

SAS Viya: Using CAS for faster inference

Model Data → **Candidate Model** → Resample Data → Fit Candidate to Resamples → Collect Parameters → And More!

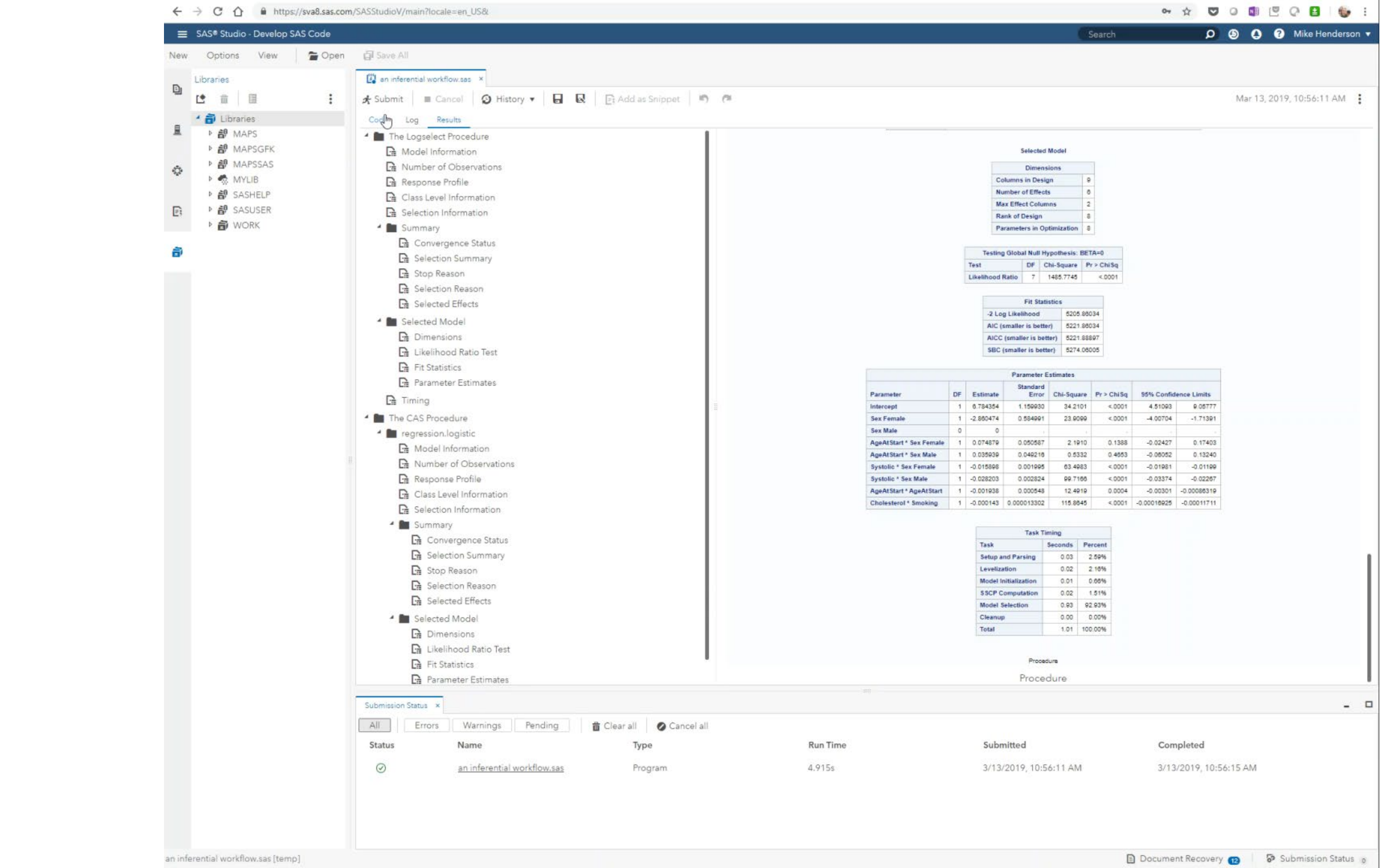
```
55   /* spec out final selected model above with lower level effects included for inf
56   proc cas;
57      logistic / table  = {name='sample'},
58            class = {'Sex','Smoking_Status'},
59            model = {
60                        clb=TRUE,
61                        target = 'Status'
62                        effects = {
63                                    {vars={'Sex','Smoking_Status','AgeAtStart','Syst
64                                    {vars={'AgeAtStart','Sex'}, interaction='CROSS'}
65                                    {vars={'Systolic','Sex'}, interaction='CROSS'},
66                                    {vars={'AgeAtStart','AgeAtStart'}, interaction='
67                                    {vars={'Cholesterol','Smoking_Status'}, interact
68                                 }
69                        },
70            display = {names={'ParameterEstimates','FitStatistics'}},
71            outputTables = {names={'ParameterEstimates'="sample_PE"}, replace=TRUE}
72            output = {casOut={name='sample_pred', replace=TRUE}, pred='Pred', resid
73   run;
```

Results from regression.logistic

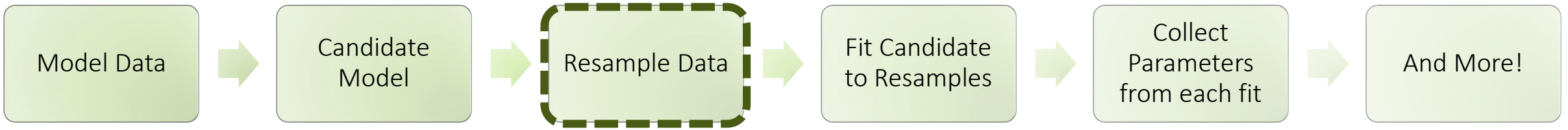
Fit Statistics	
-2 Log Likelihood	5209.08335
AIC (smaller is better)	5241.08335
AICC (smaller is better)	5241.19146
SBC (smaller is better)	5345.51448

Parameter Estimates

Parameter	DF	Estimate	Standard Error	Chi-Square	Pr > ChiSq	95% Confidence Limits	
Intercept	1	7.266858	1.309052	30.8162	<.0001	4.70116	9.83255
Sex Female	1	-2.743188	0.591510	21.5074	<.0001	-3.90253	-1.58385
Sex Male	0	0	.	.	.	.	.
Smoking_Status Heavy (16-25)	1	-0.670704	0.725376	0.8549	0.3552	-2.09241	0.75101
Smoking_Status Light (1-5)	1	-0.460460	0.837783	0.3021	0.5826	-2.10248	1.18156
Smoking_Status Moderate (6-15)	1	-0.701218	0.814542	0.7411	0.3893	-2.29769	0.89526
Smoking_Status Non-smoker	1	-0.264849	0.672919	0.1549	0.6939	-1.58375	1.05405
Smoking_Status Very Heavy (> 25)	0	0	.	.	.	.	.
AgeAtStart	1	0.040701	0.049405	0.6787	0.4100	-0.05613	0.13753
Systolic	1	-0.028239	0.002836	99.1783	<.0001	-0.03380	-0.02268
Cholesterol	1	-0.007074	0.002668	7.0314	0.0080	-0.01230	-0.00185
AgeAtStart * Sex Female	1	0.037590	0.009417	15.9334	<.0001	0.01913	0.05605
AgeAtStart * Sex Male	0	0	.	.	.	.	.
Systolic * Sex Female	1	0.012113	0.003466	12.2136	0.0005	0.00532	0.01891
Systolic * Sex Male	0	0	.	.	.	.	.
AgeAtStart * AgeAtStart	1	-0.001984	0.000551	12.9907	0.0003	-0.00306	-0.00090533
Cholesterol * Smoking_Status Heavy (16-25)	1	0.004118	0.003116	1.7473	0.1862	-0.00199	0.01022
Cholesterol * Smoking_Status Light (1-5)	1	0.005563	0.003592	2.3981	0.1215	-0.00148	0.01260
Cholesterol * Smoking_Status Moderate (6-15)	1	0.004635	0.003511	1.7425	0.1868	-0.00225	0.01152
Cholesterol * Smoking_Status Non-smoker	1	0.005818	0.002893	4.0452	0.0443	0.00014838	0.01149
Cholesterol * Smoking_Status Very Heavy (> 25)	0	0	.	.	.	.	.

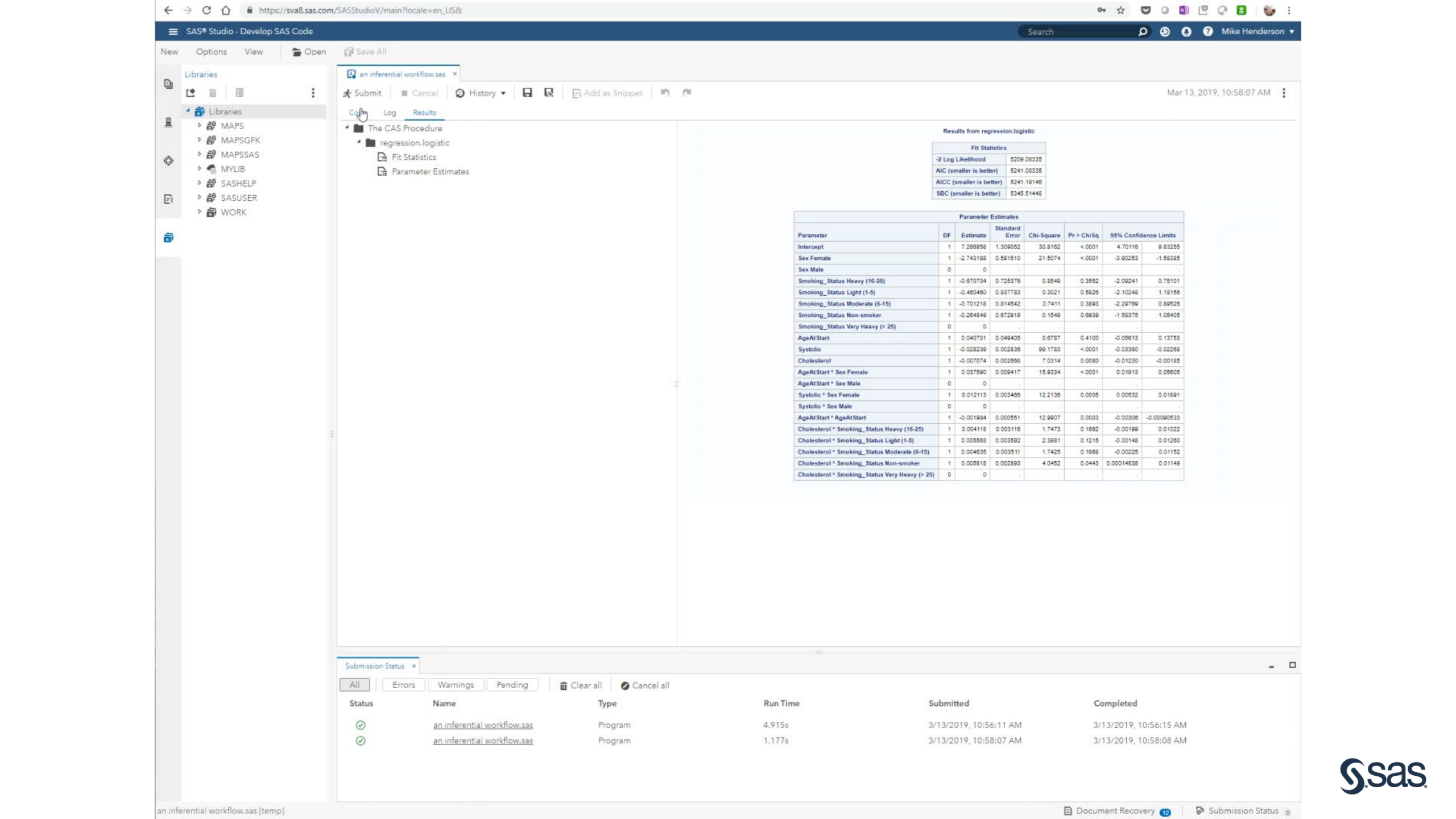

https://sva8.sas.com/SASStudioV/main?locale=en_US&
SAS® Studio - Develop SAS Code
Search
Mike Henderson
New
Options
View
Open
Save All
Libraries
MAPS
MAPSGFK
MAPSSAS
MYLIB
SASHELP
SASUSER
WORK
an inferential workflow.sas
Submit
Cancel
History
Add as Snippet
Mar 13, 2019, 10:56:11 AM
Code
Log
Results
The Logselect Procedure
Model Information
Number of Observations
Response Profile
Class Level Information
Selection Information
Summary
Convergence Status
Selection Summary
Stop Reason
Selection Reason
Selected Effects
Selected Model
Dimensions
Likelihood Ratio Test
Fit Statistics
Parameter Estimates
Timing
The CAS Procedure
regression.logistic
Model Information
Number of Observations
Response Profile
Class Level Information
Selection Information
Summary
Convergence Status
Selection Summary
Stop Reason
Selection Reason
Selected Effects
Selected Model
Dimensions
Likelihood Ratio Test
Fit Statistics
Parameter Estimates
Selected Model
Dimensions
Columns in Design 9
Number of Effects 6
Max Effect Columns 2
Rank of Design 8
Parameters in Optimization 8
Testing Global Null Hypothesis: BETA=0
Test DF Chi-Square Pr > ChiSq
Likelihood Ratio 7 1485.7745 <.0001
Fit Statistics
-2 Log Likelihood 5205.86034
AIC (smaller is better) 5221.86034
AICC (smaller is better) 5221.88897
SBC (smaller is better) 5274.06005
Parameter Estimates
Parameter DF Estimate Standard Error Chi-Square Pr > ChiSq 95% Confidence Limits
Intercept 1 6.784354 1.159930 34.2101 <.0001 4.51093 9.05777
Sex Female 1 -2.860474 0.584991 23.9099 <.0001 -4.00704 -1.71391
Sex Male 0 0
AgeAtStart * Sex Female 1 0.074879 0.050587 2.1910 0.1388 -0.02427 0.17403
AgeAtStart * Sex Male 1 0.035939 0.049216 0.5332 0.4653 -0.06052 0.13240
Systolic * Sex Female 1 -0.015896 0.001995 63.4983 <.0001 -0.01981 -0.01199
Systolic * Sex Male 1 -0.028203 0.002824 99.7166 <.0001 -0.03374 -0.02267
AgeAtStart * AgeAtStart 1 -0.001938 0.000548 12.4919 0.0004 -0.00301 -0.00086319
Cholesterol * Smoking 1 -0.000143 0.000013302 115.8645 <.0001 -0.00016925 -0.00011711
Task Timing
Task Seconds Percent
Setup and Parsing 0.03 2.59%
Levelization 0.02 2.16%
Model Initialization 0.01 0.86%
SSCP Computation 0.02 1.51%
Model Selection 0.93 92.93%
Cleanup 0.00 0.00%
Total 1.01 100.00%
Procedure
Procedure
Submission Status
All
Errors
Warnings
Pending
Clear all
Cancel all
Status
Name
Type
Run Time
Submitted
Completed
an inferential workflow.sas
Program
4.915s
3/13/2019, 10:56:11 AM
3/13/2019, 10:56:15 AM
an inferential workflow.sas [temp]
Document Recovery
Submission Status

SAS Viya: Using CAS for faster inference

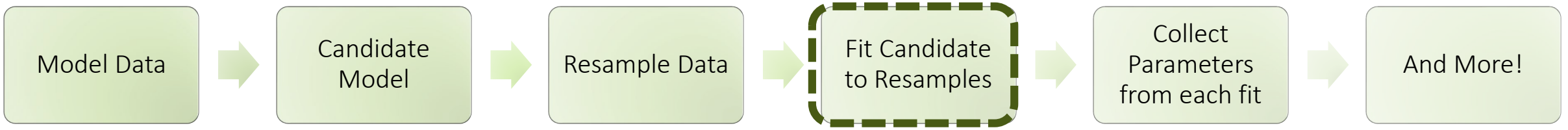



```
96   /* create bootstrap resamples */
97   proc cas;
98       builtins.actionSetFromTable / table={caslib="Public" name="resampleActionSet.sashdat"} name="resample";
99       resample.bootstrap / intable='sample' B=1000 seed=12345 Bpct=1 case='unique_case';
100  run;
```

- Easily Create Bootstrap Resamples:
  - This is an example of a user contributed action
    - https://github.com/statmike/Resampling-Methods-in-SAS-Viya
    - Contains actions for many common resampling techniques
  - With a single line of syntax you can create the desired number of resamples (B) in the size you want relative to the original sample (Bpct), specify a seed for repeatability of sampling, and even specify a variable that links rows into cases if needed.
  - The resamples are automatically evenly generated and distributed across all the computing cores of all the machines in the environment.

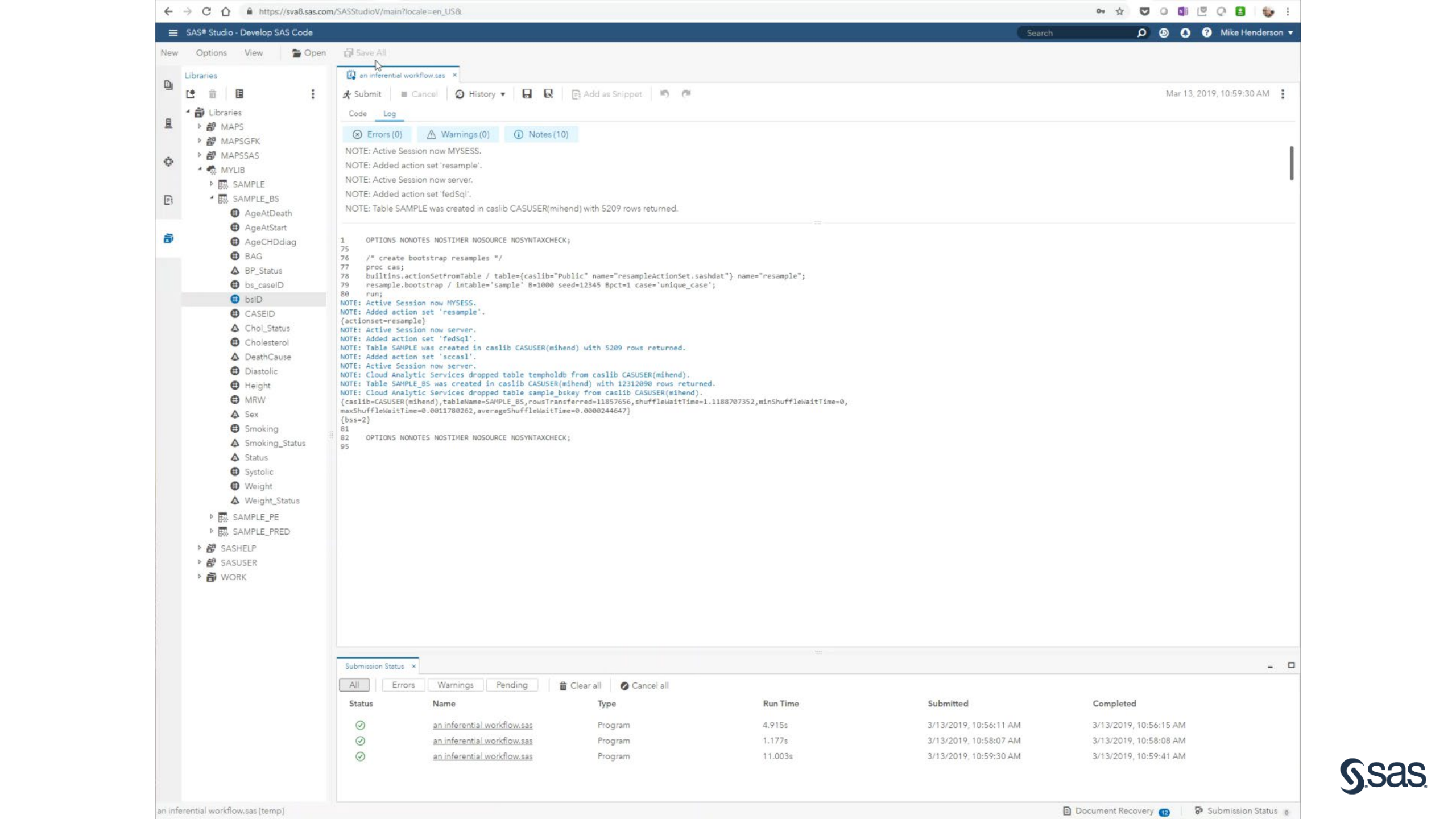

https://sva8.sas.com/SASStudioV/main?locale=en_US&
SAS® Studio - Develop SAS Code
Search
Mike Henderson
New
Options
View
Open
Save All
Libraries
Libraries
MAPS
MAPSGFK
MAPSSAS
MYLIB
SASHELP
SASUSER
WORK
an inferential workflow.sas
Submit
Cancel
History
Add as Snippet
Mar 13, 2019, 10:58:07 AM
Code
Log
Results
The CAS Procedure
regression.logistic
Fit Statistics
Parameter Estimates
Results from regression.logistic
Fit Statistics
-2 Log Likelihood 5209.08335
AIC (smaller is better) 5241.08335
AICC (smaller is better) 5241.19146
SBC (smaller is better) 5345.51446
Parameter Estimates
Parameter DF Estimate Standard Error Chi-Square Pr > ChiSq 95% Confidence Limits
Intercept 1 7.266858 1.309052 30.8162 <.0001 4.70116 9.83255
Sex Female 1 -2.743188 0.591510 21.5074 <.0001 -3.90253 -1.58385
Sex Male 0 0 . . . . .
Smoking_Status Heavy (16-25) 1 -0.670704 0.725376 0.8549 0.3552 -2.09241 0.75101
Smoking_Status Light (1-5) 1 -0.460460 0.837783 0.3021 0.5826 -2.10248 1.18156
Smoking_Status Moderate (6-15) 1 -0.701218 0.814542 0.7411 0.3893 -2.29759 0.89526
Smoking_Status Non-smoker 1 -0.264849 0.672919 0.1549 0.6939 -1.58375 1.05405
Smoking_Status Very Heavy (> 25) 0 0 . . . . .
AgeAtStart 1 0.040701 0.049405 0.6787 0.4100 -0.05613 0.13753
Systolic 1 -0.028239 0.002836 99.1763 <.0001 -0.03380 -0.02268
Cholesterol 1 -0.007074 0.002668 7.0314 0.0080 -0.01230 -0.00185
AgeAtStart * Sex Female 1 0.037590 0.009417 15.9334 <.0001 0.01913 0.05605
AgeAtStart * Sex Male 0 0 . . . . .
Systolic * Sex Female 1 0.012113 0.003466 12.2136 0.0005 0.00532 0.01891
Systolic * Sex Male 0 0 . . . . .
AgeAtStart * AgeAtStart 1 -0.001984 0.000551 12.9907 0.0003 -0.00306 -0.00090533
Cholesterol * Smoking_Status Heavy (16-25) 1 0.004118 0.003116 1.7473 0.1862 -0.00199 0.01022
Cholesterol * Smoking_Status Light (1-5) 1 0.005563 0.003592 2.3981 0.1215 -0.00148 0.01260
Cholesterol * Smoking_Status Moderate (6-15) 1 0.004535 0.003511 1.7425 0.1868 -0.00225 0.01152
Cholesterol * Smoking_Status Non-smoker 1 0.005818 0.002893 4.0452 0.0443 0.00014838 0.01149
Cholesterol * Smoking_Status Very Heavy (> 25) 0 0 . . . . .
Submission Status
All
Errors
Warnings
Pending
Clear all
Cancel all
Status
Name
Type
Run Time
Submitted
Completed
an inferential workflow.sas
Program
4.915s
3/13/2019, 10:56:11 AM
3/13/2019, 10:56:15 AM
an inferential workflow.sas
Program
1.177s
3/13/2019, 10:58:07 AM
3/13/2019, 10:58:08 AM
an inferential workflow.sas [temp]
Document Recovery
Submission Status

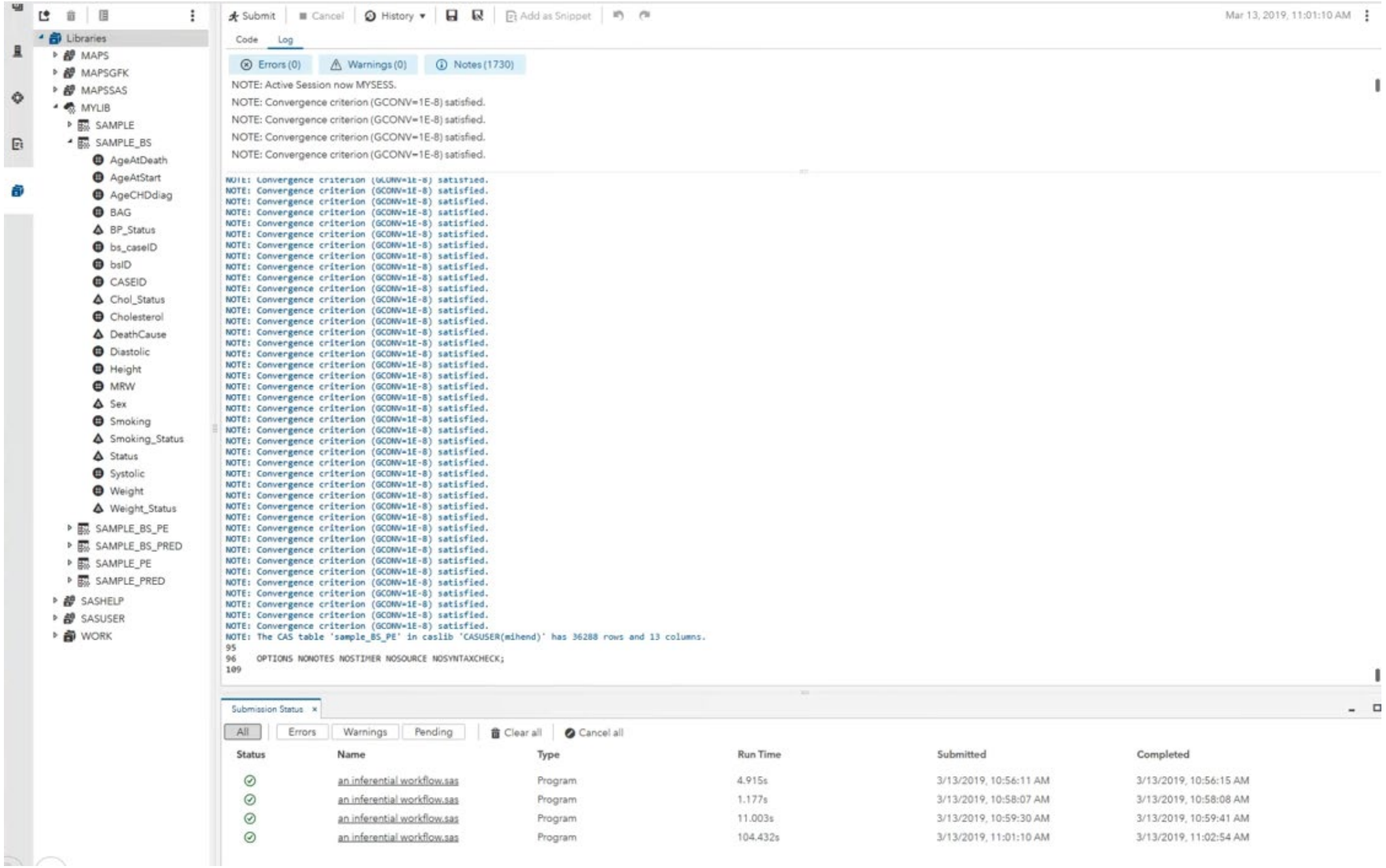
SAS Viya: Using CAS for faster inference



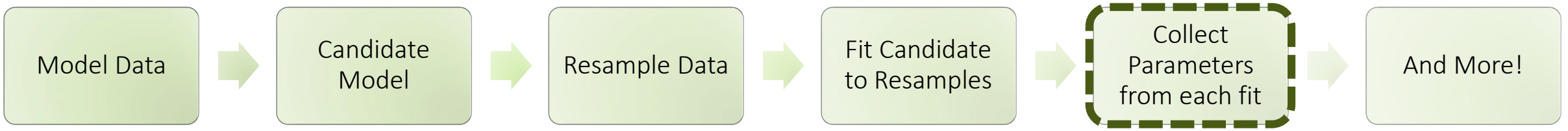

```
/* analyze/train each bootstrap resample with the same model effects selected on the full sample data */
proc cas;
   logistic result=myresult / table  = {name='sample_bs', groupBy='bsID'},
         class = {'Sex','Smoking_Status'},
         model = {
                  clb=TRUE,
                  target = 'Status'
                  effects = {
                           {vars={'Sex','Smoking_Status','AgeAtStart','Systolic','Cholesterol'}},
                           {vars={'AgeAtStart','Sex'}, interaction='CROSS'},
                           {vars={'Systolic','Sex'}, interaction='CROSS'},
                           {vars={'AgeAtStart','AgeAtStart'}, interaction='CROSS'},
                           {vars={'Cholesterol','Smoking_Status'}, interaction='CROSS'}
                        }
                  },
         partByVar = {name="bag",train="1",test="0"},
         outputTables = {names={'ParameterEstimates'="sample_BS_PE"}, groupByVarsRaw=TRUE, replace=TRUE},
         output = {casOut={name='sample_bs_pred', replace=TRUE}, pred='Pred', resid='Resid', copyVars={"caseID","bsID","bag"}};
run;
```

- Specify groupBy in order to fit the same model to each of the B bootstrap resamples
  - Automatically distributes the computation to all the computing cores of all machines in the environment to complete the task in parallel rather than sequential


https://sva8.sas.com/SASStudioV/main?locale=en_US&
SAS® Studio - Develop SAS Code
Search
Mike Henderson
New
Options
View
Open
Save All
Libraries
an inferential workflow.sas
Submit
Cancel
History
Add as Snippet
Mar 13, 2019, 10:59:30 AM
Code
Log
Errors (0)
Warnings (0)
Notes (10)
NOTE: Active Session now MYSESS.
NOTE: Added action set 'resample'.
NOTE: Active Session now server.
NOTE: Added action set 'fedSql'.
NOTE: Table SAMPLE was created in caslib CASUSER(mihend) with 5209 rows returned.
1 OPTIONS NONOTES NOSTIMER NOSOURCE NOSYNTAXCHECK;
75
76 /* create bootstrap resamples */
77 proc cas;
78 builtins.actionSetFromTable / table={caslib="Public" name="resampleActionSet.sashdat"} name="resample";
79 resample.bootstrap / intable='sample' B=1000 seed=12345 Bpct=1 case='unique_case';
80 run;
NOTE: Active Session now MYSESS.
NOTE: Added action set 'resample'.
{actionset=resample}
NOTE: Active Session now server.
NOTE: Added action set 'fedSql'.
NOTE: Table SAMPLE was created in caslib CASUSER(mihend) with 5209 rows returned.
NOTE: Added action set 'sccasl'.
NOTE: Active Session now server.
NOTE: Cloud Analytic Services dropped table tempholdb from caslib CASUSER(mihend).
NOTE: Table SAMPLE_BS was created in caslib CASUSER(mihend) with 12312090 rows returned.
NOTE: Cloud Analytic Services dropped table sample_bskey from caslib CASUSER(mihend).
{caslib=CASUSER(mihend),tableName=SAMPLE_BS,rowsTransferred=11857656,shuffleWaitTime=1.1188707352,minShuffleWaitTime=0,
maxShuffleWaitTime=0.0011780262,averageShuffleWaitTime=0.0000244647}
{bss=2}
81
82 OPTIONS NONOTES NOSTIMER NOSOURCE NOSYNTAXCHECK;
95
Libraries
MAPS
MAPSGFK
MAPSSAS
MYLIB
SAMPLE
SAMPLE_BS
AgeAtDeath
AgeAtStart
AgeCHDdiag
BAG
BP_Status
bs_caseID
bsID
CASEID
Chol_Status
Cholesterol
DeathCause
Diastolic
Height
MRW
Sex
Smoking
Smoking_Status
Status
Systolic
Weight
Weight_Status
SAMPLE_PE
SAMPLE_PRED
SASHELP
SASUSER
WORK
Submission Status
All
Errors
Warnings
Pending
Clear all
Cancel all
Status
Name
Type
Run Time
Submitted
Completed
an inferential workflow.sas
Program
4.915s
3/13/2019, 10:56:11 AM
3/13/2019, 10:56:15 AM
an inferential workflow.sas
Program
1.177s
3/13/2019, 10:58:07 AM
3/13/2019, 10:58:08 AM
an inferential workflow.sas
Program
11.003s
3/13/2019, 10:59:30 AM
3/13/2019, 10:59:41 AM
an inferential workflow.sas [temp]
Document Recovery
Submission Status

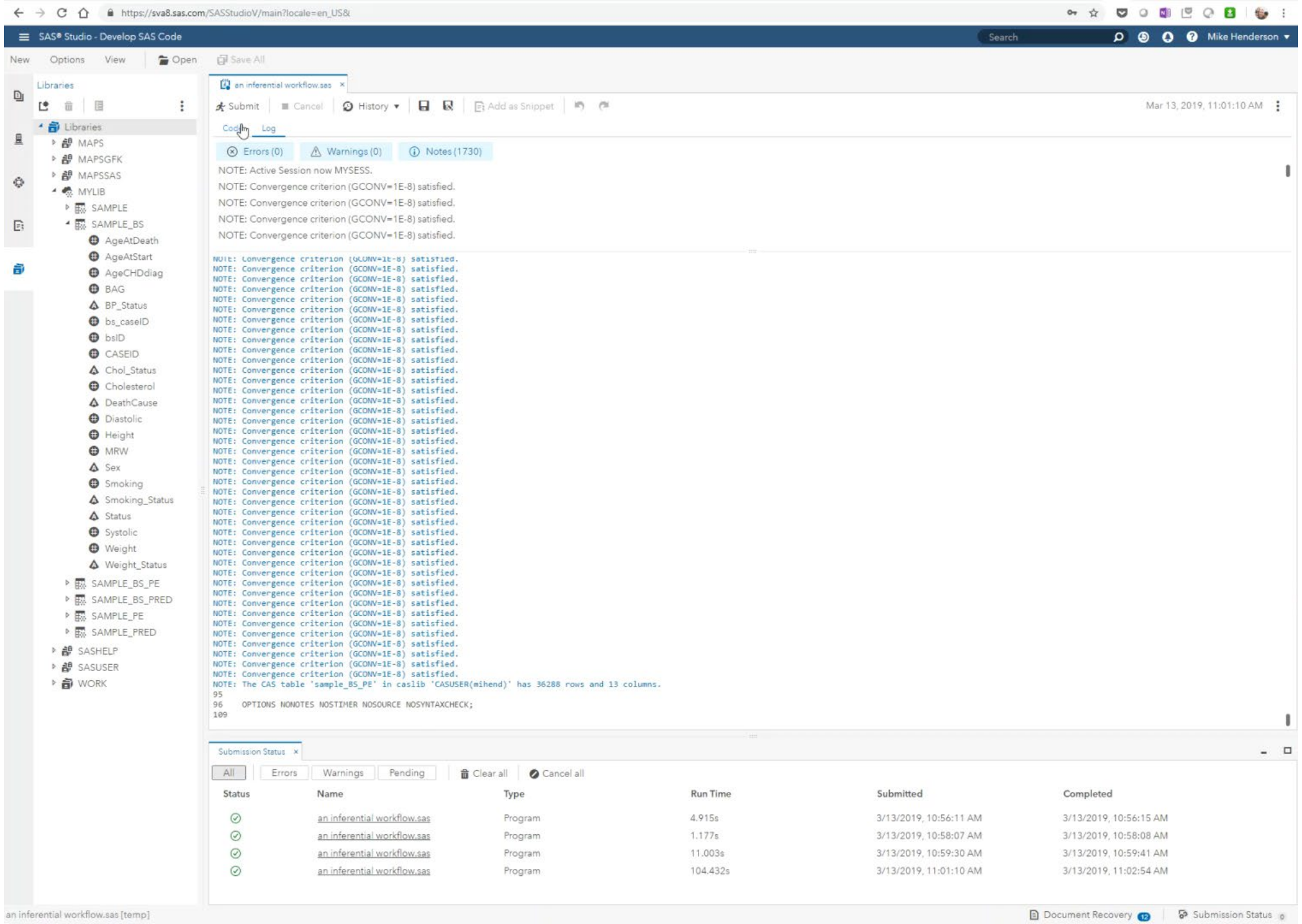


SAS Viya: Using CAS for faster inference



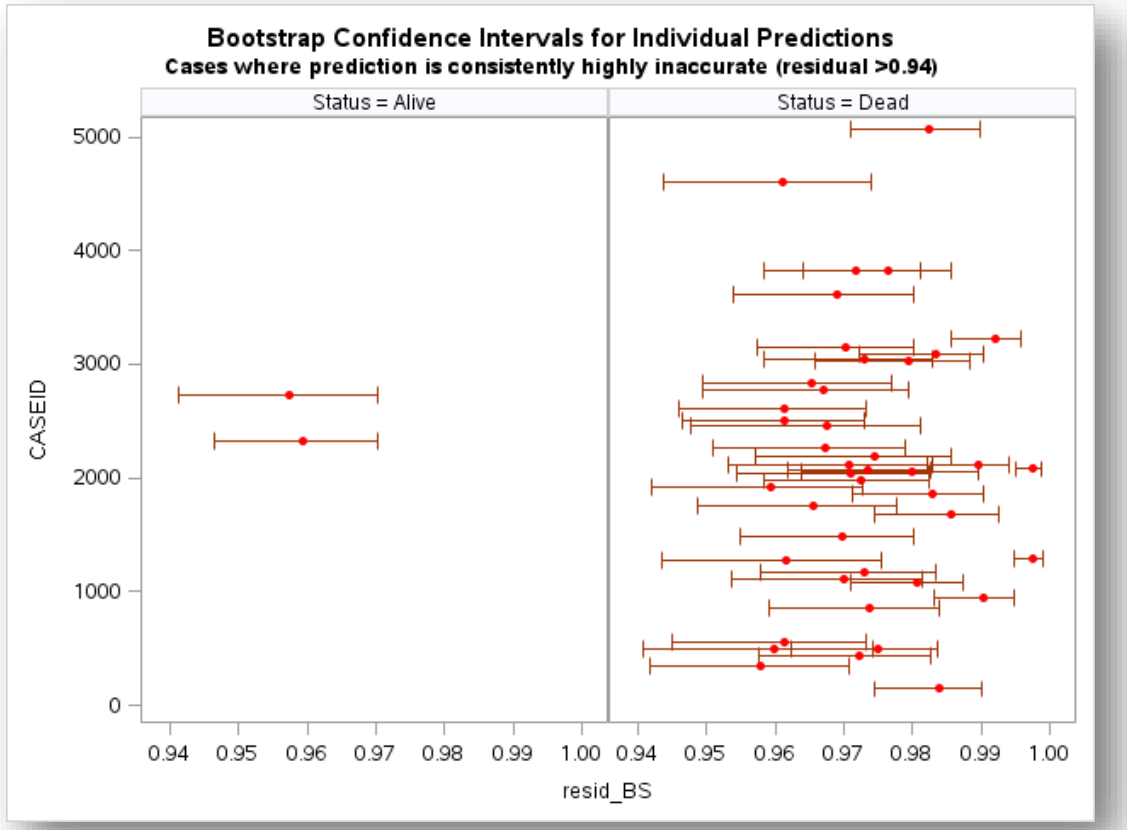

```
164  /* create percentile intervals for the bootstrap samples and merge with full model CI
165          this action expects table intable_BS_PE and intable_PE */
166  proc cas;
167      resample.percentilePE / intable='sample' alpha=0.05;
168  run;
169
170  /* plot the parameter effects and CI from the full sample data with BS percentile intervals on top */
171  data sample_BS_PE_PLOT;
172      set mylib.sample_pe_pctCI;
173  run;
174
175              title "Evaluate Parameter Estimates with 95% CI (percentile)";
176              proc sgplot data=sample_BS_PE_PLOT;
177                  where lowerCL is not null;
178                  scatter y=Parameter x=Estimate / xerrorlower=LowerCL xerrorupper=UpperCL markerattrs=(symbol=circle size=9 color=red) legendlabel='Full Data';
179                  scatter y=Parameter x=BS_Estimate / xerrorlower=BS_LowerCL xerrorupper=BS_UpperCL markerattrs=(symbol=circlefilled size=6 color=green) legendlabel='Bootstrap';
180                  *xaxis grid;
181                  yaxis display=(nolabel);
182                  refline 0 / axis=x;
183              run;
```

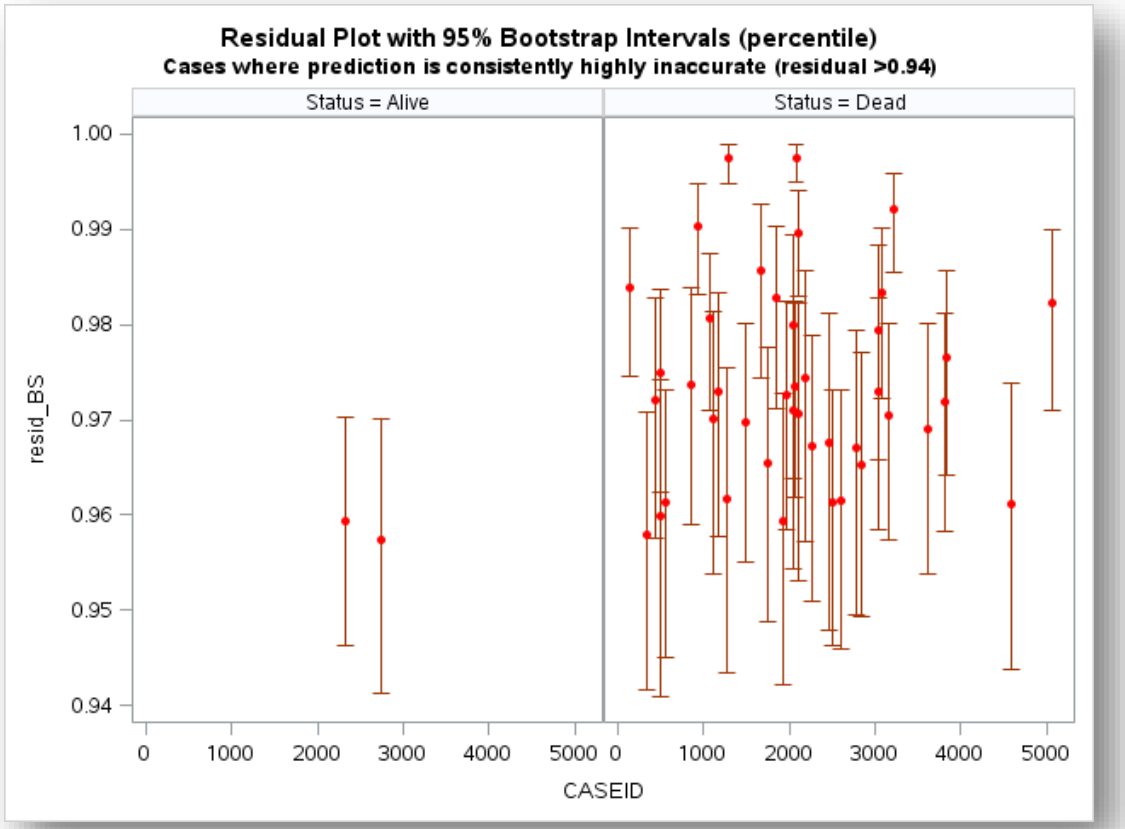
- Another user contributed action brings the Parameter Estimates together from all the resamples and the original full data
- A data step moves the data from CAS to SAS 9.4 so that…
- Use Proc SGPLOT to visualize the Parameter Estimates with both types of confidence intervals
- This fit generalizes very well with both intervals types being almost identical

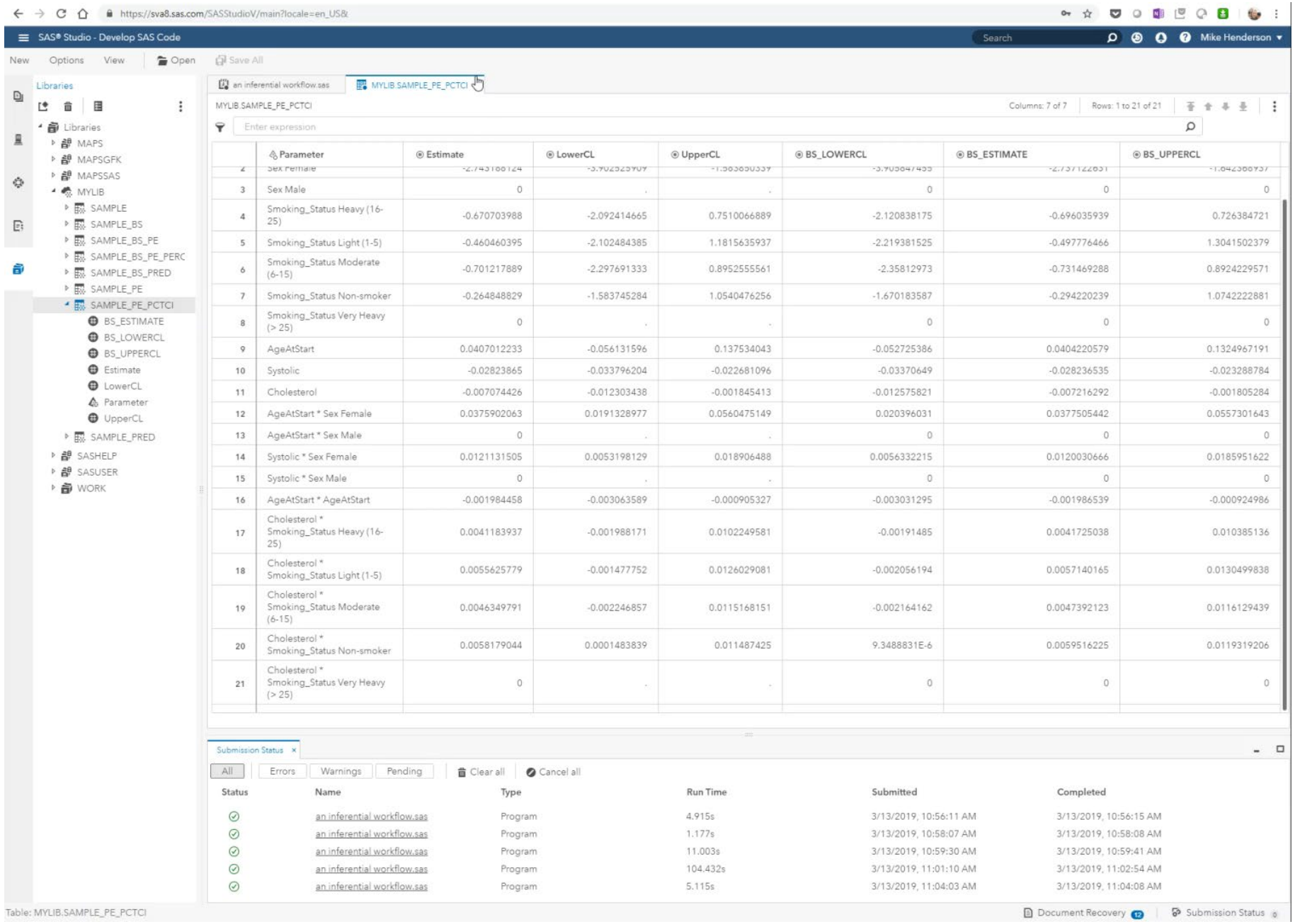

https://sva8.sas.com/SASStudioV/main?locale=en_US&
SAS® Studio - Develop SAS Code
Search
Mike Henderson
New
Options
View
Open
Save All
Libraries
Libraries
MAPS
MAPSGFK
MAPSSAS
MYLIB
SAMPLE
SAMPLE_BS
AgeAtDeath
AgeAtStart
AgeCHDdiag
BAG
BP_Status
bs_caseID
bsID
CASEID
Chol_Status
Cholesterol
DeathCause
Diastolic
Height
MRW
Sex
Smoking
Smoking_Status
Status
Systolic
Weight
Weight_Status
SAMPLE_BS_PE
SAMPLE_BS_PRED
SAMPLE_PE
SAMPLE_PRED
SASHELP
SASUSER
WORK
an inferential workflow.sas
Submit
Cancel
History
Add as Snippet
Mar 13, 2019, 11:01:10 AM
Code
Log
Errors (0)
Warnings (0)
Notes (1730)
NOTE: Active Session now MYSESS.
NOTE: Convergence criterion (GCONV=1E-8) satisfied.
NOTE: Convergence criterion (GCONV=1E-8) satisfied.
NOTE: Convergence criterion (GCONV=1E-8) satisfied.
NOTE: Convergence criterion (GCONV=1E-8) satisfied.
NOTE: Convergence criterion (GCONV=1E-8) satisfied.
NOTE: The CAS table 'sample_BS_PE' in caslib 'CASUSER(mihend)' has 36288 rows and 13 columns.
95
96 OPTIONS NONOTES NOSTIMER NOSOURCE NOSYNTAXCHECK;
109
Submission Status
All
Errors
Warnings
Pending
Clear all
Cancel all
Status
Name
Type
Run Time
Submitted
Completed
an inferential workflow.sas
Program
4.915s
3/13/2019, 10:56:11 AM
3/13/2019, 10:56:15 AM
an inferential workflow.sas
Program
1.177s
3/13/2019, 10:58:07 AM
3/13/2019, 10:58:08 AM
an inferential workflow.sas
Program
11.003s
3/13/2019, 10:59:30 AM
3/13/2019, 10:59:41 AM
an inferential workflow.sas
Program
104.432s
3/13/2019, 11:01:10 AM
3/13/2019, 11:02:54 AM
an inferential workflow.sas [temp]
Document Recovery
Submission Status

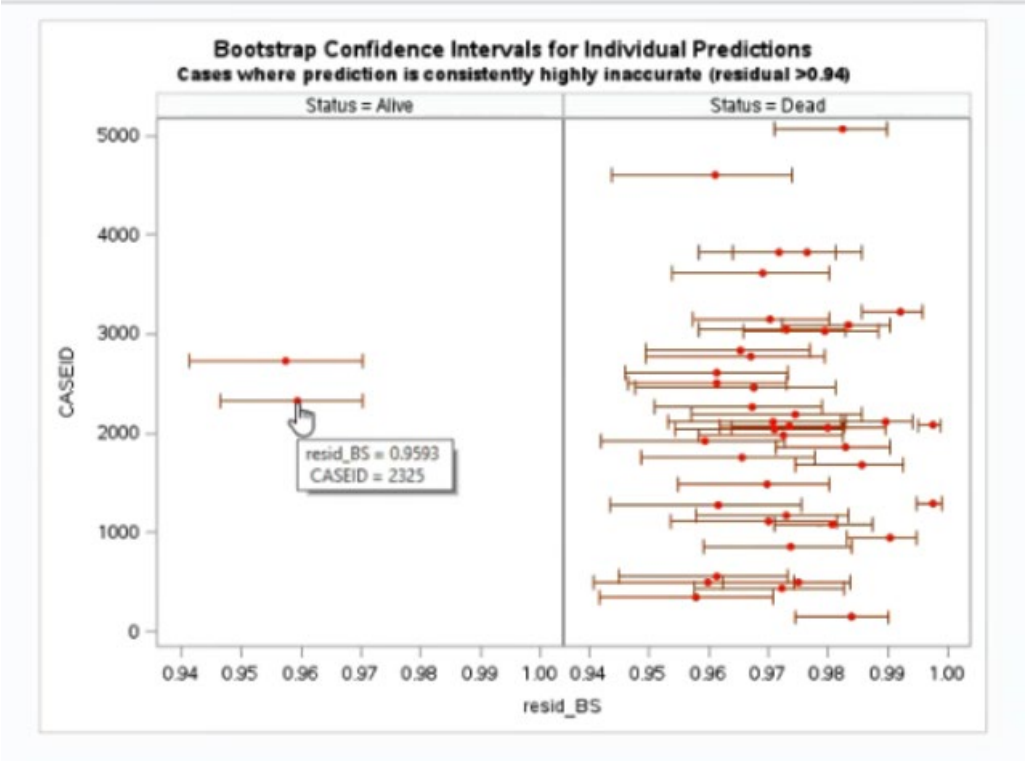
SAS Viya: Using CAS for faster inference

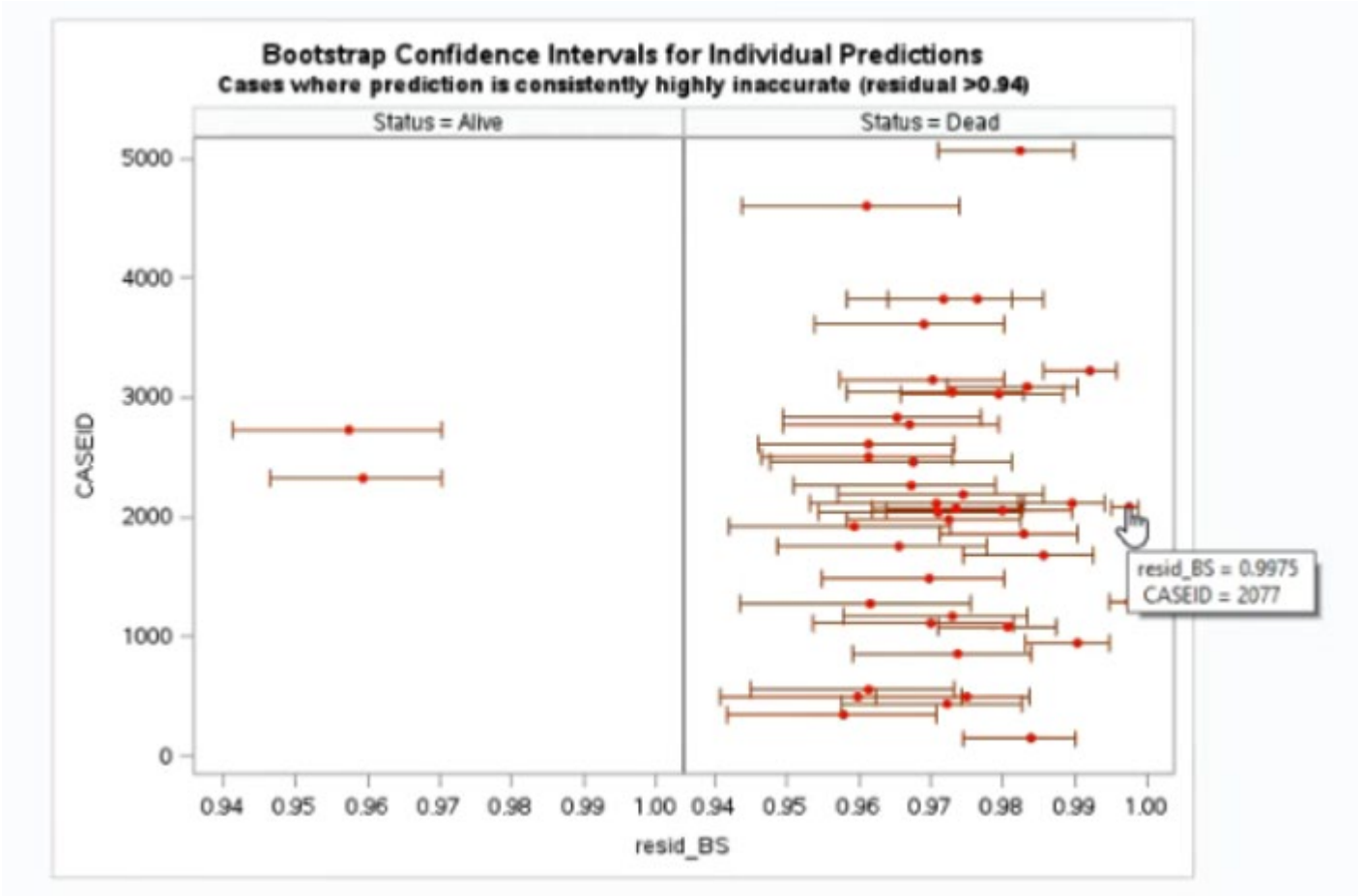
Model Data → Candidate Model → Resample Data → Fit Candidate to Resamples → Collect Parameters from each fit → And More!





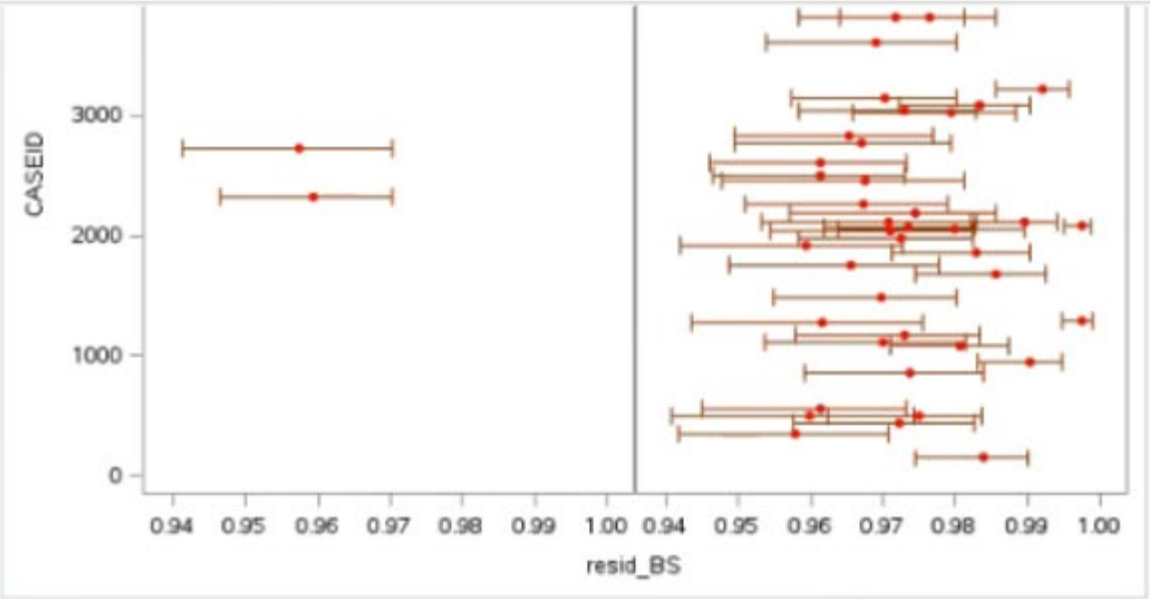


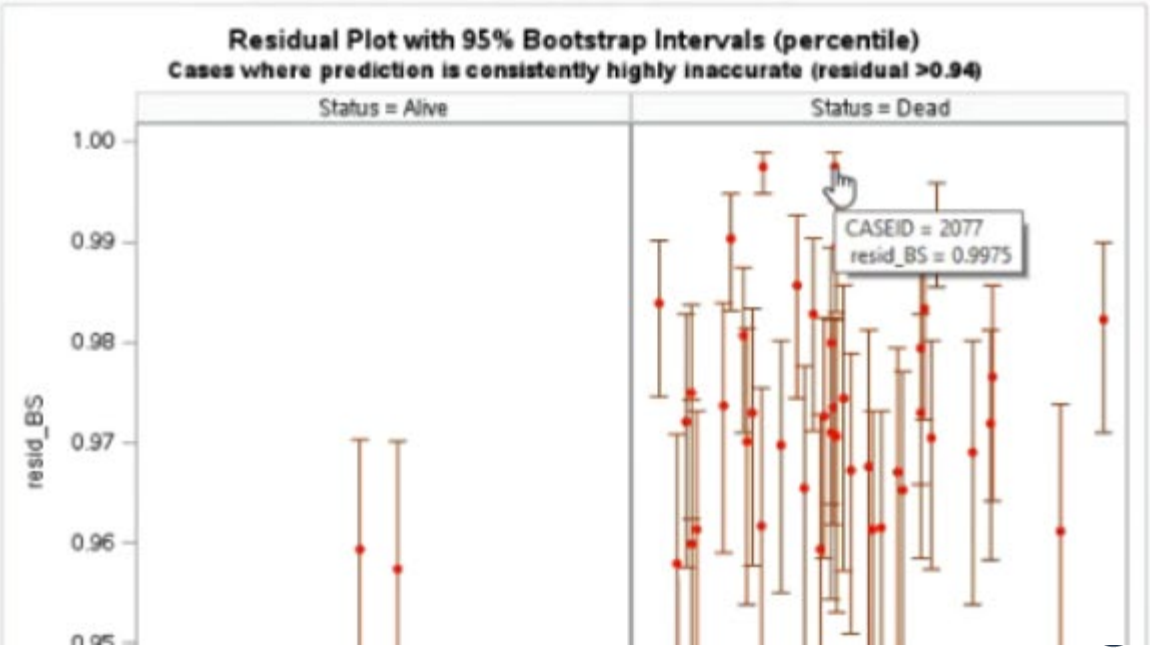




Code Log Results Output Data

- The CAS Procedure
  - percentile.percentile
    - Output CAS Tables
- The SGPanel Procedure
  - The SGPanel Procedure
- The SGPanel Procedure
  - The SGPanel Procedure





Questions ?

SAS Team

Sherrine Eid, Mike Henderson, Jim Box,
Alex Ford, Mark Lambrecht, Toru Tsunoda, Stijn Rogiers.

sas.com