

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



10th–13th November
RAI Amsterdam
The Netherlands

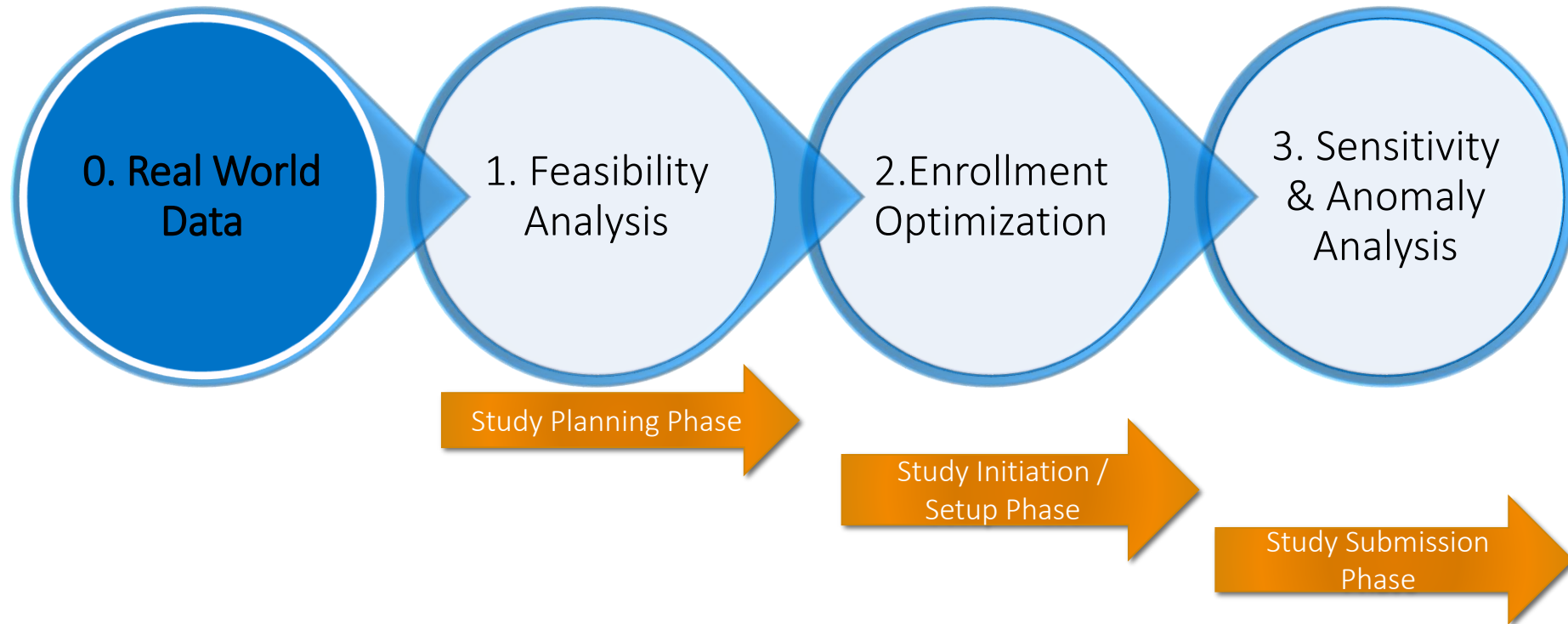
EU 2019

The Clinical Data
Science Conference

Sherrine Eid – MPH, Principal Industry Consultant, RWE, Epidemiology/Biostatistics, SAS
Stijn Rogiers – Principal Industry Consultant, Global Health and Life Sciences Practice



Scope

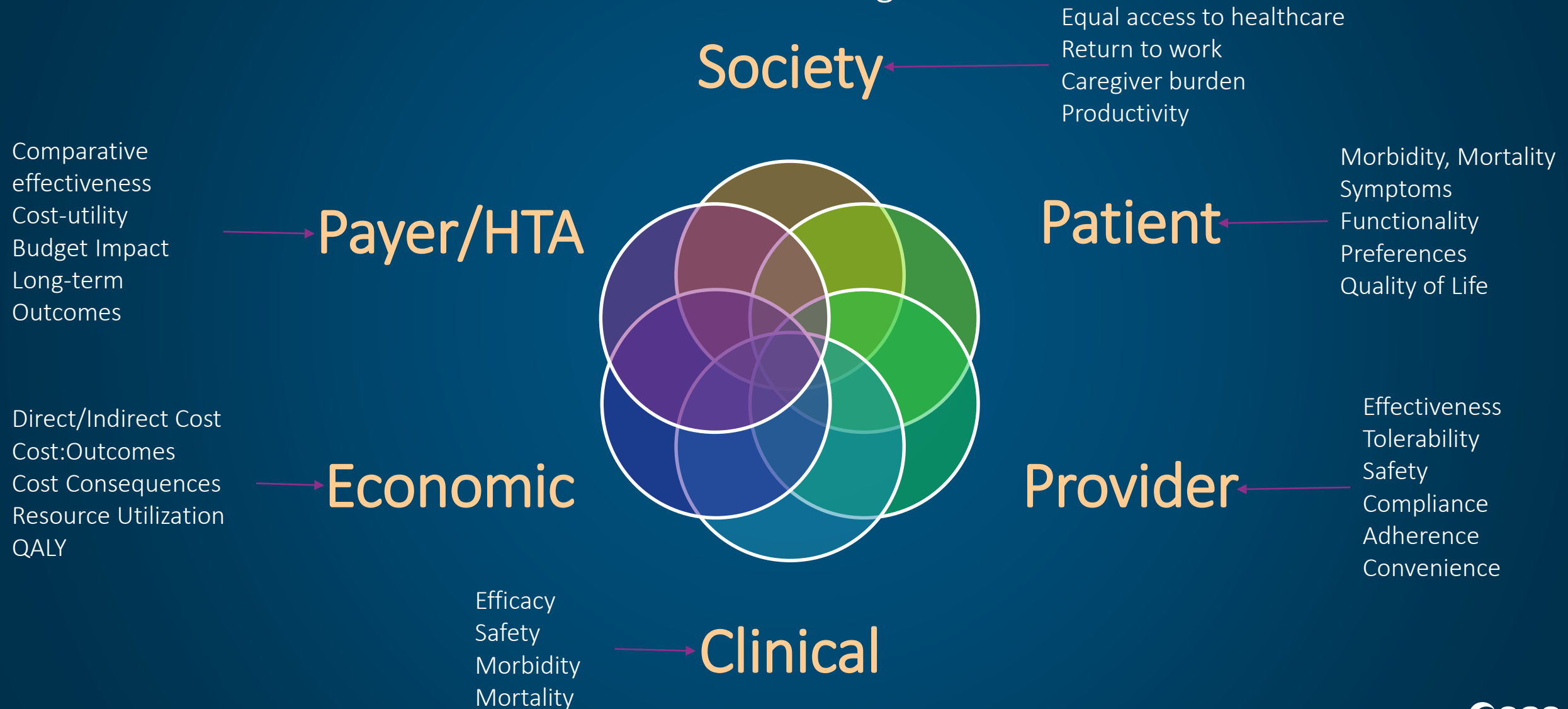




RWE is a hot topic today ...

Maximize product value: Beyond efficacy and safety

Value Message



Company Confidential – For Internal Use Only

Recent Industry Activities

US FDA

[Framework for FDA's Real World Evidence Program \(Dec 2018\)](#)

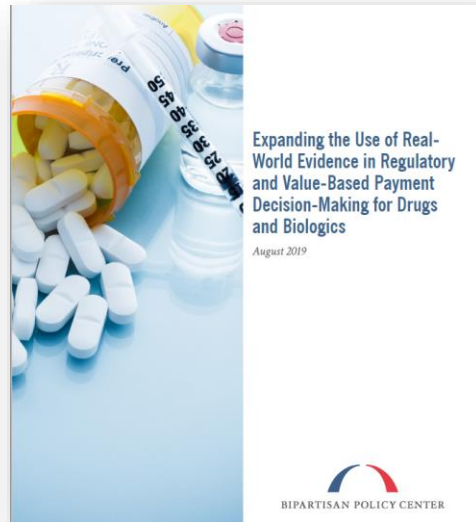
National Academies of Science (US)

- [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/>

Recent Industry Activities



Big Data – Challenges and Opportunities

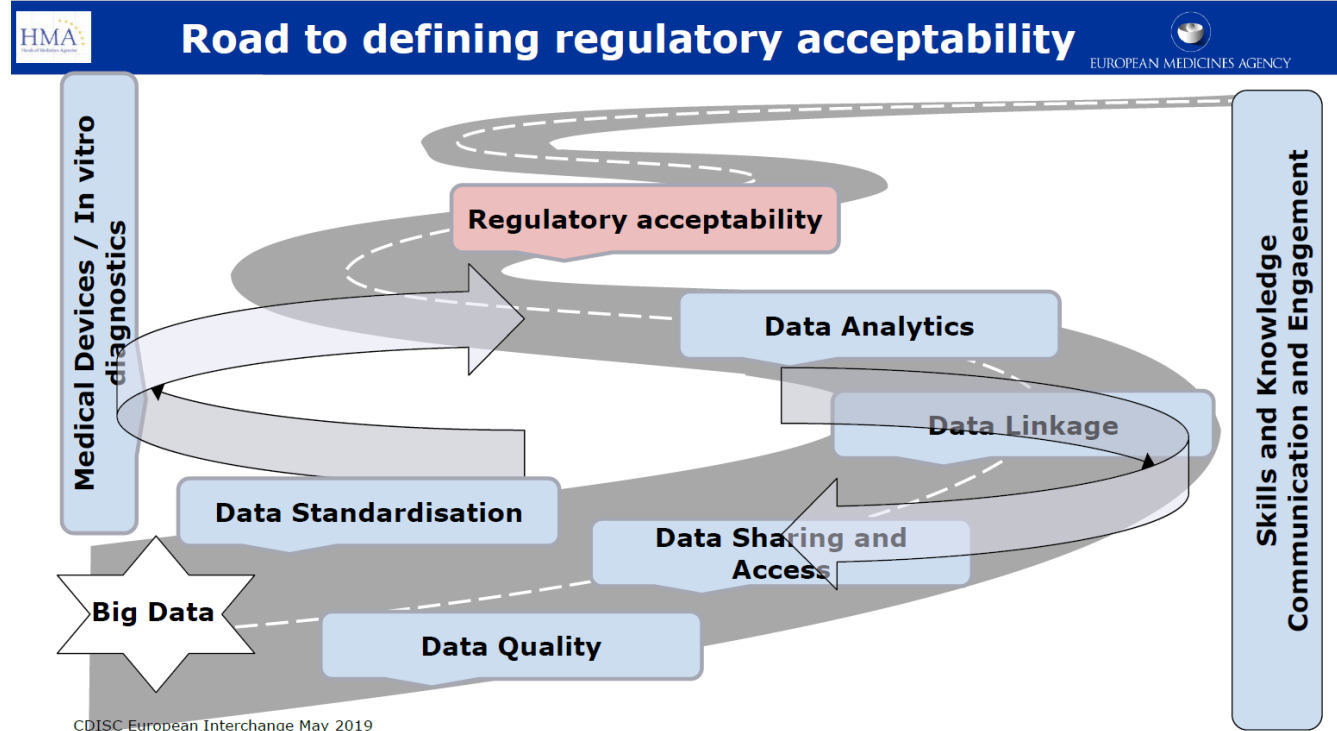
Moving forward with recommendations from the HMA-EMA Joint Big Data taskforce

Dr Alison Cave
Principal Scientific Administrator
Pharmacovigilance and Epidemiology Department

CDISC European Exchange May 2019



An agency of the European Union



How to ensure that novel analytic methods are fit for decision-making [Share](#)

News 02/10/2019



The past decade has seen the increased generation and availability of new data sources such as real-world evidence, as well as patient-level data from completed randomised clinical trials. While these data provide an opportunity to learn more about a medicine's benefits and risks, and can complement the main body of evidence coming from randomised clinical trials, they will not necessarily translate into credible evidence for regulators and other decision-makers in the absence of adequate statistical methods to extract, analyse, and interpret them.

In an article published in Clinical Pharmacology & Therapeutics, regulators and academics explain how proper methodological validation can ensure the credibility of these data sources and allow authorities to rely on them to draw reliable scientific conclusions. The article is co-signed by a number of EMA staff members, academics

<https://www.ema.europa.eu/en/news/how-ensure-novel-analytic-methods-are-fit-decision-making>

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)



PRO

CLAIMS

RX

COA

MORTALITY

mHEALTH

CONSUMER

EMR

LABS

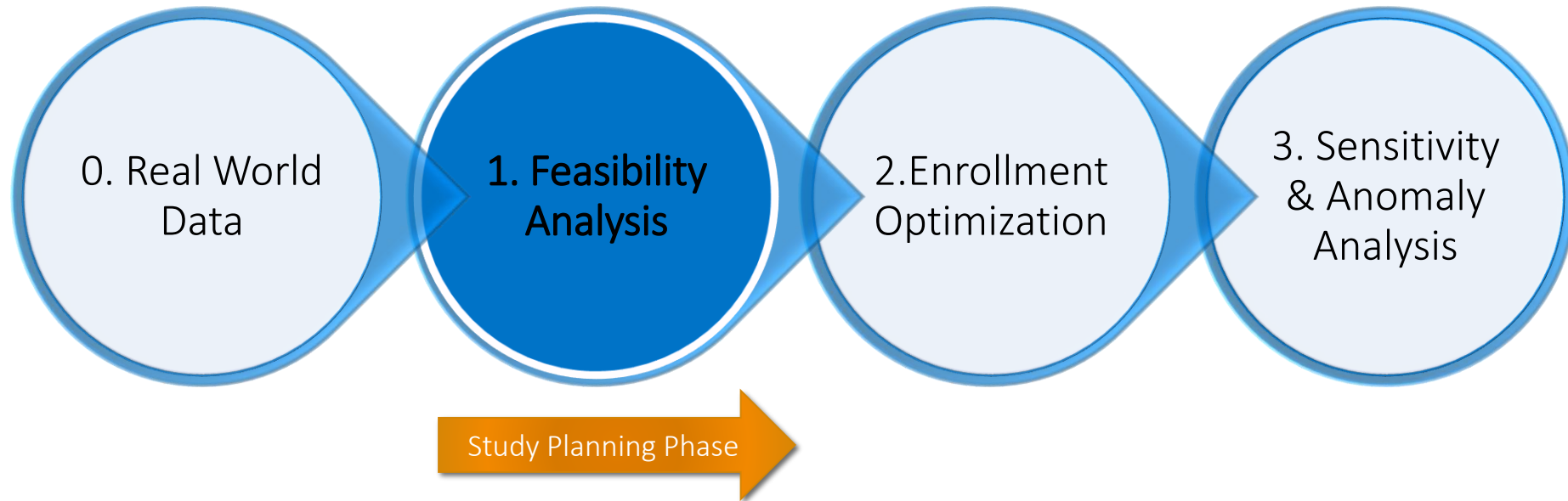
REGISTRY

SAFETY

DIGITAL

SOCIAL MEDIA

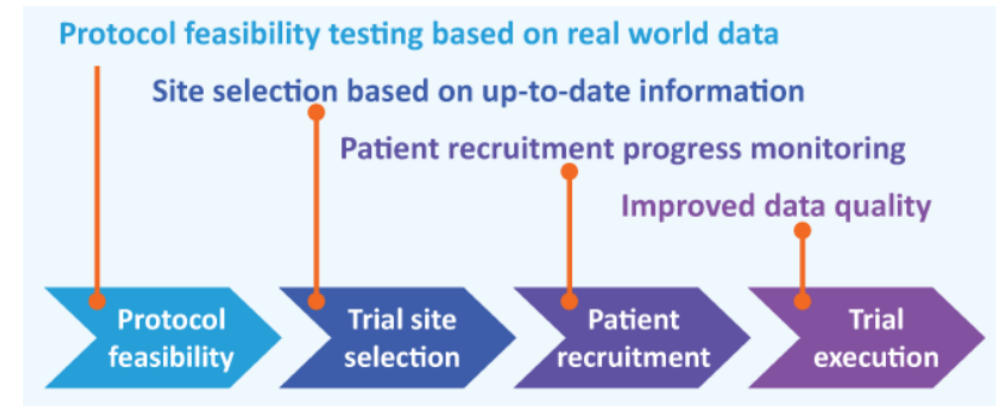
Scope



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

Hypothesis Generation

(example study / protocol)

Study Planning Phase

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)

ClinicalTrials.gov Identifier: NCT03702244

⚠ 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.

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

Profiling the data

Home

Projects (20)

Add-ins

Export Data Templates

Index Event Definitions

Data Sources

100K_ICD10_SEA_RWE

2.3M_CMS_ICD09_HAF44_20190121

50K_CMS_ICD10_HAF44_RWE_20181023

800K member commercial data

800K_CMS_ICD09_HAF44_SEA_20181025

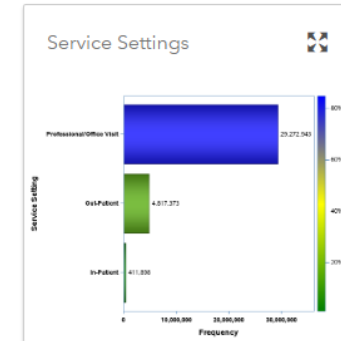
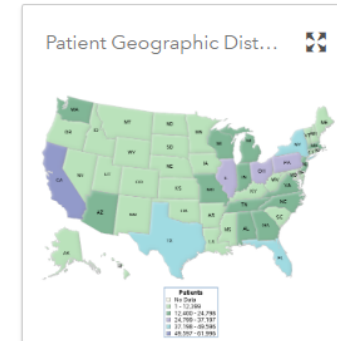
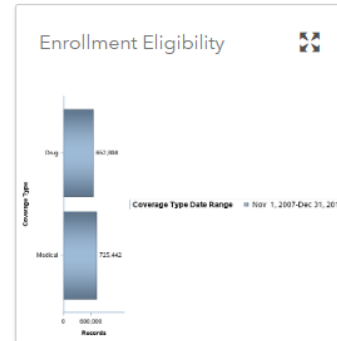
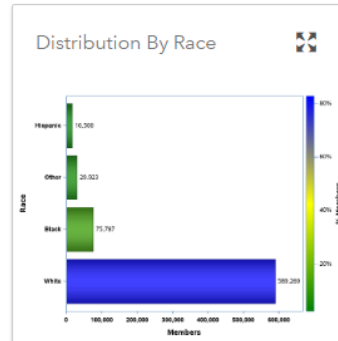
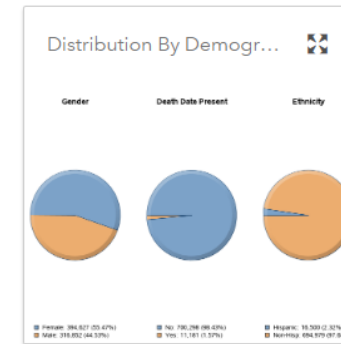
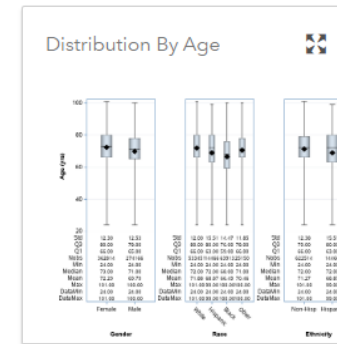
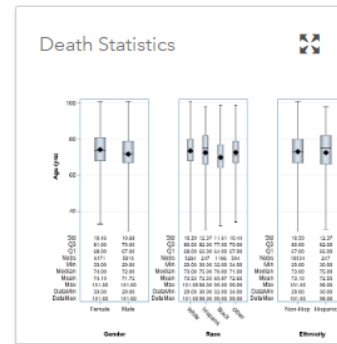
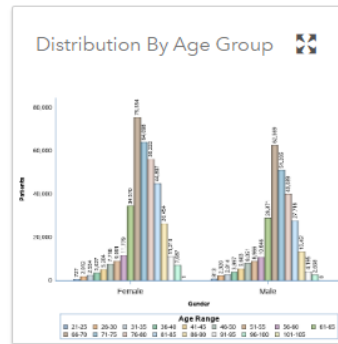
800K_CMS_ICD10_HAF44_RWE_20181025

System Utilities

Filter

Profiling Report for 800K_CMS_ICD10_HAF44_RWE_20181025

Member Information



Codes

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

Search

Diagnoses - Dx

Diagnosis Related Groups

Prescriptions - Rx

Procedures - Px

Inclusion/Exclusion Criteria

AND OR NOT

ACS Pts

Expression label: ACS Pts (based on "Diagnoses - Dx")

Select relative index event option: Use earliest date identified

Codes

Codes to be used by the query:

Select new codes...

☐	Code	Value Set	Description	Type	Vocabulary
☐	I240		ACUTE CORONARY THROMBOS...	DX	ICD-10
☐	4111		INTERMEDIATE CORONARY SYN...	DX	ICD-9
☐	41181		ACUTE CORONARY OCCLUSIO...	DX	ICD-9
☐	I248		OTHER FORMS OF ACUTE ISCH...	DX	ICD-10
☐	I249		ACUTE ISCHEMIC HEART DISEA...	DX	ICD-10

OK

Cancel

Creation 'Index Event' Definition

Acute coronary syndrome (ACS) Patients – Cont'd

Select Codes: DX

☒ Specify individual codes from:

☒ ICD-10 Diagnosis Codes (1)
 ☒ ICD-9 Diagnosis Codes

☐ Use codes in selected value set

Showing 1 codes out of 1 matching codes.

☐	Vocabulary	Code	Code Description	Chapter	Section	Level 3	Level 4
☐	ICD-10	I249	ACUTE ISCHEMIC HEART DISEASE, UNSPECIFIED	DISEASES OF TH...	ISCHEMIC HEART DISEASES (I20-I25)	OTHER ACUTE ISCHEMIC HEART DISEASES	ACUTE ISCHEMIC...

Select Codes: DX

☒ Specify individual codes from:

☒ ICD-10 Diagnosis Codes (7)
 ☒ ICD-9 Diagnosis Codes (2)

☐ Use codes in selected value set

Showing 9 codes out of 9 matching codes.

☐	Vocabulary	Code	Code Description
☐	ICD-9	411B1	ACUTE CORONARY THROMBOSIS NOT RESULTING IN MYOCARDIAL INFARCTION
☐	ICD-10	I240	ACUTE CORONARY THROMBOSIS NOT RESULTING IN MYOCARDIAL INFARCTION
☐	ICD-9	4111	INTERMEDIATE CORONARY SYNDROME
☐	ICD-10	I2102	ST ELEVATION (STEMI) MYOCARDIAL INFARCTION OF ANTERIOR WALL
☐	ICD-10	I2121	ST ELEVATION (STEMI) MYOCARDIAL INFARCTION OF INFERIOR WALL
☐	ICD-10	I2101	ST ELEVATION (STEMI) MYOCARDIAL INFARCTION OF ANTERIOR WALL
☐	ICD-10	I2109	ST ELEVATION (STEMI) MYOCARDIAL INFARCTION OF ANTERIOR WALL
☐	ICD-10	I2119	ST ELEVATION (STEMI) MYOCARDIAL INFARCTION OF INFERIOR WALL
☐	ICD-10	I2111	ST ELEVATION (STEMI) MYOCARDIAL INFARCTION OF INFERIOR WALL

Only show codes with these values:

☐ ACUTE CORONARY THROMBOSIS NOT RESULTING IN MYOCARDIAL INFARCTION (count = 1)
 ☐ INTERMEDIATE CORONARY SYNDROME (count = 1)
 ☐ OTHER (count = 1)
 ☐ ST ELEVATION (STEMI) MYOCARDIAL INFARCTION OF ANTERIOR WALL (count = 3)
 ☐ ST ELEVATION (STEMI) MYOCARDIAL INFARCTION OF INFERIOR WALL (count = 2)
 ☐ ST ELEVATION (STEMI) MYOCARDIAL INFARCTION OF OTHER SITES (count = 1)

Level 4	Level 5	Level 6
OTHER	ACUTE CORONA	
ACUTE CORONA		
INTERMEDIATE C		
ST ELEVATION (ST...	ST ELEVATION (ST...	
ST ELEVATION (ST...	ST ELEVATION (ST...	
ST ELEVATION (ST...	ST ELEVATION (ST...	
ST ELEVATION (ST...	ST ELEVATION (ST...	
ST ELEVATION (ST...	ST ELEVATION (ST...	
ST ELEVATION (ST...	ST ELEVATION (ST...	

Cohort Generation

Selection 'Index Event' Definition

Search

▲ ▼

► Demographics

▼ Diagnoses - Dx

▼ Diagnosis Related Groups

▼ Laboratory Results

► Chemistry

► Hematology

ICD-10 Diagnosis Codes

MS-DRG Diagnosis Codes

LOINC Codes

94461

765

84868

1 Cohort

2 Analysis Variables

3 Add-ins

Index Event

ACS Pts

n=98,661 14%

	Name ▲	Description	Default Base...	Default Follo...	Owner
☒	ACS Pts	Newly Diagnosed ACS Pts	-12	36	sheidz
☐	Atrial Fibrillation Patients	Newly diagnosed Afib patients	-6	60	sheidz
☐	Breast Cancer Pts		-6	18	ancoom

Custom Baseline (months): * ▼ -12 ▲ Custom Follow-up (months): * ▼ 60 ▲

Apply

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

AND

OR

NOT

➡

🗑️

↶

▼

▲

Apply

▼

✓

☐

18+

n=98,661

Expression label: (based on "Age at Index Event")

AGE_AT_INDEX_DT Value:

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 $\geq 50\%$), or known EF $\leq 40\%$ 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

☒ ☐ NOT (CABG OR CV Testing for CAD OR GFR OR Hx of Obstructive CAD)

CABG

CV Testing for CAD

GFR

Hx of Obstructive CAD

Expression label: (based on "LOINC Codes")

Choose test type:

☐ Determine whether the selected tests were performed

☒ Compare results from the selected tests with a specific numeric value

☐ Compare results from the selected tests with a specific character value

☐ Show test results within a range category

< ▼

Value:

Cohort Generation

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

Inclusion/Exclusion Criteria n=

AND OR NOT     

✓ Codes

Codes to be used by the query: Select new codes...   

☐	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: 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 $\geq 50\%$), or known EF $\leq 40\%$ or moderate to severe valvular or congenital cardiac disease
- Contraindications to cCTA including but not limited to estimated creatinine clearance (GFR) < 45 ml/min

CABG

Chest Pain

CV Testing for CAD

GFR

Hx of Obstructive CAD

Expression label: (based on "Diagnoses - Dx")

Codes

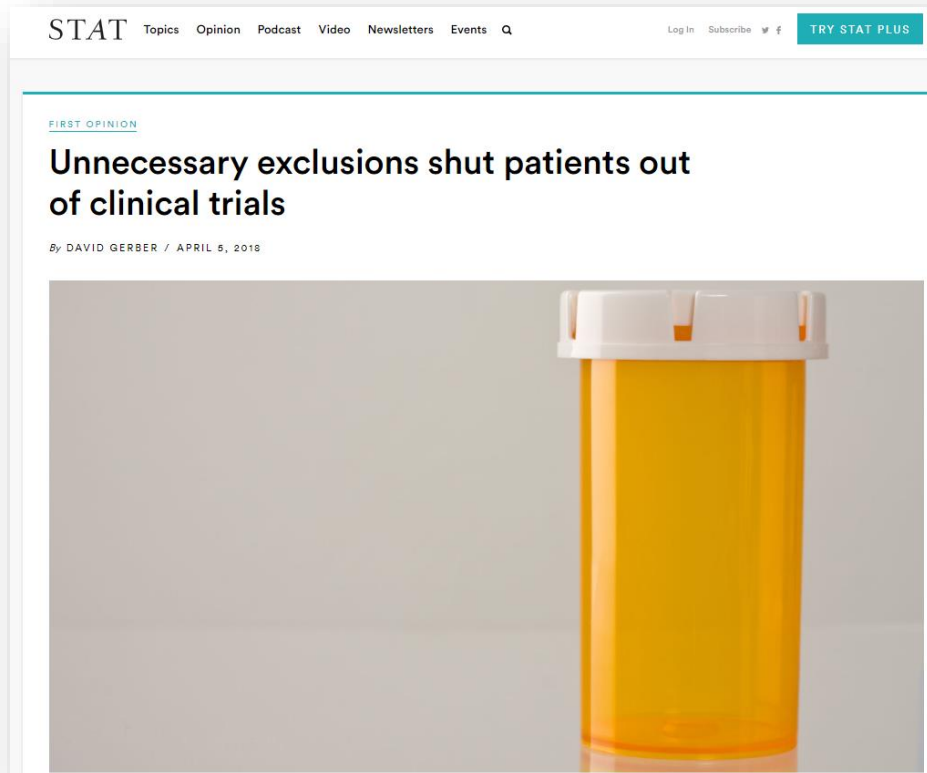
Codes to be used by the query:

Select new codes... + - i

☐	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

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 ?

1 Cohort

2 Analysis Variables

3 Add-ins

Index Event

ACS Pts

n=98,661

14%

Inclusion/Exclusion Criteria

n=19,243

3%

AND

OR

NOT

Apply

>

☒

☐

18+

n=98,661

✓

☒

☐

NOT (CABG OR CV Testing for CAD OR GFR OR Hx of Obstructive CAD)

"12.314" -->

n=19,243

CABG

CV Testing for CAD

GFR

Hx of Obstructive CAD

Expression label:

CABG

(based on "Procedures - Px")

>

Codes

>

Options

Cohort Variables

Filters

Name	Count / %
18+	98,661 / 100%
NOT (CABG OR CV Testing for CAD OR GFR OR Hx of Obstructive CAD)	19,243 / 20%

User-defined Analysis Variables

☐	Name	Count / %
☐	Chest X Ray	79 / 0%
☐	Heart Failure	62,679 / 64%
☐	(Cardiac Angiography OR MI)	46,723 / 47%
☐	(Hyperlipidemia Labs OR Hyperlipidemia Dx)	79,547 / 81%

Analysis Variables

Select Pre-defined Analysis Variables

Home > ACS Hypothesis Testing > ACS Pts 800k_ICD10 Relaxed Criteria

Name: ACS Pts 800k_ICD10 Relaxed Criteria
Description: Created by Sherrine Eid on 9/3/2019, 11:07:02 AM
Date Range: Nov 1, 2007 - Dec 31, 2011

Data source: 800K_CMS_ICD10_HAF44_RWE_20181025 n=711,479
Data source description: 800K_CMS_ICD10_HAF44_RWE_20181025
Enrollment: None n=636,980 : 89.53%

Final Cohort Count: n=19,243
Type: Index Event Cohort

Actions

1 Cohort — 2 Analysis Variables — 3 Add-ins

Filter

Predefined Variables

- Demography Variables
 - AGEGROUP_10: Age Group (0-9, 10-19, 20-29, 30-39, 40-49, ...)
 - AGEGROUP_2: Age Group (<65, >=65)
 - AGEGROUP_4: Age Group (0-17, 18-44, 45-64, 65+)
 - AGEGROUP_5: Age Group (0-17, 18-39, 40-54, 55-64, 65+)
 - AGEGROUP_6: Age Group (0-17, 18-34, 35-44, 45-54, 55-64, 65+)
 - CENSUS_STATE: Census State Classification
 - ETHNICITY: Ethnicity
 - GENDER_CD: Gender
 - RACE: Race
 - ZIP: Zip Code
- Diagnosis Variables
 - AFIB: Atrial Fibrillation
 - AFLUTTER: Atrial Flutter
 - AKIDNEYFAIL: Acute Kidney Failure
 - AMI: Acute Myocardial Infarction

Postprocess Variables

Predefined Variables Risk Factors

Name	Description	Remove
AGEGROUP_10	AGEGROUP_10: Age Group (0-9, 10-19, 20-29, 30-39, 40-49, ...)	
GENDER_CD	GENDER_CD: Gender	
RACE	RACE: Race	
ZIP	ZIP: Zip Code	
ETHNICITY	ETHNICITY: Ethnicity	

Apply

Analysis Variables

Risk Factors

1 Cohort — 2 Analysis Variables — 3 Add-ins

Postprocess Variables Apply

Predefined Variables Risk Factors

Minimum frequency percent threshold: Time Period:

☐	Name	Description
☐	CMS-HCC	CMS_HCC
☐	CHARLSON	Charlson Risk Factor Set
☐	CCS	Clinical Classification Software Risk Factor Set
☒	ELIXHAUSER	Elixhauser Risk Factor Set
☐	STANDARD	Standard

Model & UI development

Use SAS (industry) and company specific models

The image displays three overlapping screenshots of the SAS Add-in Builder interface, illustrating the workflow for creating a custom add-in.

Left Screenshot (Properties Tab): Shows the configuration for a "Binary Regression Predictive Model". The Name is "Binary Regression Predictive Model". The Description is "Binary Regression Predictive Model-V1 - Uploaded: 1/10/2019, 2:42:26 PM". The Instructions section includes a note about the LOGISTIC procedure and HTML tags allowed. The Input data is set to "Cohort". The Test with project is "Opioids". The Test with cohort is "None". The Types of cohorts available for this add-in are "Population Cohorts" and "Index Event Cohorts". The Category is "Patient-level Outcomes Predictive Model". The Visibility is set to "Public (allow others to see and run this add-in)".

Middle Screenshot (Code Tab): Shows the SAS macro code for the add-in. The macro is named "rwe_addin_binary_predmodel_main". The code includes comments about the macro's purpose, parameters, and macro calls. The macro calls include "checkVarExist", "checkNonMissing", "rwe_util_license_checks", "rwe_extract_cohortdata", "rwe_util_get_dataset_pointer", "rwe_db_connect_info", "rwe_util_delete_dataset", "rwe_util_license_checks", "rwe_util_add_message", "rwe_util_load_global_macro_vars", "rwe_aa_util_member_rf_score", "rwe_addin_pred_bintarget_model", "rwe_util_write_status_json", and "rwe_create_rf_label". The code also includes a disclaimer and a history section.

Right Screenshot (User Interface Tab): Shows the user interface design for the add-in. The "Create an interface for setting run-time parameters" checkbox is checked. The Controls section includes a "Text" control, a "Checkbox" control, a "Single Selection" control, and a "Multiple Selection" control. The User Interface - Work Area section includes a "Data and Input Variables" tab, a "Model Effects and Selection" tab, and a "Model Output Options" tab. The "Model Output Options" tab is selected, showing options for "Create predicted response data set with score information", "Select statistics to display", "Partial correlation", "Generalized R Square", "Hosmer and Lemeshow goodness-of-fit", "Influence diagnostics", and "Display default plots".

Model & UI development

Cohort Characterization Report

Home > ACS Hypothesis Testing > ACS Pts 800k_ICD10 Relaxed Criteria

Name: ACS Pts 800k_ICD10 Relaxed Criteria
 Description: Created by Sherrine Eid on 9/3/2019, 11:07:02 AM
 Date Range: Nov 1, 2007 - Dec 31, 2011

Data source: 800K_CMS_ICD10_HAF44_RWE_20181025 n=711,479
 Data source description: 800K_CMS_ICD10_HAF44_RWE_20181025
 Enrollment: None n=436,988 : 69.53%

Final Cohort Count: n=19,243
 Type: Index Event Cohort

Actions

1 Cohort — 2 Analysis Variables — 3 Add-ins

Cohort Characterization Sherrine Eid Sep 3, 2019, 5:18:40 PM

Delete Run

Description (optional)

Report

☒ Demographic Distribution
☒ Geographic Distribution
☒ Top Diagnosis, Procedure and Drug Code Values
 Number of Top Code Values: 20
☐ Analysis Variable Distribution

cohort_demographic_smry
 cohort_characterization - PDF RTF
 cohort_demographic_smry - HTML
 cohort_diagnosis_smry - HTML
 cohort_drug_smry - HTML
 cohort_procedure_smry - HTML

ACS Pts 800k_ICD10 Relaxed Criteria
 Cohort Type: Index Event
 Cohort Overview

Report Generated 03 Sep 2019: 11:15:04

General Information					
Cohort	Cohort Type	Project	Data Source	Patients	
ACS Pts 800k_ICD10 Relaxed Criteria	Index Event	ACS Hypothesis Testing	800K_CMS_ICD10_HAF44_RWE_20181025	19,243	

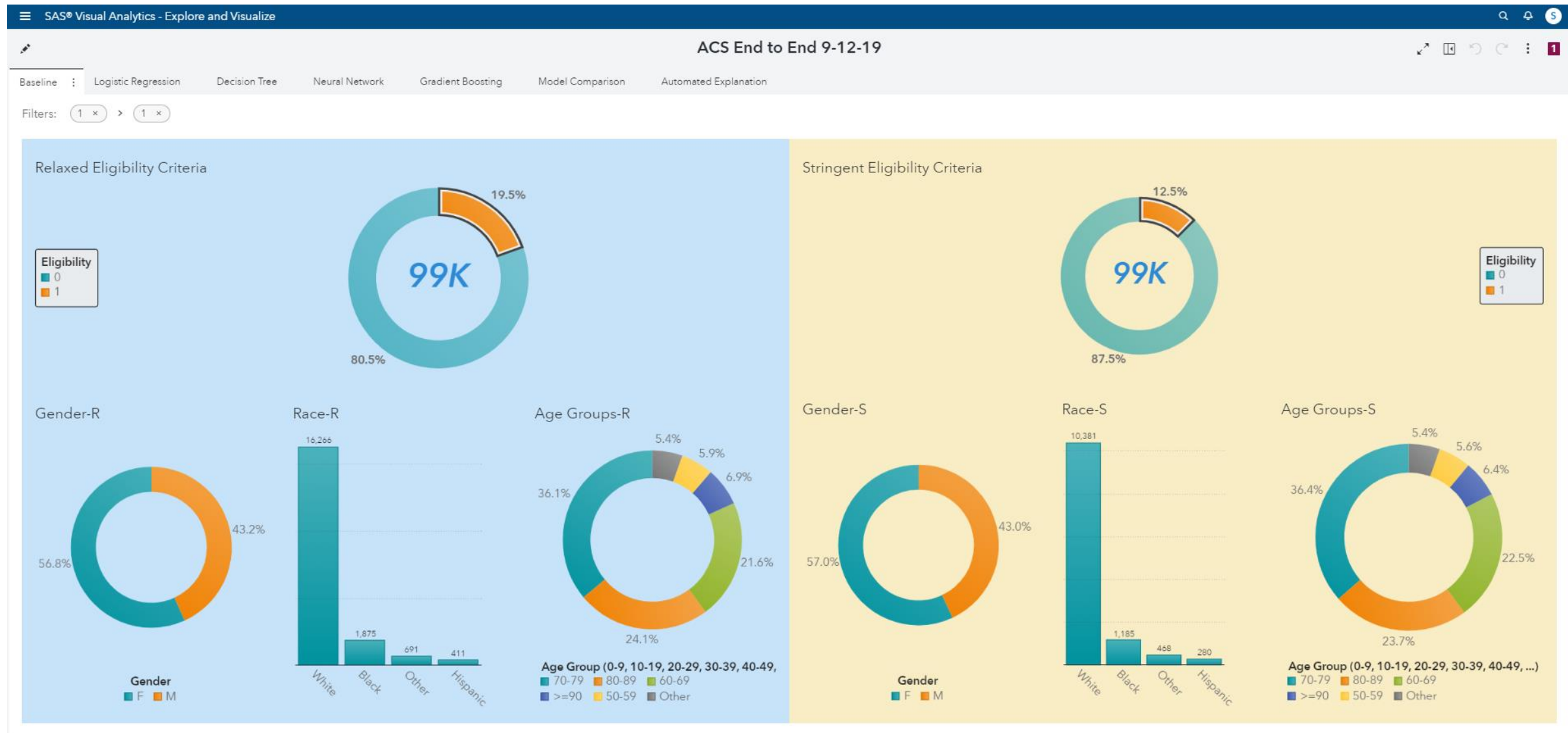
Study Start Date	Study End Date	Age at Start of Study				Deaths During Study	
		Min	Max	Mean	Median	Deaths	%
November 01, 2007	December 31, 2011	23	100	72	73	230	1.20%

*Age Calculated using the Study Date and Computed Date of Birth

Gender	Race	Patients	Mortality	Age at Beginning of Study	
				Min	Max
Female	Black	1,138	8	24	100
	Hispanic	257	2	25	99
	Other	378	8	26	98
Male	White	9,160	98	24	100
	Black	737	11	23	100
	Hispanic	154	4	24	98
	Other	313	6	26	99
	White	7,106	93	24	100

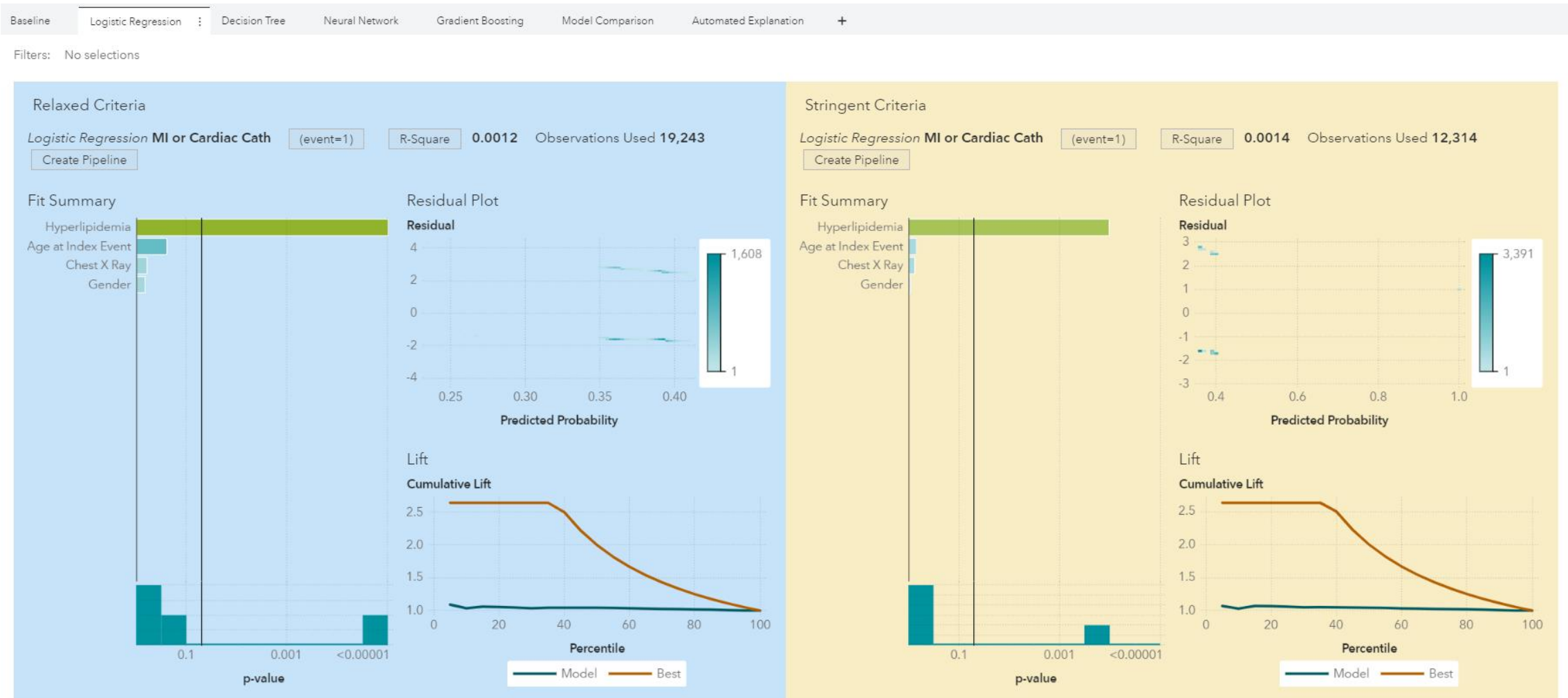
Remark: existing customer SAS and R program library can be imported if already developed.

Visualization / Exploration



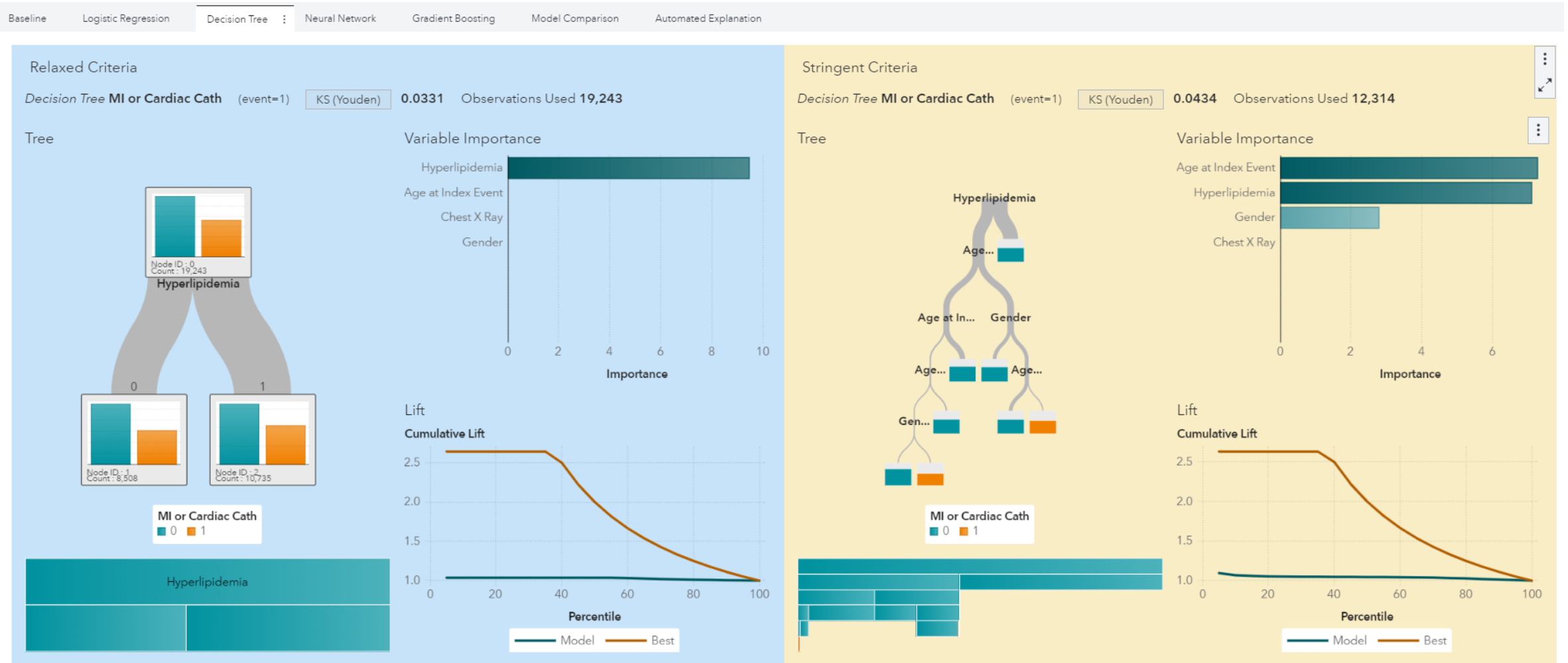
Data Mining & Machine Learning

Logistic Regression



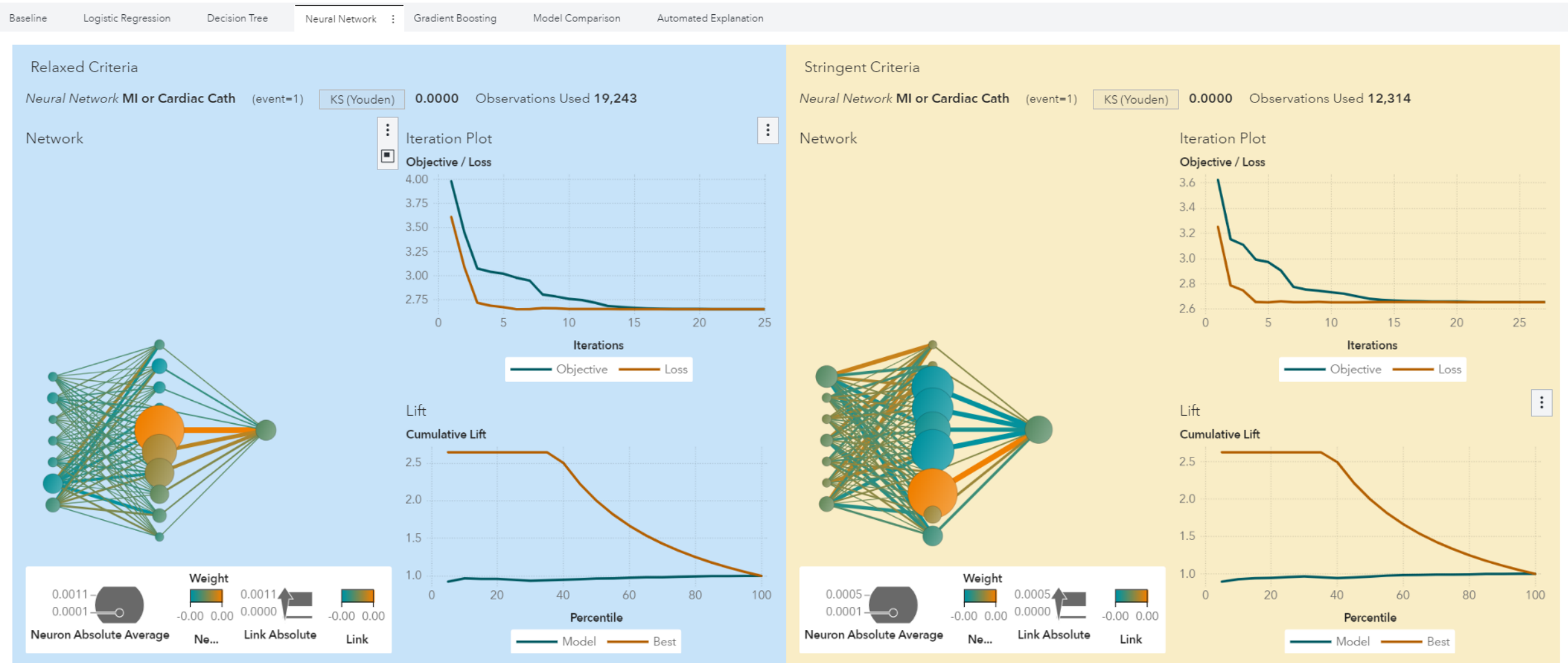
Data Mining & Machine Learning

Decision Tree



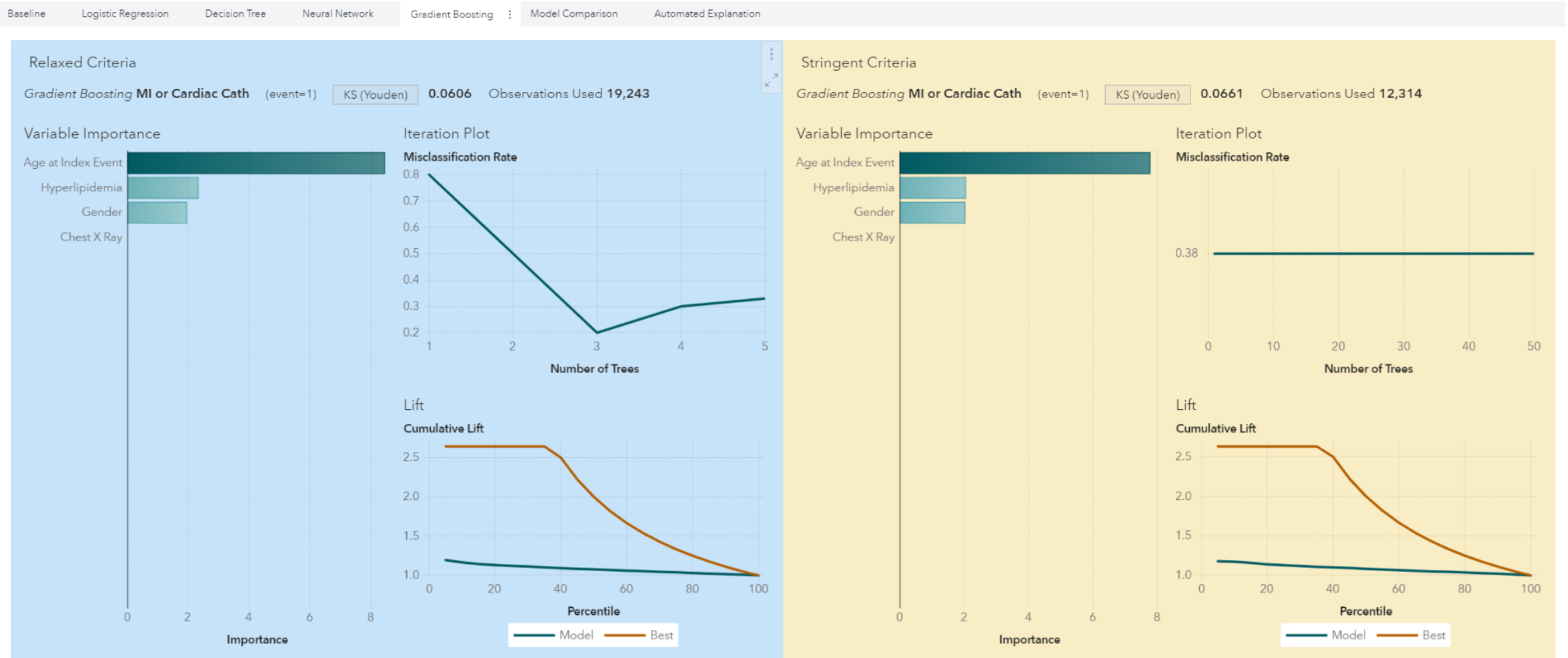
Data Mining & Machine Learning

Neural Network



Data Mining & Machine Learning

Gradient Boosting



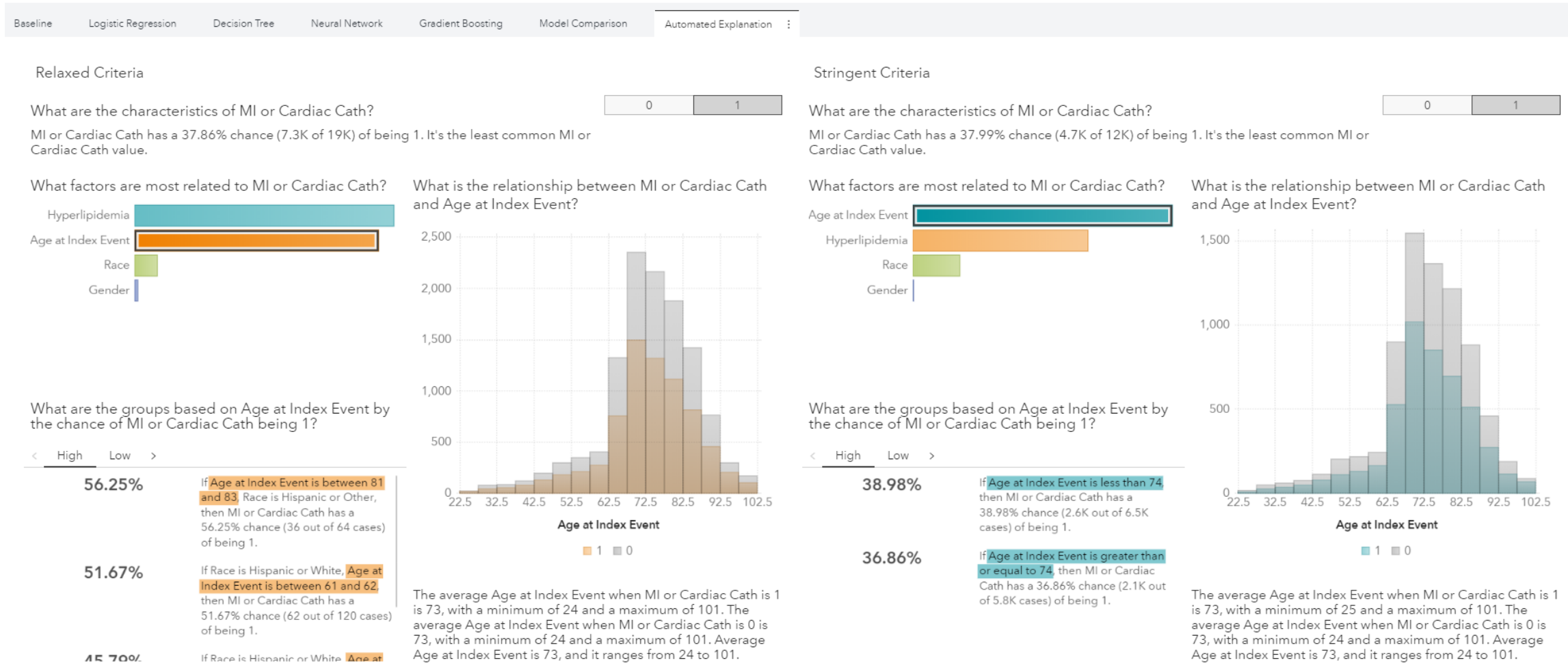
Data Mining & Machine Learning

Model (automatic) Comparison

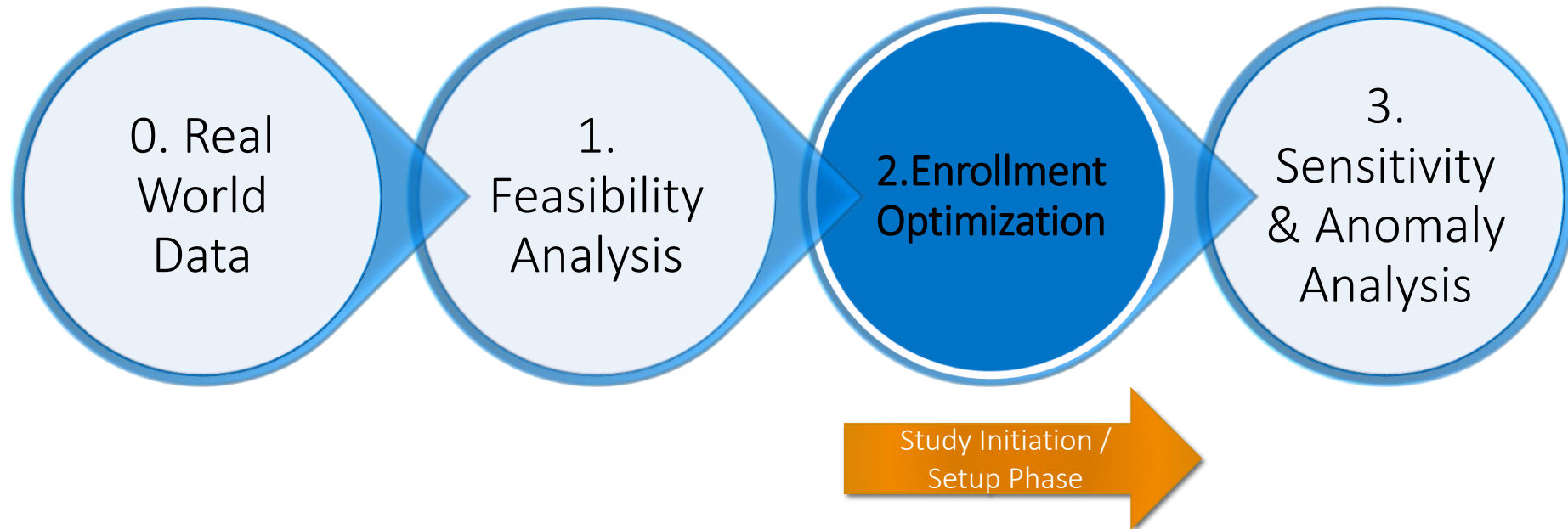


Data Mining & Machine Learning (NLG)

Automated Explanation



Scope



Enrollment Optimization

Phuse EU connect 2018

Paper TT08

Simulating Enrollment Plans

Jim Box, SAS Institute, Cary, NC, USA

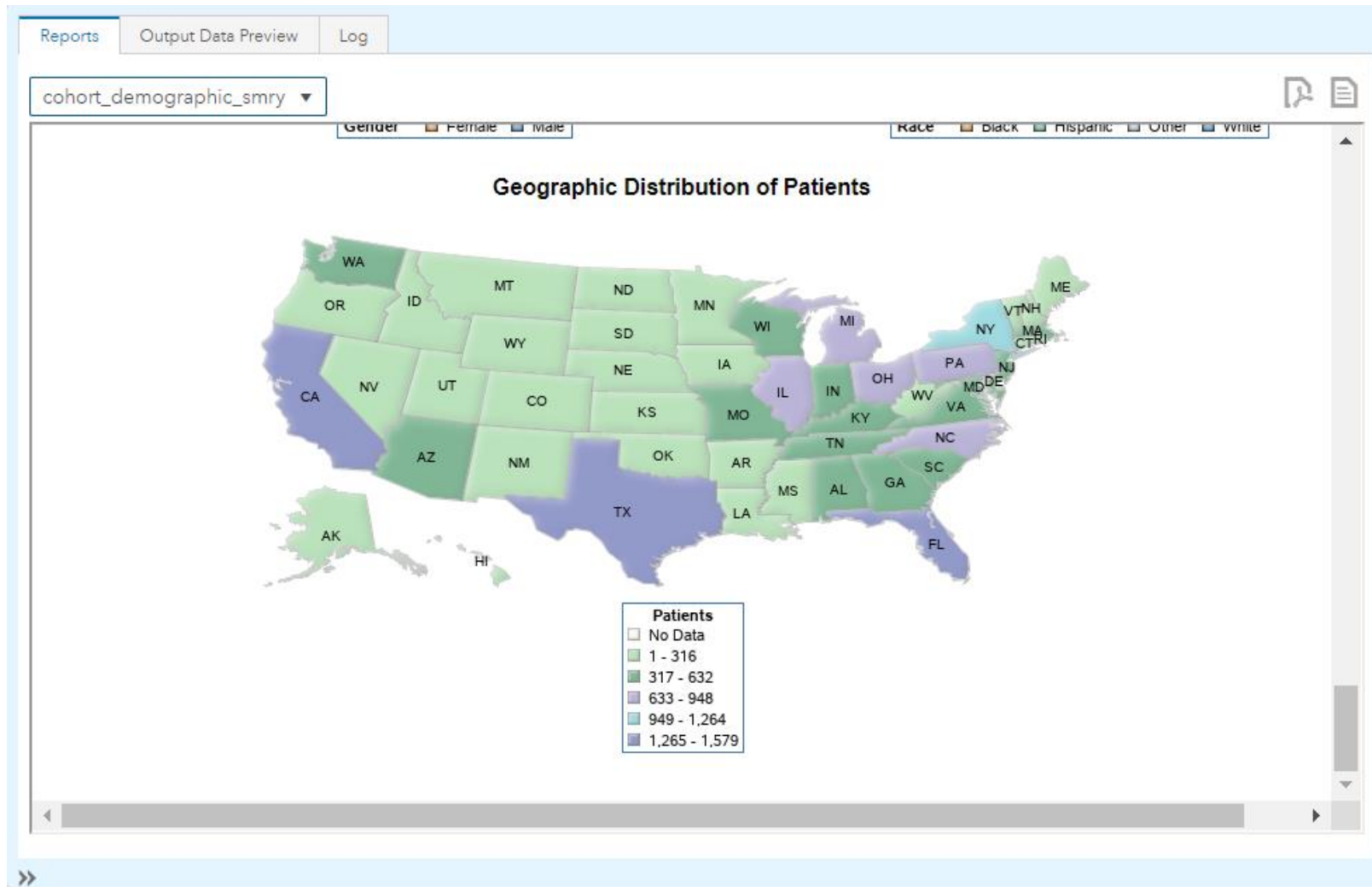
ABSTRACT

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

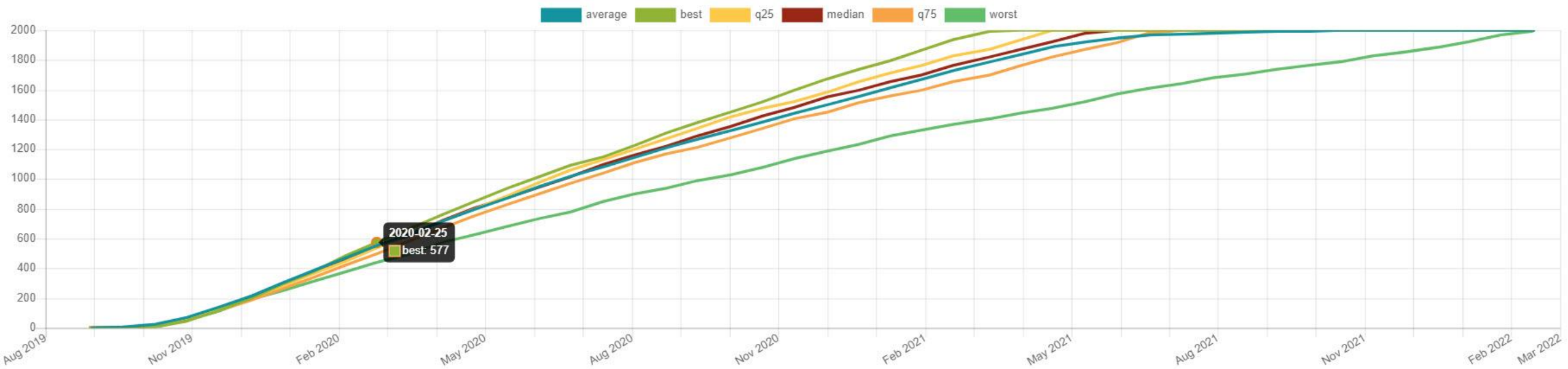
State/Site Selection



Enrollment Optimization

Relaxed Criteria

Projected Enrollment (Quantiles)



Enroll Duration (Months)



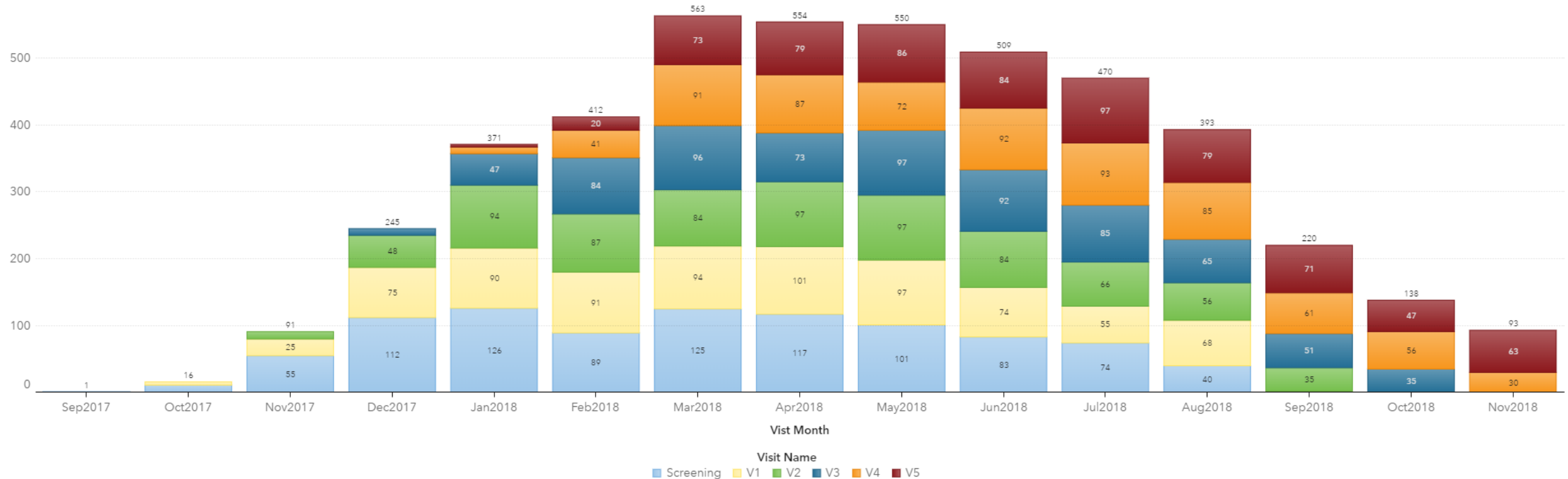
Enrollment Optimization

Monthly Visits Projection

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
----------------------	---------------------------	-------------------------	--	-----------------------	--------------------------	---------------------	---------------------	---------------------------	---------------------------

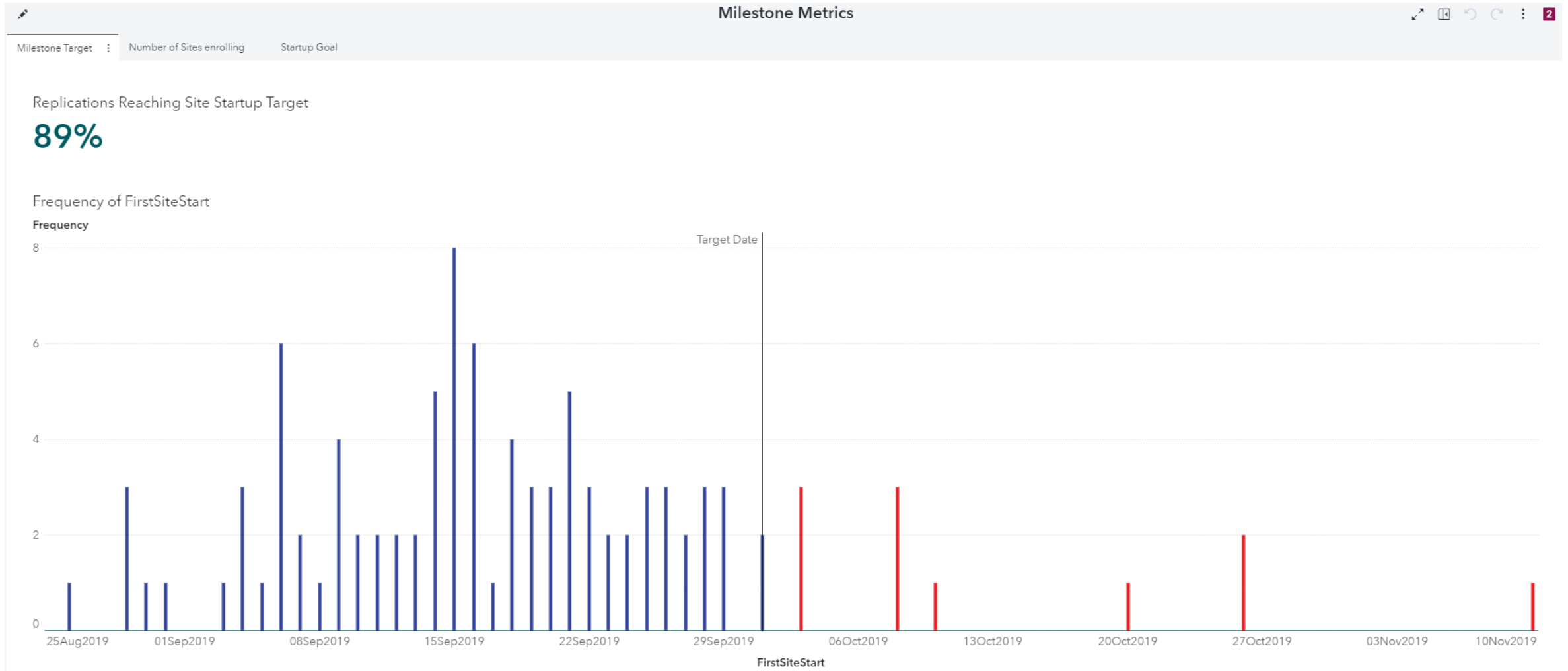
Screening	V1	V2	V3	V4	V5
-----------	----	----	----	----	----

Visit Count

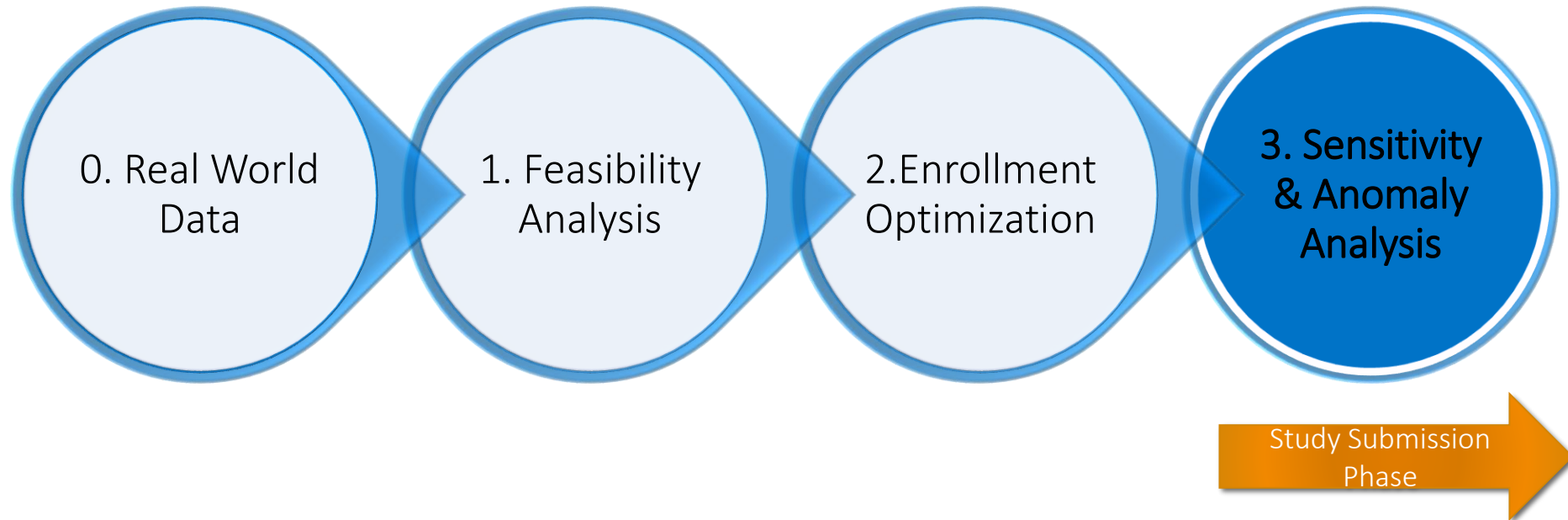


Enrollment Optimization

Milestone Metrics

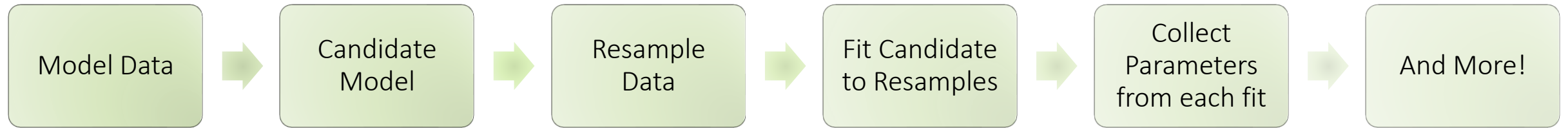


Scope



SAS Viya: Using CAS for faster inference

Review a workflow

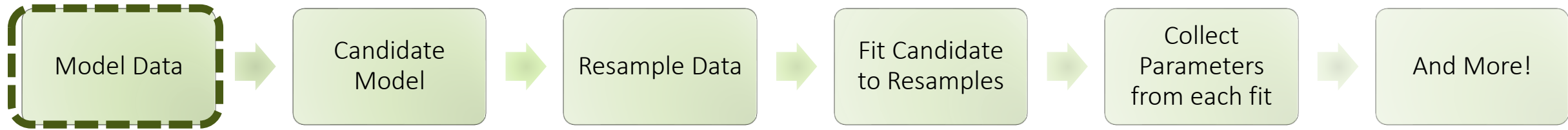


CAS = Cloud Analytics Services

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

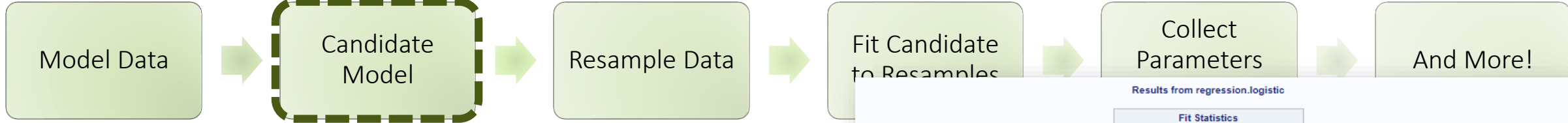
- Viya CASL syntax

- From SAS Interface with PROC CAS;
 - From other languages with SWAT library:
[Python](#), Java, LUA, [R](#), REST
 - More flexible model specification!

```
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;
```

```
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;
```

SAS Viya: Using CAS for faster inference



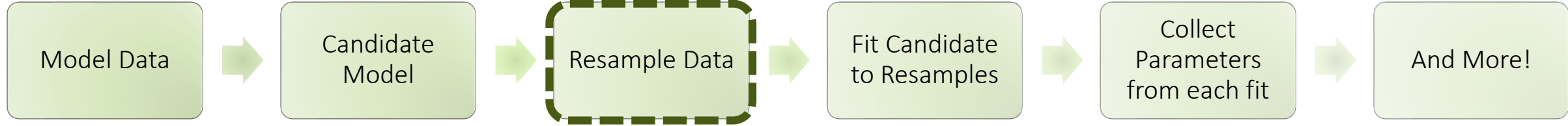
```
55  /* spec out final selected model above with lower level effects included for info
56  ⊖ proc cas;
57      logistic / table = {name='sample'},
58          class = {'Sex','Smoking_Status'},
59          model = {
60              clb=TRUE,
61              target = 'Status'
62              effects = {
63                  {vars={'Sex','Smoking_Status','AgeAtStart','Systolic'}, interaction='CROSS'},
64                  {vars={'AgeAtStart','Sex'}, interaction='CROSS'},
65                  {vars={'Systolic','Sex'}, interaction='CROSS'},
66                  {vars={'AgeAtStart','AgeAtStart'}, interaction='CROSS'},
67                  {vars={'Cholesterol','Smoking_Status'}, interaction='CROSS'}
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='Resid'},
73  run;
```

Results from regression.logistic

Fit Statistics	
-2 Log Likelihood	5209.08335
AIC (smaller is better)	5241.08335
AICC (smaller is better)	5241.19148
SBC (smaller is better)	5345.51448

Parameter Estimates						
Parameter	DF	Estimate	Standard Error	Chi-Square	Pr > ChiSq	95% Confidence Limits
Intercept	1	7.268858	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	0.1882	-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	0.1888	-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	.	.	.	.

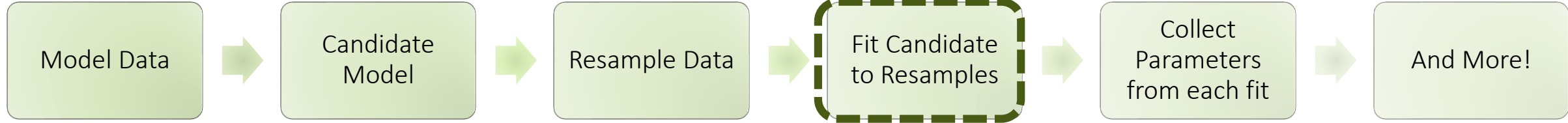
SAS Viya: Using CAS for faster inference



```
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.

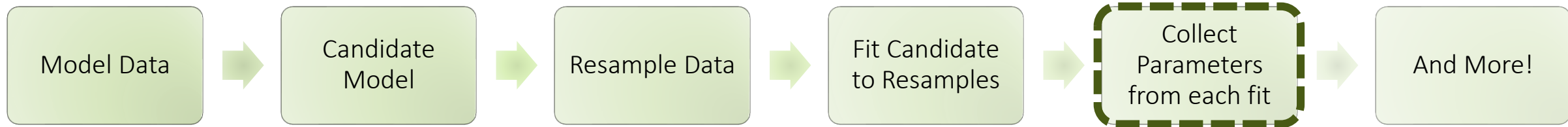
SAS Viya: Using CAS for faster inference



```
123  /* analyze/train each bootstrap resample with the same model effects selected on the full sample data */
124  Ⓣ proc cas;
125      logistic result=myresult / table = {name='sample_bs', groupBy='bsID'},
126      class = {'Sex','Smoking_Status'},
127      model = {
128          clb=TRUE,
129          target = 'Status'
130          effects = {
131              {vars={'Sex','Smoking_Status','AgeAtStart','Systolic','Cholesterol'}},
132              {vars={'AgeAtStart','Sex'}, interaction='CROSS'},
133              {vars={'Systolic','Sex'}, interaction='CROSS'},
134              {vars={'AgeAtStart','AgeAtStart'}, interaction='CROSS'},
135              {vars={'Cholesterol','Smoking_Status'}, interaction='CROSS'}
136          }
137      },
138      partByVar = {name="bag",train="1",test="0"},
139      outputTables = {names={'ParameterEstimates'="sample_BS_PE"}, groupByVarsRaw=TRUE, replace=TRUE},
140      output = {casOut={name='sample_bs_pred', replace=TRUE}, pred='Pred', resid='Resid', copyVars={"caseID","bsID","bag"}};
141  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

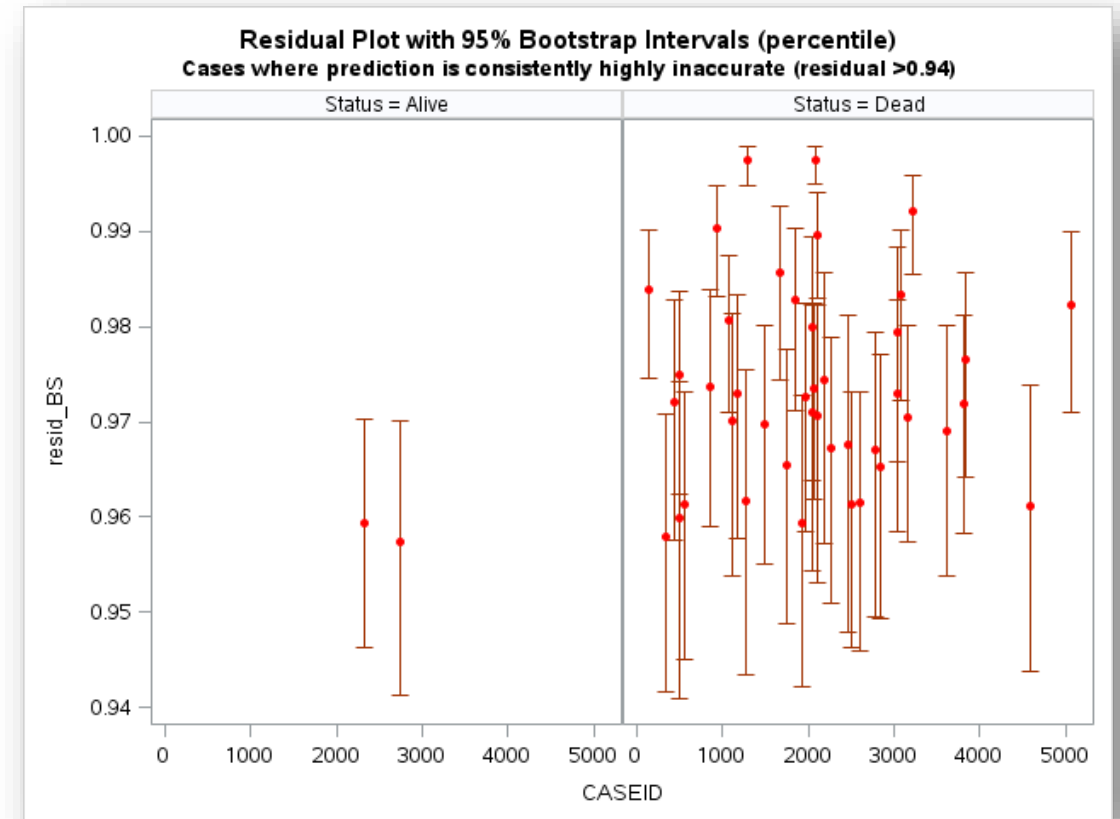
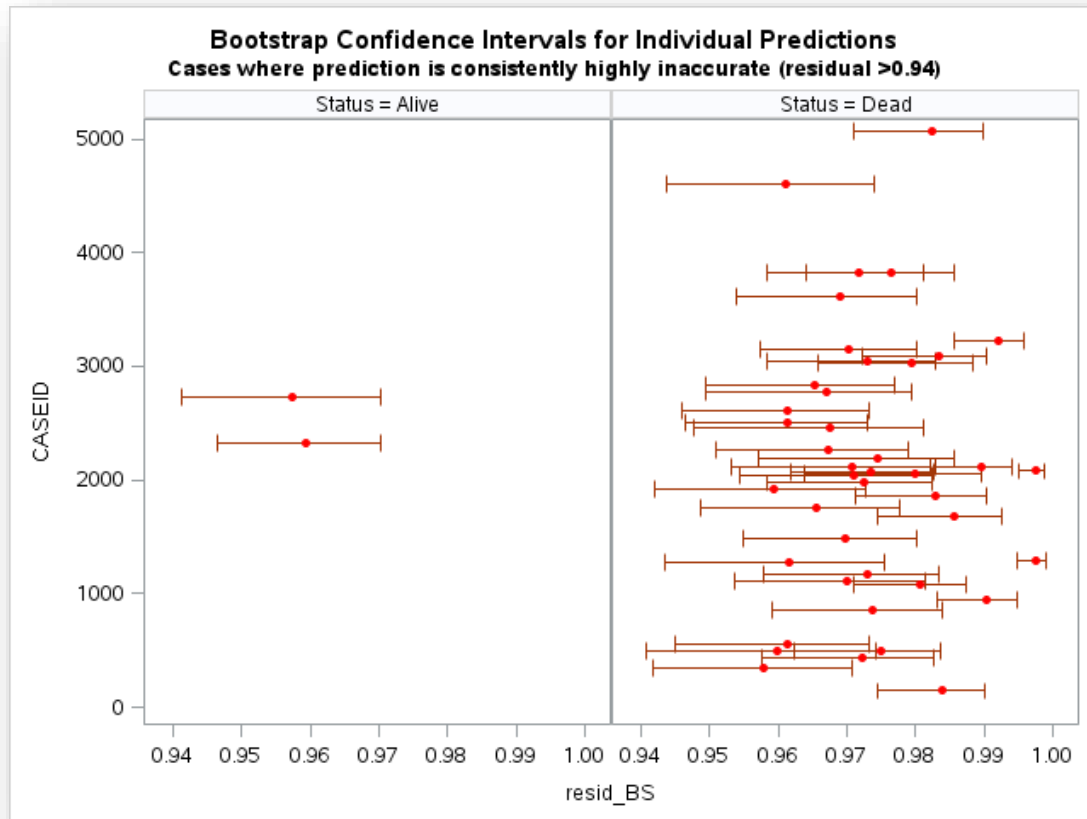
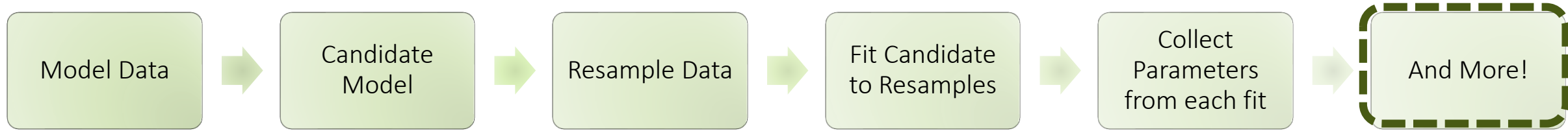
SAS Viya: Using CAS for faster inference

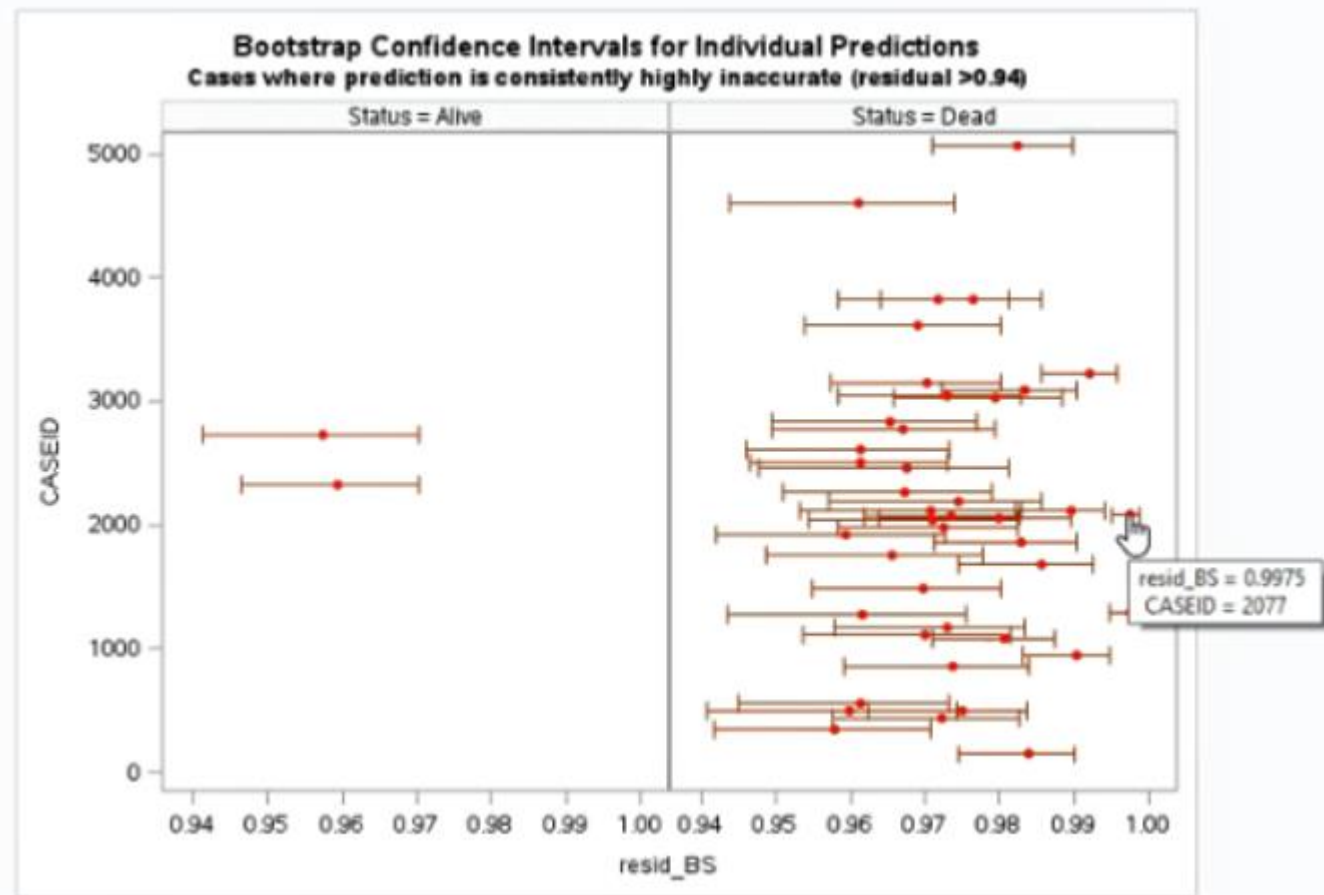
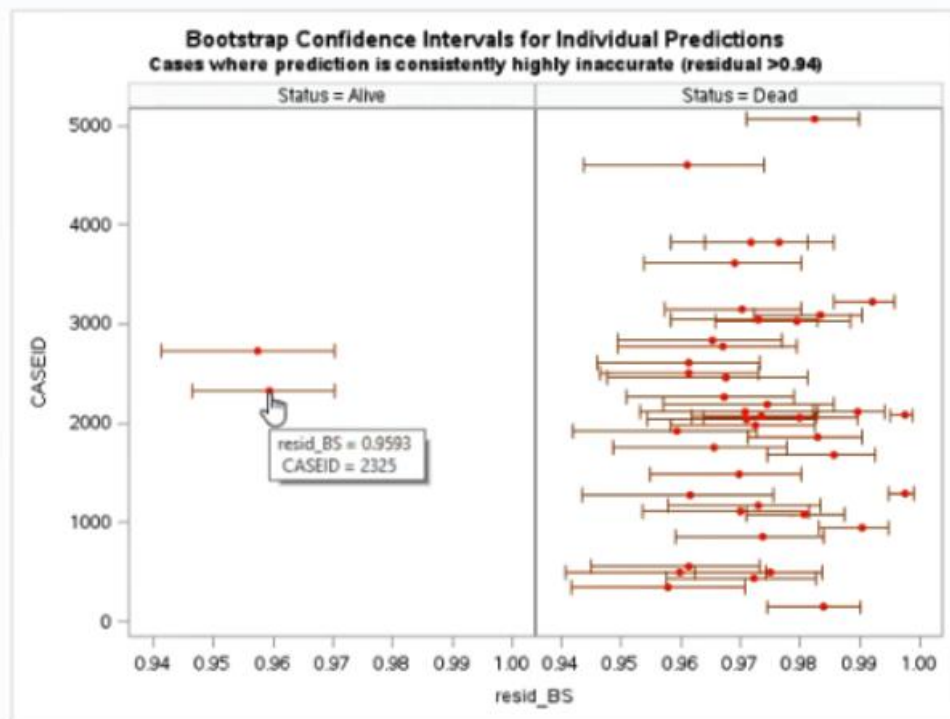


```
164      /* create percentile intervals for the bootstrap samples and merge with full model CI
165         this action expects table intable_BS_PE and intable_PE */
166  @ proc cas;
167      resample.percentilePE / intable='sample' alpha=0.05;
168      run;
169
170      /* plot the parameter effects 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;
```

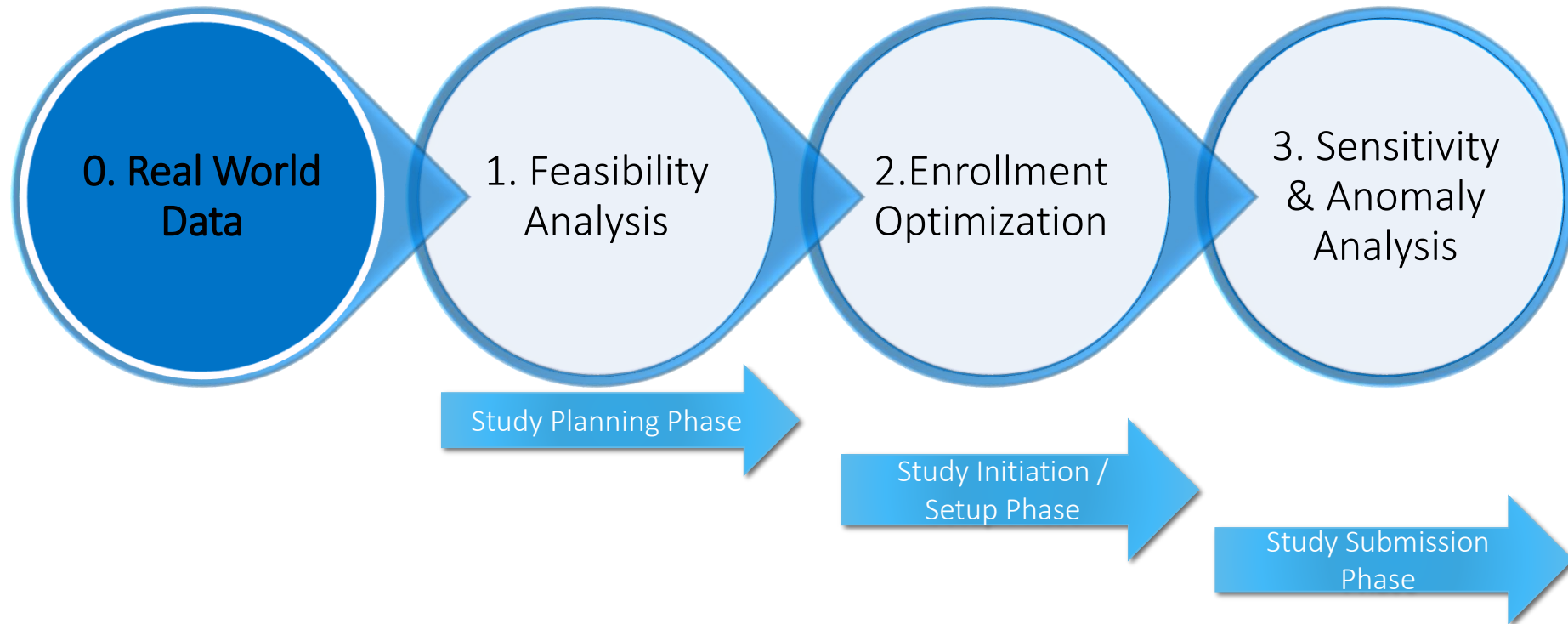
- Another user contributed action brings the Parameter Estimates together from all the resamples and the original full data
- A data step moves the data from CAS to SAS 9.4 so that...
- Use Proc SGPLOT to visualize the Parameter Estimates with both types of confidence intervals
- This fit generalizes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

SAS Viya: Using CAS for faster inference





Scope





Questions ?

SAS Team

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

sas.com